Organized by:

Hellenic Society of Medicinal Chemistry

Cheminformatics and QSAR Society

18th European Symposium on Quantitative Structure – Activity Relationships

"Discovery Informatics & Drug Design"

19-24 September 2010
Rhodes-Greece

An EFMC sponsored event

www.euroqsar2010.gr

ΕΠΕΝΔΥΟΥΜΕ ΣΤΗ

Ο ορισμός της ζωής αποτελεί ακόμη και σήμερα μια πρόκληση για επιστήμονες και φιλοσόφους. Εμείς στην **ELPEN** δεν ορίζουμε τη ζωή, όπως στο λεξικό, σαν το αντίθετο του θανάτου, αλλά σα μια διαρκή διαδικασία εξέλιξης που ανυψώνει τον άνθρωπο σε υπέρτατη αξία.

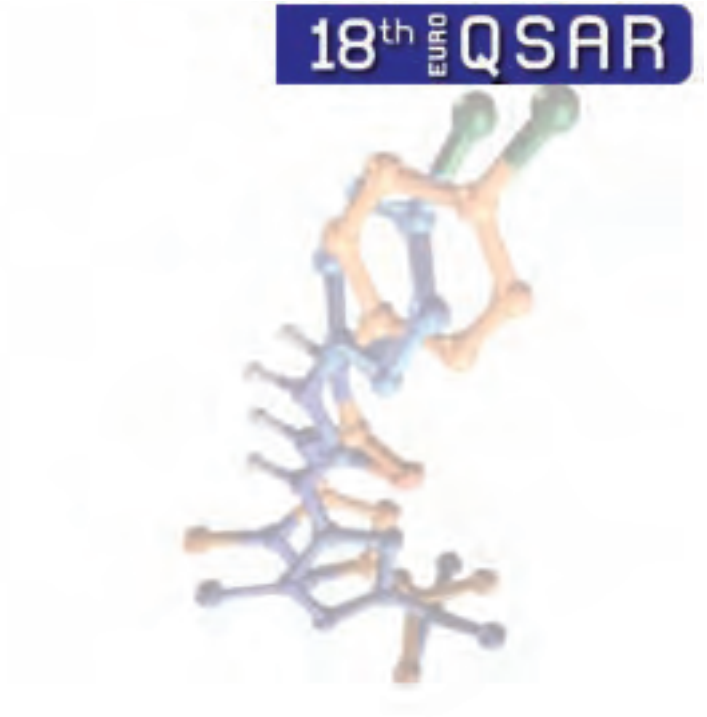
Γι' αυτό εργαζόμαστε διαρκώς στηρίζοντας αυτή τη διαδικασία με τον άριστο τρόπο:

- Σε επιστημονικό επίπεδο **επενδύουμε στην έρευνα και στην τεχνολογία** για την ανάπτυξη καινοτόμων φαρμάκων που βελτιώνουν τη ζωή όλων μας.
- Σε οικονομικό επίπεδο, **δημιουργούμε προστιθέμενη αξία** με την απασχόληση και την οικονομική ευρωστία που μας διακρίνει στην πρώτη θέση της Ελληνικής Φαρμακοβιομηχανίας, στη δύσκολη οικονομική συγκυρία του σήμερα.
- Και σε κοινωνικό επίπεδο, μοιραζόμαστε το **όραμα μιας καλύτερης ζωής** με ευθύνη και αξιοπρέπεια.

Μεγαλώνουμε διαρκώς με γνώμονα την ανθρώπινη αξία.

ELPEN

Φροντίδα για τον άνθρωπο


18th EURO QSAR

Introduction

Dear Participants

The Organizing Committee cordially welcomes you to the **18th European Symposium on Quantitative Structure-Activity Relationships** and to the majestic island of Rhodes. The 18th EuroQSAR continues the tradition of holding biannual meetings in different European countries since 1973. This is the first time that this event is hosted in Greece and we feel honoured to have you as our guests. During the next five days, we will gather to discuss and share ideas on the latest scientific and technological advances in QSAR as tools for the discovery of new, safer and more efficacious drugs. The impact of informatics in all its flavours – chemo-informatics, bio-informatics, pharmaco-informatics – in achieving this goal is duly reflected in the Symposium's title 'Discovery Informatics in Drug Design'.

Chemical space navigation and virtual screening in the face of increased biological complexity and an ever-growing number of relevant targets, counter-targets and ADME/physicochemical properties, constitute major challenges and are among the topics that will be explored extensively during the Symposium. Emphasis is also given on predictive toxicology and risk assessment, while a special session is devoted to agrochemical research. Apart from the lectures, a large number of posters will stimulate discussions on new methodologies and their use on a wide range of application domains.

The Organizing Committee did not spare any effort to set up a Symposium which, we hope, will meet your expectations and leave a lasting imprint on your personal and professional lives. To this point, we would like to express our most sincere gratitude to our sponsors who supported our efforts through their generous financial and scientific contributions. The **18th EuroQSAR Symposium** is now open and we would like to encourage you to participate actively in the scientific discussions and social events, and help create a vibrant, stimulating and congenial atmosphere inside and outside the sessions.

Scientific events – and EuroQSAR's in particular – are places where new friendships are forged, old ones rekindled, innovative ideas are sprung, and collaborations take root. We do hope that you will enjoy the conference as well as the sun and brightness of the island of Rhodes.

With our best wishes for a pleasant stay,

Anna Tsantili-Kakoulidou – Dimitris K. Agrafiotis

Organizing Committee

- **Anna Tsantili-Kakoulidou, Symposium Chair**
  (University of Athens)
- **Dimitris Agrafiotis, Co-Chair**
  (Johnson & Johnson Pharmaceutical Research & Development, USA and University of Athens, Greece)
- **Dimitra Hadjipavlou-Litina**
  (University of Thessaloniki, Greece)
- **Olga Igglessi-Markopoulou**
  (National Technical University of Athens, Greece)
- **Thomas Mavromoustakos**
  (University of Athens, Greece)
- **Emmanuel Mikros**
  (University of Athens, Greece)
- **Tudor Oprea**
  (University of New Mexico, USA)
- **Isidore Rigoutsos**
  (Thomas Jefferson University, USA)
- **Haralambos Sarimveis**
  (National Technical University of Athens, Greece)

Local Supporting Organizing Team

- **A. Afantitis**
- **K. Giaginis**
- **V. Gikas**
- **C. Koukoulitsa**
- **C. Kontogiorgis**
- **G. Melagraki**
- **E. Pontiki**
- **F. Tsopelas**

Scentific Committee

D. Agrafiotis	USA, Greece
M. Akamatsu	Japan
E. Aki	Turkey
M. Botta	Italy
A. Carotti	Italy
G. Cruciani	Italy
G. Ecker	Austria
P. Goodford	U. K.
H.- D. Holtje	Germany
G. Klebe	Germany
H. Kubinyi	Germany
T. Langer	Austria
T. Lundstedt,	Sweden
P. Macheras	Greece
R. Mannhold	Germany
Y. Martin	USA
T. Oprea	USA
V. Palyulin	Russia
I. Rigoutsos	USA
F. Sanz	Spain
C. Selassie	USA
N. de Souza	India
B. Testa	Switzerland
H. Timmerman	The Netherlands
A. Tropsha	USA
A. Tsantili-Kakoulidou	Greece
I. Yalcin	Turkey
H. Van de Waterbeemd	France
D. Winkler	Australia

Program at a glance

Sunday 19/9/2010	Monday 20/9/2010	Tuesday 21/9/2010	Wednesday 22/9/2010	Thursday 23/9/2010	Friday 24/9/2010
18:30-19:00 Opening of the Symposium	8:45-9:30 H. Waldmann	8:45-9:30 P. Kenny	8:45-9:30 S. Brunak	8:45-9:30 F. Sanz	9:00-9:45 E. Martin
19:00-20:00 Inaugural Lecture H. Kubinyi	9:30-9:50 T. Schulz-Gasch	9:30-9:55 Y. Martin	9:30-9:50 C. Laggner	9:30-9:55 T. Langer	9:45-10:05 S. Ajmani
21:00 Welcome cocktail	9:50-10:10 G. Walter	9:55-10:15 F. Burden	9:50-10:10 A. Linusson	9:55-10:15 B. Seebeck	10:05-10:25 V. Polyutin
	10:10-10:30 A. Macchiarulo	10:15-10:35 D. Manallack	10:10-10:30 R. Benigni	10:15-10:35 R. Cramer	10:25-10:45 G. Spyroulias
	10:30-11:00 Coffee break	10:35-11:00 Coffee break	10:30-10:50 Coffee break	10:35-11:00 Coffee break	10:45-11:05 Coffee break
	11:00-11:45 O. Ecker	11:00-11:45 H. van de Waterbeemd	10:50-11:35 P. Gramatica /E. Papa	11:00-11:45 J. Mestres	11:05-11:50 R. Zhou
	11:45-12:05 I. Jabeen	11:45-12:05 M. Lobell	11:35-12.00 A. Tropsha	11:45-12:05 O. Nicolotti	11:50-12:15 D. Agrafiotis
	12:05-12:25 G. Hessler	12:05-12:25 R. Scherer	12:00-12:25 S. Boyer	12:05-12:25 G. van Westen	12:15-12:40 T. Oprea
	12:25-12:45 V. Sharmagoundaran	12:25-12:45 H. Gao	12:25-12:45 J. Devillers	12:25-12:45 E. Ali	12:40-13:15 Closing
	12:45-14:00 Lunch	12:45-14:00 Lunch	12:45-13:05 O. Raevsky/ V. Polotsky	12:45-14:00 Lunch	13:15-14:00 Farewell Party
	14:00-16:00 Poster session I	14:00-16:00 Poster session II	13:05-13:45 Lunch	14:00-14:45 K.J. Schleifer	
	16:00-16:45 A. Nicholls	16:00-16:45 B. Testa	14:30 Excursion to Lindos	14:45-15:05 F. Peruccio	
	16:45-17:05 D. Filimonov	16:45-17:05 G. Vistoli		15:05-15:25 G. Lange	
	17:05-17:30 Coffee break	17:05-17:25 Coffee break		15:25-15:45 R.J. Mathöfer	
	17:30-17:50 F. Buchwald	17:25-18:10 F. Macheras		15:45-16:05 Coffee break	
	17:50-18:10 C. Luscombe	18:10-18:30 A. Kiss		16:05-16:25 I. Kouskoumvekaki	
	18:10-18:30 D. Jianan	20:30 Cultural event		16:25-16:45 B. Wendt	
				16:45 - 17:05 F. Morlaud	
				17:05-17:40 Y. R. Tahir	
				17:40-18:30 CQS Meeting	
				21:00 Symposium Dinner	

General Information

Symposium Secretariat

The Official Symposium Secretariat is Zita Congress & Travel. The Symposium Secretariat is located at a central point of Exhibition Center.
It will operate from Sunday, 19th September 2010 from 08.00 and throughout days and hours of Symposium.

Speakers Preview Room

Speakers Pre View Room will be located close to the secretariat. Make sure that you give your presentation to the PREVIEW ROOM at least two hrs before, to be copied on the Master Computer. For morning sessions presentations you are kindly requested to deliver your presentations from the previous day. Data files must be on CD-Rom or a USB stick.

Posters

Posters will be mantled on Monday morning, September 20th and stay till Friday morning, September 24th.

Internet Corner

Internet Corner with free access to Internet is available ay the designated area of the Exhibition Hall. The Internet Corner will be open during days and hours of Symposium.

Certificates of Attendance

Certificates of Attendance will be issued by the Symposium Secretariat.

The Organizing Committee of the 18th EuroQSAR Symposium acknowledges the following sponsors for their generous contribution to the Symposium's success

Local Authorities

THE PREFECTURE OF DODECANESE
MUNICIPALITY OF RHODES
4th EPHORATE OF BYZANTINE ANTIQUITIES

Gold Sponsors

inte:ligand
Your partner for in-silico drug discovery.

accelrys®

Johnson-Johnson
PHARMACEUTICAL RESEARCH
& DEVELOPMENT

SUNSET MOLECULAR

"Publishers of the WOMBAT databases"
www.sunsetmolecular.com

Tripos
A Certara Company

MEDIT

Silver Sponsors

OpenEye

SCHRÖDINGER

Sponsors

Pfizer

Lilly

Leadscope

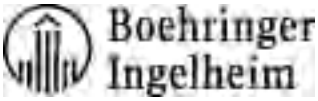
UMETRICS

ELPEN

we care for people

RAFARM

Sponsors



Contributor

FINAL PROGRAM

Sunday 19/9/2010

8:00-18:00 Registration

18:30-19:00 **Opening of the Symposium**

Chair person: Bernard Testa

19:00-20:00 Inaugural Lecture
THE LONG ROAD FROM QSAR TO VIRTUAL SCREENING
Hugo Kubinyi

21:00 **Welcome Cocktail**

Monday 20/9/2010

CHEMICAL SPACE NAVIGATION AND VIRTUAL SCREENING
Chair person: Tudor Oprea

8:45-9:30 Plenary Lecture
CHARTING BIOLOGICALLY RELEVANT CHEMICAL SPACE
Herbert Waldmann
Max Planck Institute, Dortmund, Germany

9:30-9:50 Oral Presentation
QSEARCH: A NEW METHOD FOR DE NOVO LIGAND DESIGN
Tanja Schulz-Gasch
Hoffmann-La Roche, Basel, Switzerland

9:50-10:10 Oral Presentation
LIGANDSCOUT: MORE ACCURACY FOR PHARMACOPHORE-BASED VIRTUAL SCREENING
Gerhard Wolber
InteLigand, Innsbruck, Austria

10:10-10:30 Oral Presentation
RECOVERING DESIGN STRATEGIES OF GPCRS MODULATORS FROM EXPLORATIONS OF THE CHEMICAL SPACE
Antonio Macchiarulo
University of Perugia, Italy

10:30-11:00 ***Coffee Break***

TARGETS-TRANSPORTERS-ANTITARGETS

Chair persons: Angelo Carotti - Emmanuel Mikros

11:00-11:45 Plenary Lecture
LIGAND- AND STRUCTURE-BASED APPROACHES FOR TARGETING DRUG TRANSPORTER
Gerhard Ecker
University of Vienna, Austria

11:45-12:05 Oral Presentation
STEREOSELECTIVE INTERACTION OF BENZOPYRANO[3,4-b][1,4]OXAZINES WITH P-GLYCOPROTEIN
Ishrat Jabeen
University of Vienna, Austria

12:05-12:25 Oral Presentation
IDENTIFICATION AND APPLICATION OF ANTITARGET ACTIVITY HOTSPOTS TO GUIDE COMPOUND OPTIMIZATION
Gerhard Hessler
Sanofi-Aventis, Frankfurt, Germany

12:25-12:45 Oral Presentation
BIOPHYSICS-BASED LIBRARY DESIGN: DISCOVERY OF 'NON-ACID' INHIBITORS OF S1 DHFR
Veerabahu Shanmugasundaram
Pfizer, USA

12:45-14:00 ***Lunch***

14:00-16:00 ***POSTER SESSION I***

CHEMINFORMATICS IN DRUG DESIGN
Chair person: Dimitris Agrafiotis

16:00-16:45 Plenary Lecture
INFORMATION THEORY AND QSAR
Anthony Nicholls
OpenEye Scientific Software, USA

16:45-17:05 Oral Presentation
LOCAL CORRESPONDENCE CONCEPT IN BIO- AND CHEMINFORMATICS
Dimitri Filimonov
Russian Academy of Medical Sciences, Moskow, Russia

17:05-17:30 ***Coffee Break***

17:30-17:50 Oral Presentation
USING LOCAL MODELS TO IMPROVE QSAR PREDICTIVITY
Fabian Buchwald
Technische Universität München, Germany

17:50-18:10 Oral Presentation
THE USE OF DESIGN OF EXPERIMENTS TO DEVELOP EFFICIENT ARRAYS FOR SAR AND PROPERTY EXPLORATION
Chris Luscombe
GlaxoSmithKline Research Medicines Centre, U.K.

18:10-18:30 Oral Presentation

ANALYSIS AND COMPARISON OF 2D FINGERPRINTS: INSIGHTS INTO DATABASE SCREENING PERFORMANCE USING EIGHT FINGERPRINT METHODS

Duan Jianxin

Schrödinger GmbH, Germany

Tuesday 21/9/2010

MOLECULAR DESCRIPTORS IN QSAR

Chair persons: Cynthia Selassie- Dimitra Hadjipavlou- Litina

8:45-9:30 Plenary Lecture

HYDROGEN BONDING AND MOLECULAR DESIGN

Peter Kenny

9:30-9:55 Key Note Lecture

TAUTOMERISM, THE FORGOTTEN MOLECULAR DESCRIPTOR

Yvonne Martin

Martin Consultant, USA

9:55-10:15 Oral Presentation

ROBUST SPARSE FEATURE/DESCRIPTOR SELECTION FOR QSAR

Frank Burden

CSIRO Molecular and Health Technologies, Australia

10:15-10:35 Oral Presentation

THE *p*Ka DISTRIBUTION OF SCREENING COMPOUNDS - APPLICATION TO DRUG DISCOVERY

David Manallack

Monash University, Australia

10:35-11:00 ***Coffee Break***

IN SILICO PHYSCHEM PROFILING AND ADMET

Chair persons: Raimund Mannhold – Panos Macheras

11:00-11:45 Plenary Lecture

REAL – TIME IN SILICO PHYSCHEM AND ADMET SUPPORT USING AUTOQSAR

Han van de Waterbeemd

11:45-12:05 Oral Presentation
UNDERSTANDING THE BLOOD BRAIN BARRIER: OPTIMIZATION STRATEGIES FOR CNS PENETRATION AND DISTRIBUTION
Mario Lobell
Bayer Schering Pharma, Wuppertal, Germany

12:05-12:25 Oral Presentation
MULTI-*p*H QSAR: REGRESSION ANALYSIS SENSITIVE ENOUGH TO DETERMINE THE TRANSITION-STATE pKa OF HUMAN BUCCAL ABSORPTION
Robert Scherrer
BIOpKa, USA

12:25-12:45 Oral Presentation
MULTI-PARAMETER OPTIMIZATION AND IN SILICO MODELING IN LEAD OPTIMIZATION
Hua Gao
Pfizer, USA

12:45-14:00 ***Lunch***

14:00-16:00 ***POSTER SESSION II***

ASSESSING DRUG SAFETY AND EFFICACY THROUGH ADME PREDICTIONS
Chair person: Han van de Waterbeemd

16:00-16:45 Plenary Lecture
THE BIOCHEMISTRY OF DRUG METABOLISM – WHICH ARE THE IMPORTANT REACTIONS AND ENZYMES?
Bernard Testa
University Hospital Centre, Lausanne, Switzerland

16:45-17:05 Oral Presentation
COMBINED IN SILICO APPROACHES FOR DRUG DESIGN AND PHARMACOKINETIC OPTIMIZATION OF A SET OF CARNOSINE ANALOGUES AS POTENT AND SELECTIVE CARBONYL QUENCHERS
Giulio Vistoli
University of Milan, Italy

17:05-17:25 ***Coffee Break***

17:25- 18:10 Plenary Lecture
COMPUTATIONAL-REGULATORY DEVELOPMENTS IN THE PREDICTION OF ORAL DRUG ABSORPTION
Panos Macheras
University of Athens, Greece

18:10-18:30 Oral Presentation
IS PREDICTING ACTIVE TRANSPORT NECESSARY TO PREDICT BIOAVAILABILITY?
Albin Kristl
University of Ljubljana, Slovenia

20:30 ***Cultural event***

Wednesday 22/9/2010

QSAR IN THE ERA OF BIOLOGICAL COMPLEXITY
Chair persons: Eric Martin - Esin Aki

8:45-9:30 Plenary Lecture
DISEASE SYSTEMS CHEMICAL BIOLOGY AND TOXICOGENOMICS
Søren Brunak
Technical University of Denmark, Denmark

9:30-9:50 Oral Presentation
TARGET IDENTIFICATION FOR BEHAVIORAL SCREENING HITS USING A CHEMICAL SIMILARITY METHOD
Christian Laggner
University of California, USA

9:50-10:10 Oral Presentation
CHARACTERIZATION AND MAPPING OF LIGAND-BINDING CAVITIES IN PROTEINS
Anna Linusson
Umeå University, Sweden

10:10-10:30 Oral Presentation
IN SILICO APPROACHES AND IN VITRO AND IN VIVO MUTAGENICITY ASSAYS: ALTERNATIVES TO THE CARCINOGENICITY BIOASSAY
Romualdo Benigni
Istituto Superiore di Sanità, Italy

10:30-10:50 ***Coffee Break***

PREDICTIVE TOXICOLOGY AND RISK ASSESSMENT
Chair persons: Vladimir Palyulin - Haralambos Sarimveis

10:50-11:35 Plenary Lecture
RODENT TOXICITY STUDIES ON PERFLUORINATED CHEMICALS FOR REACH
Paola Gramatica/ Ester Papa
University of Insubria, Italy

11:35-12.00 Key Note Lecture
NOVEL APPROACHES TO CHEMICAL TOXICITY PREDICTION RELYING ON THE ENTIRE STRUCTURE-IN VITRO-IN VIVO DATA CONTINUUM
Alex Tropsha
University of North Carolina, USA

12:00-12:25 Key Note Lecture
ASSESSING REACTIVE METABOLITE RISK IN DRUG DISCOVERY USING A WEIGHT-OF-EVIDENCE APPROACH
Scott Boyer
AstraZeneca, Sweden

12:25-12:45 Oral presentation
EVALUATION OF THE OECD QSAR APPLICATION TOOLBOX FOR PREDICTING THE BIODEGRADABILITY OF CHEMICALS
James Devillers
Centre de Traitement de l'Information Scientifique, France

12:45-13:05 Oral presentation
CLASSIFICATION AND REGRESSION-BASED QSAR OF ACUTE CHEMICAL RODENT TOXICITY
Oleg Raevsky / Vladimir Poroikov
Russian Academy of Sciences, Moskow, Russia

13:05-13:45 ***Lunch***

14:30 ***Excursion***

Thursday 23/9/2010

PHARMACOINFORMATICS AND PHARMACOPHORES
Chair person: Gerhard Ecker

8:45-9:30 Plenary Lecture
INTEGRATIVE PHAMACOINFORMATICS APPROACHES IN THE PREDICTION OF CLINICAL OUTCOMES OF DRUGS
Ferran Sanz
IMIM - Universitat Pompeu Fabra, Barcelona, Spain

9:30-9:55 Key Note Lecture
PHARMACOPHORES - VERSATILE TOOLS TO BRIDGE THE GAP BETWEEN STRUCTURE-BASED AND LIGAND BASED APPROACHES
Thierry Langer
Prestwick Chemical, Inc., France

9:55-10:15 Oral Presentation
FROM ACTIVITY CLIFFS TO TARGET-SPECIFIC SCORING MODELS AND PHARMACOPHORIC HYPOTHESIS
Birte Seebeck
University of Hamburg, Germany

10:15-10:35 Oral Presentation
TEMPLATE-CONSTRAINED FRAGMENT ALIGNMENT (TCFA)
Richard Cramer
Tripos Int., USA

10:35-11:00 ***Coffee Break***

MULTI-TARGET / MULTI- OBJECTIVE QSAR
Chair person: Ferran Sanz

11:00-11:45 Plenary Lecture
LIGAND-BASED APPROACHES TO IN SILICO PHARMACOLOGY: BENCHMARKS AND APPLICATIONS
Jordi Mestres
IMIM - Universitat Pompeu Fabra, Barcelona, Spain

11:45-12:05 Oral Presentation
ENHANCING MOLECULAR DESIGN VIA A MULTI-OBJECTIVE APPROACH
Orazio Nicolotti
Università di Bari, Italy

12:05-12:25 Oral Presentation
PROSPECTIVELY VALIDATED PROTEOCHEMOMETRIC MODELS OF HIV REVERSE TRANSCRIPTASE AS A TOOL IN LEAD OPTIMIZATION AGAINST MULTIPLE TARGETS
Gerard van Westen
LACDR, Leiden, The Netherlands

12:25-12:45 Oral Presentation
COMPUTATIONAL DRUG DESIGN STUDIES ON ANTITUMORAL ACTIVE HETEROCYCLIC COMPOUNDS
Esin Aki
Ankara University, Turkey

12:45-14:00 ***Lunch***

COMPUTATIONAL STRATEGIES IN AGROCHEMICAL RESEARCH
Chair persons: Ismail Yalcin - Klaus-Jürgen Schleifer

14:00-14:45 Plenary Lecture
CHALLENGES IN AGROCHEMICALS DESIGN
Klaus-Jürgen Schleifer
BASF, Germany

14:45-15:05 Oral Presentation
NEW LEADS FINDING IN AGROCHEMISTRY: A COMPUTATIONAL CHEMISTRY CHALLENGE
Francesca Perruccio
Syngenta Crop Protection, Switzerland

15:05-15:25 Oral Presentation
HYDE SCORING OF PROTEIN LIGAND COMPLEXES
Gudrun Lange
Bayercropscience, Frankfurt, Germany

15:25-15:45 Oral Presentation
INHIBITION OF *EIMERIA TENELLA* CDK-RELATED KINASE 2: FROM TARGET IDENTIFICATION TO LEAD COMPOUNDS
Richard J. Marhöfer
Intervet Innovation GmbH, Schwabenheim, Germany

15:45-16:05 ***Coffee Break***

DATABASE MINING
Chair persons: Alex Tropsha - Thomas Mavromoustakos

16:05-16:25 Oral Presentation
BACK TO THE ROOTS – BENEFITS AND LIMITATIONS CONCERNING THE IN SILICO INTEGRATION OF NATURAL PRODUCTS IN DRUG DISCOVERY
Irene Kouskoumvekaki
Technical University of Denmark, Denmark

16:25-16:45 Oral Presentation
CAPTURING SAR-TRENDS FROM CHEMOGENOMICAL SPACES
Bernd Wendt
Elara Pharmaceuticals GmbH, Heidelberg, Germany

16:45-17:05 Oral Presentation
MINING EXHAUSTIVELY THE PROTEIN DATA BANK ENABLES COMPUTATIONAL FRAGMENT-BASED DRUG DESIGN
Fabrice Moriaud
Medit SA, France

17:05-17:40 Short oral presentation of selected posters
FOUR PRESENTATIONS (TWO FROM POSTER SESSION I & TWO FROM POSTER SESSION II), 7 MIN EACH

17:40-18:30 **MEETING OF CHEMINFORMATICS AND QSAR SOCIETY**
Chair: Tudor Oprea

21:00 ***Symposium Dinner***

Friday 24/9/2010

NEW TOOLS AND APPLICATIONS
Chair persons: Jordi Mestres - Vladimir Poroikov

9:00-9:45 Plenary Lecture
ITERATIVE KINASE MEDIUM-THROUGHPUT SCREENING (IKMTS) WITH 2D PROFILE-QSAR AND 3D SURROGATE AUTOSHIM ENSEMBLE DOCKING
Eric Martin
Novartis, USA

9:45-10:05 Oral Presentation
INVESTIGATION OF THE STRUCTURAL REQUIREMENTS FOR MULTI-KINASE INHIBITION USING GQSAR METHOD
Subhash Ajmani
NovaLead Pharma, India

10:05-10:25 Oral Presentation
UNDERSTANDING THE SELECTIVITY OF ORGANOPHOSPHORUS INHIBITORS OF SERINE ESTERASES
Vladimir Palyulin
Moscow State University, Russia

10:25-10:45 Oral Presentation
STRUCTURE-ACTIVITY RELATIONSHIP OF ARKADIA RING FINGER E3 UBIQUITIN LIGASE THROUGH NMR SPECTROSCOPY
George Spyroulias
University of Patras, Greece

10:45-11:05 ***Coffee Break***

NEW AVENUES IN QSAR
Chair persons: Thierry Langer - Yvonne Martin

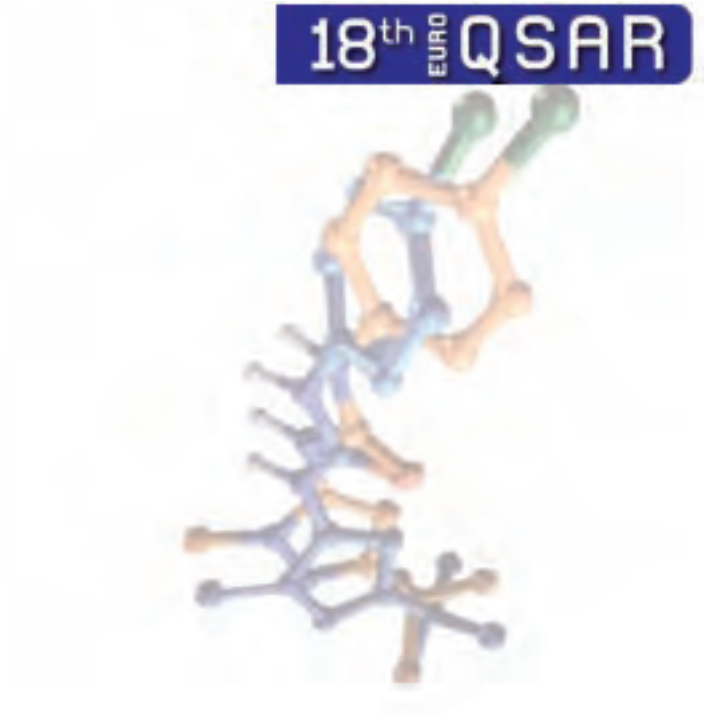
11:05-11:50 Plenary Lecture
LIGAND-RECEPTOR BINDING AFFINITY PREDICTIONS WITH LINEAR RESPONSE METHODS AND FREE ENERGY PERTURBATION CALCULATION
Ruhong Zhou
IBM, Thomas J. Watson Research Center, USA

11:50-12:15 Key Note Lecture
VISUAL (Q)SAR: SAR MAPS, SCAFFOLD TREES AND R-CLIFFS
Dimitris Agrafiotis
Johnson & Johnson, USA

12:15-12:40 Key Note Lecture
COMPUTER-AIDED DRUG REPURPOSING
Tudor Oprea
University of New Mexico, USA

12:40-13:15 ***Closing of the Symposium***

13:15-14:00 ***Farewell party***


18th EUROQSAR

ABSTRACT BOOK


18th EURO QSAR

INAUGURAL LECTURE

INAUGURAL LECTURE

THE LONG ROAD FROM QSAR TO VIRTUAL SCREENING

Hugo Kubinyi
BASF SE and University of Heidelberg, Germany (retired)

Chemoinformatics and molecular modeling have reached a high level: QSAR, pharmacophore searches, docking and scoring, and ADMET predictions aim to contribute to the discovery of new drugs. However, there is still a gap between the achieved results and the medicinal chemists' needs in lead discovery and optimization. In Hermann Hesse's virtual land Castalia, *intellectual efforts have no purpose other than the preservation and advancement of intellectual foundations of culture and humanity* ... [they] *engage in an intellectual exercise, which aims at connecting scientific and cultural values within a formal framework of mathematics and music* ("The Glass Bead Game"). Does this also describe the current situation in QSAR, molecular modeling and virtual screening? QSAR has the problem of insufficient external predictivity, resulting from inappropriate model selection and validation; modeling often does not adequately consider molecular properties and their influence on pharmacophore definitions; virtual screening and docking suffer from the lack of reliable scoring functions; ADMET predictions often fail because of the complexity of absorption and metabolism, especially the formation of toxic metabolites. Future computational chemistry research should focus on solutions to these problems, instead of keeping to the beaten tracks.

PLENARY LECTURES

CHARTING BIOLOGICALLY RELEVANT CHEMICAL SPACE

Herbert Waldmann

Max-Planck-Institut für molekulare Physiologie, Department of Chemical Biology, Otto-Hahn-Str. 11, D-44227 Dortmund, and TU Dortmund, Fachbereich 3, Chemische Biologie, Germany

Relevance to nature is one of the most important criteria to be met by compound classes for chemical biology and medicinal chemistry research. The underlying frameworks of natural products (NPs) provide evolutionary selected chemical structures encoding the properties required for binding to proteins, and their structural scaffolds represent the biologically relevant and prevalidated fractions of chemical space explored by nature so far. In order to make compounds available that encode biological relevance, efficient methods are in demand that allow to chart and navigate biologically relevant chemical space and to populate it with appropriate comnpound collections

Biology oriented synthesis (BIOS) builds on these arguments. It employs core structures delineated from NPs as scaffolds of compound collections and creates focussed diversity around a biologically prevalidated starting point in vast structural space. BIOS, therefore, builds on the diversity created by nature in evolution and aims at its local extension in areas of proven biological relevance. Consequently BIOS offers a conceptual alternative to other guiding strategies for library design which for instance are based on mechanistic considerations, sequence or structure homology or on the creation of chemical diversity.

In the lecture the trains of thought leading to the BIOS concept will be detailed, including the development of a logic to identify, map and navigate biologically relevant chemical space with the ultimate goal to allow for prospective assignment of bioactivity.

PLENARY LECTURE 2

LIGAND- AND STRUCTURE-BASED APPROACHES FOR TARGETING DRUG TRANSPORTERS

Gerhard F. Ecker
University of Vienna, Department of Medicinal Chemistry, Althanstraße 14, 1090 Wien, Austria

More than 400 membrane transporters, organized in two superfamilies, have been annotated in the human genome. Many of them are involved in drug transport, thus contributing to pharmacokinetic, safety and efficacy profiles of drugs and drug candidates. P-glycoprotein (P-gp), which has been discovered more than 30 years ago and is considered one of the paradigm transporters, has been extensively studied to shed more light on the molecular basis of drug transport across membranes. One of the striking features of P-gp is its broad substrate specificity, often denoted as polyspecificity or promiscuity. This property, together with the lack of structural information, rendered ligand design quite challenging. Within this talk, our lead optimization attempts in the class of propafenone analogues will be described. Applying classical QSAR studies combined with artificial neural networks led to inhibitors active in the low nanomolar range. New scaffolds were identified by self-organising maps and pharmacophore modeling. Last but not least structure-based studies using a homology model based on the recently published mouse P-gp structure gave first insights into the molecular basis of drug/transporter interaction.

PLENARY LECTURE 3

INFORMATION THEORY AND QSAR

Anthony Nicholls
OpenEye Scientific Software, USA

QSAR, invented by Albert Einstein (Ref.1) in 1901, is perhaps the oldest molecular modeling technique. Over the course of the last 109 years it has grown greatly in the diversity of approaches, the variety and quality of molecular properties and the scope of application, but always as a method to predict something beyond the feasibility of precise calculation. Strangely, though, given the considerable industry, there are still issues in assessing the reliability and extensibility of a model. The procedures that are employed, for instance cross-validation, y-scrambling and even prospective prediction, are all seriously flawed and for much the same reason- molecules are not "i.i.d", i.e. "independent, identically distributed", a prerequisite for most statistical techniques. But this very problem is also a solution. The real measures of a model should not involve treating molecules as random entities, rather approaches that measure the information content relative to background models that incorporate the nature of molecular similarity. Viewed at a distance, all this is saying is that we should judge a model based on what it tells us that we did not already know. The field of scoring functions for molecular docking offers useful analogies here. Many such functions are only successful to the extent binding affinity correlates with simpler characteristics, such as molecular weight or buried area, i.e. They are not truly predictive in their own right. This view of model assessment not only makes clear the problems inherent in traditional techniques, but suggest how information theory might be properly applied to both assessing models and constructing them in the first place. Such methods address over-parameterization, model-selection bias, incorporation of prior knowledge, domain applicability and prediction confidence. Examples will be given and our research program based on these ideas described.

Ref. 1: Annalen der Physik 4 (1901) pg. 513ff

PLENARY LECTURE 4

HYDROGEN BONDING AND MOLECULAR DESIGN

Peter W. Kenny

I will first demonstrate that molecular electrostatic potential (MEP), when calculated appropriately, may be used to predict both hydrogen bond acidity and basicity. Calculated MEP values provide insight into molecular recognition and examples of DNA base bioisostere evaluation will be used to illustrate use of these calculations in a molecular design context. I'll also show how these calculations can be used to quantify effects of complexation (cooperativity) on hydrogen bonding. Hydrogen bonding is also relevant to lipophilicity. In particular, the difference, ΔlogP, between octanol/water and alkane/water logP values reflects strength of hydrogen bonding and contributions of hydrogen bond acceptors to this property can be predicted from MEP. Some of this work has been published recently (Kenny, JCIM 2009, 49, 1234-1244; Toulmin, Wood & Kenny, JMC 2008, 51, 3720-3730).

PLENARY LECTURE 5

REAL-TIME IN SILICO PHYSCHEM AND ADMET SUPPORT USING AUTOQSAR

Han van de Waterbeemd
Small Molecule Drug Discovery Consultant, France

Most critical for successful drugs are its physicochemical and ADMET properties and these are therefore among the most important challenges for the medicinal chemist [1-3]. Much information is now available from medium- to high-throughput physchem and ADMET in vitro assays, either in the public domain (e.g. PubChem) or in in-house databases. Such data are increasingly being modelled and used in predictive chemistry [4].

In the past QSAR modelling was a slow process and often of limited use to medicinal chemistry projects. New pipelining technology now makes it easier to build and update QSAR models fully automatically so that such models can use the latest available data to produce robust local and global predictive in silico physchem and ADMET models [5-7]. Drug discovery teams have therefore now access to the best and latest models which can be updated daily, weekly, or monthly. Automation also ensures that all data is being used for all projects. Thus, academic QSAR modelling has evolved to industrial scale and is going through a true revival.

Using time-series simulations of real data we have demonstrated that regular updating and using so-called correction libraries make the models over time increasingly predictive by minimising the predictive RMSE [6-8]. This technology has been called autoQSAR [9] and offers real-time support to small molecule drug discovery.

References

[1] H. van de Waterbeemd, B. Testa (eds.), Drug Bioavailability, 2nd Ed. Wiley-VCH, Weinheim, Germany, 2009.
[2] H. van de Waterbeemd, Physicochemical properties in drug profiling. In: Molecular Drug Properties, R. Mannhold (ed.), Wiley-VCH, Weinheim, Germany, 2008, pp. 25-52.
[3] B. Testa, H. van de Waterbeemd (eds.), ADME-Tox Approaches, Vol. 5 of Comprehensive Medicinal Chemistry, 2nd Ed., J.B. Taylor, D.J. Triggle (series eds.), Elsevier, 2007.
[4] H. van de Waterbeemd and E. Gifford, ADMET in silico modelling: Towards prediction paradise? Nature Rev.Drug Discov. 2003, 2, 192-204.
[5] J. Cartmell, S. Enoch, D. Krstajic, D.E. Leahy, J.Comput.Aided Mol.Des. 2005, 19, 821-833.
[6] S. L. Rodgers, A.M. Davis, H. van de Waterbeemd, Time-series QSAR analysis of human plasma protein binding data, QSAR Comb.Sci. 2007, 4, 511-521.
[7] S.L. Rodgers, A.M. Davis, N.P. Tomkinson, H. van de Waterbeemd, QSAR modelling using automatically updating correction libraries: Application to a human plasma protein binding model, J.Chem.Inf.Model. 2007, 47, 2401-2407.
[8] S.L. Rodgers, A.M. Davis, N.P. Tomkinson, H. van de Waterbeemd, An analysis of the predictivity of ADME QSAR models over time, in preparation.
[9] H. van de Waterbeemd, Improving compound quality through in vitro and in silico physicochemical profiling, Chem.Biodiv. 2009, 6, 1760-1766.

PLENARY LECTURE 6

THE BIOCHEMISTRY OF DRUG METABOLISM – WHICH ARE THE IMPORTANT REACTIONS AND ENZYMES?

Bernard Testa[1], Alessandro Pedretti[2] and Giulio Vistoli[2]

[1]Dept of Pharmacy, University Hospital Centre, Rue du Bugnon, CH-1011 Lausanne, Switzerland

[2]Dipartimento di Scienze Farmaceutiche "Pietro Pratesi", Facoltà di Farmacia, Università degli Studi di Milano, Via Mangiagalli, I-20133 Milano, Italy

This lecture is not about QSAR, not even about SAR. It is about one component of the biological context these technologies explore, namely drug metabolism. In concrete terms, our objective here is to offer an overview of the relative significance of reactions and enzymes in the biotransformation of medicinal compounds and other xenobiotics.

A meta-analysis of the literature was carried out, the first conclusions of which form the main body of the lecture. A total of 797 papers were selected from three primary journals, Chemical Research in Toxicology (25%), Drug Metabolism and Disposition (51%), and Xenobiotica (24%) according to preset, objective criteria. These papers reported experimental metabolic data on 1017 different substrates, yielding a total of 5993 different metabolites, 180 of which were reported to be pharmacologically active, and 472 to be reactive or potentially toxic.

Each metabolite was classified according to two criteria, namely the metabolic reaction and the enzyme(s) that generated it [1, 2]. Reactions considered included, e.g., Csp3-oxidations, Csp2- and Csp-oxidations, alcohol-carbonyl redox reactions, tertiary amine oxidations, N-oxide reductions, S atom oxidations, ester hydrolyses, amide hydrolyses, epoxide hydration, glucuronidations, sulfonations, N-acetylations, and reactions involving glutathione or Coenzyme A. Enzymes considered included, e.g., CYPs, FMOs, dehydrogenases, hydrolases, UGTs, SULTs, NATs, MTs, and GSTs. The data were compiled in an ad hoc application based on the ACCESS software. The results confirm the primary role of CYP-catalyzed oxidations and glucuronidations in drug and xenobiotic metabolism, but they also show the non-negligible significance of several other reactions and enzymes.

1. B. Testa & S.D. Krämer. The Biochemistry of Drug Metabolism, Vol. 1 - Principles, Redox Reactions, Hydrolyses. Wiley-VCH, Weinheim, Germany, 2008, 319 pages.

2. B. Testa & S.D. Krämer. The Biochemistry of Drug Metabolism, Vol. 2 - Conjugations, Consequences of Metabolism, Influencing Factors. Wiley-VCH, Weinheim, Germany, 2010, 588 pages.

PLENARY LECTURE 7

COMPUTATIONAL-REGULATORY DEVELOPMENTS IN THE PREDICTION OF ORAL DRUG ABSORPTION

Panos Macheras
Laboratory of Biopharmaceutics and Pharmacokinetics
Faculty of Pharmacy, University of Athens, Athens, Greece

Early prediction of human intestinal absorption is important in the selection of potential orally administered drugs. Computational models for prediction of the fraction of dose absorbed in humans, F, which describes the extent to which an active drug substance is absorbed and becomes available to the general circulation has been developed in a plethora of studies. As early as 1989, the absorption potential (1) of a drug was found to have a sigmoidal relationship with F. Since then various studies have shown that physicochemical descriptors of molecules such as lipophilicity, polar surface area and hydrogen bond descriptors correlate with human intestinal absorption. In parallel, various attempts in estimating F have been reported in the literature e.g. multiparameter equations, mechanistic and QSAR/QSPR models, genetic programming, artificial neural networks and machine learning classification. All studies, except those based on statistical analysis, rely on the basic presupposition that the extent of intestinal absorption, F is mainly dependent on the solubility of drug, which drives the dissolution rate in the gastrointestinal (GI) fluids, and the rate of passive drug transport across the intestinal membrane.
In the same vein, the biopharmaceutics classification system (BCS) (2) and the relevant FDA guideline (3) rely on the same presupposition and classify drugs in four categories according to their aqueous solubility and permeability. However, the recently developed biopharmaceutics drug disposition classification system (BDDCS) (4) revealed not only the poor predictability of permeability estimates for the extent of absorption but also the major role of transporters for the GI uptake of drugs. These findings, if coupled with the role of solubility in the reaction limited model of dissolution (5) and the ubiquitous presence of supersaturated solubility-dissolution phenomena in the GI lumen, call for a more physiologically relevant, dynamic-reactive consideration of GI absorption. Aspects of this approach will be presented in this talk.

References

1. Macheras PE, Symillides MY. 1989. Toward a quantitative approach for the prediction of the fraction of dose absorbed using the absorption potential concept. Biopharm Drug Dispos. 10, 43-53.
2. G. L. Amidon, H. Lennernäs, V. P. Shah, and J. R. Crison. 1995. A theoretical basis for a biopharmaceutic drug classification: The correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm. Res. 12, 413–420.
3. Guidance for Industry. Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. FDA/CDER. Washington D.C. August (2000).
4. Wu CY, Benet LZ. 2005. Predicting drug disposition via application of BCS: transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. Pharm Res. 22, 11-23.
5. Dokoumetzidis, A., Papadopoulou, V., Valsami, G., Macheras, P., 2008. Development of a reaction-limited model of dissolution: Application to official dissolution tests experiments. Int. J. Pharm., 355, 114-125

PLENARY LECTURE 8

DISEASE SYSTEMS CHEMICAL BIOLOGY AND TOXICOGENOMICS

Søren Brunak
Center for Biological Sequene Analysis, Technical University of Denmark, Lyngby, Denmark and The Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

In the emerging area of systems chemical biology and toxicogenomics there is an increasing need for developing network-based approaches in order to understand the relationship between chemical action and many genes, either disease susceptibility genes or other genes which may related to side effects.
Protein-protein interaction (PPI) networks is one approach which can be used to study the systemic properties and identify additional genes that may play major roles in modulating chemical response i.e. to drugs, environmental chemicals and xenobiotics in general. Often such networks are integrated with other types of data, typically from the molecular level, but phenotypic data stemming from text mining of patient records can also improve our knowledge of disease-disease and drug-sideeffect relationships. Together different types of data contribute to the understanding of the underlying molecular mechanisms of drugs and how they might perturb biological pathways and generate side effects and adverse effects.
Here, we will discuss how integration of large and diverse sources of information i.e. from molecular, cellular and phenotypic data associated to small molecules can lead to a generic disease chemical biology systems. It can improve our in silico evaluation of approved drugs for repurposing as well as in the investigation of chemicals related to anti-targets, adverse drug events and toxicity. Such systems will be illustrated through examples like antidepressants and (daily exposure) environmental chemicals.

PLENARY LECTURE 9

RODENT TOXICITY STUDIES ON PERFLUORINATED CHEMICALS FOR REACH

Barun Bhhatarai, Paola Gramatica
Department of Structural and Functional Biology (DBSF), University of Insubria, 21100 Varese, Italy

Perfluorooctane sulfonic acid (PFOSH) and Perfluoro-octanoic acid (PFOA) are long chain perfluorinated chemicals (PFCs) which are categorized by US-EPA and EU-REACH as toxic chemicals[1]. Other structurally similar PFCs exist, that are widely distributed in environments[2] due to their excessive use. Some of them are classified as 'emerging pollutants'[1] and their environmental and health related toxicities are subjects of concern. This raises alarms for other excessively used PFCs on which very few experimental data on environmental and bio-toxicity are available3 and moreover, toxicity profiles are found different for types of animals and species used. These compounds are studied as one of the four classes of compounds of high concern under European Union FP-7 funded project CADASTER.

The QSAR approach, which use is also suggested in the new European legislation REACH[4] to reduce animal testing, is here applied to predict and understand the toxicity of PFCs by modeling inhalation (LC_{50})[5] and Oral (LD_{50}) data on rodents: *Rattus* (Rat) and *Mus* (Mouse). Training and test set compounds were prepared on the available experimental data by splitting using: Self Organizing Map (SOM) and random selection through activity sampling[6]. The training sets were used to derive statistically robust models based on both splitting criteria. The models were then externally validated for their predictivity on test sets. The common set of descriptors selected in both splitting was then used in full models to predict a total of 376 PFCs, including those in REACH preregistration list[7]. The structural applicability domain and the reliability of predictions were verified.

The Rat and Mouse endpoints were predicted by each model for the studied compounds, out of which 204 compounds were found within the AD of all four models. In addition, cumulative toxicity study of the compounds was performed using principal component analysis (PCA): 30 compounds, all perfluorinated, were prioritized as most important for experimental toxicity analysis by CADASTER partners. The descriptors involved, the similarities and the differences observed will be discussed.

These models are helpful in calculating the toxicity profiles of untested molecule that will help to fill the existing data gaps in risk assessment. The results demonstrated the use and importance of predictive QSAR models highlighting the utility of QSAR for REACH.

(Financial support by European Union through the project CADASTER FP7-ENV-2007-1-212668)

References:

1. a) http://www.epa.gov/ oppt/pfoa/pubs/related.htm
   b) http://europa.eu/rapid/pressReleasesAction.do?reference=IP/06/1479
2. Renner R. *Environ. Sci. Technol.* **2003**, *37*, 201A.
3. a) http://www.syrres.com, b) http://www.ncbi.nlm.nih.gov/pubmed/
4. http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm
5. Bhhatarai, B., Gramatica P. *Chem. Res. Toxicol.*, **2010**, 23, 528–539.
6. Gramatica, P. QSAR *Comb. Sci.* **2007**, 26, 694-701
7. European Chemical Agency (ECHA): http://apps.echa.europa.eu/preregistered/pre-registered-sub.aspx

PLENARY LECTURE 10

INTEGRATIVE PHAMACOINFORMATICS APPROACHES IN THE PREDICTION OF CLINICAL OUTCOMES OF DRUGS

Ferran Sanz
Research Programme on Biomedical Informatics (GRIB)
IMIM - Universitat Pompeu Fabra
Barcelona, Spain

The final aim of most (Q)SAR studies is the computer-aided prediction of clinical outcomes of drugs by means of the (quantitative) modeling of the relationships existing between the chemical, structural and biological features. This challenging objective requires the use of a appropriated datasets and the application of sophisticated modeling strategies.
The datasets required for the aforementioned modeling have to integrate molecular data, as well as a wide range of biological information resulting from in vitro and in vivo experiments, along with clinical observations. The computational models to be developed require multi-dimensional and often multi-scale approaches. The systems biology perspective has frequently to be taken into account.
Several examples will be shown such as the predictive modelling of the antipsychotics side-effects (1), the chemo- and bioinformatics substantiation of possible adverse drug effects detected by pharmacoepidemiological methods (2), and the prediction of the cardiotoxicity by multi-scale modeling (3).

References

(1) Selent J, Bauer-Mehren A, Lopez L, Loza MI, Sanz F, Pastor M. A novel multilevel statistical method for the study of the relationships between multireceptorial binding affinity profiles and in vivo endpoints. Mol Pharmacol 2010; 77: 149-58.
(2) www.alert-project.org
(3) Obiol-Pardo C, Gomis-Tena J, Saiz J, Pastor M, Sanz F. A cardiotoxicity prediction system based on multiscale modeling. Tox Lett. 2010; 196 (Suppl. 1): S254

PLENARY LECTURE 11

LIGAND-BASED APPROACHES TO IN SILICO PHARMACOLOGY: BENCHMARKS AND APPLICATIONS

Jordi Mestres
Chemogenomics Laboratory, Research Unit on Biomedical Informatics (GRIB), Municipal Institute of Medical Research (IMIM) and University Pompeu Fabra, Barcelona, Catalonia, Spain

In the last five years, a wave of new approaches to estimating the affinity profile of molecules on multiple targets has been developed. The current flourishing of these methods is a direct consequence of the important progress experienced by some coordinated initiatives dedicated to data collection, classification, and storage of both pharmacological data for ligands and structural information for proteins.
In particular, the development of ligand-based approaches to target profiling has benefited enormously from the construction of publicly available annotated chemical libraries that incorporate pharmacological data from bibliographical sources into traditional chemical repositories. Some of the most visible results obtained with these methods have been the identification of new targets for old drugs [1,2], opening an avenue for anticipating drug side-effects but also for drug repurposing. In the near future, these methods should have also a big impact in chemical biology, and the first application to probing an entire protein family was recently reported [3].
In spite of the numerous successful applications published to date, a comprehensive assessment of the true ability of those approaches to predict an entire drug-target interaction matrix is still missing. Only recently, some ligand-based methods were tested for their ability to predict a rather small interaction matrix composed of 13 antipsychotic drugs and 34 protein targets (31 receptors and 3 neurotransmitter transporters) [4]. Accordingly, this presentation will focus on presenting our most recent efforts in benchmarking ligand-based approaches to target profiling.

[1] Keiser et al. Nat. Biotechnol. 25 (2007) 197-206
[2] Keiser et al. Nature 462 (2009) 175-181
[3] Areias et al. Bioorg. Med. Chem. 18 (2010) 3043-3052
[4] Vidal and Mestres. Mol. Inf. 29 (2010) 543-551

PLENARY LECTURE 12

CHALLENGES IN AGROCHEMICALS DESIGN

K.J. Schleifer
BASF SE, Ludwigshafen, Germany

Design of active ingredients is a multidimensional task. Sufficient target activity in combination with high bioavailability and no (or at least low) toxicological behaviour are prerequisites for promising candidates. In this respect, pharmaceutical and agrochemical companies need to address the same issues. Considering the target activity one can generally conclude that "target is target" regardless of whether the respective organism is a human or a pest. In detail however, drugs modulate activity of dysfunctional proteins or their functional counterparts aiming to reduce the pathogenic effects. In contrast to that agrochemicals modulate vitally important proteins in order to keep harmful organisms under control.
Bioavailability of drugs usually addresses the human patient with regard to a preferred oral application. Agrochemicals have to be divided into herbicides, fungicides and insecticides controlling weed, harmful fungi and insect pests. For each of these harmful organisms different absorption, distribution, metabolism and excretion routes have to be considered leading to quite different physicochemical properties of lead compounds.
Based on the above-mentioned similarities and differences, the presentation will give an overview on commonly applied computational approaches in pharmaceutical and agrochemical companies as well as several unique challenges relevant just for the design of potent crop protection compounds.

PLENARY LECTURE 13

ITERATIVE KINASE MEDIUM-THROUGHPUT SCREENING (ikMTS) With 2D PROFILE-QSAR AND 3D SURROGATE AutoShim ENSEMBLE DOCKING

Eric Martin, Prasenjit Mukherjee, David Sullivan
Novartis Institutes for BioMolecular Design, Emeryville, CA, USA

Experimental high-throughput screens of large (1.5 million) compound collections typically take 6 months and costs $1,000,000. The computational alternative of virtual screening by conventional high-throughput docking suffers from 3 important limitations: 1) it requires a protein structure of the target, 2) it is slow, 3) and it does a very poor job of predicting activity.

As a different alternative, we have created 2 new kinase-specific virtual screening methods that correlate with affinity far better than traditional QSAR or docking. They combine a modicum of medium throughput IC_{50} training data (MTS) for the new kinase target, with a large database of historical kinase activities. The 2D "Profile QSAR" method builds conventional fragment-based QSARs for over 70 kinases, then uses the predicted activities as chemical descriptors for a PLS model of the new kinase. The resulting profile QSAR models, each now based on over 700,000 data points, predict affinity much better than the conventional QSARs on which they were based.

The 3D Surrogate Ensemble AutoShim method uses the MTS data to adjust pharmacophore "shims" in a standardized "Universal Kinase Surrogate Receptor" (UKSR) ensemble of 16 diverse kinase crystal structures. We spent 5 months pre-docking our 1.5 million compound archive into the UKSR, extracting the docking scores and shim interactions for up to 100 poses in each of the 16 structures. Using these stored docking results, AutoShim can now be "shimmed" on MTS data for any new kinase target, to produce highly predictive, target-specific, scoring functions, and immediately score the 1.5 million pre-docked compounds, alleviating those 3 limitations of conventional docking: 1) accurate activity predictions for 1.5 million compounds, 2) in hours instead of weeks, 3) without a crystal structure.

Iterative Profile-QSAR and Surrogate AutoShim screens are now routinely performed at Novartis. They have been successfully applied to 7 active Novartis projects, with external $R^2 = 0.35 – 0.7$ and enrichments of 20x to 40x.

With kinase enzymatic activity prediction in hand, we have now successfully extended these methods in 3 directions: 1) predicting kinase cellular activity, 2) predicting kinase selectivity, and 3) adapting to the protease protein family.

PLENARY LECTURE 14

BINDING AFFINITY PREDICTIONS WITH LINEAR RESPONSE METHODS AND FREE ENERGY PERTURBATION CALCULATIONS

Ruhong Zhou,[1,2]

[1] IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA

[2] Department of Chemistry, Columbia University, New York, NY 10027, USA

In this talk, I will present some of our recent works on ligand-receptor (HIV-1RT) and influenza antigen-receptor and antigen-antibody binding affinity predictions with a new Linear Response Method based on continuum solvent models and Free Energy Perturbation calculations. For the HIV-1RT ligand-receptor binding, various linear response schemes are presented and discussed on this relatively large binding set, consisting of a total of 57 ligands. For a training subset of 40 ligands (20 nevirapine and 20 HEPT analogues), our linear response method gives an RMS error of 0.89 kcal/mol with a correlation coefficient r^2 of 0.74. The leave-one-out cross validation results also shows a very encouraging RMS error of 1.00 kcal/mol, with a correlation coefficient r^2 of 0.69. The further blind tests on seven mostly high-potent candidates (not originally in the training subset) also show very high accuracies. Finally, six new ligands are designed for optimal binding based on our predictions. The binding mechanism of this HIV-1RT receptor is also analyzed in detail. For the influenza antigen-receptor and antigen-antibody binding, the goal is to identify which mutations on the viral glycoprotein hemagglutinin (HA) might cause its receptor binding specificity to switch from avian to human, and which mutations might cause its escape of antibody neutralization, through large scale Free Energy Perturbation calculations. It is found that a single mutation T131I on H3N2 HA can decrease the HA-Ab binding affinity by 5.2 $\pm$ 0.9 kcal/mol, in excellent agreement with experimental results. It is also predicted that a double mutation, V135S & A138S, might switch the H5N1 HA binding specificity from avian to human (2.6 kcal/mol increase in its binding to human receptor), thus allowing the virus to gain a foothold in human population. Detailed analyses also reveal molecular picture of the antigen-antibody and antigen-receptor binding mechanisms.


18th EUROQSAR

KEY NOTE LECTURES

KEY NOTE LECTURE 1

TAUTOMERISM, THE FORGOTTEN MOLECULAR DESCRIPTOR

Yvonne Martin
Martin Consultant, USA

Quantitative forecasts of potency depend on an accurate calculation of molecular properties, which in turn depend on the correct structure representation. For molecules that can tautomerize, the structure entered into a database is not necessary the (or the only) structure that one needs to consider. This talk will discuss some strategies to solve this problem.

KEY NOTE LECTURE 2

NOVEL APPROACHES TO CHEMICAL TOXICITY PREDICTION RELYING ON THE ENTIRE STRUCTURE-IN VITRO-IN VIVO DATA CONTINUUM

Alexander Tropsha
Laboratory for Molecular Modeling, Carolina Center for Computational Toxicology and Center for Environmental Bioinformatics, University of North Carolina at Chapel Hill, Chapel Hill, NC U.S.A.

Recent advances in high-throughput screening of chemical compounds of interest to regulatory agencies such as EPA and FDA afforded a wealth of data that require new computational approaches to link chemical structure, in vitro assay results, and adverse effects of chemicals. We advance a predictive QSAR modeling workflow that relies on effective statistical model validation and implements both chemical and biological (i.e., in vitro assay results) descriptors of molecules to develop *in vivo* chemical toxicity models. We have developed two distinct methodologies for in vivo toxicity prediction utilizing both chemical and biological descriptors. In the first approach, we employ biological descriptors directly in combination with chemical descriptors for model building; obviously, this approach does not fully rely on computed molecular properties as it requires the knowledge of experimentally determined biological descriptors to assess potential toxicities of new compounds. Our second modeling approach employs the empirical relationship between *in vitro* and *in vivo* data as part of the two-step hierarchical modeling strategy. Initially, compounds are partitioned into (two or more) classes defined by patterns of *in vitro – in vivo* relationships and a binary (or multi-class) QSAR model using chemical descriptors only is built to distinguish these classes. In the second step, class-specific conventional QSAR models are built, also using chemical descriptors only. Thus, this hierarchical strategy affords two-step external predictions using chemical descriptors only in each step. We will present the results of applying both strategies to several experimental datasets. Our studies suggest that utilizing *in vitro* assay results as additional biological descriptors in QSAR studies affords models with prediction accuracy superior to the outcome of conventional QSAR modeling as well as predictors of in vivo effects developed using in vitro biological responses only. We will discuss how our models can be used to prioritize compound selection for experimental toxicity studies.

KEY NOTE LECTURE 3

ASSESSING REACTIVE METABOLITE RISK IN DRUG DISCOVERY USING A WEIGHT-OF-EVIDENCE APPROACH

Scott Boyer[1], Lucija Peterlin-Masic[2], Ola Spjuth[1], Lars Carlsson[1]
[1]AstraZeneca R&D, 43183 Molndal, Sweden
[2]Faculty of Pharmacy, University of Ljubljana, 1000 Ljubljana, Slovenia

The formation of reactive metabolites during the biotransformation of drugs is considered an unnecessary safety risk in drug discovery programs. To minimize the risk of reactive metabolite formation early drug design decisions methods are required that not only alert chemists of a potential reactive metabolite hazard, but also of the possibility that it will actually be formed – a risk assessment. Ideally this warning would be generated from the chemical structure without the necessity of synthesis and experimental testing. The formation of reactive intermediates depends on a number of factors. They are 1) the presence of a substructure that will yield a chemically reactive species upon biotransformation 2) the possibility that a metabolic event will actually occur in that particular substructure such that the reactive species will be formed in that molecule and 3) that the metabolic rate of the molecule will be such that enough of the reactive species will be formed to pose a risk. This combination of factors makes the use of a simple QSAR model impractical. Instead we have employed a weight-of-evidence approach to the problem of reactive metabolite risk assessment by the combination of computational tools to address the above factors individually and then combined them to assess, using actual reactive metabolite screening data, the conditions dictated by the individual tools that will indicate high and low risk for reactive metabolite formation. Assessment of hazardous substructures is by simple SMARTS matching using a database of examples from literature and from in-house data. The probability of biotransformation within that substructure to yield a reactive species is preformed using the site-of-metabolism tool MetaPrint2D. Finally, the rate of metabolism is assessed in a number of ways, including physico-chemical properties, rate-of-metabolism QSARs or actual biotransformation rate data. The results of the risk assessment validation against glutathione trapping data suggests a set of conditions that if met, predict positive and negative glutathione trapping results with greater than 80% accuracy. These results suggest that the combination of tools to address reactive metabolite risk can provide design teams with an assessment of real risk rather than the simple hazard alert of a structural alert.

KEY NOTE LECTURE 4

PHARMACOPHORES - VERSATILE TOOLS TO BRIDGE THE GAP BETWEEN STRUCTURE-BASED AND LIGAND BASED APPROACHES

Thierry Langer
Prestwick Chemical, Inc.
Blvd Gonthier d'Andernach
67400 Illkirch/Strasbourg, France

In silico or virtual screening has gained considerable impact for the efficient discovery of novel bioactive compounds in modern pharmaceutical research. The concept of chemical feature-based pharmacophore models has been established as state-of-the-art technique for characterizing the interactions between a macromolecule and a ligand. The results of numerous case studies have been published, clearly indicating the merits of this approach for efficient hit discovery [1].

While in ligand-based drug design, feature-based pharmacophore creation from a set of bioactive molecules is a frequently chosen approach; structure-based pharmacophores are still lacking the reputation to be an alternative or at least a supplement to docking techniques. Nevertheless, screening using 3D pharmacophores as filters bears the advantage of being faster than docking. Additionally, it transparently provides the user with relevant information that is used by the screening algorithms to characterize the ligand-macromolecule interaction.

At Inte:Ligand GmbH, LigandScout [2] has been developed, as a rapid and efficient tool for automatic interpretation of ligand-protein interactions and subsequent transformation of this information into 3D chemical feature-based pharmacophore models As an extension of this approach, we have introduced parallel pharmacophore-based screening as an innovative in silico method to predict the potential biological activities of compounds by screening them with a multitude of pharmacophore models.[3] Using LigandScout, the entire Protein Databank has been processed, and a pharmacophore database of structure-based pharmacophore models for all targets of potential interest for drug development has been generated, in addition to ligand-based models for targets that lack information about their 3D structure.

We present an overview of this technology together with the results of an application example employing a set of antiviral compounds that were submitted to in silico activity profiling using a subset of the Inte:Ligand pharmacophore database. The results of the screening experiments show a clear trend towards correct prediction of activity profiles. In addition, using our approach one is able to obtain information about binding of the ligands under investigation also to 'anti'-targets, such as enzymes of the cytochrome P450 family[4], or to the hERG channel. Thus, off-target activity can be determined easily, giving support to the medicinal chemists in their hit-to-lead and lead optimization studies.

References

[1] T. Langer, Mol. *Inform.* **29**, 470-475 (2010)
[2] G. Wolber, T. Langer, *J. Chem. Inf. Model.* **45**, 160-169 (2005)
[3] T. M. Steindl, et al., *J. Chem. Inf. Model.* **46**, 2146-2157 (2006)
[4] D. Schuster, et al., *Curr. Drug Discov.* Technol. **3**, 1-48 (2006)

KEY NOTE LECTURE 5

VISUAL (Q)SAR: SAR MAPS, SCAFFOLD TREES, AND R-CLIFFS

Dimitris Agrafiotis
Johnson & Johnson, Pharmaceutical Research and Development, USA

Chemoinformatics has been historically a quantitative science, with most of the effort focusing on building predictive models relating chemical structure to some molecular property of interest. While such models have proven useful in a number of settings, the complexity of the phenomena that are typically being modeled makes it impossible for them to be error-free. Indeed, quantitative models can often be wrong, and a few incorrect predictions are usually enough to shake the confidence of the discovery scientists and their willingness to utilize them to drive their projects forward. Indeed, our experience suggests that discovery teams are more interested in practical tools that give them convenient access to all their relevant project data and allow them to visualize the underlying multi-dimensional SAR in ways that make it easy to identify general trends and exceptions. Here, we present a family of three closely-related tools, scaffold explorer, multi-dimensional SAR maps, and single R-group polymorphisms (SRP's) or R-cliffs, which help the medicinal chemists organize, understand, and exploit SAR. These tools mirror the way in which medicinal chemists design drugs, namely through the iterative optimization of scaffolds and the substituents around those scaffolds until the desired chemical and pharmacological profile has been achieved. A distinguishing element of our approach is that the core organizing principles are not derived by automated means but by the medicinal chemists themselves, thus striking an optimal balance between automation and flexibility. Collectively, these tools allow the user to get a "bird's-eye" view of the chemical space spanned by a particular data set, map and aggregate any physicochemical property or biological activity of interest onto the individual scaffolds and/or substituents, understand scaffold, R-group, and individual substituent effects across multiple biological dimensions, and quickly distinguish promising parts of the molecule from less interesting or problematic ones. By focusing on intuition vs prediction and on insight vs quantitative modeling, these tools can greatly enhance the way in which SAR are understood and exploited in a real, practical sense.

KEY NOTE LECTURE 6

COMPUTER-AIDED DRUG REPURPOSING

Tudor Oprea
University of New Mexico, School of Medicine, New Mexico, USA

The emergent sector of academic drug discovery encompasses basic sciences, translational medicine and drug repurposing (or repositioning). Finding new uses for old drugs is a viable strategy, already embraced by the pharmaceutical industry. Current drug repurposing efforts focus on identifying novel modes of action, but not in a systematic manner. With intensive data mining, processing and curation, we applied bio- and chem- informatics tools to assemble DRUGS, a database of 3,837 unique small molecules and 1,700 proteins that are likely to function as drug targets and antitargets (i.e., associated with adverse drug reactions, ADRs). The use of *in silico* technologies to suggest novel uses for approved drugs will be discussed for several drugs, together with perspectives for building a CADR platform


18th EURO QSAR

ORAL PRESENTATIONS

ORAL PRESENTATION 1

QSEARCH: A NEW METHOD FOR DE NOVO LIGAND DESIGN

Tanja Schulz-Gasch[1], Tobias Lippert[2], Wolfgang Guba[1], Olivier Roche[1], Matthias Rarey[2]
[1]Pharmaceutical Research, Chemistry, Technologies & Innovation, F. Hoffmann-La Roche Ltd., CH-4070 Basel
[2]Center for Bioinformatics, University of Hamburg, Bundesstrasse 43, D-20146 Hamburg

De *novo* ligand design[1,2] and scaffold hopping[3,4] are two related approaches that complement the HTS process for the generation of novel starting points in the drug design process. The technique requires structural information about protein-ligand interactions or a hypothesis of the bioactive ligand conformation to finally build pharmacophore constraints relevant for binding.
We recently developed a novel approach for de novo generation of drug-like molecules from chemical fragment spaces[5,6] without direct context to known actives called "Qsearch". The simplest query input is a pharmacophore definition with pharmacophore-type features (aliphatic, aromatic, donor, acceptor, cation and anion) in combination with ligand shape (inclusion shape). The query can be further refined, for example with a fixed starting fragment, addition of directionality of pharmacophore features, SMARTS inclusion spheres, ranges for molecular properties, etc. This query setup requires the bioactive conformation of a single known active or aligned multiple known actives. The large solution space for all possible combinations of fragments prevents from fully exhaustive searching. We thus use a set of heuristics to restrict the solution space to relevant portions and stochastic search algorithms to create a diverse set of molecules obeying the query.
The approach has been validated with different design scenarios where we could demonstrate that Qsearch is able to re-discover topologies of known actives.

References:

1. Todorov, N. P.; Alberts, I. L.; Dean, P. M. *Compr Med Chem II* **2006**, 4, 283-305.
2. Mauser, H.; Guba, W. Recent developments in de novo design and scaffold hopping. *Curr Opin Drug Discov Devel*. **2008**, 11, 365-374.
3. Böhm, H.-J.; Flohr, A.; Stahl, M. Scaffold Hopping. *Drug Discovery Today: Technologies* **2004**, 1, 217-224.
4. Maass P., Schulz-Gasch T., Stahl M., Rarey M. Recore: A Fast and Versatile Method for Scaffold Hopping Based on Small Molecule Crystal Structure Conformations. *J. Chem. Inf. Model.* **2007**, 47, 390-399.
5. Rarey, M., Stahl, M. Similarity Searching in Large Combinatorial Chemistry Spaces. *J. Comput. Aided Mol. Des.* **2001**, 15, 497-520.
6. Degen, J., Wegscheid-Gerlach, C., Zaliani, A., Rarey, M. On the Art of Compiling and Using 'Drug-Like' Chemical Fragment Spaces. *ChemMedChem* **2008**, 3, 1503-1507.

LigandScout: MORE ACCURACY FOR PHARMACOPHORE-BASED VIRTUAL SCREENING

Gerhard Wolber[1,2], Fabian Bendix[2], Thomas Seidel[2], Göhkan Ibis[2] and Yasmin Aristei[2]

[1]Freie Universität Berlin, Institut für Pharmazie, Königin-Luise-Str. 2+4, 14195 Berlin, Germany and
[2]Inte:Ligand GmbH, Mariahilferstrasse 74B/11, 1070 Vienna, Austria

Virtual screening using three-dimensional arrangements of chemical features (3D pharmaco1phores) has become a well-known and established method in computer-aided drug design. Although frequently used, considerable differences exist in the interpretation of these chemical features and their corresponding 3D overlay algorithms. We have recently developed an efficient and accurate 3D alignment algorithm based on a pattern recognition technique [1]. In the presented work, we extended this algorithm to be used for high-performance virtual database screening and investigate, whether applying this geometrically more accurate 3D alignment algorithm improves virtual screening results over conventional incremental n-point distance matching approaches.

[1] G. Wolber, A. Dornhofer, and T. Langer. Efficient overlay of small molecules using 3-D pharmacophores. J. Comput.-Aided Mol. Design, 20(12): 773-788 (2006)

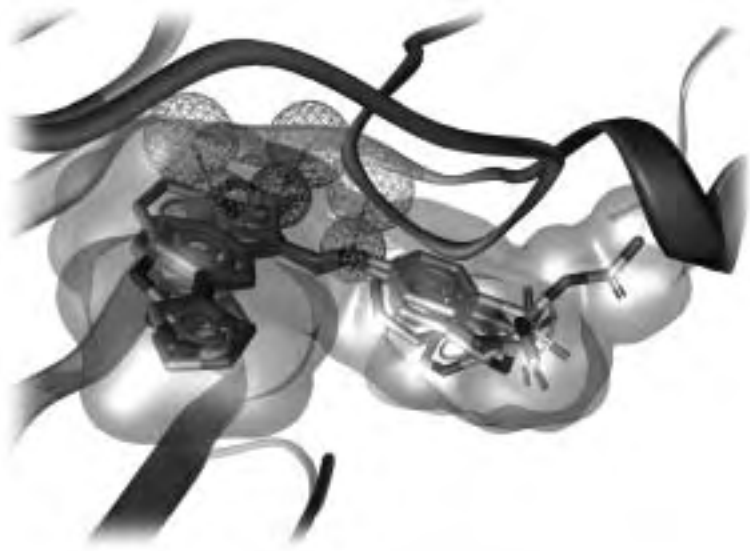


Fig. 1. Structure-based 3D pharmacophore derived from PDB complex 1KE7 (cyclin-dependent kinase 2) with the original ligand and virtual screening hits accurately identified and aligned using the presented virtual screening approach.

RECOVERING DESIGN STRATEGIES OF GPCRs MODULATORS FROM EXPLORATIONS OF THE CHEMICAL SPACE

Antonio Macchiarulo[1], Chiara Custodi[1], Roberto Nuti[1], Tudor Oprea[2], Roberto Pellicciari[1]

[1]Dipartimento di Chimica e Tecnologia del Farmaco, University of Perugia. Perugia, Italy; [2]University of New Mexico, School of Medicine, Albuquerque, New Mexico, USA

In the last decade, sequencing of human genome and the routinely use of high-throughput and virtual screening for lead identification, have considerably expanded the regions of the chemical space where G-protein coupled receptor (GPCR) modulators lie [1-5]. This offers new opportunities to understand how agonists and antagonists for this superfamily of receptors work. With this aim, we have developed a study based on the construction of decision trees that, identifying molecular properties able to distinguish GPCR agonists from antagonists, could provide novel design strategies to medicinal chemistry for drug discovery. The space covered within the WOMBAT database was taken as the GPCR modulators dataset source, whereas either 2D or 3D MOE descriptors were used to map this chemical space. In particular, four families of decision trees were built using 2D structural descriptors, 2D connectivity and shape indices, 3D molecular shape and dipole descriptors, 3D VolSurf descriptors, respectively. The analysis concerned the classification of agonist and antagonist modulators for 11 receptors belonging to family A of GPCRs.
Although the results show that distinct properties feature agonists and antagonists of selected family A GPCR modulators, a clear classification of agonists and antagonists could not always be achieved, in agreement with the current notion of the complexity of GPCR modulation [6]. As a further consideration, the results show a mixed performance of 2D and 3D descriptors, with modulators of specific GPCRs being well classified selectively by 2D (e.g. adrenergic and histaminergic modulators) or 3D descriptors (dopaminergic and cannabinoid modulators), and modulators of other GPCRs being well classified by 2D and 3D descriptors (e.g. cholinergic, prostanoid modulators). While instrumental to get insights into the design strategies of GPCRs modulators, this work provides novel clues on the molecular mechanisms that may underlie the modulation of this class of receptors.

References

1. Hopkins A.L., Groom CR. The druggable genome. *Nat Rev Drug Discov.* **2002**, 1, 727-30.
2. Lipinski C., Hopkins A. Navigating chemical space for biology and medicine. *Nature*, **2004**, 432, 855-861.
3. Bleicher K. H., Bohm H. J., Muller, K. & Alanine, A. I. Hit and lead generation: beyond high-throughput screening. Nat. Rev. *Drug. Discov.*, **2003**, 2, 369-78.
4. Bajorath J. Integration of virtual and high-throughput screening. *Nat. Rev. Drug. Discov.*, **2002**, 1, 882-94.
5. Shoichet B. K. Virtual screening of chemical libraries. *Nature*, **2004**, 432, 862-5.
6. Millar R.P., Newton C.L. The year in G Protein-Coupled Receptor research. Mol. *Encocrinol.*, **2010**, 24, 261-74

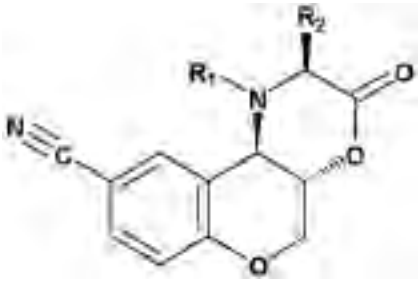
STEREOSELECTIVE INTERACTION OF BENZOPYRANO[3,4-b][1,4]OXAZINES WITH P-GLYCOPROTEIN

I. Jabeen[a], P. Wetwitayaklung[a], F. Klepsch[a], P. Chiba[b], G. F. Ecker[a]

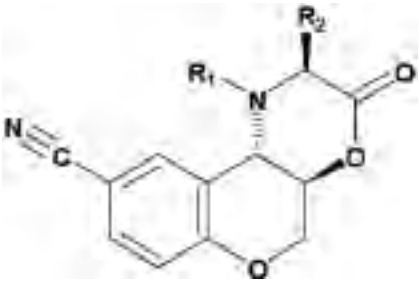
[a]Department of Medicinal Chemistry, University of Vienna, Althanstraße 14, 1090, Vienna, Austria

[b]Medical University of Vienna, Institute of Medical Chemistry Waehringerstraße 10, 1090, Vienna, Austria

The ATP-binding cassette efflux transporter P-glycoprotein (P-gp) is a membrane-bound protein associated with the multidrug resistance (MDR) phenotype. P-gp is highly promiscuous in its substrate and inhibitor interaction profile and thus only few reports on stereoselectivity of ligand-protein interaction have been published. To further probe the stereoselectivity of P-gp, we synthesized a series of enantiomerically pure Benzopyrano[3,4-b][1,4]oxazines and tested their ability to inhibit P-glycoprotein mediated daunomycin efflux in multidrug resistant CCRF vcr 1000 cells[1].



(a)



(b)

GRID-independent molecular descriptors (GRIND)[2,3] analysis was performed to study the main structural determinants for drug P-gp interaction. GRIND studies show that two H-bond acceptors at a distance of 4.80-5.20 Å are beneficial for high biological activity and that shape descriptors also play an important role. Furthermore, 3 pairs of stereoisomers were docked into a homology model of P-gp. Through Agglomerative Hierarchical Cluster analysis of the consensus RMSD matrix based on the common scaffold of the ligands, two clusters were identified. One of them contains only compounds of configuration (a) and the other one contains additionally one compound with configuration (b). Analysis of the protein-ligand interaction pattern showed different interaction behavior of the two types of stereoisomers, mainly related to amino acids Tyr 307 and Phe 343.

This work was supported by the Austrian Science Fund, grant F03502 and by the Higher Education Commission Pakistan (HEC)

References:

1. Chiba, P.; Ecker, G.; Schmid, D..; Drach, J.; Tell, B.; Gekeler, V. *Mol. Pharmacol.* **1996**, 49, 1122 -1130
2. Pastor, M.; Cruciani, G.; McLay, lain. ; Pickett, S. ; Clementi, S. *J. Med. Chem.* **2000**, 43,3233-3243
3. Durán, A.; Comesaña, G and Pastor M. *J. Chem Inf Mod.* **2008**, 48(9), 1813-1823

IDENTIFICATION AND APPLICATION OF ANTITARGET ACTIVITY HOTSPOTS TO GUIDE COMPOUND OPTIMIZATION

Gerhard Hessler, Clemens Giegerich, Friedemann Schmidt, Li-hsing Wang, Karl-Heinz Baringhaus

LGCR Structure, Design and Informatics, Building G838, Sanofi-Aventis Deutschland GmbH, Industriepark Höchst, 65926 Frankfurt, Germany

The similarity principle has driven Medicinal Chemistry for more than 100 years. This concept is in particular applied for the identification of new lead structures in ligand-based virtual screening and has stimulated chemogenomics-based lead finding strategies. The extension of this similiarity principle to guide antitarget design offers new opportunities in compound optimization.

Often, pairs of molecules with minor structural changes exhibit significant activity differences (active-inactive). We extended this concept of *activity cliffs* [1] based on *matched molecular pairs* [2] to antitargets (e.g. hERG, CYP,..) in order to derive design principles for multidimensional compound optimization of a given lead series. Databases containing descriptive pairs of molecules with steep SAR for several antitargets were generated as starting point for subsequent similarity-based comparisons to compounds of interest. Molecules suffering from an undesired antitarget activity are then compared to pairs of molecules in the antitarget database via 3D similarity-based methods. The corresponding alignment of a molecule of interest to a pair of (antitarget) reference compounds then yields hotspots for structural modification of that molecule to decrease antitarget activity which can be turned into new synthesis proposals. This allows a fast and efficient mining of compounds in large data sets of antitarget activity in order to derive optimization guidance of lead series early in lead optimization.

The details of the established procedure will be discussed as well as some case studies.

References:

[1] Review in: J. Bajorath, L. Peltason, M. Wawer, R. Guha, M. S. Lajiness, J. H. van Drie, *Drug Discovery Today* **2009**, 14, 698.

[2] A. G. Leach, H.D. Jones, D. A. Cosgrove, P. W. Kenny, L. Ruston, P. MacFaul, J. M. Wood, N.Colclough, B. Law, *J. Med. Chem.* **2006**, 49, 6672.

BIOPHYSICS-BASED LIBRARY DESIGN: DISCOVERY OF "NON-ACID" INHIBITORS OF S1 DHFR

Veer Shanmugasundaram*, Kris Borzilleri, Jeanne Chang, Boris Chrunyk, Mark Flanagan, Seungil Han, Melissa Harris, Brian Lacey, Richard Miller, Parag Sahasrabudhe, Ron Sarver, Holly Soutter, Jane Withka

Antibacterials Chemistry/Discovery Technologies, Pfizer PharmaTherapeutics Research & Development, Groton, CT - 06333, USA

Methicillin-resistant Staphylococcus aureus (MRSA), the causative agent of many serious nosocomial and community acquired infections, and other gram-positive organisms can show resistance to trimethoprim (TMP) through mutation of the chromosomal gene or acquisition of an alternative DHFR termed "S1 DHFR" To develop new therapies for health threats such as MRSA, it is important to understand the molecular basis of TMP resistance and use that knowledge to design and develop novel inhibitors that are effective against S1 DHFR. This presentation will highlight and illustrate an effort using a multi-pronged biophysics based strategy that utilizes NMR, thermodynamic, kinetic, structural, computational and medicinal chemistry information in developing an understanding of the mechanism of resistance in S1 DHFR as well as using this prospectively in drug discovery. Specifically this presentation will illustrate the following.

References:

1. Thermodynamic assessment of TMP binding to SA WT and S1 DHFR and the use of thermodynamic profiles generated from ITC data to understand synergistic effects of NADPH binding
2. Crystal structures of several WT, S1, single-point mutations that elucidate key, subtle, distal structural changes in protein-ligand complexes
3. HSQC NMR experiments that illustrate the flexibility of SA WT and S1 DHFR and use of these dynamics studies to generate design hypotheses
4. SPR experiments that measure kinetics of binding of key control compounds and design ideas for optimizing slow off-rates
5. Thermal stabilization data to confirm conclusions drawn from kinetic, thermodynamics and NMR dynamic data
6. Computational studies using WaterMap to develop an understanding of mechanism of resistance through mutations that are distal to the binding site
7. Computational biophysics-based library design efforts using several hypotheses based on experimental data (*vide supra*) that resulted in the discovery of "non-acid" inhibitors of S1 DHFR.

LOCAL CORRESPONDENCE CONCEPT IN BIO- AND CHEMINFORMATICS

D. Filimonov, T. Gloriozova, A. Lagunin, A. Zakharov, B. Sobolev, A. Dmitriev, V. Poroikov
Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences,
Pogodinskaya Str., 10, Moscow, 119121, Russia

Similarity approaches are widely used in cheminformatics and bioinformatics. We propose an improvement of the similarity principle, the "local correspondence concept", which is based on the fact that most biological activities of organic compounds are the result of molecular recognition, which in turn depends on the correspondence between particular atoms of the ligand and the target. According to this, local similarity between atoms in molecules or amino acids in proteins may by considered by itself, without aggregation into general similarity. However, in most of the existing (Q)SAR methods, the local correspondence concept is used to an insufficient degree, although it provides a platform to recognize and to predict various characteristics of organic molecules and proteins.

We shall discuss the implications and applications of this concept for (Q)SAR in the areas of both cheminformatics and bioinformatics. We have developed a consistent system of descriptors called Multilevel, Reaction Multilevel, and Quantitative Neighborhoods of Atoms (MNA, RMNA, and QNA, respectively) and have implemented them in several SAR/QSAR/QSPR modeling approaches. For instance, MNA descriptors have been employed for predicting biological activity spectra of organic molecules in the PASS software [1] for more than 15 years. PASS has been used by many scientific groups that lead to the discovery of NCEs in different therapeutical fields, including anxiolytics, anticonvulsants, cognition enhancers, NSAIDs, analgesics, antiviral and antibacterial agents [1].

For QSAR/QSPR modeling we have proposed a novel QNA based 'Star Track' approach [2] where, in accordance with the local correspondence concept, any molecule is represented as a set of points in the two-dimensional space of QNA descriptors. The estimate of the compound's property is calculated as the average value of the QNA descriptor's function for each of the atoms in the molecule. We developed the software GUSAR based on this approach. It has been shown that in the majority of cases the accuracy and predictivity of GUSAR models appeared to be better than for the reference QSAR methods [2].

Finally, in the area of bioinformatics, we have developed a new method for annotation of amino acid sequences, which is based on the calculation of the similarity score for each of the positions in the amino acid sequence. These similarity scores are used directly, without any calculation of the total similarity of sequences, as input data for a PASS-like classifier. The method was tested using leave-one-out cross-validation for three data sets covering 58 enzyme classes, and 100% accuracy was achieved for the majority of these classes, which was equivalent or superior to popular HMMer and SVM-Prot approaches [3]. It was also shown that the predictions of protein kinase substrates were suitable for finding of previously missed links in signaling pathway networks.

The local correspondence concept is a new cheminformatics and bioinformatics paradigm, which allows describing biomolecules and predicting their target properties such as bioactivity and function. It also affords an intuitively obvious interpretation of models in terms of contributions of particular atoms in a molecule or a site in an amino acid sequence.

Acknowledgements. This work was partially supported by the European Commission FP7 grant 200787 "OpenTox", ISTC grant 3777 and RFBR grant N 09-04-01281.

References:

[1] Filimonov D.A., Poroikov V.V. Probabilistic approach in activity prediction. In: *Chemoinformatics Approaches to Virtual Screening. Eds. Alexandre Varnek and Alexander Tropsha*. **2008**, Cambridge (UK): RSC Publishing, 182-216; http://www.ibmc.msk.ru/PASS

[2] Filimonov D.A., Zakharov A.V., Lagunin A.A., Poroikov V.V. QNA based 'Star Track' QSAR approach. *SAR and QSAR Environ. Res.*, **2009**, 20 (7-8), 679-709.

[3] Alexandrov K., Sobolev B., Filimonov D., Poroikov V. (2008). Recognition of protein function using the local similarity. *J. Bioinform. Computat. Biol.*, **2008**, 6 (4), 709-725

USING LOCAL MODELS TO IMPROVE QSAR PREDICTIVITY

Fabian Buchwald, Tobias Girschick, Madeleine Seeland and Stefan Kramer
Institut für Informatik I12, Technische Universität München, Germany

In a recent paper, Benigni and Bossa [1] found that local QSAR models can produce results that are mechanistically interpretable and compare favorably with the known limits in reproducibility of the experimental systems. However, many existing large databases cannot be used directly to build local QSAR models, because they contain diverse sets of non-congeneric structures. We present a novel QSAR approach that automatically detects groups of structures for local QSAR modeling. The algorithm combines clustering and classification or regression for making predictions on chemical structure data. A structural clustering procedure is applied as a preprocessing step, before a (local) model is learned for each relevant cluster. Instead of using only one global model (classical approach), we use weighted local models for predictions of query instances dependent on cluster memberships.
The obtained clusterings are overlapping and non-exhaustive. In other words, a compound can be member of more than one cluster or member of no cluster at all. The dataset is divided into several parts representing the different structural classes. Compared to global models that generally try to capture everything at once, local models are high-quality models of small regions of the chemical space that often have the advantage of being much more interesting and understandable to a domain expert. Those local models reflect the classical approach to QSAR, where only a small set of highly relevant compounds is taken into account when building a model for a specific endpoint.
The approach is evaluated together with standard statistical QSAR algorithms on various datasets. The results show that in most cases the application of local models significantly improves the predictive power of the derived QSAR models compared to the classical approach. In summary, the new combined approach is interesting both theoretically as a new synthesis of clustering and QSAR model learning and practically as a new method for making predictions in pharmacological and toxicological applications.

References

[1] Benigni, R., Bossa, C. (2008) Predictivity of QSAR, Journal of Chemical Information and Modeling, 48(5):971–980.

Acknowledgements

This research was partly funded by the EU FP7 project OpenTox - An Open Source Predictive Toxicology Framework (Health-F5-2008-200787) and the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety.

THE USE OF DESIGN OF EXPERIMENTS TO DEVELOP EFFICIENT ARRAYS FOR SAR AND PROPERTY EXPLORATION

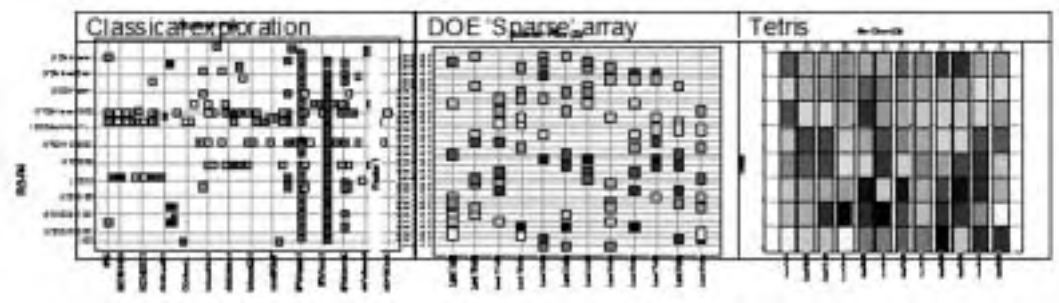
Chris Luscombe, Nick Barton,Tony Cooper, Heather Hobbs, Stephen Pickett, Darren Green

Computational Chemistry, GlaxoSmithKline, Stevenage, UK

Two or three dimensional chemical arrays are a ubiquitous and powerful tool in the armory for the rapid exploration of chemical space in Drug Design programmes. Such arrays are often the result of extensive analysis and design work leading to an optimal selection of monomers to explore chemical diversity, molecular properties and other pre-calculable parameters. These "complete" designs however are highly resource intensive.

Experimental Design approaches are well established for the optimization of multi-factor experimental procedures, such as synthetic reaction conditions but in these situations the factors are almost always continuous in nature. We propose that Design of Experiments (DOE) based approaches can be applied to exploratory chemical array scenarios where the full (M x N) array cannot be synthesized for practical reasons. By treating each monomer in the array as a categorical factor of the design, a balanced fractional (sparse) array design can be generated. Once synthesized and measured, the results can statistically analyzed to assess the additivity of SAR and then determine the contribution of the monomers to potency, selectivity and other properties of interest. This novel approach can be successfully used to understand and exploit the SAR of a late stage lead optimization programme.

This presentation will highlight the advantages and limitations and impact of using the DOE based array approach, drawing on examples from drug discovery lead optimization programmes. Additionally we will highlight a particular design (nicknamed the Tetris array) which is 'chemist' friendly and thus allows these investigational approaches to be carried out efficiently from a synthetic perspective.



ORAL PRESENTATION 10

ANALYSIS AND COMPARISON OF 2D FINGERPRINTS: INSIGHTS INTO DATABASE SCREENING PERFORMANCE USING EIGHT FINGERPRINT METHODS

Jianxin Duan*#, Madhavi Sastry, Steven L. Dixon, Jeffrey F. Lowrie and Woody Sherman
Schrödinger Inc. 120 West 45th Street, 17th Floor, New York, NY 10036-4041, USA
*Schrödinger GmbH, Dynamostr. 13, 68161 Mannheim, Germany

Virtual screening is a widely used strategy in modern drug discovery and 2D fingerprint similarity is an important tool that has been successfully applied to retrieve active compounds from large datasets. However, it is not always straightforward to select appropriate fingerprint method and associated settings for a given problem. Here, we applied eight different fingerprint methods, as implemented in the new cheminformatics package Canvas, on a well validated dataset covering five targets. The fingerprint methods include Linear, Dendritic, Radial, MACCS, MOLPRINT2D, Pairwise, Triplet, and Torsion. We find that most fingerprints have similar retrieval rates on average; however, each has special characteristics that distinguish its performance on different query molecules and ligand sets. For example, some fingerprints exhibit a significant ligand size dependency whereas others are more robust with respect to variations in the query or active compounds. In cases where little information is known about the active ligands, MOLPRINT2D fingerprints produce the highest average retrieval actives. When multiple queries are available, we find that a fingerprint averaged over all query molecules is generally superior to fingerprints derived from single queries. Finally, a complementarity metric is proposed to determine which fingerprint methods can be combined to improve screening results.
A more systematic virtual screening study has also been conducted to investigate the interrelation between eight fingerprinting methods, eleven atomtyping schemes, seven bit scaling rules, and four similarity metrics. In total, 24,068 virtual screens were performed to assess the effectiveness of each combination of options to identify active ligands in a database screen performed on 11 pharmaceutically relevant targets. Significant variations in enrichments were observed with all explored parameters. In general, fingerprints such as MOLPRINT2D and Dendritic that contain information about local environment beyond simple linear paths outperformed other fingerprint methods. Atomtyping schemes with more specific information were generally superior to more generic atomtyping schemes. With the best identified settings, enrichment factors across all targets could be improved considerably. No single combination of settings performed optimally on all targets and therefore we provide recommendations to improve enrichments based on different requirements

ROBUST SPARSE FEATURE/DESCRIPTOR SELECTION FOR QSAR

Frank Burden & David Winkler
CSIRO, Clayton, Victoria, Australia and Medicinal Chemistry and
Drug Action, Pharmaceutical Sciences, Monash University, Australia

QSAR modelling relies on experimental or calculated and/or numerical molecular descriptors and since there are very many thousands of these, it is often necessary to reduce them during the modelling process to a realistic level for evaluation. Most attempts [1] make use of PCA or PLS to reduce the dimension of the descriptor space (but not the number of descriptors); use descriptor correlations to remove unwanted descriptors; use a forward inclusion or backward elimination. We present two new and objective methods for descriptor reduction applied to a variety of datasets and molecular descriptors. The first method is an expectation maximisation algorithm applied to multiple linear regression (MLREM) [2] and it will be shown that this is a simple addition to a standard MLR algorithm using a Laplacian prior. The second method is a non-linear regression uses a similar methodology as applied to a Bayesian regularised artificial neural network (BRANNLP) [3] which is an enhancement of our published BRANN method [4]. The descriptors are all calculated from the molecular structures and comprise in-house atomistic, Burden eigenvalues, binned charges together with 2D and 3D sets provided by the software programs Dragon[5] and Adriana [6].

We have examined several data sets from the literature and applied the MLREM and BRANNLP methods to them. The first is a set of of 112 polymers and the Normalised Metabolic Activity (NMA) of foetal rat lung fibroblasts [7] on the polymer surface where we show that with the choice of appropriate descriptors and variable reduction we can produce a parsimonious model which performs better than those previously attained. The second set of the melting points of 717 ionic liquids [8] which was modelled in total and in three subsets containing respectively 126 pyridimium bromides, 384 imidazolium and benzoimidazolium bromides and 207 ammonium bromides. The results from our modelling again shows that appropriate descriptors and a non-linear variable reduction, using only 23 descriptors, produces results superior to those reported. Other data sets, including microarray analysis, will also be used to illustrate the power of the methods.

Our conclusion is that attention must be paid equally to both to the descriptors used in the modelling and the modelling method itself. We show that MLREM and, especially, BRANNLP can produce results superior and more parsimonious than other reported models.

A brief outline of the MLREM and BRANNLP methods will be given.

References:

[1] Livingstone,D.J. Salt,D.W. Variable Selection - Spoilt for Choice? *Reviews in Computational Chemistry*, 21, (2005) 287-348

[2] Burden, F.R.; Winkler D.A. Optimum QSAR Feature Selection using Sparse Bayesian Methods, QSAR *Comb Sci.* 28, (2009). 645-653,

[3] Burden, F.R.; Winkler,D.A. An optimal self-pruning neural network that performs nonlinear descriptor selection for QSAR, QSAR *Comb. Sci.* 28, (2009 1092-1097),

[4] Burden, F.R.; Winkler,D.A. Robust QSAR Models Using Bayesian Regularized Artificial Neural Networks, J. *Med. Chem.*, 42, (1999) 3183-3187

[5] Talete sri, DRAGON, Version 5.5 - (2007)- http://www.talete.mi.it

[6] Adriana, Molecular Networks, Henkestraße 91, 91052 Erlangen, Germany

[7] Vladyslav Kholodovycha, Jack R. Smithb, Doyle Knightc, Sascha Abramsonb,Joachim Kohnb, William J. Welsha, *Polymer* 45 (2004) 7367–7379.

[8] Varnek,A..Kireeva,N., Tetko.I.V. Baskin, I.I. Solov'ev,V.P, *Exhaustive QSPR Studies of aLarge Diverse Set of Ionic Liquids: How Accurately Can We Predict Melting Points?* J Chem Inf Model.,47 (2007),1111-22

ORAL PRESENTATION 12

THE pKa DISTRIBUTION OF SCREENING COMPOUNDS – APPLICATION TO DRUG DISCOVERY

David Manallack
Monash Institute of Pharmaceutical Sciences
Monash University, 381 Royal Parade, Parkville, 3052, Australia

Analyses of the acid ionization constant (pK_a) values of drugs has provided valuable insights into overall profile of pharmaceuticals that contain ionizable groups [1,2] This work provided three levels of analysis on contemporary drugs to firstly highlight those compounds that contain ionizable groups and to split these compounds into acids, bases and ampholytes, etc. Finally, a detailed analysis was provided of both single acid and single base containing compounds. Examination of these distributions demonstrated that they correspond to our current wisdom concerning the biopharmaceutical characteristics of drugs with particular regard to CNS and non-CNS agents.

The aim of these previous studies was to understand the acid/base profile of established drugs. Armed with this knowledge then it can be applied to drug discovery in a similar manner to the guidelines of Lipinski and co-workers [3]. This is particularly pertinent when we consider that the acid ionization constant (pK_a) is extremely important with regard to drug discovery and the biopharmaceutical character of drugs. Indeed, the pKa has a profound effect on the physicochemical properties of a drug controlling properties such as lipophilicity and solubility which in turn affect pharmacokinetics. As such, the pharmaceutical industry remains acutely aware of ionization issues for their drug discovery and development work.

The current study has examined the overall acid/base profile and pK_a distribution of compounds that are commercially available for screening purposes. Once again three levels of analysis were applied to identify ionizable groups, determine the proportions of acids and bases and finally to calculate and examine the distribution of pK_a values for single acids and single bases. The results clearly demonstrated that the profiles observed for drugs [1,2] were not matched by the available screening compounds. Indeed, the profile of single bases showed a paucity of compounds with pKa values above 7.0. These findings have major significance as they call into question the overall composition of screening compound collections. The results will be discussed in the light of the functional groups involved and how organic chemists prefer to synthesize compounds. Lead-like versus drug-like character will be examined particularly when we consider that we require hits and/or leads from a screening collection. Potentially, the business of drug discovery could be improved if we take into account acid/base distributions when selecting screening compounds. With this in mind, rough guidelines will be presented suggesting the overall makeup of acids and bases for a screening compound collection.

References:

1. Manallack, D.T. The pK_a Distribution of Drugs: Application to Drug Discovery. *Perspect. Medicin. Chem.* **1**, 25-38 (2007).
2. Manallack, D.T. The Acid-Base Profile of a Contemporary Set of Drugs: Implications for Drug Discovery, *SAR QSAR Env. Res.* **20**, 611-655 (2009).
3. C.A. Lipinski, F. Lombardo, B.W. Dominy, and P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Del. Rev.* **46**, 3-26 (1997).

ORAL PRESENTATION 13

UNDERSTANDING THE BLOOD BRAIN BARRIER: OPTIMIZATION STRATEGIES FOR CNS PENETRATION AND DISTRIBUTION

Mario Lobell*, Andreas Reichel
Bayer Schering Pharma, Global Drug Discovery, Germany

In the past tuning of logBB (log of the concentration ratio in brain and blood) has been used as a major criterium for the optimization of oral CNS drugs and also for non-CNS drugs which aim to minimize CNS exposure in order to reduce the occurance of CNS mediated side effects. This old paradigm has neglected to recognize that it is not the overall drug concentration in the brain but the free unbound concentration which can interact with the CNS target and excert a pharmacological effect. Binding to brain tissues and plasma proteins need to be considered as additional factors for deriving meaningful drug optimization strategies. Occurance and extend of active transport processes at the blood brain barrier (BBB) is another aspect which needs to be taken into consideration and integrated into such strategies.
Two sets of oral drug optimization strategies will be presented, one for CNS drugs and one for non-CNS drugs. These strategies integrate data from in silico, in vitro and in vivo experiments. Bayer Schering Pharma's *in silico* models for the prediction of log f_u(plasma), log f_u(brain) and logBB will also be presented.

MULTI-pH QSAR: II. REGRESSION ANALYSIS SENSITIVE ENOUGH TO DETERMINE THE TRANSITION-STATE pK_a OF HUMAN BUCCAL ABSORPTION

Robert A. Scherrer

BIOpKa White Bear Lake, MN 55110, USA

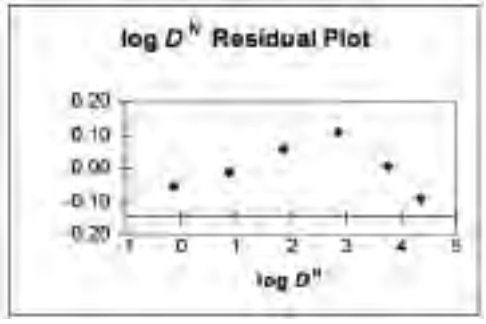
Multi-pH QSAR analysis can distinguish between neutral and ionized species, and even include both species in the same correlation.[1] This is not possible at a single pH, except for the special circumstance when there is sufficient variation in (log P^N – log P^I) within the series. We have now extended multi-pH QSAR to the analysis of individual compounds. Here the focus shifts to the model, since there are no changing parameters in the "series". This is illustrated by examining the rate of human buccal absorption of an acid and a base at six pH values.[2] The initial result for the acid, 4-n-hexylphenylacetic acid (1), is a correlation with log D^N, r^2 0.9812, but with a curious plot of residuals (Figure 1). Iterative reanalysis of the data gives an almost perfect correlation at the "effective" pK_a of 5.3 (aqueous pK_a 4.31, biolipid pK_a (unknown), octanol pKa 8.33):

$$\log K_{abs.rate} = 0.36(\pm 0.005)\log D^N - 2.09(\pm 0.015)$$

$$n = 6 \quad r^2 = 0.9994 \quad s = 0.015 \quad F = 6261$$

We interpret the "effective" or "flux" pKa to be the transition-state pK_a for absorption as the compound passes from an aqueous phase to a biolipid environment.[3] For the base, propranolol, we found transition-state pK_a 8.7 (aqueous pK_a 9.53, octanol pK_a 6.95). Iterative analysis improved correlation of the buccal absorption rate with log D^N from r2 0.9884 to 0.9982. In summary, we believe that multi-pH QSAR can be a valuable tool for drug design and development, not only for tracking ionized and neutral species, but for insight into the models as well.

Figure 1. Residuals for 1; pK_a 4.31



References:

1) (Near final stage of preparation for submission to ACS Medicinal Chemistry Letters.) Scherrer, R. A.; Leo, A. Multi-pH QSAR: A Simple Method to Differentiate Neutral and Ionized Species.

2) Schurman, W.; Turner, P. *J. Pharm. Pharmacol.* **1978**, 30, 137.

3) Scherrer, R. A. The Unstirrable Water Layer Is Unstirrable because It Does Not Exist. *Chem. Biodiversity* **2009**, 6, 1917-1922.

ORAL PRESENTATION 15

MULTI-PARAMETER OPTIMIZATION AND IN SILICO MODELING IN LEAD OPTIMIZATION

Hua Gao
Pfizer, Global Research and Development, USA

In silico modeling, especially in silico ADMET modeling has been playing an important role in drug discovery, especially during lead optimization phase. The confidence and willingness of applying in silico models in MPO in lead optimization and drug candidate seeking stages to balance potency, clearance, physico-chemical properties as well as safety endpoints have been significantly increased due to better modeling methods. Expensive in vitro and in vivo experiments could be complemented with in silico predictions to increase R&D productivity and reduce R&D cost. The quality of designed compounds has been improved. In this presentation, a successful application of in silico models in predicting human dose in discovering opiate receptor antagonists will be discussed.

ORAL PRESENTATION 16

COMBINED IN SILICO APPROACHES FOR DRUG DESIGN AND PHARMACOKINETIC OPTIMIZATION OF A SET OF CARNOSINE ANALOGUES AS POTENT AND SELECTIVE CARBONYL QUENCHERS

Giulio Vistoli, Alessandro Pedretti, Cristina Marconi, and Giancarlo Aldini
Dipartimento di Scienze Farmaceutiche "Pietro Pratesi", Facoltà di Farmacia, Università degli Studi di Milano, Via Mangiagalli, 25, I-20133 Milano, Italia

Reactive carbonyl species (RCS) are important cytotoxic mediators and represent a novel drug target since supposed to play a pathogenic role in several diseases including renal, hepatic, neurodegenerative diseases, diabetes, and atherosclerosis.[1] Taking RCS and carbonylation damage as a new drug target, we recently found that L-carnosine (β-alanyl-L-histidine), an endogenous dipeptide present in mM concentrations in some tissues, is a selective detoxifying agent of RCS,[2] which is actively absorbed by hPepT1, but is rapidly hydrolyzed in human serum by carnsosinase, a specific dipeptidase. Hence, the rational design of new carnosine analogues should (1) increase the quenching activity of carnosine, maintaining its specificity (2) confer plasma stability against human serum carnosinase, and (3) conserve an optimal recognition by hPepT1. The computational approaches proved successful in each mentioned step affording valuable tools in guiding the rational design of novel derivatives. In particular, (1) quanto-mechanical calculations, conformational analyses and physicochemical predictions cooperated to design more active and specific carnosine derivatives,[3] while (2) homology modeling techniques and docking simulations were exploited to generate reliable models for human serum carnosinase[4] and hPepT1[5], whose experimental structures were not yet resolved, in order to conveniently rationalize and predict the plasma stability and the active absorption of novel derivatives. When the introduced modifications were so as to prevent the active transport, the marked hydrophilicity of carnosine analogues was modulated by designing ester- and carbamate-based prodrugs whose hydrolysis was analyzed in silico by docking simulations with the major human carboxylesterases (hCES1 and hCES2)[6]. Globally, this study shows how drug design principles and homology modeling techniques can be synergically exploited to enhance both ligand activity and pharmacokinetic properties by simulating the main biological targets influencing the carnosine bioavailability. And indeed, such computational approaches supported the design, synthesis and biological evaluation of a significant number of carnosine derivatives which can be subdivided in two main classes of (1) compounds which are quite different compared to carnosine and are endowed by a greater activity even conserving an adequate selectivity (as seen in arylderivatives)3 and (2) compounds which are strictly related to carnosine (e.g. D-carnosine derivatives), conserve a comparable activity and reach a significant bioavailability. The biological properties of such novel derivatives as evaluated by suitable in vivo models will be discussed in detail.

References:

1. Aldini G, Dalle-Donne I, Maffei Facino R, Milzani A, Carini M. Med. Res. Rev. 2007; 6:817-68
2. Aldini G, Carini M, Beretta G, Bradamante S, Maffei Facino R Biochem. Biophys. Res. Commun. 2002; 298: 699-706
3. Vistoli G, Orioli M, Pedretti A, Regazzoni L, Canevotti R, Negrisoli G, Carini M, Aldini G ChemMedChem. 2009, 4:967-75
4. Vistoli G, Pedretti A, Cattaneo M, Aldini G, Testa B J Med Chem. 2006; 49:3269-77
5. Pedretti A, De Luca L, Marconi C, Negrisoli G, Aldini G, Vistoli G. ChemMedChem. 2008, 3:1913-21
6. Vistoli G, Pedretti A, Mazzolari A, Testa B Bioorg Med Chem. 2010, 18:320-9

IS PREDICTING ACTIVE TRANSPORT NECESSARY TO PREDICT BIOAVAILABILITY?

A. Kristl, S. Žakelj,

Faculty of Pharmacy, University of Ljubljana, Aškerčeva 7, 1000 Ljubljana, Slovenia

Development of algorithms for the prediction of drug bioavailability inevitably includes mental, mathematical and experimental "dissection" of drug absorption procedure, like drug release from the pharmaceutical dosage form and the transport/absorption of the dissolved drug from the intestinal lumen to the systemic circulation. Drug permeability through the intestinal mucosa is thus one of the processes that should be very well defined and its prediction should be experimentally verified. A lot of published data are available on this subject and many institutions have their own in-house data sets for drug permeability obtained by different experimental models. While the "in vivo" methods have a clear advantage of being more realistic, the "in vitro" methods allow simple (i.e. high throughput) measurements of the drug transport in the secretory direction to be compared with the permeability measurement in the absorptive direction.

The active transport can influence the "in vitro" permeability either in the absorptive, the secretive, or in both directions. There might be some additional uncertainty because the active transport observed "in vitro" may not always be relevant to "in vivo" situation and vice-versa. It is understood that the active transport usually does not influence the bioavailability of very high permeable drugs, while the drugs with low and intermediate permeability can be significantly affected.

Although various "in silico" and "in vivo" models are highly valuable prediction tools, one must be very careful, because the active transport is so widespread and it should never be neglected. The algorithms for prediction of drug bioavailability, based on the permeability data obtained by these models, are usually correct for the drugs which are not actively transported. We therefore determined the apparent permeability coefficients (Papp) through the isolated segment of rat jejunum 'in vitro' for a set of high and low permeability drugs with the aim to estimate the influences of active transport on their permeability and absorption (bioavailability). Only two of presented drugs (antipyrine and carbamazepine) in the table have a ratio (R) between the transport in the secretory (serosal to mucosal, S-to-M) and absorptive (mucosal to serosal, M-to-S) direction close to unity.

Drug	Papp [10^{-6} cm/s] M-to-S AVG ± SD	% absorbed or bioavailability*	Permeabily classification by CDER	Papp [10^{-6} cm/s] S-to-M AVG ± SD	R
Propranolol	11.1 ± 2.3	90*	High	21.0 ± 2.3	1.9
Naproxen	37.3 ± 5.2	99*	High	18.8 ± 1.0	0.5
Antipyrine	28.7 ± 3.2	100	High	29.5 ± 4.8	1.2
Theophylline	15.5 ± 0.7	96*	High	28.2 ± 1.2	1.8
Carbamazepine	15.8 ± 0.2	92*	High	18.4 ± 1.5	1.2
Atenolol	3.7 ± 1.2	50	Low	6.5 ± 0.8	1.8
Methyldopa	2.3 ± 0.3	70*	Low	4.3 ± 0.7	1.8
Ranitidine	4.0 ± 1.1	50-60	Low	12.9 ± 0.1	3.2
Furosemide	5.8 ± 3.0	64*	Low	27.9 ± 4.7	4.7
Hydrochlorothiazide	4.4 ± 0.8	65-82	Low	12.8 ± 0.6	2.9

All these drugs are recommended permeability standards in the CDER's (FDA, Food and Drug

Administration) Guidance for industry on BCS (Biopharmaceutical Classification System) based biowaivers (1). Anyhow, the same guidance also states that "in vivo" or "in situ" animal models and "in vitro" methods, such as those using cultured monolayers of animal or human epithelial cells, are considered appropriate only for passively transported drugs. On the basis of experimentally determined Papp values in both directions and fraction of the dose absorbed we will discuss the effective impact of active transport on drug bioavailability and on its prediction. The aspect of choosing proper permeability standards for drug classification and biowaivers will also be evaluated.

References:

1. Guidance for Industry, Center for Drug Evaluation and Research (CDER), Aug. 2000
http://www.fda.gov/cder/guidance/index.htm

ORAL PRESENTATION 18

TARGET IDENTIFICATION FOR BEHAVIORAL SCREENING HITS USING A CHEMICAL SIMILARITY METHOD

C. Laggner,[1] D. Kokel,[2] V. Setola,[3] A. Tolia,[4] M. J. Keiser,[1] D. Minor,[4,5] B. L. Roth,[3,6] R.T. Peterson,[2] B. K. Shoichet[1]

[1] Dept. of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA
[2] Dept. of Medicine, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA, USA
[3] Dept. of Pharmacology, University of North Carolina Chapel Hill School of Medicine, Chapel Hill, NC, USA
[4] Dept.s of Biochemistry and Biophysics, and Cellular and Molecular Pharmacology, California Institute for Quantitative Biosciences, University of California, San Francisco, CA, USA
[5] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[6] Division of Medicinal Chemistry and Natural Products, University of North Carolina Chapel Hill School of Medicine, Chapel Hill, NC, USA

A library of 25,000 molecules was screened in the Peterson lab for behavioral effects in zebrafish embryos in the photomotor response (PMR) assay.[1] For every compound, patterns of behavioral traits were recorded; together these make up a 'behavioral barcode'. These barcodes can be used to identify novel psycho-active molecules. 1,627 hits where statistically distinct from the untreated controls and caused various patterns of increased activity or were sedating. Along with uncharacterized synthetic compounds from commercial screening libraries, many known neuroactive compounds were recovered, consistent with the effectiveness of the assay. Of particular interest to us were those compounds for which at least no neuroactive or even no specific target interactions at all had been reported so far. Clustering by behavioral and chemical similarity led to the identification of several interesting groups of related compounds. Using the chemoinformatic Similarity Ensemble Approach (SEA),[2,3,4] we predicted possible targets for the hit molecules: by comparing each hit molecule against sets of compounds annotated for a broad panel of distinct macromolecular targets, chemical similarities were reported quantitatively as expectation values. Experimental testing in cell-based assays confirmed several of the most promising predictions. For instance, a group of monoaminooxidase inhibitors with nanomolar IC_{50} values was identified.[1] Successful phenocopying of the behavioral effects with known, chemically unrelated ligands of the predicted targets strenghtens our hypothesis that the shown protein-ligand interactions also occur in the zebrafish orthologues and are responsible for the observed phenotype. In summary, combining a high-throughput phenotypic screen with bio- and cheminformatic approaches enabled not only the identification of novel psychoactive ligands but also several of their targets. The ligand-target information that emerged provides leads for neuroactive drug discovery.

References:

1) Kokel, D. et al. Rapid behavior-based identification of neuroactive small molecules in the zebrafish. Nat. Chem. Biol. 6, 231-237 (2010).
2) Keiser, M.J. et al. Relating protein pharmacology by ligand chemistry. Nat. Biotechnol. 25, 197–206 (2007).
3) Hert, J. et al. Quantifying the relationships among drug classes. J. Chem. Inf. Model. 48, 755–765 (2008).
4) Keiser, M.J. et al. Predicting new molecular targets for known drugs. Nature 462, 175-181 (2009).

ORAL PRESENTATION 19

CHARACTERISATION AND MAPPING OF LIGAND-BINDING CAVITIES IN PROTEINS

Anna Linusson[1]*, C. David Andersson[1], Brian Y. Chen[2]

[1]Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden and Computational Life Science Cluster (CLiC), KBC, Umeå University, SE-901 87 Umeå, Sweden.

[2] Howard Hughes Institute, Department of Biochemistry and Molecular Biophysics, Center for Computational Biology and Bioinformatics, Columbia University, 1130, St. Nicholas Avenue, New York, NY 10032, USA

In this study we have characterised and mapped ligand-binding cavities of proteins based on principal component analysis of cavity properties (related mainly to size, shape, polarity and charge), without direct geometric comparison of the protein structures. This approach provides information on the similarities, and dissimilarities, of binding cavities that can be used in structure-based design. For example, mutations of proteins, between-species differences and flexibility upon ligand-binding can influence the binding cavities of proteins dramatically and thus are important for molecular docking and virtual screen experiments. The results show that information obtained by our approach of mapping ligand-binding cavity variations can complement information on protein similarity obtained from sequence comparisons, which is exemplified by successful predictions of serine proteases. The presented strategy is applied to select targets for design of virtual screening experiments, investigation of binding pockets influence on docking performance and to explore the influence of side-chain flexibility for structure-based design of ligands binding to acetycholinesterase.

ORAL PRESENTATION 20

IN SILICO APPROACHES, AND IN VITRO AND IN VIVO MUTAGENICITY ASSAYS: ALTERNATIVES TO THE CARCINOGENICITY BIOASSAY

Romualdo Benigni
Istituto Superiore di Sanità, Italy

In vivo mutagenicity studies, shortly followed by carcinogenicity, pose high demand for test-related recourses: therefore, the development and use of estimation techniques such as (Q)SARs, read-across and grouping of chemicals, might have a huge saving potential for these endpoints. Structure-Activity Relationships paradigm provides a wide range of tools.
Some are coarse-grain approaches such as Structural Alerts (SA): these have a crucial role in risk assessment, for: a) description of sets of chemicals; b) preliminary hazard characterization; c) formation of categories; d) generation of subsets of congeneric chemicals to be analyzed subsequently with quantitative (QSAR) methods; e) priority setting .
On the other side, there are fine-tuned QSARs for congeneric classes of chemicals. Good quality, local QSARs for mutagenicity and carcinogenicity, when challenged for their predictivity in respect to real external test sets were 70 to 100 % correct in their external predictions.
A crucial issue is that of the uncertainty of the modeling approaches. More properly, their uncertainty should be compared with that of the competing experimental tests. For example, the ability of SAs to predict rodent carcinogenicity is of the same order of the Ames test (around 65% accuracy). Equally illuminating is the fact that the external predictivity of good local QSARs (70 to 100 % accuracy) is of the same order of the reported inter-laboratory variability of the Ames test (85%).
Thus, uncertainties are proper to both modeling and experimental systems. The crucial issue is that of exploiting and combining –at their best- both methods.

EVALUATION OF THE OECD QSAR APPLICATION TOOLBOX FOR PREDICTING THE BIODEGRADABILITY OF CHEMICALS

J. Devillers[1] and P. Pandard[2]

[1]CTIS, 3 Chemin de la Gravière, 69140 Rillieux La Pape, France.

[2]INERIS, BP2, F-60550 Verneuil en Halatte, France

Biodegradation is an important mechanism for eliminating xenobiotics by biotransforming them into simple organic and inorganic products. Two types of aerobic biodegradation can be distinguished. The primary biodegradation which denotes a simple transformation not leading to a complete mineralization. The biodegradation products are specifically measured from chromatographic methods and the results are expressed by means of kinetic parameters such as biodegradation rate constant and half-life. The ultimate (or total) biodegradation totally converts chemicals into simple molecules such as CO_2 and H_2O. Faced with the ever growing number of chemicals available on the market, quantitative structure-property relationship (QSPR) models are increasingly used as surrogates of the biodegradation tests. Such models have great potential for a quick and cheap estimation of the biodegradation potential of chemicals.

The OECD QSAR Application Toolbox, which was developed for facilitating the practical use of (Q)SAR approaches in regulatory contexts, includes six different models for predicting the biodegradability of organic substances. They are based on different endpoints, methodologies and/or statistical approaches. The aim of this study was to estimate the usefulness of these six models. First, the models were critically analyzed to estimate whether they satisfied the Setubal/OECD principles for (Q)SAR validation. Then, they were tested on more than 400 chemicals, extracted from notification dossiers, and hence, never used for the design or the validation of models. A critical analysis of the results was performed and the advantages and limitations of each model were clearly identified. Proposals for increasing the performances of the biodegradation module of the OECD QSAR Application Toolbox were made.

Acknowledgment

The authors are grateful to the French Ministry of Ecology, Energy, Sustainable Development and Sea (MEEDDM) for financial support.

ORAL PRESENTATION 22

CLASSIFICATION- AND REGRESSION-BASED QSAR OF ACUTE CHEMICAL RODENT TOXICITY

O.A.Raevsky [a], V.Yu.Grigor'ev , [a] E.A.Modina a and A.P.Worth [b]

[a] Institute of Physiologically Active Compounds of Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia;

[b] Institute for Health and Consumer Protection, European Commission - Joint Research Centre, Via Enrico Fermi 2749, 21027 Ispra (Va), Italy

One of the aims of the EU chemicals legislation REACH is to reduce the amount of animal testing through the use of alternative methods that are not based on vertebrate animal testing, for example, *in vitro* methods, qualitative or quantitative structure-activity relationship methods, or information from structurally related substances (grouping or read-across) [1]. Acute *mammalian* toxicity is a very important toxicological endpoint which reflects a range of complex phenomena associated with the non-specific interactions of chemicals with membranes as well as the specific binding with a large number of biological targets inside organisms. As result of the complex and incomplete understanding of the mechanisms involved, the number of publications devoted to QSAR modeling of acute *mammalian* toxicity is rather small [2]. A recent detailed analysis of 150 QSAR models of acute toxicity to the mouse and the rat from published articles led to the conclusion that many available models have limited utility for the purpose of LD50 predictions due to their poor or modest statistical quality or because they were obtained from limited data sets [3].

In the present study, a data set of acute toxicity to mice following intravenous injection was selected from the SYMYX Toxicity Database [4]. The selected experimental toxicity values in mg/kg were converted to $\log(1/LD_{50})$ (LD_{50} *in mmol/kg*) units and used as a quantitative measure of acute toxicity. A total of 10,241 neutral organic compounds had LD_{50} values in this data base and were included in the studied data set. The selected data contain LD_{50} values for mice of different sex, age, strain, and experimental conditions. These variables have a influence on toxicity and result in a variability in $\log(1/LD_{50})$ values of at least ± 0.50. Four modelling approaches have been proposed and applied in the development of QSARs for acute intravenous toxicity to mice:

1. The Arithmetic Mean Toxicity (AMT) modelling approach [5]. The AMT approach is based on the use of one or a few pairs of nearest structural neighbours. Each pair contains a chemical with a higher descriptor value and with a smaller descriptor value compared with the chemical of interest. The arithmetic mean of the toxicity values of the pairs are considered as the toxicity of chemical of interest. The toxicity of the chemical of interest was not included in the development of the AMT model. For chemicals having at least four neighbours with $Tc \geq 0.40$:

$\log(1/LD_{50})_{exper} = 0.06(\pm 0.01) + 0.92(\pm 0.01) * \log(1/LD_{50})cal.\ AMT$

$n=7085,\ R^2=0.557,\ sd=0.43,\ Q^2=0.577,\ sd_{cv}=0.43,\ F=9672.4$

For chemicals from the REACH list of Pre-Registered Substances (PRS)list having four neighbours in the data set with $Tc \geq 0.40$:

$\log(1/LD_{50})_{exper} = -0.01(\pm 0.02) + 1.02(\pm 0.02) * \log(1/LD_{50})cal.\ AMT$

$n=1198,\ R^2=0.596,\ sd=0.51,\ Q^2=0.594,\ sd_{cv}=0.52,\ F=1764.8$

2. The Super Overlapping Cluster/Regression (SOCR) QSAR modelling approach. The model is based on three steps: 1. selection of structurally similar chemicals for each chemical-of-interest (clustering); 2. construction of QSARs for each cluster without the chemical-of-interest; 3. the application of constructed QSARs for toxicity estimation to chemicals-of-interest. The equation parameters for 5489 chemicals with

Tc≥0.40 and a mutual descriptor regression coefficient threshold at the level r≤0.5 are presented below:
$\log(1/LD_{50})_{exper} = 0.07(\pm 0.01) + 0.89(\pm 0.01) * \log(1/LD_{50})_{cal.\ SOCR}$
$n=5489$, $R^2=0.538$, sd $=0.45$, $Q^2=0.537$, $sd_{cv} =0.45$, $F=6388.0$
3. Mode of Action (MOA) based application for toxicity estimation of non specific chemicals (bilinear model):
$\log(1/LD_{50})_{exper} = -1.28(\pm 0.02) + 0.81(\pm 0.01) * MlogP - 1.155(\pm 0.018) * \log(0.0028442MP+1)$
$n=3821$, $R^2=0.669$, sd $=0.25$, $Q^2=0.668$, $sd_{cv} =0.25$, $F=3850.0$
4. A consensus model based on the mean values predicted by the AMT and SOCR approaches for chemicals with Tc≥0.70.
$\log(1/LD_{50})_{exper} = 0.00(\pm 0.02) + 1.05(\pm 0.03) * \log(1/LD_{50})_{cal.(AMT+SOC)/2}$
$n=158$, $R^2=0.918$, sd$=0.17$, $Q^2=0.901$, $sd_{cv} =0.20$, $F=1753.8$

Acknowledgement
This investigation was supported by the International Science & Technology Centre (Project #3777).

References
[1] Regulation (EC) 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning REACH;
[2] I.Tsakovska, I.Lessigiarska, T.Netzeva and A.Worth, *A Mini Review of Mammalian Toxicity (Q)SAR Models,* QSAR & Combinatorial Science 27 (2008), pp. 41-48;
[3] J.Devillers and H.Devillers, Prediction of acute mammalian toxicity from QSARs and interspecies correlations, *SAR & QSAR in Environmental Research,* **2009**, 20, 467-500;
[4] SYMYX Toxicity Database. Available at htttp://www.symyx.com/ /products/databases/index.jsp.
[5] O.A. Raevsky, V.Yu.Grigor'ev, E.A. Modina and A.Worth, Prediction of Acute Toxicity to Mice by the Arithmetic Mean Toxicity (AMT) Modelling Approach, *SAR & QSAR in Environmental Research,* **2010**, 21, №2 3/4.

ORAL PRESENTATION 23

FROM ACTIVITY CLIFFS TO TARGET-SPECIFIC SCORING MODELS AND PHARMACOPHORIC HYPOTHESIS

Birte Seebeck[1], Markus Wagener[2] and Matthias Rarey[1]
[1]University of Hamburg, Center for Bioinformatics Hamburg (ZBH), Bundesstrasse 43, 20146 Hamburg, Germany
[2]MSD, Molecular Design & Informatics, 5340BH Oss, The Netherlands

Modeling structure-activity relationships still remain a challenge in the drug discovery process. It underlies the two principles that specific molecular structure leads to biological activity and molecules with similar structures will possess similar activities. The question is how do compounds which are structurally similar, but yet have very different activities fit into this schema? The so-called term "activity cliffs" was introduced by Maggiora to describe such cases [1]. Activities cliffs are often the critical point why machine learning based techniques fail to derive a predictive model. Although on the first view activity cliffs seem to be the pitfalls of QSAR modeling techniques, they are very valuable for identifying key aspects of an SAR. Activity cliffs are puzzling in purely ligand-based design approaches, they can be much better understood if the protein structure is available. Here, we present a first approach to use this valuable information of activity cliffs in a structure-based design scenario; on the one hand to visualize key interactions in the active site of a target protein, on the other hand to derive a target-specific scoring model and a pharmacophoric hypothesis.

Using protein-ligand interaction energies in QSAR approaches has also been shown in the comparative binding (COMBINE) analysis by Ortiz et al. and was successfully applied on various targets over the last decade [2]. A method to identify and quantify activity cliffs was proposed by Guha and Van Drie by introducing the structure-activity landscape index (SALI) [3]. SALI is based on the principle to summarize SAR data sets by looking to pairs of compounds to identify pairs that are most similar in structure but have the largest change in activity. Considering pairs of compounds rather than looking to compounds in isolation is an intuitive way to analyze a structure-activity relationship in order to identify structural motifs which cause a change in activity.

Based on the principles of SALI we developed a structure-based approach of analyzing activity cliffs by using interaction energies from protein-ligand complexes. Interaction energies are calculated for each single receptor atom and a certain interaction type, based on a standard empirical scoring scheme [4]. For the calculation of the structure-based SALI values, we use the tanimoto similarity between receptor-atom-wise and interaction-type-based binary fingerprints to characterize the SAR landscape.

Based on the relative frequencies a receptor atom is involved in activity cliff events, we demonstrate an activity cliff "hot spot" visualization in the active site of the protein for different interaction types (i.e. H-bonds, ionic interactions, aromatic interactions and hydrophobic contacts). This approach supports the medicinal scientist to elucidate the key interacting atoms of the active site and facilitates developing a pharmacophoric hypothesis. In detail, the visualization enables the study of receptor atoms that are predominantly found in interactions with the active compounds in comparison to the less or non-active compounds with respect to activity cliff events at a certain SALI cutoff level. In early project phases, it can be applied to examine the first SAR trends for lead identification. In addition to the activity cliff "hot spots" visualization, we present another visualization mode that allows evaluating the frequencies and variances of interactions in the data set, independently from the activity cliff events. This mode is a valuable extension

to identify receptor atoms that almost always show interactions but going along with a low variance in the data set. From a machine learning point of view, no information can be drawn from those interactions in terms of the considered data set, however, they still might have an important contribution to the overall affinity.
We further present an approach to derive a target-specific scoring model and a pharmacophoric hypothesis on the basis of the structure-based analysis of activity cliffs. The target-specific scoring model was tailored by introducing additional weight factors on a per receptor atom basis to the standard empirical scoring function. The selection of pharmacophoric constraints is assisted by the activity-cliff-based evaluation of interaction frequencies in combination with the analysis of the independent or correlated occurrence of key interactions to derive logical constraint expressions. The target-specific scoring models and pharmacophoric constraints have been validated in enrichment experiments on independent external test sets. The results show an improved enrichment in comparison to the standard score for various protein targets.

References:

[1] Maggiora G. M., J. Chem. Inf. Model., 2006, 46: 1535.
[2] Ortiz A.R., Pisabarro M.T., Gago F. and Wade R.C., J. Med. Chem., 1995, 38(14): 2681-91.
[3] Guha R. and Van Drie J.H., J. Chem. Inf. Model., 2008, 48(3): 646-658.
[4] Stahl M. and Rarey M., J. Med. Chem. 2001, 44(7): 1035-42.

TEMPLATE-CONSTRAINED FRAGMENT ALIGNMENT (TCFA)

Richard Cramer
Tripos International, USA

The two general approaches for generating the alignments necessary in 3D-QSAR are the well-established "classical" approaches based on receptor or pharmacophoric information and the recently introduced topomer protocol. Template-constrained fragment alignment (TCFA) is a third approach, to be described publicly for the first time. With TCFA, the scientist provides side chain fragmentation instructions for the training set structures, just as in the topomer approach, but the 3D geometries of these fragments now become "templates" for the construction of the 3D-QSAR alignments, essentially by copying the coordinates for the template atoms wherever they sufficiently match atoms in the structure being aligned, and reserving the topomer protocol for the unmatched atoms. TCFA combines the 3D-shape specificity and control afforded by the "classical" alignment procedures with the productivity of the topomer alignment procedure. The TCFA protocol is also useful in similarity-only searching by yielding only hits that can overlay well with a template shape.

ORAL PRESENTATION 25

ENHANCING MOLECULAR DESIGN VIA A MULTI-OBJECTIVE APPROACH

Orazio Nicolotti, Ilenia Giangreco, Teresa Fabiola Miscioscia, Angelo Carotti
Dipartimento Farmaco-Chimico, Università di Bari
Via Orabona 4, 70125, Bari, Italy

The successful design of new biologically active molecules is upon the control of a number of often conflicting objectives. In this regard, a few examples are molecular diversity and drug-likeness in designing combinatorial libraries; steric similarity and energy in pharmacophore generation; statistical accuracy and chemical desirability in QSAR modelling; scores and ligand displacements in the selection of docking poses; optimization of ADME properties in drug discovery.
Despite this real-life evidence, standard methods are generally based on single-objective optimization assuming that even for highly complex problems only one single optimal solution exists. The present work describes a multi-objective optimization (MOOP) technique that is aimed at finding trade-offs among various objectives and, thus, at identifying not one but a family of equivalent optimal models. In particular, the herein proposed approach enabled the non-deterministic derivation of a number of equivalent 3D-QSAR models accounting simultaneously for two independent objectives: the first biased toward best correlation among docking scores and biological affinities and the second minimizing the distance from a properly established crystal-based binding mode.
The MOOP strategy was successful challenged on a large series, very well known to QSAR practitioners, of 3-amidinophenylalanine inhibitors of serine proteases.
Selected on the basis of the concept of Pareto dominance, the derived 3D QSAR equivalent models demonstrated that a) trade-off binding conformations can actually occur on the basis of the ligand capability to effectively engage diverse regions of the protein binding site; b) molecular selectivity was better approached and interpreted by means of the pool of equivalent models; c) trade-off models were successful in enrichment experiments of virtual screening.

References:

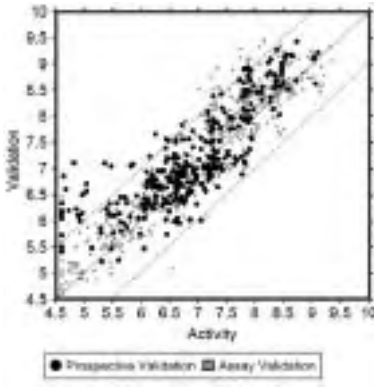
• Böhm, M.; Stürzebecher, J.; Klebe, G. *J. Med. Chem*. **1999**, 42, 458-477
• Nicolotti, O.; Gillet, V. J.; Fleming, P.; Green, D. *J. Med. Chem.* **2002**, 45, 5069-5080
• Nicolotti, O.; Miscioscia, T. F.; Carotti, A.; Leonetti, F.; Carotti, A. *J. Chem. Inf. Model.* **2008**, 48, 1211-1226
• Nicolotti, O.; Giangreco, I.; Miscioscia, T.F.; Carotti A. *J. Chem. Inf. Mod.* **2009**,49, 2290-2302
• Nicolotti, O.; Giangreco, I.; Miscioscia, T.F.; Carotti A. *QSAR & Comb. Sci.* **2009**, 28, 861-864
• Nicolotti, O.; Giangreco, I.; Miscioscia, T.F.; Convertino, M.; Leonetti, F.; Pisani, L.; Carotti A. *J. Comput. Aid. Mol. Design Pub.on line: Feb*. **2010**

ORAL PRESENTATION 26

PROSPECTIVELY VALIDATED PROTEOCHEMOMETRIC MODELS OF HIV REVERSE TRANSCRIPTASE AS A TOOL IN LEAD OPTIMIZATION AGAINST MULTIPLE TARGETS

G.J.P. van Westen[1], J.K. Wegner[2], A.P. IJzerman[1], H.W.T. van Vlijmen[2], A. Bender[1]
[1]Division of Medicinal Chemistry, LACDR, Einsteinweg 55, 2333 CC, Leiden, The Netherlands
[2] Tibotec BVBA, Generaal de Wittelaan 11B, 2800 Mechelen, Belgium

Chronic HIV infection requires patients to follow a strict drug regimen for their entire life. The onset of resistance of the virus towards medication is minimized by combining three or more drugs in Highly Active Anti Retro-viral Therapy (HAART). However, the high mutation rate of HIV can still lead to the onset of resistance. Therefore an HIV inhibitor should be active against a broad spectrum of mutants; however, testing all possible combinations of lead compounds on all possible mutants is virtually impossible. Proteochemometric modeling [1] adds a target description, based on physicochemical properties of the binding site, to conventional, ligand-based bioactivity models. Therefore proteochemometric models can potentially extrapolate the activity of compounds to novel, related targets, such as in this case HIV mutants. Our proteochemometric models [2] are based on Scitegic circular fingerprints on the compound side and on a customized protein fingerprint on the target side which is based on a selection of physicochemical descriptors obtained from the AAindex database. Using these descriptors we constructed a proteochemometric model of 450 non-nucleoside reverse transcriptase inhibitors and 14 point mutated HIV reverse transcriptase mutants, these sequences can be considered 14 individual patients. In total a pEC_{50} value was available for about 60 % of the ligand-mutant pairs. With our model we predicted pEC_{50} values for the missing 40% of the pairs which were prospectively validated.



In total 317 data points were experimentally validated after we had predicted their values with our model. When our model predictions were compared with the experimental pEC50 values, we obtained an R^2 of 0.69, an R_0^2 of 0.69 and an RMSE of 0.62 log units. In the same experiment 518 compound – sequence pairs were retested to characterize the assay error, which gave an assay concordance R^2 of 0.88, an R_0^2 of 0.88 and an RMSE of 0.49 log units. In the same experiment, conventional QSAR performed poorly with an R^2 of 0.33, an R_0^2 of 0.31 and an RMSE of 0.96 log units.

We conclude that the proteochemometric model developed here is able to reliably extrapolate the activity of compounds to new targets. This ability makes it a useful tool to steer the development of novel drugs by virtually screening ten tot hundred fold more compounds than would be possible with conventional screening tools like HTS. At the same time it has proven to be far more reliable than conventional computational tools like QSAR. This approach can also be applied on predicting the activity of ligands on related receptors such as the GPCR family (which is also subject of our ongoing work).

References:

1. Freyhult E, Prusis P, Lapinsh M, Wikberg JE, Moulton V, Gustafsson MG (2005) Unbiased descriptor and parameter selection confirms the potential of proteochemometric modelling. BMC Bioinformatics 6:50-64.
2. Doddareddy MR, van Westen GJP, van der Horst E, Peironcely JE, Corthals F, IJzerman AP, Emmerich M, Jenkins JL, Bender A (2009) Chemogenomics: Looking at biology through the lens of chemistry. Statistical Analysis and Data Mining 2:149-160.

ORAL PRESENTATION 27

COMPUTATIONAL DRUG DESIGN STUDIES ON ANTITUMORAL ACTIVE HETEROCYCLIC COMPOUNDS

Esin Aki, İsmail Yalçin, İlkay Yildiz, Tugba Ertan-Bolelli, Kayhan Bolelli, Serap Yilmaz

Ankara University, Faculty of Pharmacy, Pharmaceutical Chemistry Dept., Tandogan 06100 Ankara, Turkey

In the recent years, most effective therapeutically active drugs has been discovered and put into market by the use of computer aided drug design studies.

Eukaryotic topoisomerases is widely used in the anticancer drug development, that they are essential for several cellular processes such as replication, transcription and chromosome condensation. Antitumor activity is related to the formation of protein-concealed DNA strand breaks, resulting in the stabilization by the drug of an intermediary complex of the Topo II reaction.

The rational design of antitumoral active new heterocyclic compounds possessing benzoxazole, benzimidazole, benzothiazole and oxazolopyridine fused ring systems has been studied by our research team by testing their DNA Topoisomerase I and II inhibitory activities.

Some 3D-QSAR studies like COMFA (Comparative Molecular Field Analysis) and CoMSIA methods for the lead optimisation using the Sybyl 6.6 Software at SGI workstation were performed.

Three-dimensional pharmacophore hypotheses generated by Catalyst/HipHop from Accelrys in SGI Workstation were also used at our work. Molecules were edited using the Catalyst 2D/3D visualizer which Catalyst automatically generated conformational models for each compound using the Poling Algortihm.

For the lead optimisation and the generation, the results obtained from these molecular modelling studies including the docking procedure have been discussed.

References:

[1] J.C. Wang, Ann. Rev. Biochem. 65, 635-692, 1996.
[2] G.L. Chen, L. Yang, T.C. Rowe, B.D. Halligan, K.M. Tewey, L.F., J. Biol. Chem. 259, 13560-13566, 1984.
[3] K.H. Cho, J.M. Pezzuto, J.L. Bolton, V.E. Steele, G.J. Kelloff, S.K. Lee, Eur. J. Cancer, 36, 2146-2156, 2000.
[4] K.F. Koehler, S.N. Rao, J.P. Snyder, Guidebook on Molecular Modeling in Drug Design (Cohen NC ed) pp 235-336, Academic Press, San Diego, 1996.
[5] A. Pinar, P. Yurdakul, I. Yildiz, O. T. Arpaci, N.L. Acan, E.S. Aki I. Yalcin, Biochim. Biophys. Res. Comm., 317, 670-674, 2004.
[6]R.D. Cramer III, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc., 110, 5959-5967, 1988.
[7]G. Klebe, U. Abraham, T. Mietzner, J. Med. Chem., 37, 4130, 1994.
[8] SYBYL 7.0, Tripos Associates, St. Louis, Mo., USA.
[9] Accelrys Inc. Catalyst 4.11 (2005)

ORAL PRESENTATION 28

NEW LEADS FINDING IN AGROCHEMISTRY: A COMPUTATIONAL CHEMISTRY CHALLENGE

Francesca Perruccio, Andrea Bortolato
Syngenta Crop Protection Muenchwilen AG, Schaffhauserstrasse CH-4332 Stein, Switzerland

A common strategy to investigate chemistry scaffolds is the purchase of external commercial libraries, databases and compounds collections.
Computational tools are required to investigate these compounds on offer for multiple reasons. As first step in this kind of exercise, it is required an assessment of the diversity and the novelty that external compounds could add to the existing in house knowledge. Computational approaches can be also applied to classified external scaffolds by their structural features similarity to ongoing project, and for this reason could be of interest to perform scaffold hopping or new lead generation.
The final aim is to enrich the in house database with novel chemistry and/or to select compounds to be followed up in a particular project. In this presentation implementation and results from different computational approaches used in Syngenta Stein are going to be discussed.

ORAL PRESENTATION 29

HYDE SCORING OF PROTEIN LIGAND COMPLEXES

Gudrun Lange[1], Schneider Nadine[2], Klein Robert[1], Rarey Matthias[2]
[1]Bayercropscience, AG, Industriepark Hoechst, Frankfurt am Main, Germany
[2]University of Hamburg, Hamburg, Germany

Scoring functions are used to describe the interaction between molecules such as between proteins or the binding of ligands to target proteins. They help to identify the correct pose of a ligand with known inhibition and to suggest sensible chemical modification of the ligand (Structure-Based Drug Design). In addition, they are used to score poses of putative ligands positioned into a target protein by a pose generator and to select those ligands which bind best to the target (Virtual Screening).

Since only limited success was achieved at BCS using commercially available software, we have developed a novel scoring function (HYDE) (1). HYDE is a simplified description of the underlying physics and includes two terms only: (a) the dehydration of atoms in the respective interface and (b) the H-bond energies of interacting H-bond functions. The dehydration is calculated using logP increments derived from experimental logP values of small organic molecules. The H-bond contribution of H-bond functions is calculated using their dehydration contribution and a novel dehydration term based on structural features of water. Since no other data is included, HYDE is a general applicable scoring function with a universal cut-off score above which binding is assumed not to occur.

HYDE can be used as a tool to analyze crystal structures. We will show using several structures including small or large and weakly or strongly binding inhibitors that (a) the H-bond energies derived from the logp values and the contribution of H-bonds in aqueous solution agree well with experimental observations (2), (b) the hydrophobic effect is reasonably described (3), (c) the affinity loss due to unsatisfied H-bond functions or due to H-bonds with bad geometry can be explained (4). In addition, we show that in many cases the HYDE score corresponds to the measured affinity. In particular, it becomes clear, that a single atom replacement in a protein ligand interface can lead to a significantly reduced affinity (selectivity) and that in some interfaces water molecules play a crucial role.

HYDE has been successfully used in BCS projects for structure based design of pesticides, analysis of species dependent differences, identification of targets for in-vivo active pesticides and in Virtual Screening where we were able to identify binders including in-vivo active pesticides from uncharacterized libraries with a success rate of 40-70%.

References:

(1) Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. *ChemMedChem* **2008**, 3, 885-897
(2) Fersht, A.R., Shi, J.P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J:, Blow, D. M., Brick, P., Carter, P., Waye, M.M.Y., Winter, G. *Nature* **1985**, 314, 235
(3) Reynolds, J., Gilbert, D. B., Tanford, C. *Proc. Natl. Acad. Sci. U. S .A.* **1974**, 71, 2925-7
(4) G. Lange, D. Lesuisse, P. Deprez, B. Schoot, P. Loenze, D. Bénard, J.-P. Marquette, P. Broto, E. Sarubbi, E. Mandine,.J. *Med. Chem.* **2003**, 46, 5184

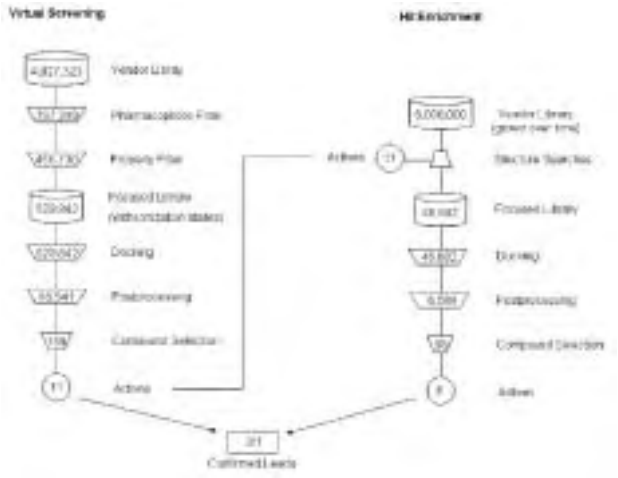
ORAL PRESENTATION 30

INHIBITION OF *EIMERIA TENELLA* CDK-RELATED KINASE 2: FROM TARGET IDENTIFICATION TO LEAD COMPOUNDS

Kristin Engels[1], Carsten Beyer[1], Maria L. Suárez Fernández[1], Frank Bender[1], Michael Gaßel[1], Gottfried Unden[2], Richard J. Marhöfer[1], Jeremy C. Mottram[3], and Paul M. Selzer[1,3,4]

[1] Intervet Innovation GmbH, Zur Propstei, 55270 Schwabenheim, Germany;
[2] Institut für Mikrobiologie und Weinforschung, Johannes-Gutenberg-Universität, Mainz, Germany;
[3] Wellcome Trust Centre for Molecular Parasitology and Division of Infection & Immunity, Faculty of Biomedical and Life Sciences, University of Glasgow, G12 8TA, UK;
[4] Eberhard-Karls Universität Tübingen, Interfakultäres Institut für Biochemie, Hoppe-Seyler-Str. 4, 72076 Tübingen, Germany;
(a) current address: Bayer Schering Pharma AG, Medicinal Chemistry, Aprather Weg 18a, 42096 Wuppertal, Germany;
(b) current address: BASF SE, Computational Chemistry & Biology, 67056 Ludwigshafen, Germany

Apicomplexan parasites encompass several human-pathogenic as well as animal-pathogenic protozoans like Plasmodium falciparum, Toxoplasma gondii, and Eimeria tenella. E. tenella is the causative agent of coccidiosis a disease of chickens, which causes tremendous economic losses to the world poultry industry. Considerable increase of drug resistance makes it necessary to develop and pursue new therapeutic strategies. Cyclin-dependent kinases (CDKs) are key molecules in the regulation of the cell cycle and are therefore prominent target proteins in parasitic diseases. Bioinformatic analysis revealed four potential CDK-like proteins of which one – E. tenella CDK-related kinase 2 (EtCRK2) – is already cloned, expressed and characterized.[1] Using the CDK specific inhibitor Flavopiridol in EtCRK2 enzyme assays and schizont maturation assays we could chemically validate CDK-like proteins as potential drug targets. An X-ray crystal structure of human CDK2 (HsCDK2) served as template to built protein models of EtCRK2 by comparative homology modeling. Structural differences in the ATP-binding site between EtCRK2 and HsCDK2 as well as chicken CDK3 have been addressed for the optimization of selective ATP-competitive inhibitors.



Virtual screening and "wet-bench" high throughput screening campaigns on large compound libraries resulted in an initial set of hit compounds. These compounds were further analyzed and characterized leading to a set of four promising lead compounds inhibiting EtCRK2.

Reference:

[1] J. H. Kinnaird, J. M. Bumstead, D. J. Mann, R. Ryan, M. W. Shirley, B. R. Shiels, F. M. Tomley (2004) EtCRK2, a cyclin-dependent kinase gene expressed during the sexual and asexual phases of the Eimeria tenella life cycle. *Int. J. Parasitol.* **34**, 683-692.

ORAL PRESENTATION 31

BACK TO THE ROOTS – BENEFITS AND LIMITATIONS CONCERNING THE IN SILICO INTEGRATION OF NATURAL PRODUCTS IN DRUG DISCOVERY

Irene Kouskoumvekaki
Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, DK-2800 Lyngby, Denmark

Current drug discovery depends largely on random screening of small compound libraries, either high-throughout screening (HTS) *in vitro*, or virtual screening (VS) *in silico*. Despite its history of success, this strategy is nowadays facing the challenge of "more investment - less outcome". Global expenditure on research has doubled since 1991, but the number of new entity drugs approved annually has fallen by 50% or even more. To change this situation, pharmaceutical companies are attempting to alter their paradigms from a random to a more rationally designed screening. One such attempt is based on the hypothesis that biologically relevant compounds may have better chances in becoming successful drugs than their synthetic counterparts. Medicinal plants have always played an essential role in the healthcare of many different cultures, and according to the World Health Organization (WHO), approximately 65% of the world's population relies mainly on plant-derived traditional medicines for their primary health care.

In this work we advocate the scope for integration of natural compounds in *in silico* drug discovery, via chemogenomics profiling of CNPD (Chinese Natural Product Database), a large database of natural compounds derived primarily from Chinese herbs. This gives an overview of the structural characteristics of bioactive natural compounds and the chemical space they occupy in relation to synthetic drugs and drug candidates and it also provides a means for evaluation of the extend of their polypharmacology. Additionally, we demonstrate the close relatedness of natural compounds from CNPD and human metabolites from HMDB (Human Metabolite Database) and we further illustrate how similarity-based networks between the two data sets could assist in the identification of novel lead structures as well as the elucidation of the molecular mechanisms of plants used in ethnopharmacology[1].

Following this, two examples of chemoinformatics-based ligand discovery from natural sources will be presented. In the first example, a PPARγ partial agonist was identified via pharmacophore-based virtual screening of CNPD, in the search for ligands with antidiabetic properties[2]. In the second example, a ligand-based chemoinformatics approach was applied for the first time in the area of bacterial biofilm inhibition, where structure-based virtual screening of CNPD revealed two compounds that abolish biofilm formation of both *S. aureus* and *S. dysgalactiae*. Such small plant components, with bacterial lifestyle altering properties are promising candidates for novel generations of anti-microbial drugs[3].

With the plethora of new emerging targets for drug discovery and the continued demand for new chemotypes, natural product libraries could serve as a new source of lead structures with biological activity.

References:

1. I Kouskoumvekaki and G Panagiotou 'Navigating the human metabolome for biomarker identification and design of pharmaceutical molecules' *submitted to J Biomed Biotechnol*

2. RK Petersen, KB Christensen, X Fretté, AN Assimopoulou, VP Papageorgiou, K Kristiansen and I Kouskoumvekaki 'Pharmacophore-based screening for the identification of novel PPAR-γ partial agonists from natural sources' *in preparation*

3. A Dürer, I Kouskoumvekaki, RB Vejborg and P Klemm 'Identification by Virtual Screening of new Generation of Bacterial Biofilm Inhibitors' *Appl Microbiol Biotechnol (2010) Epub*

ORAL PRESENTATION 32

CAPTURING SAR-TRENDS FROM CHEMOGENOMICAL SPACES

Bernd Wendt
Elara Pharmaceuticals GmbH, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Recent years have seen huge efforts in creating databases that store chemical structures with their biological activities for the public domain. With its new release Chembl[1] stores more than 520,000 individual compound records with 2.4 million records of their effects on biological systems making it one of the richest chemogenomical spaces currently available. In order to be useful for rational design of potent and selective lead structures it is necessary to extract structure-activity-relationships from this wealth of information. Mining databases the size of Chembl requires fast, robust and effective workflows.

We have developed a procedure around a new method: quantitative series enrichment analysis (QSEA)[2] that enables fully automated 3D-QSAR model creation and prediction. SAR-tables, lists of structures with associated activities used as input for the procedure, are constructed from shape similarity searches[3] using a collection of 255 marketed drugs[4] as queries. The visual output of the procedure reveals crucial information about the applicability domain of each series within a SAR-table together with insights where and how structural changes in the series might affect the biological activity. Results from mining Chembl, PubChem and ChemBank databases will be reported.

References:

[1] http:// www.ebi.ac.uk/chembldb/
[2] Wendt, B.; Cramer, R.D. J. Comp. *Aided Mol. Des.* **22**, 541-555 (2008)
[3] Jilek, R. J.; Cramer, R.D., *J. Chem. Inf. Comp. Sci.* **44**, 1221-1227 (2004)
[4] Cleves, A.E.; Jain, A.N. *J. Med. Chem*, 2006, Vol.49, 2921-2938

ORAL PRESENTATION 33

MINING EXHAUSTIVELY THE PROTEIN DATA BANK ENABLES COMPUTATIONAL FRAGMENT-BASED DRUG DESIGN

Fabrice Moriaud, Olivia Doppelt-Azeroual, Laetitia Martin, Ksenia Oguievetskaia, Artem Vorotyntsev, Stewart A. Adcock, and Francçois Delfaud
MEDIT S.A., 2 rue de Belvedere, 91120 Palaiseau, France

Based on the assumption that similar protein surfaces are likely to bind the same fragment with the same pose, an efficient protocol in computational FBDD (Fragment-Based Drug Design) is proposed.
The large volume of protein-ligand structures now available in the PDB enables applications of the protocol for diverse fragments and for any protein families.
In this work, we build a database of MED-Portions, where a MED-Portion is a new structural object encoding protein-fragment binding sites. MED-Portions are derived from mining all available protein-ligand structures with any library of small molecules. Combined with the MED-SuMo software to superpose similar protein interaction surfaces, pools of matching MED-Portions can be determined for any binding surface query.
To generate hitlike molecules from fragments, MED-Portions are combined in 3D with the MED-Hybridise toolkit. The described MED-Portion/MED-SuMo/MED-Hybridise protocol is applied to three targets: a protein kinase, a G-Protein Coupled Receptor (GPCR), and Eg5, a mitotic kinesin. The results show the potential for finding relevant compounds targeting any protein 3D structure since the occurrence of interfamily MED-Portions is 25% for protein kinase, almost 100% for the GPCR and for the hydrophobic pocket of the Eg5 allosteric pocket.

ORAL PRESENTATION 34

INVESTIGATION OF THE STRUCTURAL REQUIREMENTS FOR MULTI-KINASE INHIBITION USING GQSAR METHOD

Subhash Ajmani
NovaLead Pharma Pvt. Ltd.,
Pride Purple Coronet, 1st floor, S No. 287, Baner Road, Pune 411 045, India

The inhibition of tyrosine kinase-mediated signal transduction pathways represents a therapeutic approach to the intervention of proliferative diseases such as cancer, atherosclerosis, and restenosis. Many literature reports suggest that drugs against multiple targets may overcome many limitations of single targets and achieve a more effective and safer control of the disease. However, design of multi-target drugs presents a great challenge. The present study demonstrates application of a novel Group based QSAR (GQSAR) method to assist in understanding structural requirements for design of multi-tyrosine kinase (PDGFR, FGFR and c-SRC kinases) inhibitor. For GQSAR analysis, a wide variety of structurally diverse multi-tyrosine kinase inhibitors (225 molecules) were collected from various literature reports. Each molecule in the data set was divided into four fragments and their corresponding two-dimensional fragment descriptors were calculated. The multi-response regression GQSAR models suggest the key role of nature of polar groups and corresponding polar surface area of a terminal region (fragment 4) of the molecule in influencing inhibitory activity of all the three kinases. In addition, the GQSAR models provide other important fragment based features that can form the building blocks to guide combinatorial library design in the search for optimally potent multi-tyrosine kinase inhibitors. The developed GQSAR models could be used for screening newly designed multi-tyrosine kinase inhibitors.

ORAL PRESENTATION 35

UNDERSTANDING THE SELECTIVITY OF ORGANOPHOSPHORUS INHIBITORS OF SERINE ESTERASES

Vladimir A. Palyulin[a,b], Eugene V. Radchenko[a,b], Andrey A. Melnikov[b], Igor I. Baskin[a,b], Galina F. Makhaeva[b], Nikolay S. Zefirova,[b]

[a] Department of Chemistry, Moscow State University, Moscow, 119991 Russia

[b] Institute of Physiologically Active Compounds RAS, Chernogolovka, Moscow region, 142432 Russia

Organophosphorus compounds are valuable drugs or drug candidates for treatment of the Alzheimer's disease, schistosomiasis, glaucoma and other disorders. They are also widely used in industry and may pose safety risks due to their acute and delayed neurotoxicity. Their observed physiological effects are determined by the ability to inhibit in different extent several serine esterases: acetylcholinesterase, butyrylcholinesterase, carboxylesterase, and neuropathy target esterase, via covalent modification of serine in their active sites. Thus, understanding and controlling the 'esterase profile' of compound activity and selectivity towards four target enzymes is of great interest.

In this study, the activity and selectivity of several series of antiesterase compounds containing the phosphate, thiophosphate, methylphosphonate and phenylphosphonate moieties with different leaving groups were analyzed using a variety of QSAR approaches, including Molecular Field Topology Analysis (MFTA) [1], linear regression and back-propagation neural networks for the fragmental descriptors with labeled atoms, and Comparative Molecular Similarity Index Analysis (CoMSIA), as well as molecular modelling. The consistent and mutually complementary information from all these methods provides insight into the structural factors and features that govern the activity profile [2-3].

Based on the QSAR models, targeted structure generation [4] and virtual screening [5], a number of promising structures were identified that are expected to possess a specific esterase profile. It is shown that this approach opens the way to design potential new anti-Alzheimer drugs with optimal activity and minimal side effects.

References:

[1] Palyulin V.A., Radchenko E.V., Zefirov N.S. *J. Chem. Inf. Comput. Sci.*, **2000**, 40, 659-667.

[2] Radchenko E.V., Makhaeva G.F., Malygin V.V., Sokolov V.B., Palyulin V.A., Zefirov N.S. *Doklady Biochem. Biophys.*, **2008**, 418, 47-51.

[3] Radchenko E.V., Makhaeva G.F., Sokolov V.B., Palyulin V.A., Zefirov N.S. *Doklady Biochem. Biophys.*, **2009**, 429, 309-314.

[4] Melnikov A.A., Palyulin V.A., Radchenko E.V., Zefirov N.S. *Doklady Chemistry,* **2007**, 415, 196-199.

[5] Radchenko E.V., Palyulin V.A., Zefirov N.S., in *Chemoinformatics Approaches to Virtual Screening,* ed. by A.Varnek, A.Tropsha, RSC, **2008**, 150-181.

ORAL PRESENTATION 36

STRUCTURE-ACTIVITY RELATIONSHIP OF ARKADIA RING FINGER E3 UBIQUITIN LIGASE THROUGH NMR SPECTROSCOPY

C.T Chasapis[1], N.G Kandias[1], A. Loutsidou[1], D. Bentrop[2], V. Episkopou[3], G.A. Spyroulias[1]

[1] Department of Pharmacy, University of Patras, GR-26504, Patras, Greece.

[2] Institute of Physiology II, University of Freiburg, D-79104 Freiburg, Germany.

[3] Mammalian Neurogenesis, MRC Clinical Sciences Centre, Imperial School of Medicine, Hammersmith Hospital, London W12 0NN, United Kingdom

E3 ubiquitin ligases play a key role in the recognition of target proteins by catalyzing the covalent attachment of the ubiquitin and degradation by 26S proteasome[1]. Arkadia is the first example of an E3 ligase that positively regulates TGF–β family signaling. Arkadia has been suggested to induce ubiquitin-dependent degradation of negative regulators of TGF–β signaling, through its C-terminal RING finger domain[2].

The 68 a.a. of the Arkadia C-terminal, including the RING finger, was cloned and expressed in its zinc-loaded form, as suggested by atomic absorption (two Zn(II)), and studied through multi-nuclear and multi-dimensional NMR Spectroscopy[3]. The ^{1}H-^{15}N HSQC is characterized by well dispersed signals and the vast majority of the backbone ^{1}H-^{15}N resonances have been assigned and deposited in the BioMagResBank (Accession No. 15948). The 3D NMR solution structure of Arkadia RING finger was determined with a large number of NOE-derived constraints, while the atomic coordinates have been deposited in Protein Data Bank (i.d. 2KIZ). Additionally, titration studies monitored by NMR were also performed to probe the interaction interface of Arkadia RING and the partner E2 (UbcH5B) enzyme. The RING-E2 complex structure was also constructed through an NMR-driven docking protocol (using HADDOCK).

Additionally, mutations identified in the RING domain were studied in the light of the 3D NMR structure and putting all these data together, new non-native RING finger forms were produced through protein engineering with the aim to study the structural integrity of the RING and its ability to interact with E2 partner enzyme. The new engineered 68-a.a. RINGs were overexpressed in uniformly labeled ^{15}N form suitable for NMR studies. Among four mutated constructs three of them found to preserve their Zn-binding capacity while one abolish any metal binding ability, according to atomic absorption measurements & NMR data. Among the ^{1}H-^{15}N HSQC spectra of the other constructs, the most impressive was the one that corresponds to the construct where other potential Zinc binding residues has substituted the Zn-ligands. The ^{1}H-^{15}N HSQC exhibit dramatic changes in respect with the native RING, but preserves the signal dispersion that is typical for a structured polypeptide, suggesting along with atomic absorption measurements that the mutated polypeptide although binds two Zn(II) ion per molecule, the RING scaffold architecture have been significantly modified.

References

1. Hershko A & Ciechanover A, *Annu Rev Biochem* **1998**, 67, 425–479.

2. Mavrakis KJ, Andrew RL, Lee KL, Petropoulou C, Dixon JE, Navaratnam N, Norris DP, Episkopou V. *PLoS Biol* **2007**, 5, e67

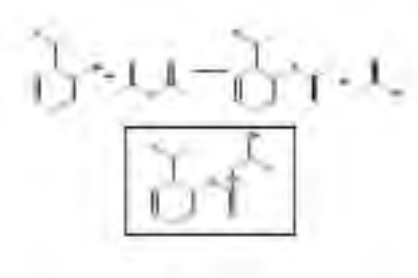
3. Kandias NG, Chasapis CT, Bentrop D, Episkopou V, Spyroulias GA. *BBRC* **2009**, 378, 498-502

POSTER SESSION I

MINING CHEMICAL REACTIONS DATA USING CONDENSED REACTION GRAPHS APPROACH

Alexandre Varnek*, Aurélie de Luca, Dragos Horvath and Gilles Marcou
Laboratoire d'Infochimie, UMR 7177 CNRS - Université de Strasbourg, France

Chemical reactions are difficult objects for chemoinformatics approaches because they involve several species of two different types: reactants and products. The "Condensed Graph of Reaction" (CGR) approach opens new perspectives in the mining of reaction databases since it allows one to transform several 2D molecular graphs describing a chemical reaction into one single graph (see Figure). Thus, a chemical reactions database can be transformed into a set of "pseudo-compounds" to which most chemoinformatics methods developed for individual molecules can be applied. Here, we discuss application of CGR approach for the (i) classification in organic chemstry and metabolic reactions, (ii) similarity search in large reaction databases and (iii) QSPR modelling of reaction parameters. In these studies ISIDA fragment descriptors extracted from CGR have been used. The advantages of CGR over conventional chemoinformatics tools will be illustrated in several "difficult" case studies.



Phenol acetylation reaction and related Condensed Graph of Reaction. Dynamical bonds marked with green and red colors correspond, respectively, to formation and breaking a single bond.

References:

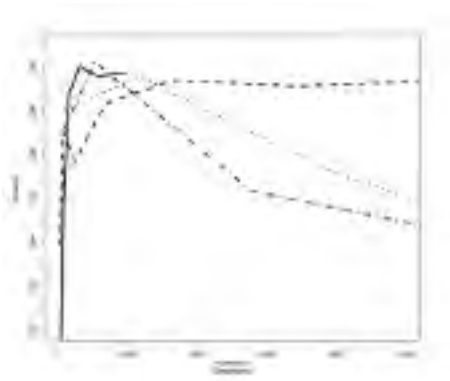
1. I. Baskin and A. Varnek
Fragment Descriptors in SAR/QSAR/QSPR Studies, Molecular Similarity Analysis and in Virtual Screening. In "Chemoinformatics Approaches to Virtual Screening", A. Varnek and A. Tropsha, Eds., RSC Publishing, 2008
2. A. Varnek, D. Fourches, D. Horvath, O. Klimchuk, C. Gaudin, P. Vayer, V. Solov'ev, F. Hoonakker, I. V. Tetko, G. Marcou
ISIDA - Platform for virtual screening based on fragment and pharmacophoric descriptors
Current Computer-Aided Drug Design, 2008, 4 (3), 191-198.

EXPLORING LINEAR SEPARABLE PROBLEMS IN CHEMISTRY

G. Marcou, I. Baskin, D. Horvath, A. Varnek*

*Université de Strasbourg, Laboratoire d'Infochimie, 4, rue Blaise Pascal, 67000 Strasbourg, France

This presentation is devoted to the study of linear separable datasets in the context of chemoinformatics. Support Vector Machines and Voting Perceptron, have been used on artificial datasets from the Directory of Usefull Decoys[1] with a set of tunable molecular descriptors from the ISIDA package[2]. In fact, these technique use a mathematical function, the kernel, that is equivalent to a projection in an abstract space crafted by the details of the kernel. In such space, a dataset can eventually be separated using only a linear function. It is shown, here, that linear separation occurs naturally on the sole basis of the molecular descriptors in use. Thus, the generation of molecular descriptors cannot be dissociated from the definition of a kernel. Some properties of these models are also investigated, in particular their control of overfitting. It is shown that the choice of the molecular descriptors plays a key role in this control. An heuristic is then proposed for early detection of the most promising models so that the most time consuming validation techniques can be applied on them in priority.



Cross-validation performances of Voting Perceptron classification models on HIVPR dataset. Dependancy over the number of descriptors depend of the type of descriptors used:

- sequences of atoms (—),
- sequences of atoms and bonds (- - -),
- augmented atoms (- -),
- augmented atoms and bonds (. . .)

References:

[1] Huang, N.; Shoichet, B.; Irwin, J. *Journal of Medicinal Chemistry* 2006, **49**, 6789–6801.

[2] Varnek, A.; Fourches, D.; Hoonakker, F.; Solov'ev V. P. *J. Comput. Aided Mol. Des*. 2005, **19**, 693-703.

PREDICTION OF THE AZEOTROPIC BEHAVIOR OF BINARY MIXTURES USING QSPR MODELS

Ioana Oprisiu, Gilles Marcou, Dragos Hovarth and Alexandre Varnek
Laboratoire d'Infochimie UMR 7177 CNRS, Université de Strasbourg, 4, rue B. Pascal, Strasbourg 67000, France

Although QSPR approach is widely used for structure-based property prediction, this method is limited to data analysis for single compounds. This work represents an extension of conventional QSPR models, attempting to predict the phase behavior of binary mixtures. Prediction of azeotrope formation is an important task in the field of process synthesis. Therefore, information about occurrence of azeotropes is absolutely necessary. Regression models for bubble point curve and classification models were developed using a data set of 102 binary azeotropes and 86 zeotropes measured at atmospheric pressure. The regression QSPR was investigated in three perspectives ("Points Out", "Mixtures Out" and "Compounds out") with respect to the creation of training and test set. Significant models were obtained having statistical parameters $R^2cv>0.89$ and $RMSE<7.6$ K. The obtained classification models are significant, with the statistical value $BA>0.75$.We demonstrated that QSPR could be applied to the mixtures, which might simplify the calculations, compared to conventional group contribution models.

POSTER PRESENTATION I-04

HIGHLY FLEXIBLE SAMPLING & DOCKING ON COMPUTER GRIDS

Dragos Horvath, Gilles Marcou, Alexandre Varnek
Laboratoire d'Infochimie, UMR 7177 CNRS - Université de Strasbourg, France

Unlike the attempts to speed up classical docking and conformational sampling calculations by GRID deployment[1], this[2, 3] project was designed as a methodological challenge, aimed to provide in-depth conformational sampling. In our opinion, using the GRID as a collection of workstations, each in charge of the docking of a subset of potential ligands relying on classical docking protocols is (albeit technically challenging) not optimal. The frustrating inaccuracy of classical docking protocols is largely due to insufficient sampling and force field parameterization. Accounting for large-scale flexibility, impossible with simple workstations, may move into the realm of the feasible if appropriate sampling protocols for computer GRIDs are developed, as attempted here.

Moreover, this is also an attempt to use the power of the GRID in order to reposition docking strategies as close as possible to the basic principles of molecular modelling. Docking is treated as conformational sampling of a system of several molecules. Therefore, the same algorithm is both able to perform ab *initio* folding of small peptides used in state-of-the-art folding simulations, and ligand docking of one – or, simultaneously, several ligands (or ligand plus a crystallographic water) – while, at the same time, rearranging flexible protein loops and/or side chains, with up to ~200 torsional and intermolecular degrees of freedom

The original GRID deployment termed "planetary" strategy is a generalization of the classical island deployment of a genetic algorithm[4], where a GRID node can be assimilated to an isolated "planet", exchanging information only with the GRID dispatcher. This is the key difference versus independent parallel deployments such as Folding@Home[1]: the dispatcher is not solely starting new remote simulations, but also interpreting incoming results and intelligently tuning new runs, in response to incoming data. A key strategic element is "panspermia": the dispatcher may randomly pick a subset of the already visited solutions from completed runs, and "seed" any new "planets" with to-date fittest genetic material or, by contrast, declare already sampled conformers as "tabu" areas.

Discussed applications of the planetary model include folding experiments of the Trp cage 1L2Y (α-helix), Trp zipper 1LE1 (β-sheet) and vilin headpiece (1VII), ligand docking into kinases with flexible hinge regions, including a mobile crystallographic water, docking into GPCRs, considering full side chain flexibility.

References:

1. V. Pande, Folding@Home, http://folding.stanford.edu/, 2008.
2. A.-A. Tantar, Conilleau, S., Parent, B., Melab, N., Brillet, L., Roy, S., Talbi, E.-G., Horvath, D., *Current Computer-Aided Drug Design,* 2008, **4**, 235-249.
3. D. Horvath, Tantar, A., Boisson, J.C., Melab, N., Roy, S., Brillet, L., Talbi, E.-G., ANR Dock - Molecular Docking on Grids, http://www2.lifl.fr/~talbi/docking/, 2008.
4. B. Parent, Kökösy, A., Horvath, D., *Soft Computing*, 2007, **11**, 63-79.

METHOD OF CONTINUOUS MOLECULAR FIELDS IN 3D QSAR

Igor I. Baskin, Nelly I. Zhokhova, Vladimir A. Palyulin, and Nikolai S. Zefirov
Department of Chemistry, Moscow State University, Moscow, 119991, Russia

Rapid development of computational techniques in the field of drug discovery calls forth the necessity to work out new methods providing the reliable prediction of various types of biological activities. Commonly used 3D QSAR CoMFA and CoMSIA methods suffer from some disadvantages. These methods are very sensitive to the size, orientation in space, and grid of lattice surrounding molecule that is used for approximating molecular fields at its nodes. [1] Instead of this, we suggest a new 3D QSAR method based on direct analysis of continuous molecular fields. This is achieved by using specially designed kernels built on continuous molecular fields.

In our study we used support vector machine regression as kernel-based statistical method. It should be pointed out that other kernel-based regression methods such as Gaussian Processes, kernel-PLS, etc. can also be used in 3D QSAR studies in conjunction with this kind of kernels.

Advantages of the Method of Continuous Molecular Fields (MCMF) were demonstrated by the example of modelling activity of inhibitors of the benzamidine type with respect to their binding affinities toward several serine proteases – thrombin, trypsin, and factor Xa. It was shown that this method provides considerably better statistical characteristics of 3D QSAR models than CoMFA and CoMSIA. [2]

References:

[1] 3D QSAR in Drug Design. Theory methods and applications. Ed. H. Kubinyi. New York: Kluwer Academic Publishers. 1997, p. 515.

[2] Zhokhova N. I., Baskin I. I., Bakhronov D. K., Palyulin V. A., Zefirov N. S. Doklady Chemistry. 2009, v. 429, Part 1, pp. 273-276.

POSTER PRESENTATION I-06

HIGH-THROUGHPUT VIRTUAL SCREENING OF PROTEINS USING GRID MOLECULAR INTERACTION FIELDS

Simone Sciabola,*[1] Robert V. Stanton,[1] James E. Mills,[2] Maria M. Flocco,[2] Massimo Baroni,[3] Gabriele Cruciani,[4] Francesca Perruccio,[5] and Jonathan S. Mason[6]

[1]Pfizer Research Technology Center, Cambridge, Massachusetts 02139, USA

[2]Pfizer Global Research and Development, Ramsgate Road, Kent CT13 9NJ, Sandwich, United Kingdom

[3]Molecular Discovery Limited, 215 Marsh Road, HA5 5NE, Pinner, Middlesex, United Kingdom

[4]Laboratory of Chemometrics, University of Perugia, Via Elce di Sotto, 10 I-60123, Perugia, Italy

[5]Syngenta, Schaffhauserstrasse, 4332 Stein AG, Switzerland

[6]Lundbeck A/S, Ottiliavej 9, DK-25000, Copenhagen, Denmark

A new computational algorithm for protein binding sites characterization and comparison has been developed, which uses a common reference framework of the projected ligand-space four-point pharmacophore fingerprints, includes cavity shape, and can be used with diverse proteins as no structural alignment is required. Protein binding sites are first described using GRID molecular interaction fields (GRID-MIFs), and the FLAP (fingerprints for ligands and proteins) method is then used to encode and compare this information. The discriminating power of the algorithm and its applicability for large-scale protein analysis was validated by analyzing various scenarios: clustering of kinase protein families in a relevant manner, predicting ligand activity across related targets, and protein-protein virtual screening. In all cases the results showed the effectiveness of the GRID-FLAP method and its potential use in applications such as identifying selectivity targets and tools/hits for new targets via the identification of other proteins with pharmacophorically similar binding sites.

TFD: TORSION FINGERPRINT DEVIATION AS A NEW METRIC TO COMPARE SMALL MOLECULE CONFORMATIONS

Tanja Schulz-Gasch*, Christin Schärfer#, Wolfgang Guba*, Matthias Rarey#

*Pharmaceutical Research, Chemistry, Technologies & Innovation, F. Hoffmann-La Roche Ltd., CH-4070 Basel, Switzerland

#Center for Bioinformatics, University of Hamburg, Bundesstrasse 43, D-20146 Hamburg, Germany

Objectivity, general applicability, and its easy, automated calculation make the assessment of relative RMSD the method of choice for the investigation of the accuracy of conformational model generators[1]. In contrast to absolute RMSD which is primarily used for docking evaluation, relative RMSD implies an additional alignment step of the molecules before the actual RMSD calculation. RMSD however suffers from some serious disadvantages, mainly because it strongly depends on molecular size and the number of rotatable bonds. It thus gives little information when applied to large datasets with diverse molecules in it. Furthermore, since it is not normalized it can result in a maximal RMSD of <2 Å for small and globular molecules, e.g. arabinose[2], whereas conformations of a molecule like NAPAP (PDB 1DWD) with 10 rotatable bonds can have RMSD that exceed 7 Å and a conformation with RMSD of ~ 2 Å would be considered as very close because it might still be able to have all crucial interactions when docked into the protein. Attempts to replace RMSD with a new metric3-6 were made but did not establish because of drawbacks, e.g. the loss of objectivity.

We recently developed a novel and objective approach for the investigation of the accuracy of conformational model generators called "TFD". It extracts and compares Torsion Fingerprints from a query molecule and generated conformations. To reduce the dependence of TFD on molecular size and number of torsions, TFD is normalized and thus allows direct conclusions about the conformation quality. In order to address the effect of the overall number of rotatable torsions in a molecule as well as the stronger effect on molecular shape of torsion deviations from central torsions compared to terminal torsions we applied a Gaussian weighting scheme for assigning penalties for torsion deviations. In order to properly assign central torsions, we also take the size of terminal substituents into consideration.

The approach has been validated and compared to relative RMSD. We demonstrate that TFD is able to overcome major disadvantages of RMSD while retaining its objectivity, general applicability and easy, automated calculation.

References:

1. Kirchmair , J.; Laggner, C.; Wolber, G.; Langer, T. J. *Chem. Info. Model.* **2005**, 45, 422-430.
2. Cole, J.C.; Murray, C.W.; Nissink, J.W.M.; Taylor, R.D.; Taylor, R. *Proteins: Struct. Funct. Bioinf.* **2005**, 60, 325-332.
3. Baber, J.C.; Thompson, D.C.; Cross, J.B.; Humblet, C. *Chem. Info. Model.* **2009**, 49, 1889-1900.
4. Kroemer, R. T.; Vulpetti, A.; McDonald, J. J.; Rohrer, D. C.; Trosset, J.-Y.; Giordanetto, F.; Cotesta, S.; McMartin, C.; Kihlen, .; Stouten, P. F. W. *J. Chem. Inf. Comput. Sci.* **2004**, 44 (3), 871–881.
5. Abagyan, R. A.; Totrov, M. M. *J. Mol. Biol.* **1997**, 268 (3), 678–685.
6. Yusuf, D.; Davis, A. M.; Kleywegt, G. J.; Schmitt, S. *J. Chem. Inf. Model* **2008**, 48 (7), 1411–1422.

AUTOMATICALLY PREDICTED SUB-POCKETS PAVE THE WAY FOR DESCRIPTOR-BASED DRUGGABILITY STUDIES

Volkamer, A.[1], Grombacher, T.[2], Rarey, M.[1]

[1]University of Hamburg, Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany

[2]Bioinformatics, Merck Serono, Frankfurter Str. 250, 64293 Darmstadt, Germany

Many drug discovery projects fail because the underlying target is found to be not druggable[1]. Predictions of the potential of disease modifying targets to be modulated by low-molecular weight compounds – referred to as druggability – is of high practical interest and can help to speed up and reduce costs in drug discovery. First methods have been presented estimating target druggability solely based on the 3D structure of the protein[2-4]. The prediction power of such methods depends on the precision of the characterization of the binding site. A multitude of methods exist for automated binding site prediction[5-9]. However, most methods developed for automated docking procedures do not explicitly focus on the definition of the volume and the boundary of the binding site. Nevertheless, a precise description is vital for correct descriptor-based druggability predictions.

For this purpose, we developed DoGSite, a new method to predict binding sites in proteins based on a Difference of Gaussian (DoG) approach which originates from image processing. In contrast to existing methods, DoGSite detects pockets which arise from merging predicted sub-pockets (Figure 1).

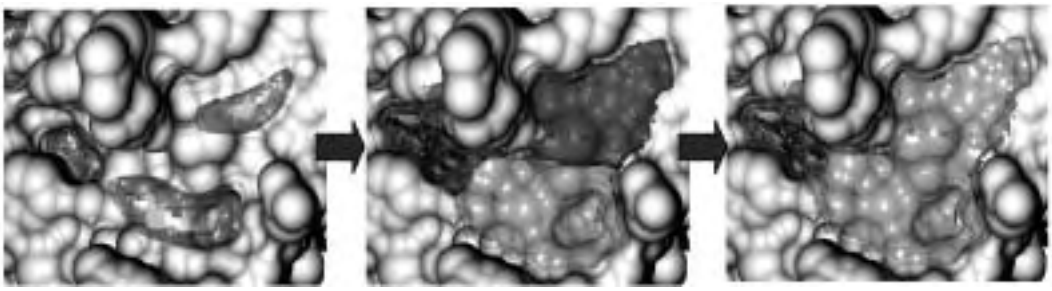


Figure 1: Binding pocket prediction by DoGSite for an protein-tyrosine phosphatase (1c83). Potential pocket kernels (green) are calculated first. Second, the kernels are dilated and produce three sub-pockets (cyan, yellow, red) that are eventually merged to one pocket (yellow). The co-crystallized inhibitor completely covers the cyan sub-pocket.

DoGSite predicts correct binding pockets in over 93% of the PDBBind[10] and the scPDB[11] data set, being in line with the best performing methods available. The majority of predicted pockets consists of more than one sub-pocket. In 90% of the cases, only one sub-pocket completely covers the co-crystallized ligand.

Since the volume of the predicted pocket has a high impact on descriptors derived for druggability studies, we introduced a more precise performance measure taking pairwise ligand and pocket coverage into account. Requiring that a predicted pocket covers at least half of the co-crystallized ligand, DoGSite yields a success rate of 92.5%. The good performance is due to a more convex shape of the pocket boundary which better resembles the true ligand binding volume. Demanding that at least a quarter of the pocket is

covered by the co-crystallized ligand diminishes the prediction success rate towards 55% highlighting the difficulty of detecting the true binding region of ligands. When considering sub-pockets coverage rates increase, yielding a success rate of 77%.
For druggability studies, such precise pocket descriptions result in more meaningful structural descriptors like binding site surface or volume. Detection of sub-pockets and good description of sub-pocket features also provide potential starting points for fragment-based drug discovery approaches. Furthermore, if several structures from one protein are known, sub-pockets may give hints about protein flexibility and induced fit conformational changes.

References:

[1] Brown, D. and G. Superti-Furga, Drug Discovery Today, **2003**. 8(23): p. 1067-1077.
[2] M. Nayal and B. Honig, Proteins, **2006**, 63(4):892-906.
[3] A. C. Cheng et al., Nat. Biotechnol., **2007**, 25(1):71-5.
[4] T. A. Halgren, J. Chem. Inf. Model., **2009**, 49(2):377-89.
[5] M. Hendlich, J. Mol. Graph Model., **1997**, 15(6):359-63.
[6] M. Weisel et al., Chem. Cent. J., **2007**, 1:7.
[7] J. An et al., Mol. Cell. Proteomics, **2005**, 4(6):752-61.
[8] A. Tripathi et al., Proteins, **2010**, 78(4):825-842.
[9] S. Henrich et al., J. Mol. Recognit., **2010**, 23: 209-219.
[10] R. Wang et al., J. Med. Chem., **2004**, 47(12):2977-80.
[11] E. Kellenberger et al., J. Chem. Inf. Model., **2006**, 46(2):717-27.

A LIGAND-BASED DRUG DESIGN METHOD EFFECTIVE FOR IDENTIFYING ACTIVE CONFORMATIONS AND 3-DIMENSIONAL PHARMACOPHORES OF CONGENERIC COMPOUNDS

Issei Doi[1, 2], Noriyuki Yamaotsu[1], Hiroaki Gouda[1], Shuichi Hirono[1]

[1] School of Pharmacy, Kitasato University, 5-9-1 Shirokane , Minato-ku, Tokyo 108-8641, Japan

[2] Discovery Laboratories, Toyama Chemical Co., Ltd., 4-1, Shimookui 2-chome, Toyama 930-8508, Japan

In computer-aided drug discovery that targets a protein whose 3-dimensional (3D) structure is not known, molecular superposition methods that make the most of a variety of ligand structures are often used to construct 3D-QSAR models. These models are based on the active conformation and the 3D-pharmacophore of relevant ligands. However, these methods were not effective in analyzing a dataset containing only congeneric compounds. Although we proposed a method that can be applied to such datasets in a previous study, it was found that molecular superpositions using only congeneric compounds were apt to cause overfit problems. In this study, we present a new method in which a few non-congeneric compounds are first superposed on each other, and then congeneric ones are superposed, to generate a molecular alignment that may be used in the following CoMFA. Furthermore, we confirmed the accuracy of the method using carbonic anhydrase II (CA II) inhibitors whose complexed structures had been determined by X-ray crystallography[1].

First, in order to generate stable conformations for a variety of CA II inhibitors, conformational analyses were carried out using an in-house software program, CAMDAS[2] (Conformational Analyzer with Molecular Dynamics And Sampling). Next, the conformers of non-congeneric compounds generated by CAMDAS were simply represented by five types of physicochemical property spheres. They were then superposed on each other using an in-house program, SUPERPOSE[3]. The conformer sets with the highest scores calculated by SUPERPOSE were extracted, and then congeneric ones were superposed on each of the conformer sets to prepare the molecular alignments used in subsequent CoMFA. Finally, CoMFA was performed for each molecular alignment and the molecular alignment with the highest q^2 value was selected to identify the binding conformation and the 3D-pharmacophore.

As a result, it turned out that the CoMFA model with the highest q^2 value could determine the binding conformation and the 3D-pharmacophore. The root mean square deviation between the binding conformation obtained using our method and the one determined by X-ray analysis was within 2Å for each relevant ligand. This suggests that our method could prove to be a useful tool in ligand-based drug design.

References

1) Supuran, C. T., *Nat. Rev. Drug Discov.*, **7**, 168-180(2008)

2) Tsujishita, H., Hirono, S., *J. Comput. Aided Mol. Des.*, **11**(3), 305-315(1997)

3) Iwase, K., Hirono, S., *J. Comput. Aided Mol. Des.*, **13**(5), 499-512(1999)

EXPLORING THE PERFORMANCE OF PHARMACOPHORE BASED VIRTUAL SCREENING

Modest von Korff*[1], Joel Freyss[1], Thomas Sander[1], Christoph Boss[2], Claire-Lise Ciana[2]

[1] Department of Research Informatics

[2] Drug Discovery Chemistry

Actelion Ltd., Gewerbestrasse 16, CH-4123 Allschwil, Switzerland

Pharmacophore based virtual screening is a commonly used methodology to investigate scaffold hopping. Often molecules developed in lead optimization programs are too large to be good seeds. These molecules are complex, exhibiting multiple pharmacophore points. But the probability of detecting biosimilars in commercially available compound collections decreases as the number of pharmacophore points grows. In this contribution, we demonstrate how automatically generated sub-pharmacophore models can be used to increase the number of virtual screening hits worth testing in biological assays. A 3D pharmacophore descriptor developed in-house was used to generate the sub-pharmacophore models [2]. The DUD dataset was used to compare the pharmacophores with docking and chemical fingerprints. The DUD is the first published dataset providing active molecules, decoys and references for crystal structures of ligand-target complexes [1]. It contains 2,950 active compounds against a total of 40 target proteins. Furthermore, the dataset contains 36 structurally dissimilar decoy compounds with similar physicochemical properties for every ligand. The ligands were extracted from the target proteins to extend the applicability of the dataset to include ligand based virtual screening. Of the 40 target proteins, 37 contained ligands that were used as query molecules for virtual screening evaluation. With this dataset, a comparison between the pharmacophores, three different chemical fingerprints and docking was done. In terms of enrichment rates, the chemical fingerprint descriptors performed better than the pharmacophores and the docking tool. After removing molecules chemically similar to the query molecules, the pharmacophores outperformed the chemical descriptors. Encouraged by these results, the sub-pharmacophores were applied to some in-house drug discovery projects. In one project, only one highly active but large molecule was available as seed. Neither the full pharmacophore model nor the chemical fingerprints were able to detect any similars in a database of seven million commercially available molecules. However, a sub-pharmacophore search resulted in the detection of hundreds of interesting molecules. These molecules were purchased and are currently undergoing biological testing. Biological assay results for this and other virtual screening experiments will be reported here.

References:

[1] Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. J. Med. Chem. 2006, 49, 6789-6801.

[2] Korff, M. v.; Freyss, J.; Sander, T. Flexophore, a new versatile 3D pharmacophore descriptor that considers molecular flexibility. J. Chem. Inf. Model. 2008, 48, 797-810.

USING NOVEL CONSENSUS SURFACE PROPERTY REPRESENTATIONS TO HELP CHOOSE THE "RIGHT" QUERY FOR "DIFFICULT" LIGAND-BASED VIRTUAL SCREENING TARGETS

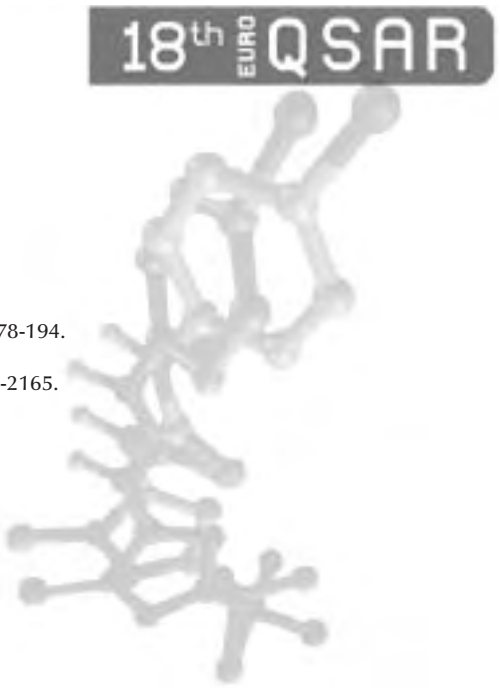
Violeta I. Pérez-Nueno[a], Vishwesh Venkatraman[a], Lazaros Mavridis[a], Antonio Carrieri[b], David W. Ritchie[a]

[a] INRIA Nancy – Grand Est, LORIA, 615 rue du Jardin Botanique, 54506 Vandoeuvre-lès-Nancy, France
[b] Dipartimento Farmaco-Chimico, Università degli Studi di Bari, Via Orabona 4, 70125, Bari, Italy

Ligand-based screening approaches have become well established in the process of computer-aided drug design [1]. However, despite their relative success, many authors have discussed the confounding problems of how to choose the initial query compounds and which of their conformations should be used [2]. Furthermore, it is increasingly the case that pharmaceutical companies have multiple ligands for a given target and these may bind in different ways to the same pocket. However, traditional shape matching approaches normally use just one conformation of a compound as the query, but it is not known a priori if this is the correct query to use to screen an entire database. For example, other compound families could also be active for the same target but they might only be found in the database if a different query conformation is used. In other words, conventional VS assumes there is only one binding mode for a given protein target. This may be true for some targets, but it is certainly not true in all cases. Several recent studies have shown that some protein targets bind different ligands in different ways [3]. The DUD dataset [4] contains 40 targets of therapeutic interest, including several targets with multiple actives which probably have different binding modes. It is therefore useful as a benchmarking set for testing these "difficult" targets.

Here we propose a novel approach based on using "multiple local spherical harmonic-based properties" calculated by PARASURF/PARAFIT [5,6] to improve VS performance for these "difficult" targets. This approach allows one or more consensus "pseudo-molecules" to be created and used as VS query structures from the known ligands by considering not only shape [7], but also the contribution of four other local properties: the molecular electrostatic potential (MEP), the local ionization energy (IEL), the local electron affinity (EAL), and the local polarizability (αL). Such pseudo-molecules can often be better VS queries than individual actives because they can capture the essential features of multiple high-affinity ligands. This approach can also be applied in cases where multiple query conformations might be required. In these cases, the pseudo-molecules produced by the consensus properties approach are used to calculate the VS ranked lists, and a data fusion technique is then applied to produce a final ranked list which aggregates the information for each of the conformations.

In this contribution, we present details of the consensus surface properties approach applied to several "difficult" DUD targets. These results are compared with conventional shape-based approaches using popular performance metrics such as receiver-operator-characteristic (ROC) plots and enrichment factors (EFs).



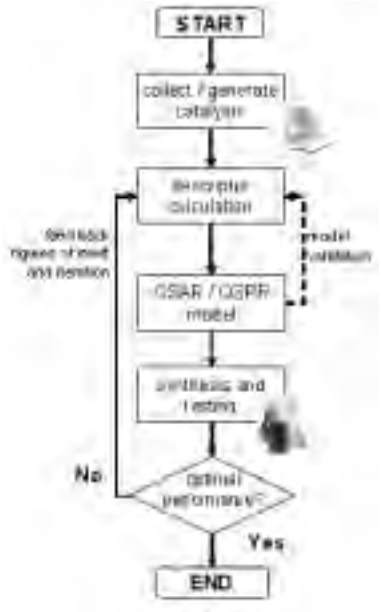
References:
1) Korff, M. et al. *J. Chem. Inf. Model.* **2009**, *49,* 209-231.
2) Hawkins, P. C. D. et al. *J. Med. Chem.* **2007***, 50*, 74–82.
3) Kellenberger, E. et a. J. Med. Chem. , **2007**, *50,* 1294-303
4) Huang, N. et al. *J. Med. Chem.* **2006***, 49*, 6789–6801.
5) Ritchie, D.W.et al. *Proteins: Struct. Func. Genet.* **2000***, 39,* 178-194.
6) Lin, J. et al. *J. Chem. Inf. Model.* **2005***, 45*, 1010-1016.
7) Pérez-Nueno, V. I. et al. *J. Chem. Inf. Model.* **2008***, 48*, 2146-2165.

INDUSTRIAL CHALLENGES IN PREDICTIVE MODELLING

Ana G. Maldonado

Rhodia Laboratoire du Futur, 178 avenue du Docteur Schweitzer F-33608 Pessac cedex, France

Predictive modelling has become a practical research tool in industrial chemistry. This is a response to the evolution of industrial processes, pharmaceutical chemistry and the increasing development of green chemistry. Predictive modelling can help us to pinpoint good regions in the chemical space, narrowing the search for the optimal molecules. But we are still far away from a program that takes substrates and products as input, and prints out the optimal formulation and the best reaction conditions. Or are we? Where is the line between dream and reality? And what are the barriers that we chemist still face?



In this lecture, these questions as well as a couple of other challenges will be addressed. In the first part of the lecture, we will explain in simple terms what predictive modeling actually is, why and when should one use it, and how it can be implemented [1]. We will present the basic requirements for good predictive models in chemistry: a reliable set of initial experimental data, an engine for generating and testing virtual libraries and robust validation protocols. Once you have these, the key task is translating the chemical information into something that a computer can understand.

In the second part of the lecture, we will show how to adapt the developed technologies to industrial needs [2]. Lacking of customization capacities is often one of the reasons why performing models does not contribute as expected to the detection and discovery of novel (and feasible!) structures in industrial projects. We will learn how to 'train' the program to propose molecules that are inside the scope of the current organic synthesis techniques, and how to "filter out" structures that are toxic, expensive or instable.

References:

[1] A. G. Maldonado, G. Rothenberg. Predictive modeling in homogeneous catalysis: A tutorial. Chem. Soc. Rev. (2010), Published on the web: 01 March 2010. DOI: 10.1039/b921393g.
http://www.rsc.org/publishing/journals/CS/article.asp?doi=b921393g

[2] A. G. Maldonado, G. Rothenberg. Predictive modelling in catalysis – from dream to reality. Chem. Eng. Proc. June (2009). http://www.science.uva.nl/~gadi/pdf_files/CEP_2009_105_26.pdf

FINDING DRUG DISCOVERY 'RULES OF THUMB' WITH BUMP HUNTING

Tatsunori Hashimoto[1], Matthew Segall[2] and Ed Champness[2]
[1] Department of Statistics, Harvard University, Cambridge, MA, USA
[2] Optibrium Ltd, Cambridge, UK

Rules-of-thumb for evaluating potential drug molecules, such as Lipinski's Rule of Five, are commonly used because they are easy to understand and translate into practice. These rules have traditionally been constructed by observation or by following simple statistical analysis. However, application of these techniques to QSAR models or early screening data often ignores the underlying statistical structure. Conversely, when machine learning algorithms are used to classify 'drug-like' molecules, they often result in black-box classifiers that cannot be modified to suit a particular target drug profile. We propose a novel hybrid approach, based on the Bump Hunting [1] method, to constructing rules-of-thumb from existing data to match a given target product profile for any therapeutic objective. These rules are easily interpretable and can be rapidly modified to reflect expert opinions before application.

[1] J.H. Friedman and N.I. Fisher Statistics and Computing (1999) **9**, pp. 123-143

DESCRIPTION AND ANALYSIS OF DRUGS AND NON-DRUGS IN MOLECULAR DESCRIPTOR SPACE

Uko Maran[1], Alfonso T. García-Sosa[1] and Csaba Hetényi[2]

[1] Institute of Chemistry, University of Tartu, Ravila 14A, Tartu 50411, Estonia

[2] Institute of Biology, Eötvös Loránd University, Pázmány sétány 1/C, 1117 Budapest, Hungary

The structural characteristics of small compounds are major determinants in the early steps of seeking for drug-like compounds from the vast chemical universe. Various threshold sets (such as Lipinski's rule-of-five and others), very often based on molecular descriptors, have been designed and extensively used for this task. It is still not completely described how known drugs' and non-drugs' areas of chemical space locate relative to each other. To obtain insight we have looked into the "maximum" molecular descriptor space available to us, and how this structural representation determines the relative position of drug and non-drug subsets.

The analyzed dataset consisted of 631 small molecules from which 311 are approved drugs and 320 are non-drugs. The selection criteria for the small molecules in the database was (i) the availability of experimental binding affinities (ΔGexp) for all compounds and (ii) both subsets had to possess a similar distribution of experimental values. The first criterion was set to select similar biological potency and the second criterion was set to keep both subsets in a similar size. Molecules were subsequently characterized with approximately 300 molecular descriptors unique for both subsets. The molecular descriptors comprised the following types: constitutional, topological, geometrical, charge distribution related, quantum chemical and calculated octanol-water partition coefficient (logP). Principal component analysis (PCA) was used for the analysis of molecular descriptor space.

The pattern exposed by PCA analysis distinguishes known drugs' and non-drugs' areas in chemical space. Whilst drugs form a well-defined dense area, non-drugs are more spread out, and there is an overlap area between subsets that identifies compounds with particular properties. Both subsets have a handful of remote points far from the origin of the subset. The first principal component describes structural characteristics related to the size of the molecules, and the second principal component is mostly related to electronic effects, described by specific molecular descriptors. The drug dataset was further divided into subsets according to distinct disease categories (nervous system, cardiovascular system, sensory organs, alimentary tract, etc. (14 in total)) and the respective individual patterns seen in the disease categories are also discussed.

THE PURPOSING OF NEW COMPOUNDS OR THE RE-PURPOSING OF OLD DRUGS BY MEANS OF MULTIVARIATE ANALYSIS ON MOLECULAR DESCRIPTORS

Antonino Lauria, Marco Tutone, Anna Maria Almerico
Dipartimento Farmacochimico, Tossicologico e Biologico
Università di Palermo, Via Archirafi 32, 90123 Palermo, Italy

Principal Component Analysis (PCA) is able to detect similarities among a set of variables (descriptors) or objects (compounds) providing a statistically reliable criterion to classify compounds on the basis of their different structural pattern against different classes of descriptor.
In this work we propose a protocol to assign a suitable biological target for a designed molecular structure by using the multivariate analysis on molecular descriptors matrix, approaching it by the ligand based point of view.
This method can be useful to hypothesize the biological target of a candidate lead compound prior to its biological evaluation or to re-purpose an old drug. The basic idea is to project it on a n-dimensional biological target space with the aim to find its well-fit biological target.
We started by choosing 47 heterogeneous biological targets, for which a consistent set of inhibitors, with available activity or binding affinity data, is known. The structures were downloaded from the bindingDB [1], and they were used as input data in CODESSA PRO program [2], where calculation of molecular descriptors was carried out. A preliminary pre-optimization was performed by LIGPREP of Schrodinger Inc [3]. A selected set of molecular descriptors, derived from geometrical and quantum-chemical structures, was computed, as available in this program. The resulting matrix, 7352 rows (structures) versus 173 columns (descriptors), was submitted to PCA. From the analysis of the significant components emerges that the 69% of the variance is comprised in the first five components. After, for each biological target, the n-dimensional barycentric coordinates were calculated.
By means of PCA on the matrix, it was possible to predict the biological target of a molecular structure by projecting it in the n-dimensional PC space. In fact, by adding the molecular descriptors values, calculated as above mentioned, as additional row to matrix, the "guessing" matrix was obtained. The "guessing" matrix enables to calculate, by PCA, the n-dimensional coordinates for the structure. The evaluation of the distance between the structure coordinates and the barycentric coordinates of each biological target could permit the biological target assignment.
With the aim to validate the protocol, a test set of known inhibitors of three biological targets (11betaHSD1, ARA1, CRFR1) was used. The results showed as the mean prediction capability of the method was $\approx$ 80%. Further applications of the protocol will be reported.

References

[1] Liu T, Lin Y, Wen X, Jorrisen RN, Gilson MK (2007) "BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities" Nucleic Acids Res. 35:D198-D201.
[2] Katrinsky AR, Lobanov VS, Karelson M (1996) Quantum-Chemical descriptors in QSAR/QSPR studies. Chem. Rev. 96:1027-1043.
[3] LigPrep, version 2.2, Schrodinger, LLC, New York, NY, 2005.

NEW PARADIGM FOR 3D/4D CLASSIFICATION OF DRUGS

Vladimir Potemkin[1], Maria Grishina[1], Athina Geronikaki[2]

[1]Chelyabinsk State Medical Academy, Vorovskogo 64, Chelyabinsk, 454092, Russian Federation

[2]Aristotle University of Thessaloniki

A biological action of a drug involves its interaction with some subjects in our organism (ferments, DNA, water, membrane, etc.). Therefore, it is necessary to find potentially active centers of molecular interactions. It is possible to determine the potentials of electrostatic, Van der Waals interactions, hydrogen bonding, etc. at each point of the molecular surface. These potentials determine a molecular field. Basing on the complementarity conception we can suppose that the field of a good drug must be complementary to the receptor. The more complementary molecule to the receptor the more active it is. Therefore, the molecular field of an ideal drug is ideally complementary to the field of the receptor. Thus we can reconstruct the field of the receptor as a complementary one to the field of the ideal drug (as its mould, its negative). Then, it is possible to recognize a new prospective drug as complementary one to the obtained mould. We do not, however, have an ideal drug. Moreover, there seems little use in creating a new drug if one already has the ideal one. What we do have, however, are real drugs and it is necessary to reconstruct the field of the receptor using the fields of these real molecules. It should be noted that the field of a real molecule can define only a part of the receptor active site. Therefore, it can be assumed that a generalized set of active molecules can completely reconstruct the receptor active site as a complementary field. We should have in view that a molecule includes active (pharmacophoric) fragments interacting with the receptor, non-active fragments that do not take part in the interaction and fragments disturbing the interaction (e.g., providing steric barriers). In addition, we must also bear in mind the conformational and tautomeric flexibility of a molecule and the flexibility of the receptor. Consequently, not only does a molecule have to adjust to the receptor, but the receptor also has to adjust to the molecule, due to their inherent flexibilities. So, the principles for 3D/4D classification are as follows:

- the geometry and the distribution of the interaction centers of an active molecule must be appropriate to the target;
- the classification method must find molecules with complementary geometry and active centers distribution provided by their conformational and tautomeric state;
- the "photograph" of a molecule must be representative. The method must recognize an active molecule, irrespective of its foreshortening, its conformational and tautomeric state.

Basing on this assumptions the new algorithm CiS ("Cinderella's Shoe") allows to model drugs and receptors flexibility is suggested. The method creates flexible pseudo-atomic receptor model and simulates a movement of drug to receptor through water and membrane. Thus, the method allows to predict bioactivity of compounds in dynamic conditions. All kinds of molecular movement (translational, rotational, vibrational, internal rotation) are taken into account. The method is used for detailed elucidation of action of dihydrofolate reductase inhibitors, monoamine oxydase, HIV-1 reverse transcriptase inhibitors. The conformational and tautomeric forms of the compounds in water and in receptor pocket are determined. The obtained conformations in receptor pockets are in a good agreement with X-ray and NMR data. Using the proposed method new potential anti-inflammators was designed. Now the designed molecules are synthesized and successfully tested in vitro and in vivo. The algorithm is used for virtual screening of more than 40 kinds of bioactivities.

The work is fulfilled with the support of SKIF – GRID Supercomputer Initiative.

References:

1) Potemkin V., Grishina M. Principles for 3D/4D QSAR classification of drugs// Drug Discovery Today., 2008, V. 13 (21/22), 952 – 959.

RECENT CADD ADVANCES ARE REPORTEDLY BENEFITING LEAD OPTIMIZATION PROJECTS

Richard D. Cramer
Tripos International: a Certara Company

During lead optimization, the candidate structures are very similar, often differing only by an R-group, and therefore exhibit similar potencies, with a pIC50 variance little greater than a log unit. Although "locally-derived" 3D-QSAR models are the CADD approach that can most meaningfully rank such similar candidates, 3D-QSAR has required tedious and somewhat subjective manual structural alignments. However, initial results from make-and-test applications of the new topomer CoMFA approach to lead optimization, combined with exceptional ease in use, are strongly encouraging. An "RGVS" capability to identify the most promising R-groups from among 10E6 synthesizable candidates could further provide critical novel ideas.
Other CADD advances for lead optimization include a docking capability that can also be "locally optimized" when multiple experimental structures with binding energies are known, and a multi-criteria optimization function adaptable to individual project goals coupled with a project-proven de novo engine.

MULTI-PARAMETER SCORING FUNCTIONS FOR LIGAND- AND STRUCTURE-BASED DE NOVO DESIGN

Uhrig U. [1] Fabian B. [1] Smith K.M. [2]. Liu Q.[3]
[1]Tripos International Munich, Germany,
[2]Tripos St. Louis/USA, [3]Tripos Shanghai, China

Successful drug discovery often requires optimization against a set of biological and physical properties. We describe de novo design studies that demonstrate successful scaffold hops between known classes of ligands for p38 MAP kinases using ligand-based and structure-based multi-parameter scoring functions coupled to the molecular invention engine Muse.
The ligand-based scoring function includes pharmacophoric and steric tuplets and structural (fingerprint based) similarity. In addition various selectivity or ADME related properties (e.g. Lipinski properties, polar surface area, activity at off-targets, etc.). can be taken into account to guide the evolution of structures meeting multiple design criteria.
The structure-based scoring function uses Surflex-Dock to pose and score invented structures inside the target's active site. In addition, a number of simple molecular properties (e.g. clogP, Lipinski properties, etc.) are used as score components to focus the design on medicinally relevant chemistries. With the ability of Surflex-Dock to start the docking process with a single or multiple placed fragments, this scoring function can be applied in fragment based drug discovery to optimize attachments onto a pre-placed substructure.

R/AUTOGRID/ADT COMBINATION AS AN ALTERNATIVE TO BUILD 3-D QSAR MODELS. METHODOLOGIES AND APPLICATIONS

Ira Musmuca, Flavio Ballante and Rino Ragno
Rome Center for Molecular Design, Dipartimento di Chimica e Tecnologia del Farmaco, "Sapienza" University, P.le A. Moro 5, 00185, Rome, Italy

Understanding the structural properties and characteristics that affect the biological activity of a drug like molecule is a key step in the ongoing process of drug design. The main objective is the prediction of novel unknown compounds on the basis of previously synthesized molecules. Classical QSAR and 3-D QSAR methods are valid strategies that have been historically applied for this purpose. The idea behind all 3-D QSAR methods is that biological activity of a ligand can be predicted by its three-dimensional structure. The well known CoMFA[1] technique and the GRID/GOLPE[2] approach are almost the solely 3-D QSAR tools widely used in the last two decades, although successfully, both these approaches are proprietary software and almost no alternatives are available.

In this work we report a classical 3-D QSAR analysis using just free and open source software in both calculation and graphical imaging of the results.

Molecular Interaction Fields (MIFs) for various probes are calculated using the Autogrid force field of the AutoDock suite.[3,4] The PLS statistical analysis are conducted by means of R programming language. Python Molecular Viewer (PMV)[5] or Autodock Tools (ADT)[5] are then used to graphically analyze the statistical results and to produce high quality and full informative images.

Our method allows the iterative generation of hundreds/thousands 3-D QSAR models and the best one is selected on the basis of the statistical r^2, q^2 and SDEP coefficients in a completely independent way avoiding the user to face tedious, time wasting and boring iterative interactions with the computer.

From the best of our knowledge this is the first time that a completely free/opensource 3-D QSAR tool is presented.

Details and comparison with either CoMFA or GRID/GOLPE applications examples will be reported.

References:

(1) Cramer III, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA) 1. Effect of Shape o Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967.

(2) GRID/GOLPE, www.moldiscovery.com, Molecular Discovery, UK.

(3) Morris, G. M.; Goodsell, D. S.; Halliday, R.S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639-1662.

(4) Huey, R.; Morris, G. M.; Olson, A. J.;. Goodsell, D. S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.* **2007**, *28,* 1145-1152.

(5) Coon, S. I.; Sanner, M. F.; Olson, A. J. Re-Usable Components for Structural Bioinformatics. In Proceedings of the 9th International Python conference. pp 157-166, March 5-8, 2001. Long Beach, California.

PREDICTING LIGAND-PROTEIN BINDING USING SVM A NOVEL CHEMOGENOMIC SCREENING METHOD

Jamel Eddine Meslamani, Didier Rognan
Structural Chemogenomics Group – Therapeutic Innovation Laboratory – UMR7200 CNRS/University of Strasbourg, France

Diverse virtual screening methods are available for selecting biologically active compounds ranging from ligand-based similarity searches and pharmacophore mapping to protein structure-based docking. Recently, machine learning algorithms have been demonstrated to perform well in binary classification models of protein-ligand association, when applied to simple fingerprints encoding information from both targets and ligands **(1,2)** with the ultimate aim of predicting novel ligands for orphan targets.
Here, we further investigate the utility of combining protein and ligand structural information to predict protein-ligand binding. A generic cavity fingerprint which enumerates pharmacophoric triplets in specific distance-ranges of all Cα binding site atoms **(3)** was used to describe target 3-D space. Scitegic ECFP4 circular fingerprint **(4)** was used to describe ligand space. The training dataset includes 87 groups of similar binding sites obtained by clustering 7078 druggable proteins **(5,6)** using the in-house program FuzCav(3) (clusters with more than 10 members were kept). Different kernels were used to separately measure ligand similarity and binding site similarity. Their tensor product efficiently separates classes in the protein-ligand feature space **(7)**.
87 binding site-specific (e.g. ATP-binding site of serine/threonine protein kinases) chemogenomic SVM models addressing 581 unique proteins and 2605 ligands were generated and 5-fold crossvalidated with an average recall value equal to 0.68 and precision value equal to 0.87. An external test set of 22080 diverse ligands was used to evaluate the models and shown that ~50% of them were predictive.
Results were compared to a SVM ligand-based approach, which describes protein space with protein name instead of the cavity fingerprint. We conclude that the chemogenomic approach outperform the ligand-based approach when studying a target with less than 40 known ligands.

1) Weill and Rognan (2009) *J. Chem. Inf. Model.,* **49**, 1049-1062
2) Geppert et al. (2009) *J. Chem. Inf. Model.*, **49**, 767-779
3) Weill and Rognan (2010) *J Chem Inf Model*, **50**, 23–135
4) Rogers and Hahn (2010) *J Chem Inf Model*, **50**, 742-754
5) Kellenberger et al. (2008) *J Chem Inf Model*, **48**, 1014-1025.
6) http://bioinfo-pharma.u-strasbg.fr/scPDB
7) Vert and Jacob (2008) *Comb Chem & High Throughput Screening*, **11**, 677-685

POSTER PRESENTATION I-21

THE USE OF STATISTICAL METHODS IN GUIDING A LEAD OPTIMISATION PROJECT

Armelle Le Gall#, Chris N Luscombe#, Aoife C Maxwell*, Stephen Swanson*
Computational and Structural Chemistry, GlaxoSmithKline Stevenage, UK
* Respiratory CEDD, GlaxoSmithKline Stevenage, UK

A primary objective of a lead optimisation project is to investigate and understand the structure-activity relationships (SAR) around a lead series and improve compounds' developability properties. This poster describes statistical techniques used at GSK to support programmes under lead optimisation.
In a typical lead optimisation cycle, molecules are explored adopting a single point strategy e.g. varying one position at a time through the making of mono-dimensional arrays. Despite its inherent benefits e.g. easier synthesis and interpretability of results, this approach suffers of a severe limitation, in that additivity of SAR is assumed a priori, therefore making this approach suboptimal in particular for the initial exploration of the physico-chemical space of a given chemical series. One way of overcoming this issue is to sample the physico-chemical space more efficiently using experimentally designed multi-dimensional-arrays, where monomers at one position are combined with several monomers at a different position. This ensures that the contribution of each monomer is assessed more reliably through multiple combinations, and therefore the additivity of SAR can be thoroughly investigated.
Here we describe the design and analysis of a balanced experimentally designed incomplete array used to investigate the SAR of a series of inhibitors of a known metalloenzyme. The array output was analysed using our proprietary SAR Toolkit to assess the additivity of the SAR using Dscore plots, and predict novel combinations through Free-Wilson analysis. The chemical diversity of this array also provided a powerful set to validate a PLS model, previously built to investigate the drop-off observed for when progressing compounds from an enzyme to a whole blood assay. The PLS model was then applied to help prioritising the design and testing of new molecules.

POSTER PRESENTATION I-22

BioDig: A MATCHED MOLECULAR PAIR (MMP) IDENTIFICATION AND ANALYSIS METHODOLOGY TO SYSTEMATICALLY MINE CHEMICAL DATA

Jameed Hussain, Ceara Rea, Jordi Munoz-Muriedas, Chris Luscombe, Gianpaolo Bravi, Stefan Senger, Zhen Yang, Xuan Hong, Darren Green
GlaxoSmithKline Research & Development Limited, Stevenage, UK

Modern drug discovery organisations generate large volumes of structure activity data. This makes it very difficult for medicinal chemists to capture all the knowledge available within it. A promising methodology that can be used to mine this chemical data to identify novel structure-activity relationships is the matched molecular pair (MMP) methodology. However, before the full potential of the MMP methodology can be utilised, a MMP identification method that is capable of identifying all MMPs in large chemical datasets on modest computational hardware is required.
The presentation will report an algorithm that is capable of systematically generating all MMPs in a chemical dataset. The algorithm is computationally efficient enough to be applied on very large datasets. The results of a proof of concept study to show how the algorithm was applied to a 2 million compound set will be presented. We will show how it has been used within GSK to systematically mine ADME data (log D, solubility and intrinsic clearance). The statistics we use to identify interesting molecular transformations will be explained. Additionally we will discuss how the MMP methodology developed at GSK has been used to create an end user tool to allow scientists to find interesting molecular transformations applicable to their input compounds or substructures. Additionally, the tool can be used to quantify the likely improvement (of the ADME property) for a user defined transformation. The future direction of the work will also be discussed. Here we will describe how the technique can be used to identify novel bioisosteres.

Spectrophores™ AND OPEN BABEL: A POWERFUL TECHNOLOGY COMBINATION TO REPRESENT LARGE CHEMICAL SPACES USING SELF-ORGANISING MAPS

Hans De Winter, Gert Thijs
Silicos NV, Wetenschapspark 7, 3590 Diepenbeek, Belgium

The theory behind the calculation of Spectrophores™ is described. A Spectrophore™ is one-dimensional descriptor that describes the three-dimensional molecular field surrounding a molecule generated by a given set of atomic properties. Spectrophores™ descriptors have been incorporated into the Open Babel cheminformatics open source domain.
The application of Spectrophores™ to represent large chemical spaces is demonstrated. The method of choice to map the large chemical space of available vendor compound libraries based on Spectrophores™ is a self-organizing map (SOM). To train the SOM, a specially tailored algorithm was developed to cope with the large compound set. The resulting SOM clearly shows that the vendor chemical space can be mapped in a meaningful organization based on its Spectrophores™ representation and that it handles compounds beyond their topological similarity. The SOM has been applied in three different problems, including:
1. overlay of vendor databases;
2. comparison of compounds with different pharmacological profiles;
3. sampling of drug-like screening libraries.

Spectrophore™ TECHNOLOGY FOR RAPID AND RELIABLE VIRTUAL SCREENING

Gert Thijs,[1] Wilfried Langenaeker,[1] Jonatan Taminau,[1] Bart De Moor,[2] Hans De Winter[1]

[1] Silicos NV, Wetenschapspark 7, B-3590 Diepenbeek, Belgium

[2] ESAT-SCD, K.U.Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

Spectrophore™ are one-dimensional descriptors calculated from a set of molecular properties that can be generated from the three-dimensional conformations of small molecules. Examples of typical molecular properties that can be converted into Spectrophores™ descriptors include, amongst others, electrostatic potentials, atomic lipophilicities, hard- and softness potentials, and the molecular three-dimensional shape. Molecules with similar three-dimensional properties, and in many cases therefore similar biological activities, will always yield similar Spectrophores™. Therefore this technology is well-suited as a rapid and accurate virtual screening tool. The power of Spectrophores™ is illustrated in the following case studies:-

1. Ligand-based virtual screening of MAO-A and MAO-B inhibitors;
2. QSPR (Quantitative Structure Property Relationships) model building of blood-brain barrier penetration;
3. Clustering with self-organising maps of the NCI AIDS antiviral screening data set;
4. Proteochemometrics modeling of protein-ligand interactions.

QSAR WORKBENCH: A NEW APPLICATION FOR AUTOMATED EXPLORATION OF QSAR MODEL SPACE

Malcom Noj
Accelrys, 334 Cambridge Science Park, Cambridge, CB4 0WN, UK

Generation of relevant QSAR models for individual Medicinal Chemistry projects is an increasingly difficult challenge for the Pharmaceutical industry. We will present a solution that significantly reduces data management cycle times while enabling the capture and sharing of best practice. This new application is based on a toolkit built with Pipeline Pilot that allows automation of QSAR model building tasks.
The underlying toolkit has been designed to guide expert users through the QSAR model building process; from initial chemistry or response normalization through descriptor calculation, model building through to publication of validated models to end users.
The modular framework means that novel methods, descriptors or analysis tools can be included with minimal effort. In addition combinatorial exploration of descriptor, model type and parameter space is easily achieved. Model validation and comparison of multiple models is driven through interactive graphical reports.
Pipeline Pilot's rich reporting capabilities have been exploited to build an example web-based QSAR Workbench application. Expert users are able to use this web-based GUI to explore model space and then select a subset of steps that can be "replayed" by non-expert users on new data series.
Automatic monitoring of each task performed by the expert user allows an audit trail to be re-traced for optimal models – this audit trail can for example be used when compiling regulatory dossiers.

THREE-DIMENSIONAL STRUCTURAL DATA MINING OF PROTEINS BASED ON GEOMETRICAL FRAGMENT SPECTRA REPRESENTATION

Hiroaki Kato
Department of Computer Science and Engineering, Toyohashi University of Technology.
1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580 Japan

Structural similarity provides us a lot of information on structure-activity and structure-property problems, and for the selection of candidate analogs as new chemicals. In the present work, to describe 3D structural information of molecules, Geometrical Fragment Spectra (GFS) method has been developed. Structural information of a molecule including its 3D geometry is represented by a weighted graph of which the nodes and edges correspond to atoms and the inter-atomic distances between them, respectively. To get a GFS representation of a molecular structure, all the possible subgraphs with the specified number of atoms are enumerated. Subsequently, every subgraph is characterized with a numerical index. For the characterization of a subgraph the authors adopt 3D Wiener Number, that is, the sum of the inter-atomic distances corresponding to the edges of the subgraphs. The histogram is defined as a GFS that is obtained from the frequency distribution of a set of individually characterized subgraphs according to the value of their characterization index. The GFS generated along with this manner is a novel representation of geometrical structural profile of a molecule. The method was also applied for 3D structural similarity analysis of proteins using the reduced representation of protein structures. In this representation, a protein structure is represented by a set of pseudo-atoms corresponding to glycine residues and 3D coordinates of their alpha-carbons. The GFS of each protein was generated for these points corresponded glycine residues in 3D space. The GFS's for 1,300 peptide chains were examined in similarity searching among them. As a result, we successfully found the proteins that have similar arrangement of glycine residues but quite different folding structure with a query protein. The presented results show the potential applicability of this method for 3D structural data mining of proteins.

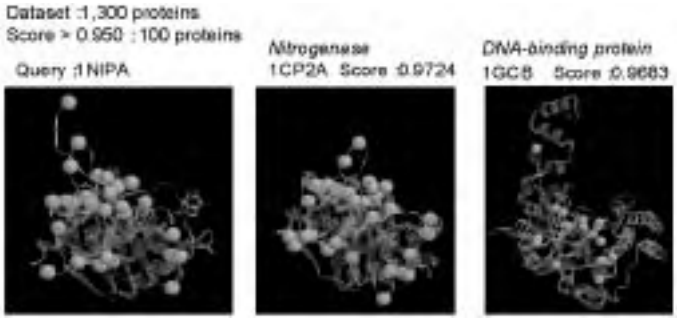


Fig. 1. Result of the 3D protein structural similarity search based on Gly-filter reduced representation with GFS method.

References

[1] Y. Takahashi, et al., *Advances in Molecular Similarity,* 2, (Eds. R. Carbo & P. Mezey), JAI Press, Greenwich, CT, 93-104 (1998).
[2] B. Bogdanov, et al., *J. Math. Chem.,* **3**, 299-309 (1989).
[3] H Kato, et al., *J. Comput. Chem. Jpn.,* **4**, 33-42 (2005).

AMBIT – CHEMINFORMATIC DATA MANAGEMENT SYSTEM

Jeliazkova Nina, Jaworska Joanna, Bruno Hubesch[3]

[1] IdealConsult, Sofia, Bulgaria

[2] Procter & Gamble, Brussels Innovation Center, Strombeek _-Bever, Belgium

[3] Cefic LRI, Brussels, Belgium

AMBIT is open source software for cheminformatics data management developed with funding from industry via a CEFIC LRI funded project. It is distributed under LGPL. The AMBIT2 software consists of a database and functional modules allowing a variety of queries and data mining of the information stored in the database. AMBIT XT is a user friendly application with a graphical user interface, based on AMBIT2 modules, and is also distributed under LGPL. AMBIT XT provides a set of functionalities to facilitate evaluation and registration of the chemicals for REACH. AMBIT XT introduces the concept of workflows, allowing users to be guided step-by-step towards a particular goal, and it provides workflows for analogue identification and PBT assessment. The software is a standalone application, with an option to install the database on a server.

Unique features of AMBIT:

- able to store and query multifaceted information about chemicals (structure, data, text)
- workflow engine to facilitating often repeated user's actions (e.g Analogue identification, PBT assessment can be applied in REACH).
- data provenance and automated quality assurance
- internal pKa calculation
- connection to KEGG* to locate pathways associated with a queried chemical

AMBIT-SMARTS: EFFICIENT SEARCHING OF CHEMICAL STRUCTURES AND FRAGMENTS

Nina Jeliazkova[1], Nikolay T. Kochev[2]

[1] Ideaconsult Ltd, 4 A. Kanchev str., Sofia 1000, Bulgaria

[2] University of Plovdiv, Department of Analytical Chemistry and Computer Chemistry, 24, Tzar Assen Str., Plovdiv 4000, Bulgaria

We present new developments in AMBIT [1] open source software package for efficient searching of chemical structures and structural fragments. AMBIT Smarts is a Java based software build on top of The Chemistry Development Kit (CDK) library http://cdk.sourceforge.net and uses SMARTS notation to define search queries. AMBIT Smarts parser implements the entire SMARTS language specification with several syntax extensions to make software compliant with modifications made by third party software packages such as OpenEye, MOE and OpenBabel. The goal of yet another open-source SMARTS parser implementation is to achieve better performance and compatibility with multiple flavours of SMARTS language, and to provide utilities for an effective application of the SMARTS parser queries in large structural databases. Typically the standard techniques for speeding up the substructure searching in a large database (e.g. pre-screening, fingerprints etc.) can be applied for queries which do not contain variable parts, expressions or wild-cards. We describe a combination of approaches which are applied for implementing a large structural database which can be effectively searched by queries utilizing the full straight of SMARTS language. AMBIT database substructure search includes two pre-screening steps, based on hashed fingerprints and predefined structural fragments, makes use of preprocessed information of important molecule features stored in the database, and finally, search results are cached for subsequent reuse.

The package is available as source code http://ambit.sourceforge.net, embedded in AmbitXT [1] standalone application and as REST web services via OpenTox API http://apps.ideaconsult.net/ambit2/. The online database consists of all compounds from REACH preregistration list, with quality labeled structures, retrieved from multiple sources.

The package has wide applicability in solving various chemoinformatics tasks. It has already been used in several projects as implementation of (MOE) SMARTS-based pKa estimation [2], in a software module for prediction of Molar refractivity, implementation of QSAR/QSPR modeling via additive schemes; within ToxTree software for toxicity prediction http://toxtree.sourceforge.net; AMBIT database with fast searching capabilities; read-across via AmbitXT, verification of coverage of four classes of compounds of interest in FP7 CADASTER project by REACH preregistration list; FP7 OpenTox project ToxPredict web application http://toxpredict.org. The future plans include integrating AMBIT Smarts parser within The CDK library.

References:

[1] Jeliazkova N., Jaworska J., Worth A., Open source Open Source Tools For Read Across And Category Formation, in "In Silico Toxicology: Principles and Applications (Issues in Toxicology)", ed. Marc Cronin, Royal Society of Chemistry, in print (Oct 2010)

[2] Adam C. Lee, Jing-yu Yu, and Gordon M. Crippen, J. Chem. Inf. Model, 48 (10), pp 2042–2053

Acknowledgements: AMBIT software was developed within the framework of CEFIC LRI project EEM-9 and extended under subsequent CEFIC LRI contract for developing AmbitXT.

MAPPING DRUG ARCHITECTURE BY MoStBioDat - RAPID SCREENING OF CATECHOL SCAFFOLD

Andrzej Bak, Jaroslaw Polanski

Department of Organic Chemistry, Institute of Chemistry, University of Silesia, PL-40-006 Katowice, Poland

Computer-assisted simulations are the most progressive component of the present day chemical investigations, producing enormous amount of data. The constraint of processing and sharing such data is thought as a major impediment in the drug discovery process. Furthermore, among the steepest barrier to overcome in the high-throughput screening (HTS) studies is the restricted amount of a reliable, publicly available repositories combining the detailed drug data with the comprehensive drug target information. Only the proper dataset aggregation and unified standards of data organization enable massive *in silico* knowledge mining. Structure-based database screening is a rapidly growing and an efficient technique in the early stages of the drug development process, gaining considerably from the current progress in the computer technology. Particularly, the subsequent sampling of a virtually infinite chemical space (VCS) in order to optimize the ligand diversity of chemical libraries (VLS) with appropriate binding affinity places emphasis more on the probability field with accidentally developed drugs than on traditional principles of the rational drug discovery. In consequence, the tools and techniques for organizing and intelligently mining this information are highly desirable. MoStBioDat has been established as a uniform data storage platform and integrated extraction system with an extensive array of software tools for structural similarity measures and pattern matching which are undoubtedly essential to facilitate the drug discovery process. In the current studies we have investigated the application of MoStBioDat software platform for the massive analysis of the spatial arrangements and conformational examinations of hydroxyl groups in the catechol-containing compounds, widely regarded as the main substructure block in many antiviral inhibitors. The geometrical orientation of the hydroxyl groups seems to determinate the ability of catechol derivatives to recognize the surrounding environment by forming the inter- and intra-molecular hydrogen bonds. The detailed analysis of the torsion angles, taking into account the spatial coordinates of the hydroxyl groups and the adjacent aromatic carbon atoms has been conducted using 3D structures taken from the freely accessible repositories.

RAPID PROTEIN STRUCTURE-BASED VIRTUAL SCREENING

Olli T. Pentikäinen, Sanna Niinivehmas, Pekka A. Postila, Salla I. Virtanen
Department of Biological and Environmental Science & Nanoscience Center, P.O. Box 35, FI-40014
University of Jyväskylä, Finland

Virtual high-throughput screening (vHTS) is typically accomplished by (1) docking ligands into three-dimensional protein structure or (2) by comparing established ligands to those in molecular databases. Both of these approaches have severe problems: (1) molecular docking is relatively slow and the available scoring functions cannot efficiently separate active ligands from inactives, (2) in addition, the ligand-based methods depend heavily on the assortment of known active ligands, Therefore, if none of the known active ligands is highly potent, it is unlikely that ligand-based screening would be able to discover them either.

We have been able to enrich the success in the vHTS by combining the protein structure data with the ligand-based drug discovery methods1. In our method we create a negative image of the ligand-binding site based on the protein structure, and use this ligand-like entity to screen molecular databases. This method concentrates on the shape of the ligand-binding pocket and here, we show that the additional features can enrich the success-rate even further. Although the method is fast, the accuracy is better than with molecular docking and force field based methods. In conclusion: our results indicate that the combined structure-based vHTS strategy identifies rapidly active ligands with excellent enrichment.

Reference:

1 Virtanen SI, Pentikäinen OT. Efficient virtual screening by using multiple protein conformations described as negative images of the ligand-binding site, Journal of Chemical Information and Modeling, 2010, 50: 1005–1011.

MAKING A NETWORK OF SIMPLE GRAPH ISOMORPHIC MOLECULAR FRAMEWORKS AND STRUCTURE DATA MINING OF DRUGS

Yoshimasa Takahashi, Kyohei Nomura
Department of Computer Science and Engineering, Toyohashi University of Technology,
1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580 Japan

In the preceding works, we proposed a basic skeletal feature representation of molecules with a graph, called Non-Terminal vertex Graph (NTG) [1]. We also showed that the NTG is a useful feature of drug molecules in structure-activity analysis. In this paper, we propose an alternative approach to knowledge discovery for structure-activity relationships of drugs using the NTG. The method is based on making a network of NTG frameworks together with their drug activity information.
A non-terminal vertex graph (NTG) is defined with a vertex graph, which does not have any terminal vertex or any isolated vertex. For molecular graphs, several types of NTG can be defined with different graph representations. In the present work, we focus on the NTGs which are all isomorphic in terms of simple graph representation but they are different NTG molecular graphs.
Ohtaguro et al proposed quantitative measures, called structure difference, of the difference between molecular graphs that are isomorphic in simple graph representation [2]. Three different indices were defined to describe the structure difference. They are *vertex differences* (the number of vertices that have different atom labels), *edge differences* (the number of edges that different bond multiplicity) and *chemical structure differences* (the number of different both vertex-labels and edge-weights).
We employed them to produce the NTG network. The network is made by producing the link between the nodes (NTGs) that have a specified value of the structure difference. For the computational trial of making the NTG network, we used 12 NTGs which support 169 drug molecules. They were taken from NTG database constructed by Noto, et al [3]. Those NTGs involve in antibacterials, antineoplastics, antibacterials (AIDS) etc. With the related biological activity information, the resulted network describes the relationships among the NTGs and the effects caused by the difference. Each node of a NTG network shows a NTG. A link of the network shows a NTG-NTG relationship with the structure difference.
The NTG network can be used to get some instance-based association rules for NTG-based molecular frameworks and drug activity. The paper describes the detail of the algorithm for making the NTG network and the utility for knowledge acquisition.

References

[1] Takahashi, Y. Chemical data mining based on non-terminal vertex graph. In *Systems, Man and Cybernetics;* 2004 IEEE International Conference, 2004; Vol. 5, pp. 4583 – 4587.
[2] Ohtaguro S, Takahashi Y, Numerical Indices to Represent Structural Difference between Molecular Graphs, *The 9th Data-Mining and Statistical Science,* 2009, pp.121-124.
[3] Noto K, Takahashi Y, NTG Knowledge-based System for Candidate Structure Design in Drug Development, *The 22nd Annual Conference of the Japanese Society for Artificial Intelligence,* 2008, 2P2-13.

POSTER PRESENTATION I-32

SELF ORGANIZING MAPS FOR THE MINING AND THE MODELLING OF BIOPHARMACEUTICAL DATA

Alban Arrault, Philippe Vayer, Marc Bertrand
Technologie Servier, 27 rue Eugene Vignat, Orleans, 45000, France

Chemical space representation is of importance for managing, simplifying and mining the pharmaceutical data. Many chemical descriptors have been investigated in this way, and the most frequently used are topological and physicochemical properties based, since they are able to cluster together molecules with similar biologic or therapeutic profile.
A concern is which mathematical method is the most suitable to deal with the high dimensionality of these chemical descriptions. Principal component analysis is one of them, very useful, but the quality of the data compression evaluated by the percentage of explained variance necessitates sometimes more than the 2 or 3 first components, making the representation useless. Moreover, even when the explained variance is satisfying with the 2 first components, plotting millions of compounds becomes rapidly tricky.
Cell-based methods have been developed in this context and self-organizing maps is one able to simplify the representation of huge ensembles through the clustering of similar compounds in a common cell or neighbouring cell(s).
As an unsupervised method, it confers the opportunity to project new molecules on an existing map and to compared their topological and/or physicochemical profiles to the molecular set which has been used to build this map. Comparing different molecular descriptions (different maps) with the biological data to be mined, some structure-properties relationships can be easily pointed out just using the positioning of the molecules within the different maps. Even if the algorithm is unsupervised, a strategy able to evaluate the quality of the built map with a training/external validation set system can be elaborated.
A pseudo-supervised mode of this method, which consists in mixing one or several biological properties to the chemical description, was also experimented. Because this mode of representation is changing according to the biological properties added, it gives a better understanding of the structure property relationships, very useful in the managements of pharmaceutical research projects. However, it remains a visualisation tool in opposition to a prediction tool which has to be based on a chemical description only.
In order to get round this drawback, the opportunity to supervise the map to assess predictions was studied. This mode appears very attractive since it is able to manage chemical descriptions in relation to the biological data studied, as in a classical QSAR modelling method. In addition, this method can help the management of the model applicability domain prior to every in silico prediction.
In this study, the Burden eigenvalue descriptors (BCUT) were used for the chemical description of the molecules. The pseudo-supervised mode of our self organizing map algorithm will be illustrated with a biopharmaceutical parameter which refers to metabolic stability of the molecules.

FUZZY CLUSTERING AND ITS APPLICATIONS IN QSAR

Costel Sârbu

"Babeş-Bolyai" University, Faculty of Chemistry and Chemical Engineering, Department of Analytical Chemistry, Arany Janos Str. 11, RO- 400028 Cluj-Napoca, Romania

Cluster analysis is a large field, both within fuzzy sets and beyond it. Many algorithms have been developed to obtain hard clusters from a given data set. Among those, the *c-means algorithms* are probably the most widely used. *Hard c-means* execute a sharp classification, in which each object (sample) is either assigned to a class or not. The membership to a class of objects therefore amounts to either 1 or 0. The application of *Fuzzy sets* in a *classification function* causes this class membership to become a relative one and consequently an object can belong to several classes at the same time but with different degrees. The c-means algorithms are *prototype-based procedures,* which minimize the total of the distances between the prototypes and the objects by the construction of a target function. *Fuzzy generalized n-means* is easy and well improved tool, which have been applied in many fields of chemistry including QSAR. In this paper, different fuzzy clustering algorithms, namely hierarchical fuzzy clustering, hierarchical and horizontal samples and characteristics clustering and a new clustering technique, namely *Fuzzy hierarchical and horizontal cross-clustering* were applied to the study of several classes of compounds (polycyclic aromatic hydrocarbons, peticides) using different molecular descriptors. The clustering results are in a good agreement to the results reported in the literature. Much more, the characteristics clustering technique produces fuzzy partitions of the descriptors (properties) involved and thus is a useful tool for studying (dis)similarities between them. In this way it is possible to identify which descriptors or other physico-chemical features are responsible for the similarities or differences observed between different groups of active and inactive compounds. The results obtained indicated good performance in terms of classification and prediction for all the fuzzy clustering algorithms applied. The present assessment has shown that the fuzzy hierarchical cross-clustering with convex combination of point and linear prototypes algorithm is the superior and effective method for clustering active and inactive compounds. This conclusion has been also clearly supported by the values of some conventional cluster validity indices.

References:

1. Sârbu, C. and Pop, H. F., "Fuzzy Soft-Computing Methods and Their Applications in Chemistry" in Reviews in *Computational Chemistry*, K.B. Lipkowitz, R. Larter and T. R. Cundari (eds.), Wiley-VCH, **2004**, Chapt. 5, 249-332.

2. Sârbu, C., Zehl, Katharina and Einax, J. W., Fuzzy Divisive Hierarchical Clustering of Soil Data Using Gustafson-Kessel Algorithm, *Chemometrics and Intelligent Laboratory Systems*, **2007**, 86, 121-129.

INTEGRATION OF NEW SHAPE-BASED DESCRIPTORS FOR QSAR APPLICATIONS

Vladyslav Kholodovych[1*], Dmitriy Chekmarev[1], Jens Erik Pontoppidan Larsen[2], Oyenike Olabisi[1], Svava Ósk Jónsdóttir[3], William Welsh[1]

[1]Dept. of Pharmacology, UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey,USA

[2]Dept. of Systems Biology, Center for Biological Sequence Analysis, Technical University of Denmark, Denmark

[3]Dept. of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, Denmark

3D shape matching is an essential part for the most QSAR applications, including docking, virtual screening and pharmacophore mapping. Successful matching offers guidance into the molecular basis for similarities and differences and helps in further selection on computer-aided molecular modeling strategies.

Recently we introduced a new technique, Shape Signatures, that allows fast and easy comparison of molecules based on their molecular shape and electrostatic surface potential. The method employs a customized ray-tracing algorithm to explore the volume enclosed by the surface of a molecule, then uses the output to construct compact histograms (i.e., signature fingerprints) that encode for molecular shape, polarity, and other biorelevant properties. The method lends itself to rapid large-scale screening, and Shape Signatures databases can be created for an almost limitless number and variety of molecular entities (organic/organometallics, neutral/charged species, etc).

We have employed in several applications of virtual screening in drug discovery and was shown to be a promising tool for selecting potential lead compounds from libraries of commercially available chemicals. Additionally we utilized Shape Signatures as a new tool for generating shape (1D) and shape-property (2D) molecular descriptors for QSAR applications. Those sets of descriptors used alone and in combination with other molecular descriptors were applied for classification, regression and enrichment studies of pesticides, cardioglycosides, and ligands for nuclear receptors.

Following a brief description of the method itself, we present case studies that illustrate the capability and utility of Shape Signatures in the arena of predictive toxicology, environmental science and drug discovery.

*VK is a recipient of Alexander von Humboldt Research Fellowship wants to express his gratitude to AvH Foundation for support in conference attendance and in research

DC, VK and WJW gratefully acknowledge the support for this work provided by the USEPA-funded Environmental Bioinformatics and Computational Toxicology Center (ebCTC), under STAR Grant number GAD R 832721-010. This work was also funded by NIH -GM081394 from the National Institute of General Medical Sciences (to WJW).

ChemGPS-NP$_{Web}$: CHEMICAL SPACE NAVIGATION ONLINE

Josefin Rosén
Umetrics AB, Stortorget 21, 4tr, SE- 211 34 Malmö, Sweden

ChemGPS-NP$_{Web}$ is a web-based public tool[1-3] for comprehensive navigation and exploration in biologically relevant chemical space in terms of global mapping onto a consistent, eight dimensional map over structure derived physico-chemical characteristics. Compounds are positioned on this map using interpolation in terms of score prediction. It can assist in compound selection and prioritization; property description and interpretation; cluster analysis and neighbourhood mapping; as well as comparison and characterization of large compound datasets.
The ChemGPS-NP system includes three main components: DragonX from Talete for calculation of molecular descriptors; SIMCA-QP from Umetrics for multivariate predictions; and a web interface with a batch queue manager (Batchelor). Since the web-site came online in May 2008 it has processed 2646303 molecules from 1973 source files (numbers as of July 2010).
Here the website is introduced and its use exemplified with a study where the system differentiates between cancer inhibiting compounds based on their respective modes of action[4].
Automated predictive tools based on in-house experience are valuable methods for improving the success rates of projects. Interpretable e.g. QSAR and ADMET models built with Umetrics SIMCA-P+ software are exportable to Umetrics automated SIMCA-QP products. SIMCA-QP is a standalone multivariate prediction engine based on the same technology as SIMCA-P+. Due to its simple DLL design, communication can be handled through a COM interface, or by direct calls using C++ or VB. It can be integrated in any acquisition system or connected to a discovery database. The output can be shared with existing database and visualization systems. ChemGPS-NPWeb has its own visualization system.

References:

1. http://chemgps.bmc.uu.se (for non-commercial use only)
2. Larsson, J., Gottfries J., Muresan, S, and Backlund, A. ChemGPS-NP : tuned for navigation in biologically relevant chemical space. *Journal of Natural Products* 2007, 70, 789-794.
3. Rosén, J., Lövgren, A., Kogej, T., Muresan, S., Gottfries, J., and Backlund, A. ChemGPS-NPWeb: chemical space navigation online. *Journal of Computer-Aided Molecular Design* 2009, 23, 253-259.
4. Rosén, J., Rickardson, L., Backlund, A., Gullbo, J., Bohlin, L., Larsson, R., and Gottfries J. ChemGPS-NP mapping of chemical structures as a complement for prediction of anticancer mode of action. *QSAR and Combinatorial* Science 2009, 28, 436-446

CLUSTERING STRUCTURE DATABASES USING FREQUENT SUBSTRUCTURE MINING

Madeleine Seeland, Tobias Girschick, Fabian Buchwald and Stefan Kramer
Institut für Informatik I12, Technische Universität München

The goal of clustering a database of molecular structures is to identify groups of similar structures, such that intra-group similarity is high and inter-group similarity is low. This can serve to structure the chemical space and to improve the understanding of the data. Two principal approaches to cluster the structures of a chemical database have been considered in the literature so far: graph-based and fingerprint-based clustering [1]. Empirically, the quality of both types of clusterings have been found to be comparable [1]. Graph-based approaches, however, still suffer from the high computational cost of computing (variants of) the maximum common subgraph. In the presentation, we propose a novel graph-based clustering approach based on the frequent graph mining algorithm gSpan [2]. In the novel approach, clusters encompass all molecules that share a sufficiently large common substructure. The size of the common substructure of a compound in a cluster has to take at least a user-specified fraction of its overall size. The new algorithm works in an online mode (processing one structure after the other) and produces overlapping (non-disjoint) and non-exhaustive clusters.
Several experiments are designed to evaluate the effectiveness and efficiency of the structural clustering algorithm on various real-world data sets of molecular graphs. We show that the approach is able to rediscover known structure classes in the NCI standard anti-cancer agents [3]. Moreover a baseline comparison with a PubChem Tanimoto fingerprint-based clustering is presented. To investigate how well the algorithm scales regarding running time, we perform extensive experiments with 10,000 compounds selected from the NCI aids anti-viral screen data (http://dtp.nci.nih.gov/docs/aids/aids_data.html). In summary, our results suggest that this overlapping, non-exhaustive structural clustering approach increases cluster homogeneity compared to fingerprint-based clustering.

References

[1] Raymond, J.W., Blankley, C.J., Willett, P. (2003) Comparison of Chemical Clustering Methods using Graph- and Fingerprint-Based Similarity Measures, Journal of Molecular Graphics and Modelling 21(5): 421-433.
[2] Yan, X., Han, J. (2002) gSpan: Graph-Based Substructure Pattern Mining, Proceedings of the 2002 IEEE International Conference on Data Mining: 721-724.
[3] Boyd, M. R. (1989) Status of the NCI Preclinical Antitumor Drug Discovery Screen, Principles and Practice of Oncology Update, DeVita, V. T., Jr. and Hellman, S. and Rosenberg, S. A. (eds), vol 3, pp. 1–12. Lippincott: Philadelphia.

Acknowledgements

This research was partly funded by the EU FP7 project OpenTox - An Open Source Predictive Toxicology Framework (Health-F5-2008-200787) and the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety.

COMBINING LOCAL AND GLOBAL CHEMICAL DISTANCE MEASURES FOR QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS

Tobias Girschick, Fabian Buchwald and Stefan Kramer
Institut für Informatik I12, Technische Universität München, Germany

The task of QSAR model learning is often tackled by instance-based methods (like k-Nearest Neighbors), which are all based on the notion of chemical (dis-)similarity. Our starting point is the observation by Raymond and Willett [1] that the two big families of chemical distance measures, fingerprint-based and maximum common subgraph (MCS) based measures, provide orthogonal information about chemical similarity. While fingerprint-based measures capture the local similarity of two structures, MCS-based measures capture their global similarity. Instead of selecting a fixed one of these measures, we propose to use linear combinations of the contributions of the different distance measures instead of only one. In the presentation, we will show how optimal linear combinations can be learned directly from a given data set or transferred from a related QSAR task. If the size of the data set is sufficiently large, optimizing linear combinations of local and global distance measures can improve upon individual measures. If, however, the size of the data set is too small, it can be beneficial to transfer knowledge from a related problem, e.g. (to some extent) from mouse carcinogenicity to rat carcinogenicity. In the presentation, we will show the performance of several strategies to learn or transfer combined chemical distance measures on DHFR inhibition, ER relative binding affinity and CPDB mouse and rat carcinogenicity data.

References

[1] Raymond, J. W., and Willett, P. Effectiveness of graph-based and fingerprint-based similarity measures for virtual screening of 2D chemical structure databases. *Journal of Computer-Aided Molecular Design* (2002). 59-71.

Acknowledgements

This research was funded by the EU FP7 project OpenTox - An Open Source Predictive Toxicology Framework (Health-F5-2008-200787).

AUTOQSAR: A SYSTEM FOR AUTOMATED MODEL SELECTION FROM MULTIPLE CHEMICAL DOMAINS

David Wood[1a], Sarah Rodgers[2], David Buttar[1a], John Cumming[1a], Andy Davis[1b] and Ulf Norinder[1c]

[1] AstraZeneca,
[a] Alderley Park, Macclesfield, Cheshire, UK
[b] Charnwood, Loughborough, UK
[c] Södertälje, Sweden
[2] Accelrys, Cambridge Science Park, Cambridge, UK

It is frequently observed that global QSAR models can be biased in their predictions for certain chemical series, and significant computational chemistry resources are often required by project teams to generate local models that correct these biases. In addition to this, previous research has demonstrated that QSAR models that are regularly updated with recently measured data tend to result in more accurate predictions than static models [1] [2]. AutoQSAR was proposed to automate these modeling procedures, automatically generating local models, and ensuring that the models reflect the most recent chemistry. AutoQSAR is currently being developed for AstraZeneca by Accelrys.

Users of AutoQSAR submit tables of measured data and a global model is created for each measured property. If there are sufficient data available, AutoQSAR will also generate models for individual drug development projects and individual chemical series sub-domains. Multiple regression techniques are evaluated for each of the sub-domains, and AutoQSAR identifies the most predictive model for any given query compound.

We present a retrospective, cross project analysis of the modeling approach used by AutoQSAR for a range of measured properties, specifically logD, solubility, and human plasma protein binding. We demonstrate the improvements to prediction accuracies that would have been achieved had AutoQSAR been operational from the beginning of 2009 by comparison to static- and updating- global models. Finally we demonstrate how we have used this analysis to optimize AutoQSAR's model selection protocols.

References:

1. Rodgers et al. (2006) "Time-series QSAR analysis of human plasma protein binding data", *QSAR and Combi. Sci.*, **26**, p511-521

2. Rodgers et al. (2007) "QSAR modeling using automatically updating correction libraries: application to a human plasma protein binding model", *JCIM*, **47**, p2401-2407

AZORANGE – HIGH PERFORMANCE OPEN SOURCE MACHINE LEARNING IN A GRAPHICAL PROGRAMMING ENVIRONMENT

Jonna Stålring[1], Lars Carlsson[1], Pedro Almeida[2] and Scott Boyer[1]
[1]Global Safety Assessment, AstraZeneca R&D, Sweden
[2]EngInMotion Ltd, Portugal

Development of a quantitative structure activity relationship (QSAR) requires a numerical representation of molecular structure and the selection of a decision generator to predict a biological response. In principal, there is an infinite number of ways to represent molecular structure and the selection of machine learning (ML) algorithms is vast. Confined to a set of available descriptors, development of a QSAR model can be reduced to the high dimensional statistical problem of selecting the most representative descriptors and the most accurate ML method.

AZOrange is a platform for development of ML models, including functionality to facilitate QSAR modeling in particular. AZOrange makes several state of the art ML algorithms available in a graphical programming interface, as well as in a flexible scripting environment. AZOrange aims to be a comprehensive platform for QSAR development, providing methods for descriptor calculation, data preprocessing, model building and evaluation. In addition, the extensive scripting capabilities, in conjunction with high performance ML methods, enable exhaustive testing of the methods available in AZOrange.

AZOrange relies solely upon open source code. The Orange machine learning platform constitutes the framework into which other codes are integrated. The major machine learning component appended to Orange is the OpenCV package, interfacing Random Forests (RF), Support Vector Machine (SVM) and Artificial Neural Network (ANN) algorithms. Furthermore, as a step towards automated model development and to assure a valid selection of model parameters, AZOrange provides a generalized model parameter optimizer. Any number of parameters can be optimized simultaneously in a distributed computational environment by the pattern search algorithm in APPSPACK.

The AZOrange algorithms have been validated with benchmarks against the R RF module. The accuracy is assessed over a suite of public data sets from the UCI machine learning repository and from QSAR publications and databases. The Wilcoxon rank test determines whether a statistically significant difference in achievable accuracy exists between the algorithms. In particular, the benchmarks show the impact of the generalized model parameter optimizer on model accuracy.

A few public QSAR data sets are selected to illustrate the value of the automated multiple hypothesis testing. Tens of thousands of models, based on various descriptor sets and ML algorithms with optimized parameters, are built in one day in a distributed environment. The accuracies of the thereby obtained models are compared to those of a more traditional approach based on a few models representing a best practice selection of descriptors and ML algorithms.

AZOrange is a unique QSAR development platform by making several computationally efficient and validated ML algorithms available within the same framework. The integrated workflow, the automated model parameter optimizer, the support for distributed execution and the scripting capabilities facilitate multiple hypothesis testing, potentially increasing model accuracy beyond what is in practice achievable by other algorithms.

Acknowledgments
We would like to acknowledge the Open Source community and in particular the Artificial Intelligence Laboratory at the University of Ljubljana, without which this project had not been possible.

POSTER PRESENTATION I-40

ENABLING PARTIAL SHAPE SIMILARITY SEARCHING WITH INDEX-DRIVEN VIRTUAL SCREENING

Christin Schärfer, Matthias Rarey
University of Hamburg, Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany

Most tools for virtual high-throughput screening use a sequential screening procedure: Every compound of a given library is screened against a target protein respectively a reference ligand. For ligand-based approaches using a small molecule alignment, this requires calculating superpositions between the reference and each compound of the library to obtain a ranked list of hit compounds. As calculating molecular superpositions is computationally expensive, we developed a new tool for ligand-based virtual screening named TrixS BMI which avoids the sequential screening procedure by reducing the number of compounds to superimpose in a computationally much faster preprocessing step. This allows for sublinear runtimes with respect to the library size while still providing comparable enrichment and hit rates. Furthermore, the method allows a user-controlled partial shape match reflecting the typical scenario in complex formation that only parts of the binding molecules are involved.

TrixS BMI is an adaptation of the structure-based virtual screening tool TrixX BMI[1] and employs descriptors containing chemical and shape information, and an indexed database. An outline of the workflow could be described as follows: Conformations and descriptors are calculated in a one time effort for every compound of a given compound library. The compound descriptors are stored in an indexed database using the FastBit[2] indexing technology. Conformations are generated with the build-in TrixX Conformer Generator[3] (TCG). The database is queried by descriptors calculated from a reference ligand, resulting in a preselection of compounds for the superposition process. As each preselected compound already automatically implies a superposition with the reference ligand, the last step is only to assess these superpositions.

Most shape-based screening tools work under the basic assumption that molecules of same size and chemistry to known actives are likely to show a similar activity. These tools ignore the fact that while ligands fitting into the same binding pocket of a receptor share at least one common region of similar shape, they can differ in their overall shape. TrixS takes these considerations into account by providing different shape modes, ranging from 25% to 100% shape similarity requirement.

Screening experiments show that TrixS BMI obtains similar good enrichment and hit rates compared to other ligand-based virtual screening tools.

References:

[1] Schlosser, J.; Rarey, M. J. Chem. Info. Model. 2009, 49(4):800-809
[2] Wu, K. J. Phys.:Conf. Ser. 2005, 16:556-560
[3] Griewel, A.; Kayser, O.; Schlosser, J.; Rarey, M. J. Chem. Info. Model. 2009, 49(10):2303-2311

CONSENSUS QSAR MODELING AND DOMAIN OF APPLICABILITY: AN INTEGRATED APPROACH

Georgia Melagraki[a], Pantelis Sopasakis[a], Antreas Afantitis[a-b], Haralambos Sarimveis[a], and Olga Igglessi-Markopoulou[a]

[a] School of Chemical Engineering, National Technical University of Athens, Athens, Greece
[b] Department of ChemoInformatics, NovaMechanics Ltd, Cyprus

Consensus modelling is a term that has been used in many scientific disciplines to define methods by which a group of individuals can come to an agreement. The QSAR community has used this term for methodologies that aggregate the predictions of several QSAR models to arrive at a single prediction. Literature reports on the validity of consensus modelling approaches are quite conflicting. Many publications present advantages of consensus models: More accurate QSAR models, greater confidence in predictions, regulatory significance, improved robustness. Several other references however, have criticized consensus modelling for complexity, lack of portability, transparency and mechanistic interpretation and for not showing significant improvements over single QSAR models.

Many consensus QSAR models that have appeared in the literature use a naïve approach, that calculates the average value among all the individual model predictions. More sophisticated methods consider only the models for which the compound to be predicted falls into their domain of applicability. Alternative consensus modelling methods consider the individual model predictions as attributes in an overall multiple linear regression model, where the model coefficients play the role of weights. This way, the contribution of each individual model in the overall prediction is weighted.

In this work, we present a new approach, integrating three basic components in the process of building a QSAR model: variable selection, regression/classification, and domain of applicability. In particular, the proposed method requires a single wrapper variable selection method, a single method for defining the domain of applicability and many regression/classification algorithms depending on the type of the problem. The wrapper variable selection method is applied separately to each QSAR algorithm and produces a QSAR model which used a certain subset of features. In general, different sets of features are selected by the various QSAR models that are generated. Thus, for each QSAR model, a different domain of applicability is defined, by applying the domain of applicability method on the respective set of descriptors. For a new compound, the proposed method first calculates the individual QSAR models predictions. Then it checks for each model, if the compound falls into its domain of applicability. In the case of a negative answer, the model is not taken into account in the calculation of the aggregated prediction. If the answer is positive, a weight is produced depending on the location of the compound inside the domain of applicability. Obviously the weight becomes lower when the location of the compound is closer to the boundaries of the domain of applicability. The weights are finally normalized, so they add to 1. The normalized weights are used to produce the final aggregated prediction. The results of the application of the method to QSAR problems illustrate the advantages and limitations of the method.

REPLACING STRUCTURE-BASED PHARMACOPHORE FILTERING BY HIGH-THROUGHPUT DOCKING USING CONSTRAINTS: A CASE STUDY TARGETING A PROTEIN-PROTEIN INTERFACE

Oliver Koch[1,2], Richard J. Marhöfer[1], Leopold Flohé[2], Timo Jaeger[2], Paul M. Selzer[1]

[1] Intervet Innovation GmbH, Zur Propstei, 55270 Schwabenheim, Germany

[2] MOLISA GmbH, Brenneckestraße 20, 39118 Magdeburg, Germany

Using high-throughput docking of millions of compounds in a structure based virtual screening campaign is still not feasible. The main drawback is the lack of enough computational power to get accurate results in a reasonable time. Pharmacophore based filtering is a standard procedure to massively reduce the number of compounds for docking in the first place[1], thus reducing the computational time for docking. High throughput docking using constraints combines the strength of both worlds and it can be applied as pre-filtering step before accurate docking. The constraints give the opportunity to dock with very fast settings. Using GOLD[2], it is possible to filter for specific protein-ligand hydrogen bonding or hydrophobic interactions, protein-ligand atom distances or specific positions of substructures in the binding pocket. As advantage over pharmacophore filtering, hydrogen bond strength cut-offs can be set, all binding site properties are included and a discussion about the right conformer generation, like in the case of pharmacophore approaches, is no longer required.

As a test case, high-throughput docking using constraints was applied for targeting the thioredoxin reductase (TrxR) / Thioredoxin (Trx) System in *Mycobacterium tuberculosis*. It is part of the antioxidant system that inactivates peroxides, contributes to ribonucleotide reduction, and thus guarantees the survival within macrophages[3]. *M. tuberculosis* lacks the common glutathione system and the M.t. TrxR shows a substantial difference in sequence, mechanism and structure to the human TrxRs. This makes the TrxR a promising new target of drugs for the treatment of tuberculosis.

Two important hydrogen bonding interactions were identified at the protein-protein interface of the TrxR-Trx complex. A high-throughput docking using constraints was applied to filter the Intervet *in-house* compound library ($\sim$6.5 million compounds) for possible hits that interact with these two hydrogen bonding acceptors. This reduced the number of interesting compounds to $\sim$150000 that were redocked without any constraints and most accurate docking settings. Rescoring using the same constraints leads to $\sim$11.000 compounds for the final compound selection that was based on an automated ranking using a normalisation and consensus scoring strategy.

So far, 17 out of the first 170 tested compounds showed an activity with an IC_{50} value upto µM range. Especially four different scaffolds with a low molecular weight are promising candidates for further developments. Slight modification of the best hit lead already to an improved activity with an IC_{50} value of 1.6 µM.

References:

[1] Klebe, G., Virtual ligand screening: strategies, perspectives and limitations. 2006, Drug Discovery Today 11, p. 580 – 594

[2] Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., Taylor., R. D. Improved Protein-Ligand Docking Using GOLD. 2003, Proteins 52, p. 609-623

[3] Jaeger, T., Flohé, L., The thiol-based redox networks of pathogens: unexploited targets in the search for new drugs. 2006, Biofactors 27, p. 109-120

CHEMPROT: A DATABASE OF BIOACTIVE CHEMICALS ASSOCIATED TO A PHENOME INTERACTOME NETWORK

Olivier Taboureau, Sonny Kim Nielsen, Karine Audouze, Nils Weinhold, Daniel Edsgard, Francisco Roque, Thomas Skot Jensen, Søren Brunak and Tudor Oprea

Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Denmark

In the emerging area of systems pharmacology there is an increasing need for developing network-based approaches in order to understand the relationship between drug action and disease susceptibility genes. Via protein-protein interaction (PPI) networks one can better study the properties of biological systems. Analyses of such PPI networks have contributed to the genesis of systems pharmacology. These studies can improve our understanding of drug targets, suggest novel targets as well as new approaches for therapeutics.

Here, we will present ChemProt, a database that integrate large scale of bioactive molecules (chemogenomics), PPI data associated to tissue and disease information (phenome-interactome). We were able to assembly more of 620 000 unique chemicals with protein associations for 15000 proteins. In total, 2.100.000 chemical-protein interactions were gathered and integrated in a human protein-protein interaction network of 428.429 PPIs. Diseases and tissues specificity can then be explored through each protein complex.

Taking together, this system can provide a better understanding of the underlying molecular mechanisms of chemicals and how they perturb biological pathways. It can improve our in silico evaluation of approved drugs for repurposing, and our ability to select new compounds based on estimates against many potential targets, including those related to adverse drug events. Such systemic evaluation of certain biological pathways may be critical for the identification of additional genes that may play a role in modulating drug responses.

MINING EXHAUSTIVELY THE PROTEIN DATA BANK ENABLES COMPUTATIONAL FRAGMENT-BASED DRUG DESIGN

Fabrice Moriaud, Olivia Doppelt-Azeroual, Laetitia Martin, Ksenia Oguievetskaia, Artem Vorotyntsev, Stewart A. Adcock, and Francçois Delfaud

MEDIT S.A., 2 rue de Belvedere, Palaiseau, France

Based on the assumption that similar protein surfaces are likely to bind the same fragment with the same pose, an efficient protocol in computational FBDD (Fragment-Based Drug Design) is proposed.
The large volume of protein-ligand structures now available in the PDB enables applications of the protocol for diverse fragments and for any protein families.
In this work, we build a database of MED-Portions, where a MED-Portion is a new structural object encoding protein-fragment binding sites. MED-Portions are derived from mining all available protein-ligand structures with any library of small molecules. Combined with the MED-SuMo software to superpose similar protein interaction surfaces, pools of matching MED-Portions can be determined for any binding surface query.
To generate hitlike molecules from fragments, MED-Portions are combined in 3D with the MED-Hybridise toolkit. The described MED-Portion/MED-SuMo/MED-Hybridise protocol is applied to three targets: a protein kinase, a G-Protein Coupled Receptor (GPCR), and Eg5, a mitotic kinesin. The results show the potential for finding relevant compounds targeting any protein 3D structure since the occurrence of interfamily MED-Portions is 25% for protein kinase, almost 100% for the GPCR and for the hydrophobic pocket of the Eg5 allosteric pocket.

COMPUTER-AIDED PREDICTION OF BIOLOGICAL ACTIVITY SPECTRA FOR SUBSTANCES: VIRTUAL CHEMOGENOMICS

V. Poroikov, T. Gloriozova, A. Lagunin, D. Filimonov
Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences,
Pogodinskaya Str., 10, Moscow, 119121, Russia

The numbers of pharmacological targets and compounds available for screening have significantly arose in recent years making it impossible to test available libraries against all possible targets. Computational prediction of biological activity spectra for chemicals reduces the risk of missing useful pharmacological as well as unwanted adverse/toxic effects.

We have developed the PASS software that predicts about 4000 types of biological activity based on structural formula of drug-like organic molecules with average accuracy about 95% [1, 2]. Prediction is realized through the ligand-based design approach on the basis of structure-activity relationships established for about 260,000 drugs, drug-candidates and pharmacological agents.

PASS predicted the activity profile of 250,000 compounds from the NCI database with the hit rate increased up to ~17 times, and new pharmacological agents have been detected [3]. Application of PASS to selection of the most prospective compounds from virtual libraries significantly increases a probability of finding compounds with the required properties [4].

PASS predictions are available via Internet (http://www.ibmc.msk.ru/PASS). About forty papers have been published, in which PASS predictions for diverse chemical classes and different types of activity were confirmed by further chemical synthesis and biological testing (for review see [5]).

PASS can be used as in silico chemogenomics tool, for estimation of the most probable targets and biological effects of drug-like chemical compounds' action on these targets, thus significantly increasing the efficiency to find hits and leads with the required properties.

Acknowledgements. This work was partially supported by the CRDF grant RC1-2064, RFBR grants 05-07-90123 and 06-03-08077, INTAS grant 03-55-5218, ISTC grants 3197 and 3777, European Commission FP6 grant LSHB-CT-2007-037590 (Net2Drug) and FP7 grant 200787 (OpenTox).

References:

[1] Filimonov D.A., Poroikov V.V. Probabilistic approach in activity prediction. In: *Chemoinformatics Approaches to Virtual Screening*. Eds. Alexandre Varnek and Alexander Tropsha. **2008**, Cambridge (UK): RSC Publishing, 182-216; http://www.ibmc.msk.ru/PASS

[2] Filimonov D.A., Poroikov V.V. Prediction of biological activity spectra for organic compounds, *Russian Chemical Journal,* 2006, **50**: 66-75.

[3] Poroikov V.V., Filimonov D.A., Ihlenfeldt W.-D., et al. PASS Biological Activity Spectrum Predictions in the Enhanced Open NCI Database Browser, *J. Chem. Inform. Comput. Sci*., 2003, **43**: 228-236.

[4] Geronikaki A., Dearden J., Filimonov D., et al. Design of new cognition enhancers: from computer prediction to synthesis and biological evaluation, *J. Med. Chem*., 2004, **47**: 2870-2876.

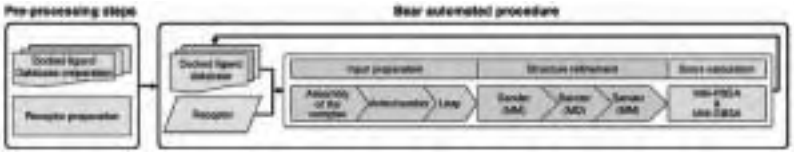
[5] Geronikaki A., Druzhilovsky D., Zakharov A., Poroikov V. Computer-aided predictions for medicinal chemistry via Internet, SAR and QSAR *in Environ. Res*., 2008, **19**: 27-38.

BEAR, A NOVEL VIRTUAL SCREENING METHODOLOGY BASED ON MOLECULAR DYNAMICS REFINEMENT AND ACCURATE BINDING FREE ENERGY ESTIMATION

M. D. Parenti, G. Degliesposti, C.Portioli and G. Rastelli
Dipartimento di Scienze Farmaceutiche, Università di Modena e Reggio Emilia, Italy

In the drug discovery process, accurate methods of computing the affinity of small molecules with a desired biological target are strongly needed. Even if, in the last years, the accuracy and efficiency of the available virtual screening algorithms have been improved, many drawbacks and limitations still exist. For example, docking methods lack a reliable simulation of both ligand and receptor flexibilities, as well as good scoring functions able to estimate ligand binding energies in reasonable agreement with experimental data. These limitations often lead to a high level of false positives or false negatives in the hit list. For that reason, it is generally agreed that docking results need to be post-processed with more accurate tools.
To this end, we developed **B**inding **E**stimation **A**fter **R**efinement (**BEAR**), a new and automated post-docking procedure for the conformational refinement of docking poses through molecular dynamics (MD) followed by accurate prediction of binding free energies using MM-PBSA and MM-GBSA[1] (Fig.1).
The BEAR performance in virtual screening was evaluated on several macromolecular targets and related sets of known ligands, determining the enrichment factors and assessing the correlation between predicted and experimental binding affinities. These analyses suggested critical improvements with respect to standard docking softwares[2,3].
Moreover, when applied in virtual screening campaigns, BEAR was able to discover novel and potent inhibitors of *Plasmodium* falciparum plasmepsin II with an impressive hit rate [4], and has been successful in identifying promising scaffolds for the design of irreversible protein kinase inhibitors[5]. Therefore, taken as a whole, the results obtained so far prospect that BEAR may become a prominent tool in the drug discovery pipeline.
The BEAR virtual screening procedure is reliable and strongly automated, and can be tailored to the needs of the end-user in terms of computational time and the desired accuracy of the results. BEAR is under constant development and validation on additional biological targets in order to further improve accuracy, automation and calculation speed.

Fig.1: The BEAR Workflow



References:
1. Rastelli G, Degliesposti G, Del Rio A, Sgobba M (2009) Chem Biol Drug Des 73:283–28
2. Ferrari AM, Degliesposti G, Sgobba M, Rastelli G (2007) Bioorg Med Chem 15:7865–787
3. Rastelli G, Del Rio A, Degliesposti G, Sgobba M (2010) J Comput Chem 31:797–81
4. Degliesposti G, Kasam V, Da Costa A, Kang HK, Kim N, Kim DW, Breton V, Kim D, Rastelli G. ChemMedChem. 2009 Jul;4(7):1164-73.
5. Del Rio A, Sgobba M, Parenti MD, Degliesposti G, Forestiero R, Percivalle C, Conte PF, Freccero M, Rastelli G. J Comput Aided Mol Des. 2010. in press. doi: 10.1007/s10822-010-9324-x

POSTER PRESENTATION I-47

IDENTIFICATION OF FREQUENT HITTER STRUCTURES AND SUBSTRUCTURES IN A COMMERCIAL GENERAL PURPOSE uHTS SCREENING COLLECTION

Christian J. Kirchhoff, Pierre Ilouga, Alexander Böcker, Birgit Sauer-Hecks
Evotec AG Schnackenburgallee 114, D-22525 Hamburg, Germany

Industrial uHTS has become a vital part to drug discovery. High numbers of potentially interesting chemical entities are routinely tested against single specific targets of interest, both in biochemical and cellular assays. Detecting frequent hitters and promiscuous compounds are a key aspect in the interpretation and analysis of screening data especially if used for deriving early SAR [1].
Evotec is a global service provider along the drug discovery value chain progressing new chemical entities identified during high-throughput screening (HTS) to clinical candidate nomination. A carefully selected maximum diverse library of more than 250,000 drug-like compounds is readily available for uHTS screening. Beside commercially available compounds, the Evotec Lead Discovery Library comprises a significant fraction of Evotec exclusive compounds (about 30%). The library also includes several target-focused libraries.
The Evotec compound collection is actively managed [2] and Evotec invests much effort to maintain highest quality standards and to keep up with the expansion of accessible chemistry space world wide. This includes continuously selection and purchase of new chemical entities to replace depleted parts of the collection as well as synthesis of novel Evotec exclusive compounds. To address the library enhancement as efficiently as possible one has to be aware of promiscuous scaffolds and substructures in the selection process to avoid unwanted compounds.
The approach at Evotec includes both monitoring of the appropriate literature and collecting feedback from the internal medicinal chemistry experts. All promiscuity information is encoded into SMARTS strings and will alert not only the purchasing team but also all members during analysis and the selection process in the HTS screening cascade. In addition historical screening data is analysed to obtain new substructure alert information [3].
Evotec is in the fortunate position to have executed more than 150 uHTS campaigns on a large variety of target classes with different assay principles and readout technologies. This includes also a significant set of counter screen exercises that experimentally detect the compound to readout interference. Readout and assay principle based interference is one of the major sources of false positives in screening.
The poster will give insight into the analysis of assay/readout type specific interference and the identification of substructures alerts recommended to be treated with care.

References:

[1] Baell JB, Holloway GA. J Med Chem. 2010 Apr 8;53(7):2719-40
New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays.
[2] Ilouga PE, Winkler D, Kirchhoff C, Schierholz B, Wölcke J. J Biomol Screen. 2007 Feb;12(1):21-32
Investigation of 3 industry-wide applied storage conditions for compound libraries.
[3] Wölcke J, Hunt N, Jungmann J, Ullmann D. J Biomol Screen. 2006 Jun;11(4):341-50
Early identification of false positives in high-throughput screening for activators of p53-DNA interaction.

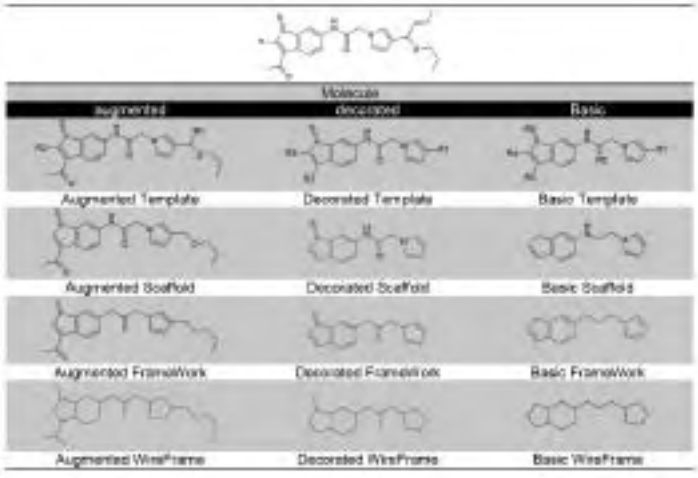
MULTI-DIMENSIONAL HIERARCHICAL SCAFFOLD ANALYSIS

Beccari A.R., Trombetta S., Lorenzi S.
Dompé S.p.A. Via Campo di Pile - 67100 L'Aquila, Italy

The analysis of chemical diversity of huge collections of commercially available, synthetically accessible, or in silico designed, compounds have become a topic of considerable interest in the drug discovery process to guide the exploration of chemistry space [1,2]. Also the problem of recognition and classification of meaningful relationships between shared chemical features and target biological activity, involves the analysis of large sets of diverse compounds, generally produced during high-throughput screening (HTS) campaigns.

To address these topics many different solutions have been developed in the last twenty years. The most relevant strategies can be divided in different groups: property based, structure based, or pharmacophoric based descriptors. Recently the scaffold based approaches have become widely implemented due to the efficiency in the clustering analysis and in the database indexing [3,4,5]. They are also generally more interpretable from a medicinal chemist prospective.

In this context we developed a new flexible algorithm that is able to generate different levels of scaffold's abstractions (scaffolds, frameworks, and wireframes) and deconstructions. This procedure generates unbiased multi-dimensional hierarchies, which are effective tools in the HTS data analysis where, due to the diversity driven composition of the screening libraries, conventional scaffold representations often are not the most suitable metrics for compound clustering.



[1] Blum L. C. et al. J. Am. Chem. Soc., 2009, 131 (25), pp 8732–8733;
[2] Chuprina A. et al., J. Med. Chem., 2010 asap;
[3] Schuffenhauer A. et al. J. Chem. Inf. Model., 2007, 47 (1), pp 47–58;
[4] Wilkens S. J. et al J. Med. Chem., 2005, 48 (9), pp 3182–3193;
[5] Wetzel S. et al Nat Chem Biol. 2009 Aug;5(8):581-3;

HYDE SCORING OF PROTEIN PROTEIN COMPLEXES

Gudrun Lange[1], Eva Vennmann[2], Nadine Schneider[2], Robert Klein[1], Matthias Rarey[2]

[1] Computing, Bayer CropScience AG, Industriepark Hoechst, G836, 65926 Frankfurt am Main, Germany

[2] Center for Bioinformatics, University of Hamburg, Bundesstrasse 43, D-20146 Hamburg

Scoring functions describe the interaction between molecules such as the binding of ligands to target proteins and the interactions between proteins. In order to test the abilities of our novel scoring function HYDE (1) to describe the interaction between proteins, we analyzed a number of different protein protein complexes. HYDE has the advantage that it consists of only two terms both of which are based on physical chemistry: (a) the dehydration of interface atoms derived from logP data of small organic compounds and (b) the H-bond energies of interacting H-bond functions. Since no other data such as binding affinities of protein ligand complexes or specific correction factors or terms are included, HYDE should be a generally applicable scoring function capable of describing the interaction between proteins.

In particular, we wanted to look at the "hot spot theory" (2) according to which only a few amino acids in a protein protein interface contribute strongly ("hot spots") while most of the interacting amino acids contribute only weakly to the overall binding. Several structures of hormone receptors with bound peptide hormones were selected from the Brookhaven data base (3) and analyzed using HYDE. These included (a) hGH-R and growth hormone for which the "hot spot theory" was originally proposed, (b) the extracellular domain of PTH-R complexed with PTH or PTH-related hormone, (c) the domains 11-13 of IGF2-R complexed with IGF-II, (d) the extracellular domain of GLP1-R complexed with extendin or GLP-1, (e) FSH-R complexed to FSH and (f) TSH-R complexed to its antibody. In all cases, we found a good correlation between the published experimental affinity changes due to the mutation of the interacting amino acids into alanine ("alanine scanning") and the predicted HYDE contributions of the respective amino acid or "in silico alanine scanning", respectively. In addition, our analysis showed that in some cases, including the interaction between the hormone growth factor and its receptor, the main contributions to binding are indeed due to a small number of interacting amino acids. However, we have also seen other cases in which many amino acids stabilize the protein protein interaction by small contributions which in turn results in a substantial overall affinity.

Protein protein interactions are involved in many important biological processes and thus they are important but extremely challenging targets for small molecule drug discovery. Discussed reasons for the limited success of finding small molecule mimics include the large size of the interacting surface and the difficulty to know which part of the interactions site needs to be mimicked by the small molecule. Our examples show that HYDE is able to pinpoint the important amino acids in the interface and predict the amount by which the amino acids forming a protein protein interface contribute to overall binding. Thus, HYDE can support the design of small molecules interfering in biological processes initiated by protein protein interactions. The results also show that a single point mutation can lead to a significantly altered affinity thus giving rise to the selectivity required in biological processes.

(5) Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. *ChemMedChem* **2008,** 3, 885-897.

(6) Clackson, T., Wells, J. A. *Science* **1995**, 267, 383-6.

(7) Berman et al., *Nucleic Acids Research* **2000**, 28, 235-242.

POST-DOCKING OPTIMIZATION WITH THE AMMOS SOFTWARE TOOL FOR IN SILICO SCREENING: ANALYSIS OF PROTEIN-LIGAND INTERACTIONS

Tania Pencheva[a], Dessislava Jereva[b], Ilza Pajeva[a], Maria A. Miteva[c]

[a] Centre of Biomedical Engineering, Bulgarian Academy of Sciences, Sofia, Bulgaria

[b] Sofia University, Sofia, Bulgaria

[c] MTI, INSERM U973 – University Paris Diderot, France

Previously we reported an open-source software tool AMMOS (Automated Molecular Mechanics Optimization tool for *in silico* Screening) for structural refinement of compound collections and energy minimization of protein-ligand complexes in virtual ligand screening. The tool performs an automatic minimization of protein-ligand complexes, based on molecular mechanics, at five levels of the protein receptor flexibility: 1) all protein atoms can move; 2) only the atoms of protein side chains can move; 3) only the protein atoms inside a sphere around the ligand can move; 4) only the atoms of the protein side chains inside a sphere around the ligand can move; 5) only the ligand atoms can move. AMMOS has been tested in relation to improvement of the enrichment after docking allowing 40 to 60% of the initially added active compounds to be found in the top 3% to 5% of the entire compound collection.

In this work we report on a further validation of AMMOS by identification and comparison of the particular protein-ligand interactions in several binding sites of different topology and physics-chemical properties explored previously [1]. The interactions of the proteins with their ligands have been analyzed and compared for the X-ray crystal structures and the docked poses and those obtained with AMMOS in all five cases of structure flexibility using the Ligand Interaction Application in MOE [2]. The preliminary results show that the implementation of AMMOS at the post-docking stage allows refinement of the protein-ligand interactions and depending on the level of flexibility restores to a different extent the interactions identified in the experimental structures. The results outline the usefulness of the protein flexibility consideration in AMMOS at post-docking optimization stage.

Acknowledgement: TP appreciates a grant from the City Hall of Paris. We acknowledge the financial supports from the INSERM institute, the University Paris Diderot and the National Science Fund of Bulgaria (grants No. DTK02/58 and No. RNF01/0110).

References

1. Pencheva T., D. Lagorce, I. Pajeva, Br. Villoutreix, M. Miteva, AMMOS: Automated Molecular Mechanics Optimization Tool for *in silico* Screening, *BMC Bioinformatics*, 2008, 9:438.

2. MOE, Chemical Computing Group Inc., Montreal, http://www.chemcomp.com.

TOMOCOMD-CAMPS AND 2D-PROTEIN BILINEAR INDICES: NOVEL BIO-MACROMOLECULAR DESCRIPTORS FOR PROTEIN RESEARCH. 1. PREDICTING PROTEIN STABILITY EFFECTS OF A COMPLETE SET OF ALANINE SUBSTITUTIONS IN ARC REPRESSOR

Sadiel E. Ortega Broche, Yovani Marrero-Ponce

Unit of Computer-Aided Molecular ''Biosilico'' Discovery and Bioinformatic Research (CAMD-BIR Unit), Faculty of Chemistry-Pharmacy, Central University of Las Villas, Santa Clara 54830, Villa Clara, Cuba

A new set of amino-acid based biomacromolecular descriptors support on a bilinear map are presented. This novel approach has been designed from a linear algebra point of view. These biochemical descriptors are based on the computation of bilinear maps on R^n in canonical basis. Protein's bilinear indices are calculated from k^{th} power of non-stochastic and stochastic graph–theoretic electronic-contact matrices, M_m^k and ${}^sM_m^k$, respectively. That is to say, the k^{th} non-stochastic and stochastic protein's bilinear indices are calculated using M_m^k and ${}^sM_m^k$ as matrix operators of bilinear transformations. Moreover, biochemical information is codified by using different pair combinations of amino-acid properties as weightings (z-values, side-chain isotropic surface area, amino-acids atomic charges and hydrophathy index). Quantitative models that discriminate near wild-type stability alanine-mutants from the reduced-stability ones in training and test series were obtained. Non-stochastic and stochastic equations permitted the correct classification of 100% (41/41) and 97.56% (40/41) of proteins in training set, respectively. Correct classification in test sets were 91.67% for both models. In order to predict Arc alanine-mutant's melting temperature (tm) and free energy differences of folding process with regard to wild-type Arc ( ), multiple linear (MLR) and piecewise regression (PLR) models were developed. The MLR model obtained by using non-stochastic bilinear indices accounted for 83% of the variance of the experimental t_m (SDEC = 3.57 °C) as long as the stochastic bilinear indices-based equation describe 83% of the tm variance (SDEC = 3.59 °C). Statistics associated to the internal (leave-one-out, bootstrapping and Y-randomization) and external validation procedures evidenced robustness, stability and suitable power ability for both models (= 0.77, Y_{BOOT} = 0.73 SDEP = 4.20 °C, a (R^2) =0.13, a (q^2) =-0.37 and = 0.80 for non-stochastic and = 0.73, Y_{BOOT} = 0.70 SDEP = 4.53 °C, a (R^2) =0.12, a (q^2) =-0.38 and = 0.62 for stochastic bilinear indices). Piecewise regression models based on non-stochastic and stochastic protein's descriptor accounted for 90% and 92% of the tm variance, respectively, for those cases in training set and explained 86% and 73% of tm variance, correspondingly, for those cases in test set. The best MLR models obtained to predict values of Arc mutants accounted for 83% of dependent-variable variance for those cases in training and test samples, respectively. The performance of this non-stochastic descriptors based model in the prediction of thermodynamic parameters for Arc-mutants ( ) was higher than the achieved by PoPMuSiC algorithm (Gilis and Rooman, 2000) for the same dataset. An analysis of the relevance of protein structural information for the numerical characterization of Arc mutants and its relationship with stability changes was performed. On the other hand, it has been demonstrated that we can use the linear discriminant analysis and piecewise models in combination to classify and predict the stability of the mutant Arc homodimers. Statistical parameters for protein's bilinear indices based models compared favorably with those of other models reported previously and developed with a similar purpose (Marrero-Ponce, Medina Marrero et al. 2004; Marrero-Ponce, Medina-Marrero et al. 2005). The models developed in this work also permitted the interpretation of the driving forces of such a folding process, indicating that topologic/topographic protein's backbone interactions control the stability profile of wild-type Arc and its alanine-mutants. The results

achieved demonstrate the ability of protein's bilinear indices to encode biochemical information related to those structural changes which significantly influence the Arc repressor stability when punctual mutations are induced.

References:

Gilis, D. and M. Rooman (2000). Protein Engineering **13**(12): 849-856.
Marrero-Ponce, Y., R. Medina-Marrero, et al. (2005). Bioorganic & Medicinal Chemistry **13**: 3003-3015.
Marrero-Ponce, Y., R. Medina Marrero, et al. (2004 Molecules **9**: 1124-1147.

POSTER PRESENTATION I-52

3D-PROTEIN BILINEAR INDICES: NOVEL BIOMACROMOLECULAR DESCRIPTORS FOR PROTEIN RESEARCHS. 1. CLASSIFYING PROTEIN DOMAINS ACCORDING TO ITS TOPOLOGY

Sadiel E. Ortega Broche[a], Yovani Marrero-Ponce[a], Néstor Cubillán Acosta[b]

[a] Unit of Computer-Aided Molecular ''Biosilico'' Discovery and Bioinformatic Research (CAMD-BIR Unit), Faculty of Chemistry-Pharmacy, Central University of Las Villas, Santa Clara 54830, Villa Clara, Cuba

[b] Laboratorio de Electrónica Molecular, Departamento de Química, Modulo II, Grano de Oro, Facultad Experimental de Ciencias, La Universidad del Zulia (LUZ), Venezuela

The success achieved in the study of molecular properties by applying numerical characterization of molecular structure and estimation of quantitative structure-property relationships (QSPR) has motivated the design of similar strategies to study similarity/dissimilarity degree among biomolecules and its properties. Although the development of new schemes to represent graphically DNA were initiated over two decades ago (Hamori, 1983), the first forms of graphical representation for proteins have been reported recently (Marrero-Ponce, Medina *et al*. 2004; Randić, 2004; Randić, Zupan *et al*. 2004; Randić, Butina *et al*. 2006). Each form of representation allows a partial characterization of protein structure. Therefore, descriptors calculated from a specific representation scheme encode only one part of chemical information. For this reasons, to construct novel graphical representations of proteins and novel protein descriptors which can provide us with new information about the structure of proteins has turned a need (Randić, Mehulic *et al*. 2009). Here, a new set of 3D protein's descriptors is presented. This novel scheme for the numerical representation of polypeptide chains has been designed from a linear algebra point of view. These 3D protein's bilinear indices are calculated by means of bilinear transformations of R^{η} elements which encode information related to protein sequences and amino acid properties. Coefficients of bilinear transformations can represent the inverse of three different Minkowski distances ($p = 1, 2$ and 3), between α-Carbon from i-th and j-th amino acids in the protein backbone, raised up to the potency n. These descriptors can be used for the numerical characterization of protein structures as a whole or fragments of it. The new descriptors were used as predictor variables in the development of linear discriminant models to classify a set of 204 protein domains according to their topologies (All α, All β, α/β and α + β). The best model obtained allowed correct classification of 92.6% of cases in the training set, and the 92.7% in prediction set. These results are comparable, and sometimes superior, to those reported by other authors (Cai, Li *et al*. 2000; Cai and Zhou, 2000; Cai, Hu *et al*. 2002; Cai, Liu *et al*. 2002; Cai, Feng *et al*. 2006; Chen, Tian et al. 2006), who have developed several strategies for the structural classification of proteins, using a similar database, but with modeling techniques and information input different.

References:

Cai, Y. D., K. Y. Feng, et al. (2006). "Using LogitBoost classifier to predict protein structural classes." Journal of Theoretical Biology **238**: 172–176.

Cai, Y. D., J. Hu, et al. (2002). "Prediction of Protein Structural Classes by a Neural Network Method." Internet Electronic Journal of Molecular Design **1**(7): 332-338.

Cai, Y. D., Y. X. Li, et al. (2000). "Using neural networks for prediction of domain structural classes." Biochimica et Biophysica Acta **1476**: 1-2.

Cai, Y. D., X. J. Liu, et al. (2002). "Prediction of protein structural classes by support vector machines." Computers and Chemistry **26**: 293-296.

Cai, Y. D. and G. P. Zhou (2000). "Prediction of protein structural classes by neural network." Biochimie **82**: 783-785.

Chen, C., Y. X. Tian, et al. (2006). "Using pseudo-amino acid composition and support vector machine to predict protein structural class." Journal of Theoretical Biology **243**: 444-448.

Hamori, E. (1983). "H curves, a novel method of representation of nucleotide series especially suited for long range DNA sequences." J. Biol. Chem. **258**: 1318-1327.

Marrero-Ponce, Y., R. Medina, et al. (2004). "Protein Quadratic Indices of the "Macromolecular Pseudograph's α-Carbon Atom Adjacency Matrix". 1. Prediction of Arc Repressor Alanine-mutant's Stability." Molecules **9** 1124-1147.

Randić, M. (2004). "2-D Graphical representation of proteins based on virtual genetic code." SAR and QSAR in Environmental Research **15**(3): 147-157.

Randić, M., D. Butina, et al. (2006). "Novel 2-D graphical representation of proteins." Chem. Phys. Lett. **419**: 528-532.

Randić, M., K. Mehulic, et al. (2009). "Graphical representation of proteins as four-color maps and their numerical characterization." Journal of Molecular Graphics and Modelling **27**: 637–641.

Randić, M., J. Zupan, et al. (2004). "Unique graphical representation of protein sequences based on nucleotide triplet codons." Chem. Phys. Lett. **397**: 247-252.

NUCLEOTIDE'S BILINEAR INDICES: NOVEL BIO-MACROMOLECULAR DESCRIPTORS FOR BIOINFORMATICS STUDIES OF NUCLEIC ACIDS. I. PREDICTION OF PAROMOMYCIN'S AFFINITY CONSTANT WITH HIV-1 Ψ-RNA PACKAGING REGION

Sadiel E. Ortega Broche[a], Yovani Marrero-Ponce[a,b,c] , Yunaimy Echeverría Díaz[a] , Ysaías Alvarado[d] , Néstor Cubillán Acosta[d], Gladys Casas Cardoso[e] ,Francisco Torrens[b], Facundo Pérez-Giménez[c]

[a] Unit of Computer-Aided Molecular ''Biosilico'' Discovery and Bioinformatic Research (CAMD-BIR Unit), Faculty of Chemistry-Pharmacy, Central University of Las Villas, Santa Clara 54830, Villa Clara, Cuba

[b] Institut Universitari de Ciencia Molecular, Universitat de Valencia, Edifici d' Instituts de Paterna, Poligon la Coma s/n, E-46071 Valencia, Spain

[c] Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Departamento de Química Física, Facultad de Farmacia, Universitat de Valencia, Spain

[d] Laboratorio de Electrónica Molecular, Departamento de Química, Modulo II, Grano de Oro, Facultad Experimental de Ciencias, La Universidad del Zulia (LUZ), Venezuela

[e] Bioinformatics Group, Informatics Research Center (CEI), Faculty of Mathematics, Physics and Computer Science, Central University of Las Villas, Santa Clara 54830, Villa Clara, Cuba

A new set of nucleotide-based bio-macromolecular descriptors are presented. This novel approach to bio-macromolecular design from a linear algebra point of view is relevant to nucleic acids quantitative structure-activity relationship (QSAR) studies. These bio-macromolecular indices are based on the calculus of bilinear maps on R^n in canonical basis. Nucleic acid bilinear indices are calculated from kth power of non-stochastic and stochastic nucleotide's graph-theoretic electronic-contact matrices, M^k and $^sM^k$ respectively. That is to say, the K^{th} non-stochastic and stochastic nucleic acid bilinear indices are calculated using M^k and $^sM^k$ as matrix operators of bilinear transformations. Moreover, biochemical information is codified by using different pair combinations of nucleotide-base properties as weightings (experimental molar absorption coefficient ϵ_{260} at 260nm and pH=7.0, first (ΔE_1) and second (ΔE_2) single excitation energies in eV, and first (f_1) and second (f_2) oscillator strength values (of the first singlet excitation energies) of the nucleotide DNA–RNA bases (Pogliani 2000). As example of this approach, an interaction study of the antibiotic paromomycin with the packaging region of the HIV-1 Ψ-RNA has been performed and several linear models have been obtained in order to predict the interaction strength. The best linear model obtained by using non-stochastic bilinear indices explains about 91% of the variance of the experimental Log K ($R = 0.95$ and $s = 0.08 \times 10^{-4}$ M^{-1}) as long as the best stochastic bilinear indices-based equation account for 93% of the Log K variance ($R = 0.97$ and $s = 0.07 \times 10^{-4}$ M^{-1}). The leave-one-out (LOO) press statistics, evidenced high predictive ability of both models ($q^2 = 0.86$ and $s_{cv} = 0.09 \times 10^{-4}$ M^{-1} for non-stochastic and $q^2 = 0.91$ and $s_{cv} = 0.08 \times 10^{-4}$ M^{-1} for stochastic bilinear indices).The nucleic acid bilinear indices-based models compared favorably with other nucleic acid indices-based approaches reported nowadays (González-Díaz, Ramos de Armas et al. 2003; González-Díaz, Ramos de Armas et al. 2003; Marrero-Ponce, Nodarse et al. 2004; Marrero-Ponce, Castillo Garit et al. 2005). These models also permit the interpretation of the driving forces of the interaction process. In this sense, developed equations involve short-reaching ($k \leq 3$), middle-

reaching ($4 \leq k \leq 9$), and far-reaching ($k = 10$ or greater) nucleotide's bilinear indices. This situation points to electronic and topologic nucleotide's backbone interactions control of the stability profile of paromomycin-RNA complexes. Consequently, the present approach represents a novel and rather promising way to theoretical-biology studies.

References:

González-Díaz, H., R. Ramos de Armas, et al. (2003). "Markovian negentropies in bioinformatics. 1. A picture of footprints after the interaction of the HIV-1 y-RNA packaging region with drugs " Bioinformatics **19**: 1-9.

González-Díaz, H., R. Ramos de Armas, et al. (2003). "Vibrational Markovian Modelling of Footprints after the Interaction of Antibiotics with the Packaging Region of HIV Type 1." Bulletin of Mathematical Biology **00**: 1-13.

Marrero-Ponce, Y., J. A. Castillo Garit, et al. (2005). "Linear indices of the 'macromolecular graph's nucleotides adjacency matrix' as a promising approach for bioinformatics studies. Part 1: prediction of paromomycin's affinity constant with HIV-1 psi-RNA packaging region." Bioorg Med Chem **13**(10): 3397-3404.

Marrero-Ponce, Y., D. Nodarse, et al. (2004). "Nucleic Acid Quadratic Indices of the "Macromolecular Graph's Nucleotides Adjacency Matrix". Modeling of Footprints after the Interaction of Paromomycin with the HIV-1 Ψ-RNA Packaging Region." Int. J. Mol. Sci. **5**: 276-293.

Pogliani, L. (2000). "From Molecular Connectivity Indices to Semiempirical Connectivity Terms: Recent Trends in Graph Theoretical Descriptors. ." Chem. Rev. **100**: 3827-3858.

MULTI-CORE MOLECULAR FIELD TOPOLOGY ANALYSIS

Eugene V. Radchenko, Vladimir A. Palyulin, Nikolay S. Zefirov
Department of Chemistry, Moscow State University, Moscow, 119991, Russia

The Molecular Field Topology Analysis (MFTA) technique [1] is intended to model the relationships between the biological activity of compounds and the parameters of their atoms and bonds. Meaningful comparison of these local molecular properties is enabled by a so-called molecular supergraph such that any training set structure can be superimposed onto it. To build a uniform descriptor set, each supergraph position is assigned a property value for a corresponding atom or bond in a structure (e.g., effective atomic charge, van der Waals radius, H-bond donor and H-bond acceptor ability, local lipophilicity and/or other parameters) while for unoccupied vertices the neutral descriptor values are used. The statistical analysis of this descriptor set (usually by means of the Partial Least Squares Regression) yields a predictive model that allows one not only to predict the bioactivity for new compounds but also to identify the structural features critical for activity (i.e., related to the local descriptors making the largest contribution to the activity). Thus, the MFTA modelling provides a useful tool for the targeted virtual screening [2] of novel promising structures.

This approach has been successfully applied to many chemical classes and kinds of bioactivity, ranging from antiviral agents to neuroreceptor ligands to irreversible inhibitors of serine esterases. However, the experience indicates some areas where improvement is desirable. In its original form, MFTA is best suited for series of congeneric structures involving a single common core and various peripheral fragments. If such 'substituent' fragments are too complex and diverse, the intuitively desirable superposition of structures may be problematic to achieve and/or significant manual intervention may be required. In addition, the ability to handle structurally different but roughly bioisosteric core fragments would be useful both to expand the applicable training set and to support predictions for related compounds as well as scaffold hopping.

In view of these goals, the MFTA technique was extended to detect reasonable core fragments (e.g., polycyclic) that may be separated by more flexible linkers. The automatic selection of the cores may be adjusted by a researcher, affecting further superimposition steps. Similarly, during prediction procedure new cores may be introduced and mapped to some of the already analyzed ones.

The extended multi-core MFTA approach was implemented in the C++ software. Its application to a number of test cases shows promising results that are presented in the paper.

References:

[1] Palyulin V.A., Radchenko E.V., Zefirov N.S. *J. Chem. Inf. Comput. Sci.*, **2000**, 40, 659-667.
[2] Radchenko E.V., Palyulin V.A., Zefirov N.S., in *Chemoinformatics Approaches to Virtual Screening*, ed. by A.Varnek, A.Tropsha, RSC, **2008**, 150-181.

POSTER PRESENTATION I-55

VALIDATION OF TWO NOVEL SCHEMES OF PARTIAL ATOMIC CHARGES CALCULATION IN 3D QSAR

Daria A. Tsareva, Dmitry I. Osolodkin, Dmitry A. Shulga, Alexander A. Oliferenko, Vladimir A. Palyulin, Nikolay S. Zefirov

Department of Chemistry, Moscow State University, Moscow, 119991, Russia

Molecular electrostatic potential (MEP) plays an important role in describing various biomolecular systems, especially in 3D QSAR, docking and virtual screening procedures. Therefore different partial atomic charges calculation schemes can affect the results, predictability and accuracy of QSAR models.

GSK-3β is a serine/threonine protein kinase which takes part in numerous signal pathways in human organism and is involved in Alzheimer and Parkinson diseases as well as type 2 diabetes.

The main aim of the work was to validate MEP-reproducing Kirchhoff Charge Model (KCM) [1] and Dynamic ElectroNegativity Relaxation (DENR) scheme [2] in terms of their applicability to 3D QSAR and compare them with eight other charge calculation schemes on the example of 5 classes of GSK-3β inhibitors and the common set of steroids.

Basing on the detailed statistical analysis of 3D QSAR models and careful comparison of the contour maps of steric, electrostatic, hydrophobic and H-bond donor and acceptor interaction fields it was shown that KCM and DENR charges are the most reliable among the others used. The estimation of charges in MMFF94 scheme is rough enough and does not allow distinguishing subtle substituent effects, this makes the contour maps of the interaction fields more complicated. The greater number of outliers was observed during the analysis of the models with MK-ESP and RESP charges than with other schemes. The application of quantum chemistry calculation (Löwdin and Mulliken charges) did not lead to a substantial improvement in CoMFA and CoMSIA modeling as compared with fast KCM and DENR schemes.

References:

[1] Oliferenko A.A., Palyulin V.A., Pisarev S.A., Neiman A.V., Zefirov N.S. *J. Phys. Org. Chem*. **2001**, 14, 355–369

[2] Shulga D.A., Oliferenko A.A., Pisarev S.A., Palyulin V.A., Zefirov N.S. *SAR QSAR Environ. Res*. **2008**, 19(1-2), 153-165

POSTER PRESENTATION I-56

COMPUTATIONAL APPROACHES FOR FRAGMENT-BASED AND DE *NOVO* DESIGN

Kathryn Loving, Ian Alberts and Woody Sherman
Schrödinger Inc. 120 West 45th Street, 17th Floor, New York, NY 10036-4041, USA

Fragment-based and de *novo* design strategies have been used in drug discovery for years. The methodologies for these strategies are typically discussed separately, yet the applications of these techniques overlap substantially. We present a review of various fragment-based discovery and de novo design protocols with an emphasis on successful applications in real-world drug discovery projects. Furthermore, we illustrate the strengths and weaknesses of the various approaches and discuss how one method can be used to complement another. We also discuss how the incorporation of experimental data as constraints in computational models can produce novel compounds that occupy unique areas in intellectual property (IP) space yet are biased toward the desired chemical property space. Finally, we present recent research results suggesting that computational tools applied to fragment-based discovery and de *novo* design can have a greater impact on the discovery process when coupled with the right experiments.

LIGAND EFFICIENCY INDICES (LEIs): MUCH MORE THAN A SIMPLE EFFICIENCY YARDSTICK

Cele Abad-Zapatero[a,b], Daniel Blasi[a], J. Wass[c]

[a]Drug Discovery Platform (PDD), Parc Científic de Barcelona, Barcelona, Spain

[b]Center for Pharmaceutical Biotechnology, University of Illinois at Chicago, 900 So. Ashland St, MBRB Building, Room 3020; (M/C870), Chicago, IL 60607-7173, USA

[c]Independent Statistics Consultant, Lake Forest, IL., 60045, USA

The concept of Ligand Efficiency (LE), relating the binding energy (DG) of a ligand towards a target to a size-related parameter (number of non-hydrogen atoms) of the ligand was introduced by Hopkins et al[1]. Variations of this definition are becoming accepted as a tool to improve lead evaluation[1] and fragment-based strategies[2] among other applications. In addition, various related definitions (i.e., 'group efficiency'; Ligand Lipophilicity Index: LLE, among others) are used along the path of preclinical drug discovery projects[3,4,5] by various groups. The use of 'Ligand Efficiency' relating the affinity (Ki, IC50 or related) to its molecular size (MW, number of non-hydrogen atoms: NHEA and others) is now well accepted.

An expanded definition was introduced in 2005 that consisted of two Ligand Efficiency Indices (LEIs): BEI and SEI, defined along similar continuous numerical scales[6]. The first related the potency to the MW of the ligand (in KiloDaltons), and the second combined the potency with the Polar Surface Area (PSA) scaled to 100 Å^2. The combined used of these two LEIs in comparing retrospectively chemical series in an optimization plane SEI-BEI has been published[7]. Recently, a study comparing the size-related ligand efficiency indexes (LE) of drugs to its original leads (N=60 for various targets) has been published[8]. The study revealed that large increases in binding efficiency from leads to drugs (sharing the same scaffold) can be achieved. Using the same data set, we have applied the combined use of SEI-BEI efficiency indices in a Cartesian plane, as well as other related indices including the lipophilicity index (LLE), to provide a more comprehensive view of the Lead-to-Drug optimization process. Details of our findings will be presented.

We acknowledge the additional data kindly provided by Dr.E. Perola.

References:

[1]L. Hopkins Andrew, R. Groom Colin, A. Alex, *Drug Discov Today,* (2004) 9, 430.
[2]Hajduk, P. J. & Greer, J. (2007). *Nat Rev Drug Discov* (2007) 6, 21.
[3]Reynolds, C.H., Bembenek, S.D., and Tounge, B.A. (2007). Bioorg Med Chem Lett 17, 4258
[4]Verdonk, M.L., and Rees, D.C. (2008). Group efficiency: a guideline for hits-to-leads
[5]Hajduk, P. J. (2006). *J Med Chem* 49, 6972.
[6]Abad-Zapatero, C. and J. T. Metz. Drug Discovery Today, 2005. 10, 464.
[7]C. Abad-Zapatero, *Expert Opin. Drug Discovery*, 2007, 2, 469.
[8]E. Perola, J. Med. Chem. 2010, 53, 2986.

ATOM-CENTRED FRAGMENTS IN QSAR MODELLING

Kühne R, Ebert R-U and Schüürmann G.*

Department of Ecological Chemistry, Helmholtz Centre for Environmental Research - UFZ, Permoserstr. 15, 04318 Leipzig, Germany

*Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Strasse 29, 09596 Freiberg, Germany

Atom-centred fragments (ACF) models work by decomposing molecules into individual atoms, and characterizing the atoms by selected atomic properties, bonds to other atoms, next neighbour atoms, and optionally more than next neighbours. They can be used as a new tool to improve model predictions, to select appropriate models from a set of models, to apply read-across approaches, and to determine the chemical domain of models. This paper focuses on ACFs for read-across models and chemical domain analyses.

In read-across approaches, ACFs can be used to run k next neighbour models. ACFs are employed to select the most similar compounds. Working examples are shown for the consistent prediction of the acute aquatic toxicity for both baseline and excess toxic compounds.

In the chemical domain analyses, ACFs are employed to examine whether a test compound is covered in the training set of a model with respect to he chemical composition. A study is presented to optimize the use of ACF for the model domain characterization of quantitative and classification models for physical-chemical properties and ecotoxicological endpoints.

This study has been financially supported by the EU projects OSIRIS (IP, contract no. 037017), CAESAR (contract no. 022674), and 2-FUN (contract no. 036976), and by the German Federal Environmental Agency (UBA).

PREDICTION OF THE PARTITION COEFFICIENT BETWEEN AIR AND BODY COMPARTMENTS FROM THE CHEMICAL STRUCTURE

Stöckl S*, Kühne R, Ebert R-U and Schüürmann G.*
Department of Ecological Chemistry, Helmholtz Centre for Environmental Research - UFZ, Permoserstr. 15, 04318 Leipzig, Germany
*Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Strasse 29, 09596 Freiberg, Germany

For PBPK modelling, partition coefficients between tissues and environmental compartments are required. A simple approach starts with the system blood/air, fat/air, and fat/blood. Employing thermodynamic relationships, one of the three coefficients can be calculated from the other two values. With respect to available human and mammal data, modelling efforts focus on the blood/air and fat/air partition coefficient. Respective data sets from literature have been collected and evaluated. The chemical domain of the sets is presented in terms of chemical structure, complexity, and polarity. Data gaps have been identified.
The blood/air and fat/air partition coefficient data set has been applied to a validation of literature models, with particular remark on the performance for specific compound classes. A new model for the blood/air partition coefficient and a preliminary model for the fat/air partition coefficient are presented.

Acknowledgements:
The study was supported by the EU projects 2FUN (contract No. 036976) and OSIRIS (IP, contract No. 037017).

pK_a PREDICTION FROM LOCAL MOLECULAR PARAMETERS

Haiying Yu,[a,b] Ralph Kühne,[a] Ralf-Uwe Ebert,[a] Gerrit Schüürmann[a,b,*]

[a] UFZ Department Ecological Chemistry, Helmholtz Centre for Environmental Research, Per1moserstr. 15, 04318 Leipzig, Germany

[b] Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Str. 29, 09596 Freiberg, Germany

The acid-base dissociation constant pK_a defines the ionization degree of organic compounds. It is an important parameter that affects their toxicity and environmental behavior. Among the present approaches to predict pKa values of organic chemicals, increment methods show a good predictability but are limited by missing values of special groups. The purpose of this study is to develop models for predicting the pKa of 574 organic bases directly from their molecular structures.

A quantum chemical method was introduced by employing electronic structure parameters. The dataset includes 137 anilines, 239 amines, and 198 heterocyclic bases, covering 17 pK_a units (-5.00 ~ 11.72). All molecular geometries were optimized using the semi-empirical AM1 Hamiltonian. Local molecular electronic characteristics including charge-limited energy, effective donor/acceptor energy and energy-limited charge were calculated from the neutral compounds in their optimized geometries. Simple models with experimental pK_a values for different subsets were calibrated by multilinear regression.

The obtained models with up to three local molecular structure descriptors of disso1cia1tion groups yield good predictive squared correlation coefficients q^2, and a superior performance as compared with existing quantum chemical models. Cross validation, scrambling, and external validation demonstrate good prediction ability and robustness of the developed models. The discussion includes consideration of solution-phase parameters such as aqueous solvation energy and its impact on the final regression models.

Acknowledgements:

The study was financially supported by China Scholarship Council, EU project OSIRIS (contract No. 037017) and Helmholtz Impulse and Networking Fund through Helmholtz Interdisciplinary Graduate School for Environmental Research HIGRADE.

LIPOPHILICITY STUDIES OF SET OF QUATERNARY AMMONIUM BROMIDES WITH NEUROMUSCULAR BLOCKAGE ACTIVITY

Jamille Moreira Moraes[1], Leonardo J. A. Siqueira[1], Antonio J. Lapa[2], Caden Souccar[2], Alberto Malvezzi[1] and Antonia T. do-Amaral[1]

[1]Department of Chemistry, Chemistry Institute, University of São Paulo, Av. Prof. Lineu Prestes, 748, 05508-900, São Paulo, SP; Brazil

[2]Department of Pharmacology, Federal University of São Paulo, Rua Três de Maio, 100, 04044-020, São Paulo, SP, Brazil

In order to differentiate various sub-types of neuronal nicotinic receptors a set of eleven [2-(4-X-benzamido)ethyl]benzyldimethylammonium bromides have been synthesized and their interaction with α7 and α4β2 nicotinic receptors analyzed.[1,2] The lipophilicity of each compound, measured in n-octanol/Trizma buffer by means of the shake-flask method,[3,4] showed log P_{app} values in the range of –1.63 to 2.76.
In this work,[5] the ion-pair concept have been applied in order to increase the accuracy of the lipophilicity measurements regarding, specially, the highly hydrophilic compounds.[6,7] For this purpose, logP measurements have been done in the presence of different concentrations of the hydrophobic sodium n-octylsulfate, as counter ion, to overcome the problems arising from their high hydrophilicity. The obtained logPapp values - extrapolated to zero concentration of the counter ion - have been analyzed and compared with their corresponding values observed without counter-ions.

Financial Support: FAPESP, CNPq

References

[1] Siqueira,L.J.A. Quantitative structure-activity relationship for a set of [2-(4-X-benzamido) ethyl] benzyldimethylammonium bromide with neuromuscular blockage activity. MSc. Thesis. University of São Paulo, São Paulo, SP, 2001.
[2] Arias, H. R.; Bhumireddy, P. & Bouzat, C. *Int. J. Biochem. Cell Biol.* 38(8):1254-76, 2006.
[3] Amaral, A.T.-do; Oliveira, A.C.; Neidlein, R.; Gallacci, M.; Caprara, L. & Miyazaki, Y. *Eur. J. Med. Chem.*, 32, 433-443, 1997.
[4] Pires, J.R.; Saito, C.; Gomes, S.L.; Giesbrecht, A.M. and Amaral, A.T.-do. *J. Med. Chem*, 44 (22), 3673-3681, 2001.
[5]Moraes, J. M., Siqueira, L.J. A; Lapa, A. J., Souccar, C., Malvezzi, A. and Amaral, A. T. –do, Lipophilicity Studies of Set of Quaternary Ammonium Bromides with Neuromuscular Blockage Activity, PhysChem and ADMET Profiling in Drug Research The 4th LogP Symposium, Zürich, Swizerland, 2009.
[6] Takács-Novák, K. & Szász, G. Pharm. Res. 16(10): 1633-8, 1999.
[7] Koufopoulou, S. A.; Pistos, C.; Giaginis, C; Tsantili-Kakoulidou, A. *Int. J. Pharm*. 316(1-2): 52-7, 2006.

OPTIMIZATION OF QUERCETIN NANOPARTICLE EMULSION PREPARATION USING EXPERIMENTAL DESIGN AND MULTIPLE LINEAR REGRESSION

Pouneh Ebrahimi[1], Fereshteh Pourmorad [2], Alireza Banai [3], Soheila Honary[2], Bahar Ebrahim-Magham [3]

[1]Gonbad Institute of Higher Education, PO BOX 163, Gonbad, Iran,

[2]Pharmaceuticl Research Center, Faculty of Pharmacy, Mazandaran University of Medical Sciences, Sari, Iran

[3]Faculty of Science, University of Payam Noor, Ardabil, Iran

Introduction:

Quercetin is one of the most important natural products with considerable therapeutic effects and because of having compatibility with human's body and less unwanted effects its consumption is increasing day by day. Unfortunately quercetin solubility is very low in aqueous solution and this causes problem in preparation of pharmaceutical products. Nano- emulsions are a kind of colloidal oil/water dispersions with very small size around 10-200 nm. Compounds like quercetin with limited solubility can be dissolved in oil/ water interphase[1-2]. In fact in such systems due to greater surface area and less surface tension, absorption of drug molecule through the biological membranes such as skin and blood brain barrier increases and then enhances bioavailability of the drug and then administrating of less amount of drug to the patient.

In this study, in order to solve problems related to quercetin solubility in water, we prepared a kind of nano-emulsions with the purpose of increasing solubility and skin absorption for their potentials in the treatment of skin disorders such as psoriasis and atopic dermatitis.

Methods:

Nano-emulsions prepared with spontaneous emulsification method [2]. First of all, optimization of experiments' condition using Box –Behnken method, one of the most suitable factorial experimental designs for modeling studies, was carried out. Experimental parameters and related levels including concentration of lecithin in ethanol (2-0.7% w/w), the concentration of Tween in ethanol (8-2% w/w) and sonication time (60-10 minutes) were studied.

In each step of experimental design, the organic phase (including Lecithin, Tween and quercetin (0.1% by w/w) in ethanol) was added to the water phase. Then mixture stands on ultrasonic bath for 15 minutes and then after removing organic solvent by rotary evaporator, the particle size was measurement by Zetasizer instrument.

Conclusion:

After determining the average size of particles in each step of the experimental design, modeling of this parameter in three experimental conditions using multiple linear regressions with SPSS software was conducted. Then, using network searching method, optimal condition for preparaing nano- particles predicted and compared with experimental results.

References:

[1] S. Tamilvanan, S. Benita, Eur. J. Pharm. Biopharm. 58 (2004) 357–368.

[2]K. Bouchemal, S. Branc.on, E. Perrier, H. Fessii, Int. J. Pharm. 280 (2004) 241–251.

POSTER PRESENTATION I-63

SIMULTANEOUS MODELING OF THE KOVATS RETENTION INDICES ON PHENYL OV STATIONARY PHASES WITH DIFFERENT POLARITY USING MLR AND ANN

Pouneh Ebrahimi[1], Eisa Fallah[2]
[1]Gonbad Institute of Higher, P. O. Box 163, Gonbad, Iran;
[2]Islamic Azad University, Babol Branch, Babol, Iran

A QSRR study was performed to develop a predictive model correlating the observed kovats retention indices of three congeneric aromatic series with their molecular structures in gas chromatography. The congeneric aromatic series included the substituted benzene, benzaldehyde and acethophenone compounds, which had been studied, previously, on six OV stationary phases with different phenyl percentages1. At first, a model was generated for six columns separately, using only calculated descriptors and MLR technique. Then a combined model, added a polarity term of stationary phase (M), was also developed for all these columns, and the result was apparently satisfactory (R^2= 0.991, F=1009.828, SE=23.98). Since the intercept had a high value in this model, the neural network back propagation algorithm was applied for comparison, and it was found that the neural network could exceed the level of the multiple regression method. The stability and validity of both models were tested by cross-validation technique and by predicted response values for the prediction set and test set. The results of the study indicated that a seven parameter equation can be utilized for prediction of retention indices of compounds on different OV stationary phases simultaneously using ANN technique, for which there are no empirical RI values.

References:

1. A. Hassani, B. Y. Meklati, Chromatographia 1992, 33, 267 –271

PREDICTION OF AQUEOUS DRUG SOLUBILITY USING COUNTER-PROPAGATING ARTIFICIAL NEURAL NETWORKS WITH AUTOMATIC ADJUSTMENT OF RELATIVE IMPORTANCE OF THE DESCRIPTORS

Slavica Erić[1], Marko Kalinić[1], Aleksandar Popović[1], Mire Zloh[2], Igor Kuzmanovski[3]

[1]Institute of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Belgrade, Vojvode Stepe 450, 11 000 Belgrade, Serbia

[2] School of Pharmacy, University of London, 29-39 Brunswick Square, WC1N 1AX London, UK

[3] Institut za hemija, PMF, Univerzitet "Sv. Kiril i Metodij", P.O. Box 162, 1001 Skopje, Republic of Macedonia

Aqueous solubility is an important property which can significantly impact the bioavailability of an orally administered drug. The use of computational models for the prediction of solubility can help the identification and elimination of compounds with inadequate solubility in the early stages of drug discovery, prior to synthesis, and can thus reduce the chance of late-stage drug development failures.

Many linear and nonlinear models for prediction of solubility have been presented in the literature. Although the success of a chosen modeling method considerably depends on the data set and the descriptors used, the data in literature suggest that nonlinear methods, such as artificial neural networks, can provide better models for the prediction of solubility than linear approaches.[1] The application of nonlinear models is however limited by their low interpretability, since such models usually represent black box models. Counter-propagation artificial neural networks (CPANN) have some advantage since, besides the good performance of the models, provide possibility for data exploratory analysis through examination of the clustering of the data in the respective CPANN layers. However, if the number of descriptors is large, the contribution of the descriptors describing different structural features affecting drugs solubility could still remain unclear. In this work, we present the development of a reliable predictive model using CPANN and novel algorithm for the automatic adjustment of relative importance (AARI)[2] of the descriptors, which provides better insight into the individual influence of descriptors on the prediction of the models, as well as their importance for the mapping of the training objects into the 2D CPANN grid.

A data set consisting of 375 drug-like molecules was used for development of the model. Our earlier study showed that the structural diversity and quality of experimental data of compounds in this data set makes it a favorable starting point for model development[3]. A large number of molecular descriptors were calculated using CODESSA software and their preselection was performed using the heuristic method. The initial set was then divided into training (n = 281) and test (n = 94) sets using the Kennard-Stone algorithm. Preselected descriptors for compounds in the training set were used as input variables. A genetic algorithm (GA) was used for determination of the optimal size and training parameters for the CPANN, whilst AARI algorithm was applied for adjustment of the relative importance of descriptors used for the construction of the model. During the optimization performed by GA, the models were validated using cross-validation. The final models were validated using test set.

AARI performed by GA helped us to find simple and interpretable model suitable for data exploratory analysis. The developed predictive model showed good performance for both the training set (r = 0.9407) and the test set (r = 0.8541).

Simplicity, interpretability and good generalization performances of developed model allow its application

for screening of diverse drug structures. Furthermore, predictive model can give guidelines for modification of structure in aim of the improvement of solubility of drugs.

References:

1) K.V. Balakin, N.P. Savchuk, I.V. Tetko, *Curr. Med. Chem.* **2006**, 13, 223-241.
2) I. Kuzmanovski, M. Novič, M. Trpkovska, *Anal. Chim. Acta* **2009**, 642, 142-147.
3) S. Erić, M. Kalinić, A. Popović, H. Makić, E. Civić, M. Bektašević, *J. Serb. Chem. Soc.* **2010**, 75, 1-13

THE POTENTIAL OF IAM CHROMATOGRAPHIC INDICES TO SIMULATE IN VITRO AND IN VIVO PERMEABILITY DATA

Demetris Vrakas[1], Costas Giaginis[1], Fotios Tsopelas[1], Niki Malaki[1], Vassilis Demopoulos[2], Anna Tsantili-Kakoulidou[1]

[1]Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, Panepistimiopolis, Zografou, Athens, 157 71, Greece

[2]Department of Pharmaceutical Chemistry, School of Pharmacy, University of Thessaloniki, Thessaloniki, Greece

ADME properties are considered in early drug discovery process so that attrition due to poor bioavailability to be reduced. In this aspect, there is an emergent demand for drug permeability assessment and up to date a variety of in vitro models of differing complexity based either on cell lines or artificial membranes are available. The development of Immobilised Artificial Membrane (IAM) surfaces to be used as stationary phases unfolded new perspectives in the application of HPLC to permeability assessment, since it combines cell membrane simulation with rapid and friendly measurements. However the increased IAM retention, observed for protonated basic compounds due to their electrostatic interactions with the phosphate anions of the IAM surface [1], raises the question whether IAM indices are suitable to predict passive diffusion of drugs or they should rather be used to simulate drug-membrane interactions. In the present study IAM retention factors, $\log k_{wIAM}$, determined at pH 7.4 and pH 5 [1] have been compared with apparent permeability data, logPerm, measured on MDCK cell lines at analogous pH [2]. Improved correlation between logPerm and $\log k_{wIAM}$ was obtained, if the fraction protonated of the basic drugs was included in the regression equation. IAM retention indices were further used to correlate % of Human Oral Absorption (%HOA). Since the pH affects absorption through the gastrointestinal track the highest $\log k_{wIAM}$ values considered as best $\log k_{wIAMbest}$ were introduced in the model. A non linear regression model was obtained if Abraham's hydrogen bond basicity parameter B was included as an additional parameter with satisfactory statistics (r^2=0.933) although inferior than the corresponding model generated using logPerm data (r^2=0.969). Considering also fraction protonated F^+ in the statistical analysis, an improved three parameter equation was obtained (r^2= 0.951), while a negative coefficient was assigned to the latter term.

References:

[1] D. Vrakas, C. Giaginis, A. Tsantili-Kakoulidou Electrostatic interactions and ionization effect in IAM retention. A comparative study with octanol- water partitioning, *J.Chromatogr. A* 1187: 67-78 (2008)

[2] V.E. Thiel-Demby, J.E. Humphreys, L.A. St John Williams, H.M. Ellens, N. Shah, A.D. Ayrton, J.W. Polli. Biopharmaceutics classification system: validation and learnings of an in vitro permeability assay. *Mol Pharm.* 6, 11-8 (2009)

THE USE OF VOLSURF APPROACH TO EXPLORE IAM AND HSA CHROMATOGRAPHIC INDICES

Catherine Koukoulitsa[1], Marios Chrysanthakopoulos[2], Demetris Vrakas[2], Costas Giaginis[2], Thomas Mavromoustakos[1], Anna Tsantili-Kakoulidou[2]

[1]Department of Chemistry, University of Athens, Panepistimiopolis Zographou 15771 Athens, Greece.
[2]Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, Panepistimiopolis, Zografou, 157 71, Athens, Greece

Biomimetic chromatography offers a promising alternative in the evaluation of both membrane permeability and the affinity for binding to proteins.
In the present study we used VolSurf descriptors to establish quantitative models of retention data obtained on an Immobilized Artificial Membrane (IAM) and on a Human Serum Albumin (HSA) stationary phase in the aim to highlight the most critical factors in retention mechanism. The data set consisted of neutral, basic, acidic compounds. IAM and HSA retention factors expressed as $\log k_w^{IAM}$ and $\log kw^{HSA}$ were obtained from references [1] and [2] respectively. A three component PLS model with $R^2 = 0.838$ and $Q^2 = 0.767$ was obtained for IAM retention using 70 original VolSurf variables with rugosity, hydrophobic descriptors, H-bond donor and flexibility having a positive effect, while descriptors of hydrophilic regions exert a negative contribution. The negative coefficient of % unionized indicates the counterbalance between ionization effect and electrostatic interactions with the charged groups on the IAM surface. For HSA retention a five component PLS model was obtained with 97 original descriptors with $R^2 = 0.831$ and $Q^2 = 0.540$. The model was governed by analogous descriptors as the IAM retention model with more pronounced negative effect of % unionized reflecting the stronger contribution of electrostatic interaction on HSA stationary phase.

Reference

1. D. Vrakas, C. Gianinis, A. Tsantili-Kakoulidou, J. Cromatogr. A, 1187 (2008) 67-78.

HINDRANCE OF TWO DIMENSIONAL PROTEIN MOVEMENTS ON TOP OF (MAGNETO)LIPOSOMES BY INSERTION OF PEGYLATED PHOSPHOLIPIDS IN THE MEMBRANE BILAYER

Nele Nuytten, Stefaan.J.H. Soenen, Lien Defour, Tine Ysenbaert & Marcel De Cuyper*
Interdisciplinary Research Centre (IRC), Laboratory of BioNanoColloids, Katholieke Universteit Leuven – Campus Kortrijk, University Campus, B-8500 Kortrijk, Belgium

Over the last few decades, tailor-made nanoparticles have been introduced in a lot of biomedical applications, e.g. as drug carriers or in medical imaging [1]. Their small sizes permit them to be injected intravenously but, triggered by binding of opsonin proteins at the particle's surface, they are rapidly cleared from the bloodstream by the immune system. In the case of phospholipid-based liposomes, however, it has been observed that particle uptake can be drastically retarded by using a small percentage of phospholipid types the polar headgroup of which is modified with poly(ethylene glycol) polymer chains. By virtue of their capacity to largely avoid uptake by the mononuclear phagocytotic system the resulting biocolloids are commonly designated as Stealth liposomes [2].
In the present work we focus on the binding of the water soluble, cationic cytochrome c (molecular weight 13 kDa) on both classical and Stealth liposomes and magnetoliposomes. The latter nanocolloids consist of a magnetizable iron oxide core enwrapped in a phospholipid bilayer coat [3]. Cytochrome *c* binding is detected spectrophotometrically at 550 nm after its oxidation by cytochrome c oxidase which is embedded in the phospholipid bilayer. Dimyristoylphosphatidylglycerol (DMPG), whether or not in the presence of distearoylphosphatidylethanolamine~PEG_{5000} (DSPE~ PEG_{5000}), is chosen as the matrix lipid molecule. With the *PEGylated lipid-deprived constructs* it is found that the catalytic activity per enzyme molecule is halved upon increasing the enzyme density in the lipid bilayer from 1 to 15, but at higher densities (up to 60 enzyme molecules per (magneto)liposome) the first-order rate constants remain constant. The drastic drop in enzyme performance observed in the lower enzyme/particle range points to the occurrence of a significant lateral diffusion of the adsorbed protein prior to its oxidation by the biocatalyst. In contrast, with *Stealth liposomes and Stealth magnetoliposomes* (both containing 16.66 mol% of DSPE~ PEG_{5000}) the calculated first-order rate constants are in the lower activity zone independent of the enzyme content suggesting that in this case a lateral displacement of the substrate on top of the particles is hindered by the PEG chains. This conclusion is further sustained by quantitative calculations based on geometrical data of the molecules involved.

Acknowledgements:
S.J.H.S. is a recipient of a research grant from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). M.D.C. is a recipient of a SBO (StrategischBasisOnderzoek) grant sponsoring project no. 80017, entitled *Integrated magnetic nanoparticle-Enabled Imaging of Therapeutic Cells – 'iMAGiNe'*.

References:
[1] Soenen SJH, Hodenius M & De Cuyper M (2009) Magnetoliposomes : Versatile innovative nanocolloids for use in biotechnology and biomedicine. *Nanomedicine* **4**, 177-191.
[2] Allen TM & Chonn, A (1987) *FEBS Lett*. **223**, 42-46.
[3] De Cuyper M, & Soenen SJH (2010) Cationic magnetoliposomes. In *Methods in Molecular Biology* (Ed. Volkmar Weissig), Humana press, Vol. 605, 97-111

Caco-2 TRAFFIC LIGHT: *IN SILICO* RISK ASSESSMENT OF DRUG CANDIDATES WITH REGARD TO Caco-2 FLUX PROPERTIES

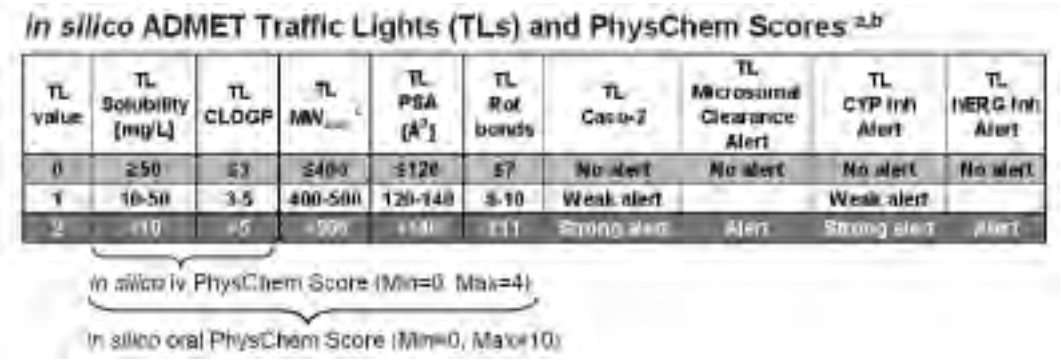
Mario Lobell, Mark-Jean Gnoth
Bayer Schering Pharma, Global Drug Discovery, Wuppertal, Germany

The Caco-2 traffic light is an *in silico* classification method for the Caco-2 flux properties of drug candidates. Each traffic light colour corresponds to a risk assessment (Green: Low risk, Yellow: Intermediate risk, Red: High risk of a liability) with regard to low Caco-2 flux. Colours are set via decision tree rules, which make use of calculated properties describing the compounds size (MW_{corr}), charge state (acidic/basic group) and hydrogen-bonding capabilities (PSA, H-bond acidity / basicity). Training has been performed with Caco-2 inhouse data for 897 drug-like compounds which are predicted with an accuracy of 76%. The prediction accuracy is 77% on a test set of 651 inhouse compounds. A set of 101 compounds with Caco-2 data from the literature is predicted with an accuracy of 79%.

IN SILICO ADMET TRAFFIC LIGHTS AND PHYSCHEM SCORES

Mario Lobell, Andreas H. Göller, Alexander Hillisch
Bayer Schering Pharma, Global Drug Discovery, Wuppertal, Germany

The *in silico* ADMET Traffic Lights (TLs) are computational tools for the prediction of ADMET related liabilities from chemical structure. The system's primary purpose is to aid medicinal chemistry experts in the analysis and prioritization of HTS hit sets with regard to potential ADMET related liabilities. Other important applications include the profiling of real and virtual compound libraries and the filtering and prioritization of commercially available compounds intended as additions to the HTS compound repository.



a) T. Wunberg, M. Hendrix, A. Hillisch, M. Lobell, H. Meier, C. Schmeck, H. Wild, B. Hinzen; Improving the hit-to-lead process: data-driven assessment of drug-like and lead-like screening hits; *Drug Discovery Today*, **2006**, 11(3/4), 175-180

b) M. Lobell, M. Hendrix, A. Hillisch, B. Hinzen, J. Keldenich, H. Meier, C. Schmeck, R. Schohe-Loop, T. Wunberg; In silico ADMET Traffic Lights as a Tool for the Prioritization of HTS Hits; *ChemMedChem*, **2006**, 1, 1229-1236

c) Corrections are applied for the occurrence of halogen atoms in order to make the corrected molecular weight (MW_{corr}) proportional to calculated molecular volume: MW_{corr} = MW - 13.8 * (No. of F-atoms) - 16.2 * (No. of Cl-atoms) - 53.7 * (No. of Br-atoms) - 89.5 * (No. of I-atoms)

IN SILICO PREDICTION OF Caco-2 CELL PERMEABILITY BY A CLASSIFICATION QSAR APPROACH

Miguel Ángel Cabrera Pérez [1], Hai Pham The [1], Isabel González Álvarez[2], Marival Bermejo Sanz[2], Inmaculada Centelles Sangüesa[3] and Teresa Garrigues Pelufo[3]

[1]Molecular Simulation & Drug Design Group. Centre of Chemical Bioactive. Central University of Las Villas. Santa Clara 54830, Villa Clara, Cuba.

[2]Department of Engineering, Area of Pharmacy and Pharmaceutical Technology, Miguel Hernández University, 03550 Sant Joan d'Alacant, Alicante, Spain.

[3]Department of Pharmacy and Pharmaceutical Technology, University of Valencia, Burjassot 46100, Valencia, Spain

One of the most important challenges of the oral drug administration is their movement across the intestinal epithelial barrier that determines the rate and extent of human intestinal absorption and ultimately affects its bioavailability. The cell-based *in vitro* models, such as Caco-2 monolayers, have been extensively used for prediction of intestinal permeability and drug absorption. A rising need for introducing absorption filters prior to design large compound libraries has turned the interest towards *in silico* prediction of Caco-2 permeability (P_{app}) using diverse types of molecular descriptors and computational techniques. Until now, many of the computational approaches developed to predict intestinal permeability has a limited practical use due to the size of the permeability datasets and their inter- and intra laboratory variability, which makes difficult to develop accuracy and predictive models. In this sense, the main goal of this study was to develop general comprehensive models that discriminate compounds with high permeability (Papp $\geq 8x10^{-6}$ cm/s) from those with moderate-poor permeability ($P_{app} < 8x10^{-6}$ cm/s) using a large dataset of 719 compounds with reported values of permeability in Caco-2 cells. More than 1400 molecular descriptors implemented in the Dragon software were calculated according to the dependence with the dimensionality of the structural representation (0D, 1D, 2D and 3D indices). Twenty one classification models based on indices generated from 20 descriptor families were performed and the best model by family of descriptors were selected considering the percentage of good classifications for training and test sets. The global accuracies of all models were ranking between 78-82%. A general model based on all molecular descriptors was developed and it classified correctly 81.31% and 80.71% for training and test sets. The statistic assumptions and application domain for the final model were verified. An external set of 9 compounds, with different permeability/solubility profiles, were predicted and assessed by experimental assay in Caco-2 cells. The potential use of the final classification model was evaluated by a virtual screening of a database of 290 compounds (not present in the permeability data) with reported values of human intestinal absorption (HIA). The model predicted 99 compounds with high permeability in Caco-2 cells and 90 of them had experimental HIA values greater than o equal to 80% for a 91% of good classification. The results of this study suggest that with the present methodology is possible to obtain models with strong predictive ability and can be used in the design of large libraries of compounds with appropriate values of permeability and to perform virtual screening in the early states of drug development.

References:

1-Norinder, U., Bergstrom, CAS, Prediction of ADMET Properties. Chem Med Chem 2006, 1, 920-37.

2-Hou T, Wang J, Zhang W, Wang W, Xu X 2006. Recent advances in computational prediction of drug absorption and permeability in drug discovery. Curr. Med. Chem. 13(22):2653-2667

POSTER PRESENTATION I-71

CLASSIFICATION AND REGRESSION MODELS TO PREDICT HUMAN ORAL BIOAVAILABILITY FROM MOLECULAR STRUCTURE: A RELATIONSHIP AMONG P-GLYCOPROTEIN, CYP3A4 AND BIOAVAILABILITY

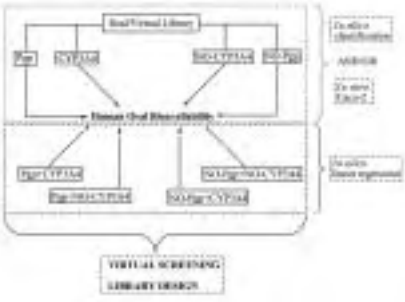
Miguel Ángel Cabrera-Pérez [1], Hai Pham The [1], Isabel González-Alvarez [2], Teresa Garriguez-Pelufo [3] and Marival Bermejo-Sanz [2]

[1]Molecular Simulation & Drug Design Group. Centre of Chemical Bioactive. Central University of Las Villas. Santa Clara 54830, Villa Clara, Cuba.

[2]Department of Engineering, Area of Pharmacy and Pharmaceutical Technology, Miguel Hernández University, 03550 Sant Joan d'Alacant, Alicante, Spain.

[3]Department of Pharmacy and Pharmaceutical Technology, University of Valencia, Burjassot 46100, Valencia, Spain

Cytochrome P450 (CYP3A4) and P-glycoprotein (P-gp) are present, at high levels, in the small intestine enterocytes. This suggests that bioavailability (F) of Pgp-CYP3A4 substrates can be reduced by both processes. To study this relationship were used 292 substrates and non-substrates of P-gp and CYP3A4, with human bioavailability values reported. Different molecular descriptors based on chemical functional groups, atom-centred fragments and molecular properties were calculated. Classification Tree (CT) Analysis were applied to identify, through simple hierarchical rules, those structural features responsible to classify compounds with poor ($F \leq 20\%$). and moderate-high ($F > 20\%$) bioavailability values in separate datasets of P-gp (CT1) and CYP3A4 substrates (CT2), and P-gp (CT3) and CYP3A4 non-substrates (CT4). The global accuracy of each model was 89%, 85%, 90% and 85%, respectively and variables such as molecular weight, polarizability, the number of double bond, electronegativity, partition coefficient, topological polar surface area, the hydrophilic factor, etc., were able to explain this biopharmaceutical property. An external validation set of 62 compounds was used to assess the classification models and the general accuracies achieved were 87%, 89%, 91% and 100% for CT1, CT2, CT3 and CT4 models, respectively. In order to explore whether better models for F can be built when compounds share some ADME properties like efflux (P-gp) and metabolism (CYP3A4), four subsets were selected from the complete dataset to develop quantitative models and better models were obtained by multiple linear regression (MLR) analysis (R^2 between 58-72 and Q^2_{LOO} between 51-61). The fourth linear models were assessed with the same external set. These studies evidenced that a combination of classification and regression models is a good method to predict human oral bioavailability where different pharmacokinetic properties are shared (See Figure). The results suggest that previous methodology may help chemist to planning the synthesis focused on compounds with increased oral bioavailability, being a useful tool during the drug discovery process.



CLASSIFICATION OF DRUGS ACCORDING TO THEIR MILK / PLASMA CONCENTRATION RATIO

Mohammad H. Fatemi*, Mehdi Ghorbanzad'e
Laboratory of Chemometrics, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran

In this work the classification of drugs according to their milk/plasma concentration ratio (M/P) was done accurately by using counter propagation artificial neural network (CP-ANN) for the first time. The features of each drug were encoded by five LFER descriptors including: the solute excess molar refractivity (E), the solute dipolarity/polarizability (S), the McGowan volume (V) and overall hydrogen bond acidity (A) and basicity (B). These descriptors were used as inputs for developing of linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), least square support vector machine (LS-SVM) and CP-ANN models to distinguish the potential risk of 154 drugs as high risk (with $M/P > 1$) and low risk (with $M/P < 1$) for lactating women. The accuracy of classification for training, internal and external test sets was 91.1%, 90.00% and 80.00%, respectively for LS-SVM model while the accuracy of classification for training, internal and external test sets was 100.00%, 100.00% and 90.00%, respectively for CP-ANN model. The total accuracy for LS-SVM and CP-ANN models in classification of drugs was 90.25% and 99.35%, respectively. Comparison among the results of these models shows the superiority of CP-ANN over others. The reliability of CP-ANN was further examined by an additional external test set, which consists of the M/P values of 6 drugs and 25 organic pollutants. The results obtained indicate the overall non error rate of 89.3% for this model. In conclusion, the results of these investigations revealed the applicability of CP-ANN model in classification of drugs based on their M/P values, using LFER parameters.

M.H. Abraham, Javi Gill , M.H. Fatemi, Europ. J. Med. Chem., 44 (2009) 2452–2458.
M.H. Fatemi, S. Garaghani, Bioorg. Med. Chem. 15(2007) 7746

POSTER PRESENTATION I-73

QSPR MODELS FOR PREDICTIONS AND DATA QUALITY ASSURANCES: MELTING POINT AND BOILING POINT OF PERFLUORINATED CHEMICALS

Barun Bhhatarai[1], Wolfram Teetz[6], Tomas Öberg[2], Tao Liu[2], Nina Jeliazkova[3], Nikolay Kochev[4], Ognyan Pukalov[4], Igor V. Tetko[5], and Paola Gramatica[1,*]

[1]QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Structural and Functional Biology (DBSF), University of Insubria, via JH Dunant 3, Varese, 21100, Italy.

[2]School of Pure and Applied Natural Sciences, Linnaeus University, SE-391 82, Kalmar, Sweden.

[3]Ideaconsult Ltd, 4 A. Kanchev str., Sofia 1000, Bulgaria;

[4]Department of Analytical and Computer Chemistry, University of Plovdiv, 24 Tsar Assen Str., Plovdiv 4000, Bulgaria.

[5]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum Muenchen - German Research Center for Environmental Health, Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany.

Quantitative Structure Property Relationship (QSPR) studies on physico-chemical properties of per- and polyfluorinated chemicals (PFCs) are presented. The experimental data for the PFCs for the Melting Point and the Boiling Point used for developing the QSPR models were selected from Syracuse PhysProp database[1] and from other literatures. The data sets were split into training, for model development, and prediction sets, for predictivity check, in two different ways: a) by random selection of response values, and b) by structural similarity verified by Self Organizing Map (SOM). This helps to propose reliable predictive models developed on the training set and externally verified on the prediction set (Test I). Multiple linear regression (MLR), partial least squares (PLS) and (ASNN) were used for QSPR modeling. Individual models based on 0D-2D dragon descriptor, E-state descriptors and fragment based descriptors and their prediction as well as consensus model predictions will be presented and compared.

In addition, the predictive performances of the developed models were verified on a blind external validation set (EV-set) prepared from experimental values available from PERFORCE database[2]. This database contains only long chain perfluoro-alkylated chemicals, particularly monitored by regulatory agencies like US-EPA[3] and EU-REACH[4]. QSPR modeling using different approaches, internal and external validation on two different prediction sets and study of applicability domain highlights the robustness and higher accuracy of the proposed models. Finally, Melting Point for additional 131 PFCs and Boiling Point for 116 PFCs for which experimental measurements are unknown were predicted, verifying their applicability domain.

Acknowledgements:
(Financial support by European Union through the project CADASTER FP7-ENV-2007-1-212668)

References:

1. http://www.syrres.com, b) http://www.ncbi.nlm.nih.gov/pubmed/
2. Hildo Krop, P de Voogt, PERFORCE 2, Task 1: Physicochemical parameters and source markers of PFAS. 2008, Interfaculty Environmental Science Department (IVAM), Universiteit van Amsterdam, ISO doc O/0818.
3. http://www.epa.gov/ oppt/pfoa/pubs/related.htm
4. Regulation (EC) No 1907/2006 of the European Parliament and of the Council (18/12/2006) concerning REACH; http://eur-lex.europa.eu/LexUriServ//LexUriServ.do?uri= OJ:L:2007:136 :0003:0280:en:PDF (accessed 12/10/ 2009).

ALLOSTERIC MEK1/2 INHIBITORS WITH IMPROVED ADMET PROPERTIES

Robert Fraczkiewicz, Robert D. Clark, David Miller, Walter S. Woltosz
Simulations Plus, Inc. 42505 10th Street West, Lancaster, CA 93534, USA

MEK1 and MEK2 are integral parts of the Ras-dependent Raf/MEK/ERK1/2 mitogen activated protein (MAP) kinase signaling pathway. As a major regulator of cell proliferation and survival, this pathway has been the focus of many targeted cancer therapy programs. We have used this therapeutic area as a proof-of-concept exercise to demonstrate a new methodology of drug lead generation and optimization using published structures and pIC50 values for allosteric MEK1/2 inhibitors. We built a QSAR model for pIC50 values and validated our predictions of potential ADME/Tox problems against literature reports. Next, we applied an R-table explosion tool to automatically create 3240 novel variations on the published inhibitors, 280 of which lay within the applicability domain of our MEK1/2 inhibition model. Filtering out those compounds expected to be less potent (IC50>10nM) and unlikely to be orally absorbed (ADMET_Risk > 2) derivatives left us with 30 compounds. We have concentrated our attention on this selected group of compounds, examining their potential water solubility, cardiotoxicity, and hepatotoxicity using appropriate predictive models. Descriptor sensitivity analysis allowed us to identify other derivatives predicted to have improved solubility and fewer toxicity problems.

POSTER PRESENTATION I-75

COMPARISON OF CYP450-MEDIATED METABOLISM PREDICTION MODELS: APPLICATION OF OPENTOX FRAMEWORK FOR PREDICTIVE TOXICOLOGY

Nina Jeliazkova, Diana Zhiryakova, Boris Aleksiev, Vedrin Jeliazkov
Ideaconsult Ltd., Sofia, Bulgaria

The cytochrome P450 plays an important role in the oxidative degradation of compounds and multiple approaches towards its modelling have been developed. Recognizing the significance of these models for predicting the P450-mediated metabolism of drugs and xenobiotics, this work aims at reviewing existing models and datasets, and comparing a selected set of models, with the help of the OpenTox framework. The OpenTox framework, developed by partners of the FP7 OpenTox project provides unified access to toxicity data, predictive models, procedures supporting validation and additional information that helps with the interpretation of the model predictions. This is achieved on two levels: i) a common information model, based on ontologies, and ii) availability of data and methods via a standardized web services interface, where every compound, data set, descriptor calculation algorithm or statistical method has an unique web address, which is used to retrieve information or initiate the calculations. The selected models are made available online and can be applied to new user supplied compounds. An exploration and comparison of the prediction results on data, extracted from publications, online databases with the in-vitro and in-vivo assays from EPA ToxCast program is performed.

Acknowlegements:
(1) OpenTox - An Open Source Predictive Toxicology Framework, is funded under the EU Seventh Framework Program: HEALTH-2007-1.3-3 Promotion, development, validation, acceptance and implementation of QSARs (Quantitative Structure-Activity Relationships) for toxicology, Project Reference Number Health-F5-2008-200787 (2008-2011). More information at www.opentox.org
(2) http://www.epa.gov/ncct/toxcast/

IN SILICO CLASSIFICATION MODELS OF CYTOCHROME P450 LIGANDS USING MACHINE LEARNING METHODS

Jin Hee Lee, Boyeon Kim, and Sun Choi*

College of Pharmacy, Division of Life & Pharmaceutical Sciences, and National Core Research Center for Cell Signaling & Drug Discovery, Ewha Womans University, Seoul 120-750, Korea

Cytochrome P450 (CYP) enzyme superfamily is involved in oxidative phase I metabolism which is a critical determinant for metabolism of various xenobiotics including drugs. It is, therefore, desirable to have models that can predict whether a compound interactions with a specific CYP isoform. There are five major CYP isoforms which are closely associated with more than 80% of the metabolism of all the pharmaceuticals in clinical use and we focused on them.

The purpose of this study is to develop *in silico* classification models that could effectively distinguish ligands from non-ligands for each CYP isoform. We performed machine learning methods, Laplacian-modified naïve Bayesian, random forest (RF), recursive partitioning (RP), and support vector machine (SVM), which are frequently used as a computational approach due to their high speed and accuracy. Data sets were collected from public domain and randomly divided into training and test sets in the ratio of 7 to 3. The classification models were built using the six descriptors which are AlogP, molecular weight, number of hydrogen bond donors, number of hydrogen bond acceptors, number of rotatable bonds, and FCFP_6. The quality of each model was evaluated by the accuracy, sensitivity, specificity, area under the receiver operating characteristic (ROC) scores, and Matthews correlation coefficient (MCC), and most of our models successfully classified ligands and non-ligands for each CYP isoform. These models could be useful to predict the CYP-mediated drug metabolism profiles in early drug discovery.

PROBABILISTIC MODEL FOR THE PREDICTION OF THE HUMAN LIVER MICROSOMAL METABOLISM REGIOSELECTIVITY

P. Japertas[a], J. Dapkunas[a,b], and A. Sazonovas[a,c]

[a] ACD/Labs, Inc., A.Mickeviciaus g. 29, LT-08117 Vilnius, Lithuania

[b] Faculty of Natural Sciences, Vilnius University, M.K.Ciurlionio g. 21/27, LT-03101 Vilnius, Lithuania

[c] Faculty of Chemistry, Vilnius University, Naugarduko g. 24, LT-03225 Vilnius, Lithuania

Cytochromes P450 are the main enzymes involved in the metabolism of drugs and other xenobiotics within human organism. In this work we present a model for in silico prediction of the most probable sites of human liver microsomal (HLM) metabolism in a molecule. The developed models calculate the probabilities of being a target of human cytochrome P450 enzymes (CYP3A4, CYP2D6, CYP2C9, CYP2C19, CYP1A2) for any atom in a molecule and allow forecasting the most probable phase I metabolites. The novel GALAS (Global, Adjusted Locally According to Similarity) modeling methodology was used for development of probabilistic models. The latter technique allows for a dynamic determination of the similarity inside model space, the subsequent corrections of the baseline predictions according to experimental values for the most similar compounds in the training set of the model and estimation of the final prediction quality.

Experimental data on HLM and cytochrome P450 metabolism for 873 compounds with >9000 different atoms (1324 metabolism sites) were used for modeling. Five baseline models were developed for five types of atoms considered in the modeling of HLM metabolism (aromatic carbon, aliphatic carbon, carbon near nitrogen, carbon near oxygen, and sulfur). Final GALAS models provide the list of all the atoms with predicted probabilities to undergo metabolic transformations in human liver microsomes.

As a result of GALAS modeling concept application each prediction of the proposed models is provided with a quantitative estimation of its quality in the form of calculated Reliability Index (RI). This quantity is shown to correlate with the prediction accuracy, as both the numbers of mispredictions and inconclusive results reduce significantly when only results of high quality (RI>0.5) are taken into account, demonstrating that RI is suitable for the assessment of the Applicability Domain of the models presented in this work. Moreover, as it is demonstrated by clear examples, the Applicability Domain of those models can be easily expanded to cover specific compound classes of user interest with the help of 'in-house' databases containing experimental metabolism data. In addition, training of the corresponding baseline models with experimental data on metabolism by individual CYP450 isoforms allowed attributing each of the predicted metabolism sites to one or more particular enzymes (CYP3A4, CYP2D6, CYP2C9, CYP2C19, or CYP1A2).

A TRAINABLE IN-SILICO SCREENING FILTER FOR VARIOUS HUMAN CYTOCHROME P450 ISOFORMS INHIBITION LIABILITY

P. Japertas[a], R. Didziapetris[a], J. Dapkunas[a,b], and A. Sazonovas[a,c]

[a] ACD/Labs, Inc., A.Mickeviciaus g. 29, LT-08117 Vilnius, Lithuania

[b] Faculty of Natural Sciences, Vilnius University, M.K.Ciurlionio g. 21/27, LT-03101 Vilnius, Lithuania

[c] Faculty of Chemistry, Vilnius University, Naugarduko g. 24, LT-03225 Vilnius, Lithuania

This study focuses on development and validation of a series of in-silico models that can distinguish between inhibitors and non-inhibitors of the cytochrome P450 isoforms 3A4, 2D6, 2C9, 2C19, and 1A2. Inhibition constant thresholds equal to 10 and 50 uM were used to classify compounds regarding CYP isoform inhibition. The initial data sets ranged from ca. 5000 to 8000 compounds for the five considered enzyme isoforms. These have been compiled from literature publications and PubChem screening database. A novel GALAS (*Global, Adjusted Locally According to Similarity*) modeling methodology was applied, utilizing predefined set of molecular fragments as descriptors. A very important feature of this modeling methodology is the possibility to quantitatively evaluate prediction quality using calculated Reliability Index (*RI*) values.

The obtained *RI* values correlate with prediction accuracy. Predictions with low *RI* are outside applicability domain of models and cannot be considered. For the predictions with acceptable RI values, the accuracy approaches 90% in all five internal test sets (20% of the corresponding initial database). All models have been further subjected to a more sophisticated external validation, using the latest data from PubChem screening program. As an example, in case of CYP3A4 inhibition, it yielded the results similar to the model testing on internal test set (88% accuracy when $RI > 0.3$). Model trainability feature was assessed in an attempt to train the CYP3A4 inhibition model based on literature dataset with PubChem library, while reserving half of the PubChem data as a test set. After adding 5% of the training library the number of test set predictions with acceptable reliability ($RI > 0.3$) was below 50% while the number of high reliability predictions ($RI > 0.5$) barely exceeded 10%. Subsequent additions of library portions gave a steady increase in these numbers reaching ca. 85% and 60% correspondingly with whole library added.

Obtained models represent the valuable computational filters in early drug discovery to identify compounds that may have unwanted cytochrome P450 inhibition liability. The GALAS modeling methodology enables fast and efficient model training, allows extending applicability domain of current models and adjusting them to screen proprietary databases for potential CYP inhibitors.

MODELING TOXICITY OF CHEMICALS TO AQUATIC ORGANISMS

P. Japertas[a], K. Lanevskij[a,b], L. Juska[a,b], R. Didziapetris[a]

[a] ACD/Labs, Inc., A.Mickeviciaus g. 29, LT-08117 Vilnius, Lithuania.

[b] Department of Biochemistry and Biophysics, Vilnius University, M.K.Ciurlionio g. 21/27, LT-03101 Vilnius, Lithuania

This study focuses on the application of recently introduced GALAS modeling methodology for the development of predictive models that would allow estimating toxicity of new chemicals to several aquatic species. Experimental data used for analysis were expressed as median lethal concentration of test compound in water (LC50) representing compounds' toxicity to fish and crustaceans. The overall data set collected from literature contained toxicities of 900 compounds to fathead minnows (*Pimephales promelas*) and almost 600 LC50 values determined for water fleas (*Daphnia magna*).

Each GALAS model consists of two parts, the first one being a global QSAR reflecting the general trends (baseline toxicity prediction), while the second part accounts for more specific effects by introducing local corrections to the baseline values based on the analysis of experimental data for similar compounds. One of the major benefits of the underlying methodology is the ability to estimate prediction reliability by the means of calculated Reliability Index values. Also, new experimental data can be added to expand applicability domain of these models without full statistical reparameterization (trainability feature).

The modeling approach utilized herein for aquatic toxicity predictions was validated by applying the same principles to develop a new model that predicts IGC50 (50% inhibitory growth concentration) to protozoan Tetrahymena pyriformis. This model was submitted as an entry for environmental toxicity prediction challenge hosted by CADASTER project. The final model derived using known IGC50 values for 644 compounds was identified among the winners achieving RMSE under 0.8 log units for prediction of blind validation set containing 120 chemicals.

OLIMPIC - OVERCOMING CURRENT LIMITATIONS IN METABOLISM PREDICTION OF INDUSTRIAL CHEMICALS

Lothar Terfloth[1], Aleksey Tarkhov [1], Johann Gasteiger [1], Melanie Bausen[2], Eric Fabian[2], Robert Finking[2], Thorsten Bernshausen[3]

[1]Molecular Networks GmbH, Henkestrasse 91, 91052 Erlangen, Germany,
[2]BASF SE, Carl Bosch Strasse 38, 67056 Ludwigshafen, Germany
[3]Archeiron, Steinbacher Strasse 9, 65620 Waldbrunn, Germany

Strategic combinations and tiered application of alternative testing methods to replace or minimize the use of animal models in risk assessment is attracting much attention. The aim of the CEFIC-LRI funded OLIMPIC project is to provide improvements in the prediction of metabolism of industrial chemicals in mammals. The holistic investigation will consider in vivo, in vitro and in silico data in order to identify discrepancies between the different methods and to enhance the predictive power of in silico models for metabolism.
In this poster we will report on the compilation of a data set on the metabolism of chemicals in rats, the measurements of S9 in vitro metabolism data for additional 30 compounds, the evaluation of the software package TIMES based on the experimental measurements, and the chemoinformatics reactivity modeling.
The metabolic degradation of the food dye Curcumin (E100) will be exemplarily shown. Four metabolites of curcumin could be detected by LC/UV and LC/MS/MS after standard incubations in rat liver S9-fraction in a cofactor (NADPH) containing buffer system at 37 °C. The metabolites result from either the O-demethylation reaction of an aromatic methoxy group or the reduction of the two alpha,beta-unsaturated carbonyl groups.
Five metabolites of curcumin are predicted by TIMES. Only the metabolite resulting from O-demethylation was detected in vitro and in silico. Neverthless, the principal functionalization reactions were predicted correctly.
The chemoinformatics reactivity modeling with MOSES Risk Assessment predicts the metabolite resulting from the O-demethylation as main metabolite. In addition, the glucuronidation or sulphation of the aromatic hydroxyl group is predicted.

APPLICATION OF CSRML IN THE TTC APPROACH AND METABOLISM PREDICTION IN THE RISK ASSESSMENT WORKFLOW

Lothar Terfloth, Bruno Bienfait, Johann Gasteiger, Thomas Kleinöder, Jörg Marusczyk, Oliver Sacher, Christof H. Schwab, Aleksey Tarkhov

Molecular Networks GmbH, Henkestraße 91, 91052 Erlangen, Germany

CSRML is a novel Markup Language Definition for Chemical Substructure Representation. Although, chemical subgraphs or substructures are quite popular and used since a long time in chemoinformatics, the existing and well established standards still have some limitations. In general, these standards are suited even for complex substructure queries, however, show some insufficiences, e.g., for the inclusion of physicochemical properties or annotation of meta information. In addition, the existing standards are not fully interconvertible and specify no validation techniques to check the semantic correctness of a query definition.

This paper proposes an approach for the representation of chemical subgraphs that aims to overcome the limitations of existing standards. The approach presents a well-structured, XML-based standard specification, the Chemical Subgraph Representation Markup Language (CSRML), that supports a flexible annotation mechanism of meta information and properties at each level of a substructure as well as user-defined extensions. Furthermore, the standard foresees a mandatory inclusion and use of test cases. In addition, it can be used as an exchange format.

The application of CSRML will be demonstrated on the identification of a *de minimis* value of chemicals «below which there would be no appreciable risk to human health», including those of unknown toxicity, based on the consideration of their chemical structures according to the threshold of toxicological concern (TTC) principle [1].

Furthermore, the CSRML format is the underlying basis of the metabolism prediction system *MOSES. Metabolism*. Due to the open and transparent format the knowledge base of the software can be easily modified and extended by the user. The evaluation scheme of the metabolism prediction system as well as the validation with publicly available data sets [2-4] will be presented.

[1] Kroes, R. et al., 2004. Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. Fd. Cosmet. Toxicol., 42: 65-83.

[2] Singh S.B. et al. 2003 A model for predicting likely sites of CYP3A4-mediated metabolism on drug-like molecules. J. Med. Chem., 46:1330-1336

[3] Lewis, D.F.V. 2003 Human cytochromes P450 associated with the phase 1 metabolism of drugs and other xenobiotics: a compilation of substrates and inhibitors of the CYP1, CYP2 and CYP3 families. Curr. Med. Chem., 10: 1955–1972

[4] Sheridan, R.P. et al. 2007 Empirical regioselectivity models for human cytochromes P450 3A4, 2D6, and 2C9. J Med Chem., 50: 3173-3184

POSTER PRESENTATION I-82

AUTOMATED KNOWLEDGE ACQUISITION AND STRUCTURE-ACTIVITY RELATIONSHIP ANALYSIS REGARDING CYTOCHROME P450 METABOLISM

Fumiyoshi Yamashita,[1] Chunlai Feng,[1] Shuya Yoshida,[1] Mitsuru Hashida[1,2]
[1]Graduate School of Pharmaceutical Sciences and [2]Institute for Integrated Cell-Material Sciences, Kyoto University, Japan

The cytochrome P450s (CYPs) are a superfamily of heme-containing mixed function oxygenases that catalyse the regio- and stereo-selective oxidation of a wide variety of xenobiotics, including drugs. Such broad substrate specificity of CYP isoforms often leads to undesirable drug-drug interactions. Competitive or noncompetitive inhibition of CYP by co-administered drugs retards the clearance of drugs from the body, leading to an unexpected rise in their blood concentrations. On the other hand, induced expression of CYP reduces or shortens the duration of pharmacological activity of the drugs by accelerating their clearance. Prediction of drug-drug interactions associated with CYP is an important issue in drug discovery and development as well as in clinical applications. While in vitro high throughput screenings using hepatocytes and microsomes are routinely conducted in drug discovery settings, useful information on drug metabolisms can be obtained from public literatures. The present study was initiated to develop a natural language processing system specializing in extracting information on CYP-chemical interactions. This system utilizes an open-source language processing system GATE, implementing the rules that identify chemical names and extract CYP-chemical interaction information in a context-based manner, in addition to a chemical names dictionary (>100,000 compounds registered). Information extraction was performed by the following steps: identification of chemical and CYP names in the text, transformation of sentences into multiple simple clauses implying each single event, and pattern matching of keyword sequences within the clauses, where the compounds were categorized in substrates, inhibitors, and inducers. Using PubMed database, approximately 2000 compounds names were obtained by the present system. When 100 PubMed abstracts regarding CYP3A4 were randomly selected to examine feasibility of the system, it was found that the present text mining system gave a high performance on extraction of chemical names (0.871 recall, 0.941 precision) and CYP-chemical interactions (0.852 recall, 0.920 precision). For about 1000 compounds that are registered in PubChem, structure-activity relationship analysis was conducted by an extended recursive partitioning method. The analysis revealed 1) CYP2C9 and CYP2C19 are similar in substrate specificity, 2) substrates and inhibitors for CYP2E1 are smaller in molecular size, 3) substrates for CYP2D6 and CYP3A4 are relatively larger, 4) many of CYP2D6 substrates are cationic, and 5) compounds that are not cationized at neutral do not tend to inhibit CYP2D6 activity. These trends can be intuitively understood by a large-scale data visualization technique “HeiankyoView”, which can represent hierarchically structured data with rectangular icons and borders.

VIRTUAL SCREENING FOR SUBSTRATE AND NON-SUBSTRATE OF EFFLUX PUMP OF *Candida albicans*

Om Prakash[1], Nidhi Puri[2], Rajendra Prasad[2] Indira Ghosh[1]

[1]School of Information Technology, [2]School of Life Sciences, Jawaharlal Nehru Univerisity, New Delhi, India

An important aspect of computational drug discovery is *virtual* screening, which is usually done in order to prioritize molecules from large database with respect to biological action. The similarity based virtual screening is widely used technique to rank the database molecules based on decreasing similarity to reference bio-active compound(s). The top-ranked molecules (the nearest neighbors) are expected to have the highest a priori probabilities of bio-activity. There are two broad categories of virtual screening techniques: ligand-based and structure-based. The ligand-based virtual screening generally use pharmacophore models or chemical similarity principle to scan a database of molecules against one active ligand structure. Whereas, the structure-based virtual screening involves docking of candidate ligands into a protein target followed by applying a scoring function to estimate the likelihood that the ligand will bind to the protein with high affinity.

In this study, we present ligand based virtual screening for the identifying probable substrate and non-substrate for two efflux pump protein from fungus *Candida albicans*. Failure of drug accumulation mediated by efflux pump proteins represents one of the prominent mechanisms of antifungal resistance in C. albicans. The two efflux pump we have investigated are CaCdr1p belonging to ATP-binding cassette protein (ABC) and CaMdr1p of Major Facilitator Superfamily (MFS) of C. albicans. The significant promiscuity level of substrate recognition by CaCdr1p and CaMdr1p, presents a challenge for assessing the molecular basis of substrate recognition. We have identified 67 compounds (referred as training set) as substrate and non-substrate for both the efflux pump proteins CaCdr1p and CaMdr1p based on filter-disk assay and cell growth comparison with respect to corresponding wild type variant from *Saccharomyces cerevisiae*.

The high structural diversity of training set (with mean intra-set 2D similarities of 0.08 and molecular weight between 68 to 1260) among the training set, posed a challenge for computational study. We have attempted two computational study: (i) for building classification model(s) and identifying molecular basis of substrate recognition of efflux pump based on theoretical descriptors (calculated using Dragon[1]) importance using Random Forest[2] machine learning approaches, and (ii) for identifying new substrates and non-substrates based on virtual screening using 3D similarity.

We have identified 52 compounds, from database of known drugs (DrugBank[3]), based on shape and electrostatic similarity to each of compounds in training set. Shape and electrostatic has been calculated using ROCS and EON tools[4]. The compounds has been classified as substrate or non-substrate on the basis of maximum of shape or electrostatics similarity to any 67 compounds in training set. Each of the newly identified 52 compounds has been procured and experimentally tested for being substrate and non-substrate properties against CaCdr1p and CaMdr1p using filter-disk assay.

It has been observed that the similarity to the training-set of compounds classes has reasonably acceptable and higher classification accuracy as compared to models built using Random Forest. The selected compounds on the basis of similarity to any of CaCdr1p and CaMdr1p substrate, has been found to be experimentally substrate with accuracy of 68% and 71% in case of CaCdr1p and CaMdr1p respectively. Whereas, the selected compounds selected on the basis of similarity to CaCdr1p and CaMdr1p non-substrate, has been found to be experimentally non-substrate with accuracy of 78% and 83% in case of CaCdr1p and CaMdr1p respectively. Whereas, the classification accuracy of model(s) built using Random

Forest and theoretical descriptors has been observed to approximately 50% in both cases. However, using Random Forest, we have found that hydrophobicity and charge descriptors could be important factors governing molecular recognition of efflux pump substrates.
It can be observed that the 3D shape and electrostatic similarity can be used as important descriptor for "selecting" (not designing) new substrates and non-substrates for CaCdr1p and CaMdr1p.

References:

[1] http://www.talete.mi.it/products/dragon_description.htm
[2] http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm
[3] Wishart DS et.al (2008) Nucleic Acids Res. 2008 Jan;36(Database issue):D901-6.
[4] http://www.eyesopen.com

DIFFERENCE OF SUBSTRATE RECOGNITION BETWEEN HUMAN CYP3A4 AND CYP2C19 BASED ON METABOLITE IDENTIFICATION AND *IN SILICO* PREDICTION

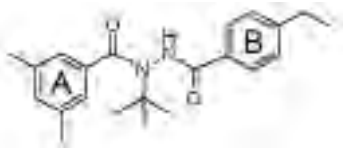
Naoki Shirotani[1], Moe Togawa[1], Shinichi Ikushiro[2], Toshiyuki Sakaki[2], Toshiyuki Harada[1], Hisashi Miyagawa[1], Masayoshi Matsui[3], Kazuki Mikata[3], Kazuhiko Nishioka[3], Nobuhiro Hirai[1], And Miki Akamatsu[1]

[1] Graduate School of Agriculture, Kyoto University,
[2]Faculty of Engineering, Toyama Prefectural University,
[3]Environmental Health Science Laboratory, Sumitomo Chemical Co., Ltd., Japan

Humans are daily exposed to many chemicals and agrochemicals as residues in food and water, or through occupational use. Metabolism and efflux systems have been developed to protect humans from xenobiotics absorbed by the body. The majority of metabolic pathways are mediated by the Cytochrome P450 (CYP) superfamily, and the genetic polymorphism is found among the CYP isozymes. Therefore, examination of the metabolic reactions by each CYP isozyme is important to clarify the substrate recognition mechanism. If the mechanism is clarified, it could be useful for not only the risk assessment of chemicals for humans, but also the development of new drugs and agrochemicals.

In this study, we identified the metabolite structures of an insecticide, Tebufenozide, as a model compound, by two human CYP isozymes: CYP3A4 and CYP2C19. CYP3A4 accounts for 30% of total CYP proteins in human liver and is a major CYP isozyme involved in drug metabolism. CYP2C19 is known for the genetic polymorphism for which about 20% of Japanese express the poor metabolizer phenotype. Tebufenozide was metabolized by the yeast-expressed human CYP3A4 and CYP2C19. Metabolites from each CYP isozyme were isolated and the structures were identified using LC/MS, NMR, and/or organic syntheses. As a result, three metabolites by CYP3A4 were identified; the methyl group of the A-ring, the ethyl group of the B-ring, and the tert-butyl group were hydroxylated. In contrast, three metabolites by CYP2C19 were identified, but in each metabolite the only ethyl group of the B-ring was modified. This result shows that the substrate-binding site of CYP3A4 is larger than that of CYP2C19. Thus, substrates can interact with CYP3A4 by various binding modes, but by limited modes with CYP2C19.

Docking simulation was carried out for in silico prediction of the metabolite structures. Since the CYP2C19 crystal structure is not clarified yet, it was modeled based on its homology with human CYP2C9. Using the crystal structure of CYP3A4 and the homology model of CYP2C19, we performed the docking of Tebufenozide with the two isozymes and compared the binding modes between them.



Tebufenozide

SUBSTRATE RECOGNITION BY P-GLYCOPROTEIN, A DRUG EFFLUX PUMP, BASED ON STRUCTURE-ATPASE ACTIVITY RELATIONSHIPS OF DIVERSE COMPOUNDS

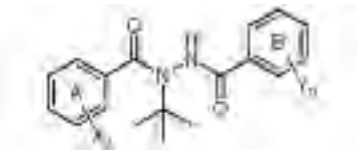
Miki Akamatsu, Satoshi Kanaoka, Yasuhisa Kimura, Yoshiaki Nakagawa, and Kazumitsu Ueda

Graduate School of Agriculture, Kyoto University, Kyoto, Japan

Humans are daily exposed to many chemicals and agrochemicals as residues in food and water, or through occupational use. Metabolism and efflux systems have been developed to protect humans from xenobiotics absorbed by the body. P-glycoprotein (P-gp) is a member of the protein family labeled ATP-binding cassette (ABC) transporters. P-gp actively transports a wide variety of drugs and xenobiotics out of cells and functions as an efflux pump. Since P-gp recognizes compounds having diverse structures as substrates, it plays an important role for multidrug resistance in the treatment of cancers. However, the mechanism of P-gp substrate recognition is complicated and still poorly understood.

In this study, we first screened diverse chemicals including agrochemicals by measuring the ATPase activity of P-gp. As a result of screening, dibenzoylhydrazine (DBH) -type insecticides such as tebufenozide and methoxyfenozide, and organophosphates, such as iprobenfos, showed relatively high ATPase activity. The activities of neonicotinoids, pyrethroids, triazines, and thiocarbamates were not so high. Among the tested steroids, progesterone showed moderate activity. Other steroids such as sex hormones (testosterone and estradiol) and an insect hormone (ecdysone) had no or very low activity. However, progesterone, testosterone, and estradiol have been reported to not be substrates of P-gp in previous reports. Since it was also reported that progesterone significantly stimulates P-gp ATPase activity, more investigation is necessary.

Next, the ATPase activity of the A-ring: 3,5-dimethyl and 2-Cl derivatives and non-substituted B-ring derivatives of DBHs was evaluated and analyzed with 3D-QSAR, CoMFA. Good CoMFA equations were obtained for 3,5-dimethyl and 2-Cl derivatives, and the combined set of the compounds. For non-substituted B-ring derivatives, no significant equation was obtained probably because of low activity of the set of compounds. Introduction of log P did not improve the equations. Based on the qualitative SAR and CoMFA result, 3,5-dimethyl groups on the A-ring were more favorable than 2-Cl groups. Regarding the B-ring, electropositive 2-substituents and bulkier 3-substituents were favorable while smaller 4-substituents were better for the activity. The QSAR results can be useful for determining the P-gp binding site of DBHs and other compounds.



DBH

MOLECULAR DOCKING OF NATURAL PHARMACOLOGICALLY ACTIVE XANTHONES WITH HUMAN SERUM ALBUMIN AND BOVINE SERUM ALBUMIN, MAIN PROTEIN TRANSPORTERS IN PLASMA

L. A. Saraiva[1,2], N. J. F. da Silveira[1], I. C. Rodrigues[1], J. M. S. Silva[2], P. G. Santos[2], M. H. Santos[2], M. P. Veloso[2]

[1]Departamento de Ciências Exatas - Universidade Federal de Alfenas (Unifal-MG) Rua Gabriel Monteiro da Silva, 700 - Centro - Alfenas/MG, Brazil.

[2]Departamento de Farmácia - Unifal-MG

Mangiferin (1,3,6,7-tetra-hidroxixanthone-C2-β-D-glucoside), a natural xanthone formed from the glycosylation of Mangiferitin (1,3,6,7-tetra-hidroxixanthone) has been described for exhibiting pharmacological effects such as: immunomodulation, anti-inflammatory activity and promotion of bone resorption[1] in parallel with a potent antioxidant activity of Mangiferitin[2].

Nevertheless, there is a lack of information regarding possible molecular mechanics of interaction which justify their biological activity. With respect to this, the present research aims to enlighten the molecular interactions involving the natural xanthones (Mangiferin and Mangiferitin) with Human Serum Albumin (HSA), protein of larger abundance in blood plasma and mainly transporter of drugs in human blood[3]. In addition to that, this study also intends to investigate the molecular interactions of those compounds with Bovine Serum Albumin (BSA). Once the latter is much easier to obtain, is thus used more often than the former in studies which promote researches involving the plasmatic serum albumin.

Since the structure of BSA is unavailable in the Protein Data Bank – PDB, an homology model was created. A BLAST search in PDB with the bovine serum albumin sequence [Swissprot sequence ALBU_BOVIN (P02769)] showed that BSA shares 75% identity with HSA (PDB Code: 1n5u). Model building of BSA was carried out using the program MODELLER[4] and the overall stereochemical quality of the final model for BSA was assessed by the program PROCHECK[5]. For the molecular docking was used the software Induced Fit Docking from the suite Schrödinger[6].

The results enable to map the interactions between the functional groups of the ligands with the specific residues belonged to the biding site of HSA and BSA. With regard to the biding Gibbs free energy (ΔG°), the proximity between the ΔG° of the complex Mangiferin-HSA (ΔG°: -30,66 KJ.mol^{-1}) and Mangiferin-BSA (ΔG°: -27,76 KJ.mol^{-1}) and the proximity ΔG° between the complex Mangiferitin-HSA (ΔG°: -31,16 KJ.mol^{-1}) and Mangiferitin-BSA (ΔG°: -30,66 KJ.mol^{-1}) revealed that BSA might be a cheaper alternative to experiments involving HSA. Moreover the little difference between the biding Gibbs free energy of the systems involved Mangiferin and Mangiferitin leads to the conclusion that the presence of the sugar residue in the Mangiferin has little influence to its molecular interaction mechanism with HSA.

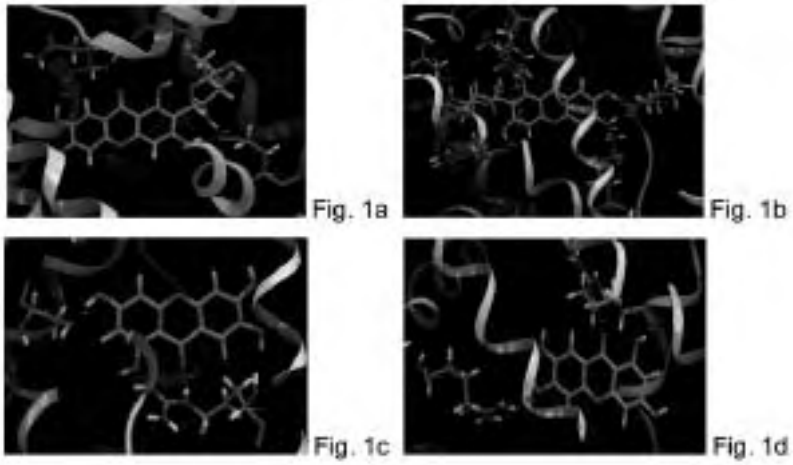


Fig1a: Mangiferin bound to HSA. with a ΔG° of -30,66 KJ.mol^{-1} (in silico). Fig 1b: Mangiferin bound to BSA. with a ΔG° of -27,76 KJ.mol^{-1} (in silico) and ΔG° of -26,15 KJ.mol-1 (experimental data)[7] Fig 1c: Mangiferitin bound to HAS with a ΔG° of -31,16 KJ.mol^{-1} (in silico). Fig 1d: Mangiferitin bound to BSA. with ΔG° of -29,49 KJ.mol-1 (in silico) and a ΔG° of -28,17 KJ.mol^{-1} (experimental data)[7].

We would like to thank to FAPEMIG, CNPq and FINEP for financial support.

References:

1. Leiro, J. M. et. al., (2003) In vitro effects of mangiferinon superoxide concentrations and expression of the inducible nitric oxide synthase, tumor necrosis factor-alfa and transforming growth factor-beta genes, *Biochem. Pharmacology*, 65, 1361–1371.
2. Sánchez, G. M. et. al. (2003) Mitos y realidades de la terapia antioxidante – VIMANG Nuevo Producto Natural Antioxidante. Minist. de la Salud Pública, 3, 13.
3. Hsu MF, Raung SL, Tsao LT, Lin CN, Wang JP (1997) Examination of the inhibitory effect of norathyriol in formylmethionyl-leucyl-phenylalanine-induced respiratory burst in rat neutrophils. *Free Radic Biol Med* 23:1035–1045
4. Sali, A., Blundell, T. L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol*. 234, 779 – 815.
5. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thorton, J. M. (1993) PROCHEK: a program to check the stereochemical quality of protein structure. *J. Appl. Cryst*. 26, 283 – 291.
6. Schrödinger Suite 2008 Induced Fit Docking protocol; Glide v .5.0; Prime v. 1.7, Maestro v. 8.5 Schrödinger, LLC, NY, 2005.
7. P. G. Santos; M. H.Santos; J. M. S. F. da Silva; Personal communication, 2010.

QSAR OF BINDING OF LIGANDS TO BSA: A QUARTZ-CRYSTAL MICROBALANCE WITH DISSIPATION MONITORING APPROACH

C. Selassie, M. Johal, R. Rawle, K. Myers and R. P. Verma
Chemistry Department, Pomona College Claremont, CA 91711, USA

Plasma-protein binding is a critical parameter of a drug like absorption, distribution, bio-transformation and excretion. Serum albumin, one of the most abundant blood plasma proteins, plays a significant role in transport, modulation and inactivation of endogenous substrates, drugs and other small ligands (1). There is a striking homology in the sequences of bovine serum albumin (BSA) and human serum albumin (HSA) thus BSA can easily be utilized as a model system for studying the interactions between ligands and serum albumin.

Several methods have been used to study the binding interactions of small, organic molecules with serum albumin. These methods include high performance affinity chromatography, quartz crystal resonant sensor techniques, solid phase micro-extraction and gas chromatography, and surface plasmon resonance technology. The main limitation of these methods is their inability to monitor structural changes of surface bound molecules in real-time and to provide specific information on adsorption processes and kinetics. In this study, quartz-crystal microbalance with dissipation (QCM-D) monitoring provides a relatively easy, highly sensitive and cost-effective way to monitor the binding interactions between a series of miscellaneous drugs and BSA

In QCM-D, a thin crystal quartz disk with metal electrodes on either side is induced to oscillate at its resonant frequency through the application of an AC current (2).Since the quartz is piezoelectric, mechanical strain on the disk creates an electrical field in the crystal. Thus any mass that is added to, or removed from the crystal causes a frequency shift, Δf which is related to the mass deposited or removed from the crystal surface. When this layer is uniformly distributed and rigid, the Sauerbrey equation can be used to relate changes in oscillation frequency to changes in mass, Δm

$$\Delta m = - \Delta f \bullet .C/n$$

where n = 1, 3, 5 ...is the overtone number and C is a constant (17.7ng cm^{-2} Hz^{-1} for a 5MHz AT-cut quartz crystal). Thus an increase in mass leads to a decrease in frequency. Changes in dissipation allows us to learn more about the viscoelastic properties of the surface formed on the crystal and thus an increase in dissipation is associated with an increase in "floppiness' and loss of the film's rigidity.

In order to create this model platform of non-specific drug binding in vivo, the gold-coated quartz crystal was functionalized with 11-mercaptoundecanoic acid and activated with N-hydroxysuccinimide (NHS) and 1-ethyl-3-{3-dimethylaminopropyl} carbodiimide hydrochloride (EDC) for attachment to BSA (3). BSA deposition was subsequently monitored in real time. The various ligands at fixed concentrations and pH = 7.0 were then passed over the BSA surface in the QCM-D flow cell and dissipation and frequency shifts were monitored.

From the QCM-D data, kinetic parameters such as k_{on}, k_{off} and K_d were obtained using non-linear regression. Desipramine and salicylic demonstrated aberrant behavior that was further analyzed using dual polarization interferometry. QSAR analyses was then carried out to ascertain the relationship between equilibrium binding constants and the physicochemical attributes of the other fifteen drugs. These results will be discussed in detail; hydrophobicity accounted for most of the variance in the data.

References:
(1) Kragh-Hansen, U. 1981. Molecular aspects of ligand binding to serum albumin. Pharm. Rev., 33:17-53
(2) Höök, F.;Ray, A.; Nordém, B.; Kasemo, B. 2001. Characterization of PNA and DNA immobilization and subsequent hybridization with DNA using acoustic-shear-wave attenuation measurements. Langmuir 17:8305-8312
(3) Chitpan, M.; Wang, X.; Ho, C.-T..; Huang, Q. 2007. Monitoring the binding processes of black tea thearubigin to the bovine serum albumin surface using quartz crystal microbalance with dissipation monitoring. J. Agric. Food Chem. 55: 10110-10116

QSAR MODELING OF CHEMICAL COMPOUNDS EFFECTS ON ANTITARGETS

A.V. Zakharov, A.A. Lagunin, D.A. Filimonov, V.V. Poroikov
Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences,
Pogodinskaya Str., 10, Moscow, 119121, Russia

Approximately 10% of new chemical entities (NCEs) show serious adverse drug reactions (ADRs) after the introduction into medical practice. The most of common ADR (e.g. hepatic toxicity, hematologic toxicity and cardiovascular toxicity) are caused by action on antitargets. To avoid the occurrence of the antitargets effects, specific studies to detect them should therefore be conducted before the NCE is launched. The QSAR modeling was used for assessment of the compounds action on antitargets in this work.
We collected the information about 4000 organic molecules with data on about 30 antitargets effects. The end-points of antitargets are based on the Ki, IC_{50} and EC_{50} values. We used MNA [1] and QNA [2] (Multilevel and Quantitative Neighbourhoods of Atoms) descriptors and Self-Consistent Regression [3] for QSAR modeling of antitargets effects. The initial data set was randomly divided onto the training and test sets in proportion 90% and 10% respectively for each antitarget activity. More than 10 QSAR models were developed for each antitarget activity and used for consensus prediction. The accuracy of prediction estimated by the leave-one-out cross-validation procedure was in the range 0.5-0.9. The obtained models were also validated using the test set. Consensus predictions for the test set were performed taking into account the applicability domain for each model. The accuracy of prediction was obtained in the range 0.5-0.9 for all antitargents effects in the test set.
Therefore, MNA and QNA descriptors with Self-Consistent Regression could be successfully used for predictive QSAR modeling of antitargets effects.

Acknowledgements. This work is partially supported by European FP7 grant # 200787 (OpenTox).

References:

[1] Filimonov D., Poroikov V., Borodina Yu., Gloriozova T. Chemical Similarity Assessment through multilevel neighborhoods of atoms: definition and comparison with the other descriptors. *J. Chem. Inf. Comput. Sci.*, **1999**, 39, 666-670.
[2] Filimonov D.A., Zakharov A.V., Lagunin A.A., Poroikov V.V. QNA based 'Star Track' QSAR approach. *SAR and QSAR Environ. Res.*, **2009**, 20 (7-8), 679-709.
[3] Lagunin A., Zakharov A., Filimonov D., Poroikov V. A new approach to QSAR modelling of acute toxicity. *SAR and QSAR in Environmental Research*, **2007**, 18, 285-298.

STRUCTURAL INSIGHTS INTO AROMATIC-AROMATIC INTERACTIONS IN HERG

Kirsten Knape[1], Peter Wolschann[1], Anna Weinzinger[2]

[1]Institute of Theoretical Chemistry, University of Vienna, Währinger Straβe 17, 1090 Vienna, Austria

[2]Department of Pharmacology and Toxicology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

HERG (human ether-a-go-go related gene) encodes the pore-forming subunit of the voltage-gated potassium channel in the heart[1]. Disturbance of hERG function due to inherited mutations, or due to side effects of drugs, has been linked to congenital long QT syndrome, which may lead to serious arrhythmia and sudden cardiac death[2].

Amino acids essential for hERG channel block (T623, S624 and V625 located at the bottom of the pore helix and residues G648, Y652 and F656, located in S6 segments) have been identified by systematic mutations, in particular by an alanine-scan[3].

We used our recently refined hERG structure[4] to study aromatic residues essential for hERG channel block in detail. In the homology model residues Y652 and F656 interact with each other via π-π stacking. Furthermore, the Y652 side-chain forms aromatic interactions with F557 from the neighbouring S5 helix. 100 ns molecular dynamics (MD) simulations revealed that the aromatic side-chains Y652 and especially F656 adopt a variety of conformations. When Y652 is mutated to alanine, the previously mentioned aromatic interactions are lost. However, MD simulations show a rotation of the F656 side-chain away from the pore axis, thereby allowing new π-π stacking interactions with F557. This may change the binding behavior of hERG channel blockers. The compensatory effect is reversable in a 20 ns MD simulation, when the Y652 side-chain is reintroduced.

The elucidation of conformational changes of aromatic-aromatic interactions in the binding site will be crucial to help interpreting hERG channel block.

References

1. Sanguinetti, M. C., Jiang, C., Curran, M. E., and Keating, M. T. (1995) Cell 81, 299–307
2. Chiang, C. E., and Roden, D. M. (2000) J. Am. Coll. Cardiol. 36, 1–12
3. Kamiya, K., Niwa, R., Mitcheson, J.S.,Sanguinetti, M.C. (2006) Mol. Pharmacol. 69, 1709-1716
4. Stary, A., Wacker, S.J., Boukharta, L., Zachariae, U., Karimi-Nejad, Y., Aqvist, J., Vriend, G., de Groot, B.L. (2010) ChemMedchem 5, 455-467

COMPUTER AIDED STRUCTURAL ANALYSIS FOR FTase INHIBITION, hERG AND TOXICITY OF ARYLTHIOPHENE DERIVATIVES AS ANTICANCER AGENTS

N.S. Hari Narayana Moorthy, Sergio F. Sousa, Maria J. Ramos , Pedro A. Fernandes*

REQUIMTE, Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Among the numerous validated targets used for the discovery of the novel molecules for cancer treatment, one well studied is farnesyltransferase (FTase). Many FTase inhibitors are in phase-I, II and III clinical trials for the treatment of cancer (tipifarnib, SCH 226734, lonafarnib, BMS-214662, L-778123 and U49). Research efforts in the last 8 years are taking place in our laboratory for the elucidation of the catalytic mechanism of FTase and the development of novel FTase inhibitors for cancer therapy. In the present investigation, we have used computational based drug design approaches to calculate the correlation between the structural properties and the FTase inhibitory activity, toxicity (predicted) and hERG (predicted) of some arylthiophene derivatives. Molecular Operating Environment (MOE), q-ADME, q-Tox, q-hERG and Statistica software were used to perform the study.

The variable analysis done by partial least square analysis shows that a polarizability descriptor (Molar refractivity) and the partial charge descriptors (PEOE) are the most important parameters for the FTase inhibitory activity. The shape index descriptors are responsible for the hERG inhibition. Pharmacophore and molar refractivity properties are needed for lethal effect (toxicity). The validated QSAR models obtained from multiple linear regression analysis reveal that the developed models are statistically significant. Distance based approaches were also used to validate the models and itw as shown that they are free from serial autocorrelation and multicollinearity at 1% significant level.

Descriptor	Description	Contribution of descriptor		
		FTase	hERG	Toxicity
BCUT_PEOE_1	Partial charge of the molecules calculated through Burden´s atomic contribution method.	-		
PEOE_VSA+0	Partial charge on the van der Waals surface of the molecule.	-		
Kier3	First kappa shape index		+	
KierA3	Third alpha modified shape index		+	
KierFlex	Kier molecular flexibility index		+	
vsa_don	Approximation to the sum of VDW surface areas of pure hydrogen bond donors			-
SMR_VSA1	The polarizability (molar refractivity) on the van der			
SMR_VSA4	Waals surface area of the molecule	+		
SMR_VSA4				-
GCUT_SMR0			-	
b_rotR	Fraction of rotatable single bonds	+		+
a_nO	Number of oxygen atoms		+	
balabanJ	Balaban's connectivity topological index	+		+

Note: + is positive contribution and – is negative contribution.

The results reveal that the polar and polarisable properties are positively contributing for the FTase inhibitory activity and negatively for the hERG and toxicity of the molecules. It is interesting to note that the topological properties (shape) and molecular flexibility of the molecules are positively correlated for FTase inhibition, hERG and toxicity. The present study confirms that the partial charge groups (negative charge) and polarisable groups (Molar refractivity) on the vdW surface area along with flexibility of the molecules are necessary for the FTase inhibitory activity and these properties reduce the toxicity and hERG interaction of the molecules. This implies that the molecular flexibility of the molecules is the common property for interaction in all targets, while presence of polar groups on the molecular surface is the determinant of the favourable (FTase) or unwanted effect (hERG and toxicity) of the molecules.

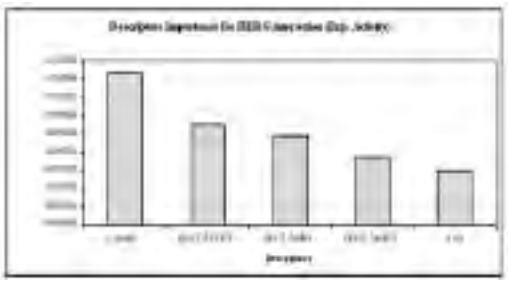
Reference:

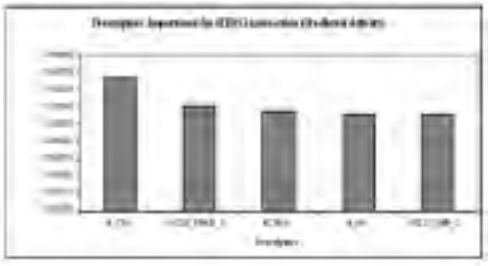
Lethu, S., Ginisty, M., Bose, D., Dubois, J., **J. Med. Chem**., 2009, 52, 6205-6208

STRUCTURAL ANALYSIS OF hERG INHIBITORS: A DRUG OPTIMIZATION STUDY

N.S. Hari Narayana Moorthy, Maria J. Ramos, Pedro A. Fernandes
REQUIMTE, Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

The human ether-a-go-go-related gene (hERG) encodes the major protein underlying the rapid delayed rectifier K^+ current in the heart. The blockade of the HERG channel has been associated with the acquired long QT syndrome (LQTS). Marketed drugs such as sertindole, grepafloxacin and terfenadine were withdrawn due to this effect; they cause ventricular tachyarrythmia and sudden cardiac death. Nowadays, this is one of the main problems in drug design along with inappropriate ADME properties, and the scientists must consider this issues while design and developing new bioactive moieties. We have made on structural analysis of some piperidinyl urea derivatives having hERG inhibitory activity, in order to design novel compounds with reduced or null hERG interactions. The QSAR analysis was performed between the HERG inhibitory activity (experimental and predicted) and the calculated structural properties, and has provided some interesting results. The partial least square analysis provided the importance of the descriptors that influenced in the activity prediction. These results are shown in the graph represented below. The distance based approaches (Cook´s and Mahalanobis distance) were also used to validate the developed models along with other validation method (test and training set). The developed models have Cook´s distance <1, the Q^2 value (>0.85) and the models are free from multicollinearity and autocorrelation.





The results show that the atom contributed descriptors (SlogP, SMR and PEOE) calculated through adjacency and distance matrix methods (GCUT and BCUT) are important for action with the hERG ion channels. The positive contribution of these descriptors suggests that the topological feature of the molecule is important for the interaction. The multiple regression analysis with other descriptors suggests that the polar charged groups on the van der Waals surface area are favourable for the interaction. The vsurf descriptors such as vsurf_wp7, vsurf_ID8, vsurf_EDmin3 show that the hydrophobic properties on the surface of the molecule are unfavourable for binding and polar volume is favourable for the interaction. The results reveal that the active site of the protein may have polar properties and the molecules should have polarizable groups for interaction. The results of the study will help us to design new biological activity moiety with reduced or free hERG interaction.

Reference:

Berlin, M., Lee, Y.J., Boyce, C.W., Aslanian, R., McCormick, K.D., Sorota, S., Williams, S.M., West Jr, R.E., Korfmacher, W., Reduction of hERG inhibitory activity in the 4-piperidinyl urea series of H3 antagonists, Bioorg. Med. Chem. Lett., 2010, 20, 2359-2364.

BIOLOGICAL NETWORKS OF DRUGS AND THEIR CLINICAL EFFECTS BASED ON ATC CLASSIFICATION SYSTEM

Sonny Kim Nielsen, Irene Kouskoumvekaki and Olivier Taboureau
Center for Biological Sequence Analysis (CBS), Department of Systems Biology, Technical University of Denmark, Denmark

Background: System pharmacology is a new way of considering drug action and side effects via multi-target activity profiling i.e. to acknowledge the fact that a drug is likely to interact with other than the primary drug protein target from the same or different protein family [1].
In this context, we tried to elucidate the relationship between clinical effects (therapeutic effects or/and adverse effects) with the drug action of FDA approved drugs using network analysis. To do so, we used the Anatomical Therapeutic Chemical (ATC) Classification System [2]. The ATC Classification System is controlled by the World Health Organization (WHO) and is used for the classification of drugs based on their clinical effect.
Methodology: DrugBank [3] is a public free database containing detailed drug information. Using the information from DrugBank we linked drugs based on shared drug target(s) and labeled them with the ATC code to create a clinical effect - drug network. A network of more than 1200 known drugs with ATC code and annotated proteins was constructed.
Results: The network shows that most of the drugs (62%) bind to more than 2 targets showing their multi-target aspect. Using the 1st level of the ATC code, which is divided into fourteen main therapeutic classes, we see that 33% of drugs are associated to more than one therapeutic effect. As an example we found that two drugs, sibutramine and citalopram that are used for the treatment of obesity and depression respectively, shared the same target, namely sodium-dependent serotonin transporter (5HTT). The serotonin transporter is responsible for the reuptake of serotonin and is known to be involved in mood control [4]. This may suggest that sibutramine could be repositioned as antidepressant and citalopram may have a therapeutic effect on eating behavior. The latter finding is in agreement with the study by McElroy L, et. al. [5], where the authors showed that citalopram can be used for the treatment of binge-eating disorder, which is an onset of obesity.
Conclusion: We propose a method, which combines existing knowledge of drug target and therapeutic effect to reveal new relationships between drugs and protein targets. This method has the potential to discover new therapeutic targets for existing drugs or explain their adverse effects.

References:

[1] Berger, S.I. & Iyengar, R. Network analyses in systems pharmacology. *Bioinformatics* **25**, 2466-2472 (2009).
[2] WHO Collaborating Centre for Drug Statistics Methodology http://www.whocc.no/
[3] Wishart, D.S. et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucl. Acids Res.* gkm958 (2007).doi:10.1093/nar/gkm958
[4] Kuzelova, H., Ptacek, R. & Macek, M. The serotonin transporter gene (5-HTT) variant and psychiatric disorders: review of current literature. *Neuro Endocrinol. Lett* **31**, 4-10 (2010).
[5] McElroy, S.L. et al. Topiramate for the Treatment of Binge Eating Disorder Associated With Obesity: A Placebo-Controlled Study. *Biological Psychiatry* **61**, 1039-1048 (2007).

SUPPORT VECTOR MACHINE FOR PREDICTION OF HEPATOTOXICANTS

Liew Chin Yee, Lim Yen Ching, Yap Chun Wei
Pharmaceutical Data Exploration Laboratory, Department of Pharmacy, National University of Singapore, Singapore

Drug-induced liver injuries (DILIs), although infrequent, is an important safety concern that can lead to fatality in patients and failure in drug developments. A majority of the DILIs are idiosyncratic and may have multiple factors [1]. Hence, support vector machine (SVM), which can be applied to problems where there is limited knowledge on the mechanisms or specific association between the toxic events and molecular properties [2], was used to build a model to predict the hepatotoxic potential of compounds. A total of 1244 pharmaceutical compounds and organic molecules were collected, with their 1D and 2D molecular descriptors calculated for model building. There are 686 positive compounds (hepatotoxicants) and 558 negative compounds (non-hepatotoxicants).

Twenty-four compounds [3] which are drugs withdrawn from the market or those with black box warning for hepatotoxicity were set aside as test compounds to examine the capability of the model to identify severe hepatotoxicants, hence 1220 compounds were available for model training. First, internal validation was done by 5-fold cross-validation and the model has an estimated average sensitivity of 72.9% and specificity of 55.9%. Out of the 24 severe hepatotoxicants, 19 (79.2%) compounds were identified correctly by the model.

The PubChem Substructure Fingerprint (FP) of the 1244 compounds were computed and analyzed. 39 FP were found to be present only in the hepatotoxicants of our data set. Some of these fingerprints are summarized in the figure below. A few of these fingerprints agreed with the drug design guideline on structural alerts for bioactivation [4] which may lead to toxicity, that is, halogenated aromatics, halogenated hydrocarbons and thiol compounds.

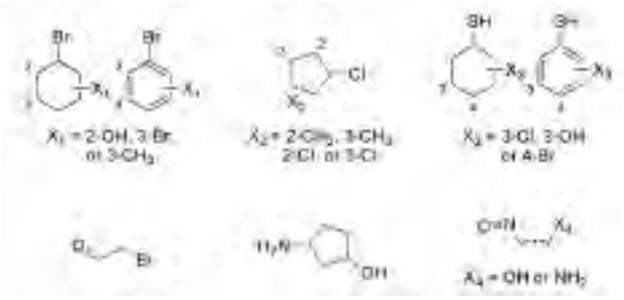


Fig. 1. Selected fingerprints that are present only in the hepatotoxic compounds in the experimental data.

References:

[1] Muster W, Breidenbach A, Fischer H, Kirchner S, Müller L, and Pähler A. Computational toxicology in drug development. Drug Discovery Today, **2008**, 13(7-8):303–310

[2] Xue Y, Yap CW, Sun LZ, Cao ZW, Wang JF and Chen YZ. Prediction of P-Glycoprotein Substrates by a Support Vector Machine Approach. J. Chem. Inf. Comput. Sci., **2004**, 44:1497-1505

[3] Walgren JL, Mitchell MD and Thompson DC. Role of metabolism in drug-induced idiosyncratic hepatotoxicity. Critical Reviews in Toxicology, **2005**, 35:325-361

[4] Guengerich FP and MacDonald JS. Applying mechanisms of chemical toxicity to predict drug safety. Chemical Research in Toxicology, **2007**, 20:344-369

REACTION BARRIERS FOR ELECTROPHILIC MICHAEL-TYPE REACTIONS: MODELLING REACTIVE TOXICITY

Denis Mulliner,[a,b] Gerrit Schüürmann[a,b,*]

[a] UFZ Department Ecological Chemistry, Helmholtz Centre for Environmental Research, Per-moserstr. 15, 04318 Leipzig, Germany

[b] Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Str. 29, 09596 Freiberg, Germany

Covalent binding of toxic chemicals to proteins and their side chains is understood to be a key mechanism in reactive toxicity. In this work the Michael-type reaction between a model nu1cleo1phile representing the protein and a series of electrophilic α,β -unsaturated esters, ketones and aldehydes, known to be reactive toxicants, was studied to develop a regression model for predicting toxicity. In chemoassays it was shown, that reactivity of α,β unsaturated carbonyl compounds towards the tripeptide glutathione (GSH) could be employed to model reactive toxicity. In the present communication, a quantum chemical approach is employed with me1than1thiole (MeSH) as model nucleophile representing the methyl thiole side chain of GSH, mimicking thiol groups of proteins as nucleophilic sites of attack. Quantum chemical rate constants of the reactions between the above-mentioned carbonyl compounds and MeSH were determined using transition state theory (TST). For this purpose ground state geometries of reactants and products as well as transition state geometries were optimized with the Hartree-Fock method, several density functional theory models, and 2nd-order Møller-Plesset per1tur1bation theory utilizing the 6 31G(d,p) and cc pVTZ basis sets. A frequency analysis was carried out for each optimized geometry. Despite large differences in absolute values between experimental solution-phase GSH rate constants and calculated gas-phase MeSH rate con1stants, the latter correlate well with the former and are able to reproduce qualitative trends.

Financial support by the European Commission through the project OSIRIS (contract No. 037017) is gratefully acknowledged.

POSTER PRESENTATION I-95

EXAMINATION OF ELECTROPHILE-PEPTIDE BINDING MECHANISMS BY QUANTUM CHEMICAL TRANSITION-STATE CALCULATIONS

Johannes A. H. Schwöbel, Judith C. Madden, Mark T. D. Cronin
School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England

The binding of xenobiotic electrophiles to biological macromolecules leads to potentially irreversible toxic effects, for example, skin sensitisation, irritation, and enhanced acute toxicity. An improved understanding of the underlying bioorganic mechanisms of toxicity, the so-called 'molecular initiating event', will allow for the development of improved chemical categories and the refinement of structural alerts in expert software systems. Such developments are important for *in silico* based risk assessment under the REACH, and other, legislation. For this reason, we applied quantum chemical DFT calculations, amongst others, in the study of the Michael addition reaction. The Michael addition reaction is one of the most common covalent mechanisms that results in toxicity. Following the reaction coordinate of a series of electrophiles binding to the particular reactive sites in a peptide allowed factors that enhance or diminish chemical reactivity to be investigated. Chemical reactivity was calculated by the activation energy obtained from transition-state calculations via application of the Eyring equation. Systematic examination of both the structure and energies enhance mechanistic understanding. In addition, it is well established that solvent effects have a large influence on reactivity under physiological conditions, thus calculations involving explicit solvent molecules, and continuum solvent models were applied. The results have been compared to experimental kinetic rate constants for peptide binding and show a good quantitative correlation between experimental kinetic rate constants and theoretical peptide binding energies. The findings form the basis of better *in silico* approaches, including QSAR to predict the toxicities of these chemicals. The funding of the EU FP6 InSilicoTox Marie Curie Project (Contract No. MTKD-CT-2006-42328) is gratefully acknowledged.

POSTER PRESENTATION I-96

DEVELOPMENT OF A CHEMICAL REACTIVITY DATABASE FOR COVALENT BINDING BETWEEN ELECTROPHILIC CHEMICALS AND BIOCHEMICAL MACROMOLECULES: OPPORTUNITIES FOR *IN CHEMICO* AND *IN SILICO* PROFILING

Johannes A. H. Schwöbel, Judith C. Madden, Mark T. D. Cronin
School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England

Covalent binding of xenobiotic electrophiles to nucleophilic endogenous biomolecules, e.g. peptides or DNA, is a common molecular initiating event, leading to potentially irreversible toxic effects such as skin sensitisation, mutagenicity, or enhanced acute toxicity. This knowledge provides the basis for the *in silico* prediction of these toxicities. The potential for a chemical to be reactive can be determined experimentally by a number of chemical tests and therefore can be captured computationally to form (Q)SARs. Providing a source of *in chemico* data for the reactivity of electrophiles with reference nucleophiles could assist in the non-animal based risk assessment of chemicals under the REACH legislation and in the application of integrated testing strategies (ITS). For this reason, we have compiled a database from a full range of *in chemico* assays containing various reactivity data of numerous electrophiles forming peptide and DNA adducts, or reference nucleophiles for the evaluation of protein binding mechanisms. This includes reactivity data, kinetic rate constants, and qualitative information regarding the adducts formed. Case studies relating to the *in silico* profiling of toxicologically relevant compounds by this database have been developed to illustrate the application of grouping and category approaches. These allow for the combination of the following information: the identification of electrophilic compounds; their mechanistic applicability domain and compound class; physical-chemical properties; reactivity data; and toxicological data (e.g. acute aquatic toxicity). This data collection will provide a means to evaluate compounds, as required by regulators and industry, and will assist in the development of sound screening tools, e.g. to refine structural alerts in expert systems for toxicity prediction.

The funding of the EU FP6 InSilicoTox Marie Curie Project (MTKD-CT-2006-42328) is gratefully acknowledged.

IN SILICO PREDICTION OF CHEMICAL MUTAGENICITY WITH SUBSTRUCTURE PATTERN RECOGNITION

Shen J., Cheng F., Li W., Tang Y.*
School of Pharmacy, East China University of Science and Technology,
130 Meilong Road, Shanghai 200237, China

Mutagenicity is one of the important toxicology end points, which is highly concerned for both industry and regulators. Mutagenicity is considered as an early alert for carcinogenicity and can be evaluated via *in vitro* Ames tests. However, performing Ames tests for a large set of compounds is relatively resource expensive, and the throughput capacity is nothing compared to the high throughput activity assay and combinatorial synthesis. Recent years, *in silico* techniques have been widely and successfully applied in this field. There are several ongoing EC-funded projects for *in silico* toxicity modeling in Europe, i.e. CAESAR, OSIRIS and CHEMPREDICT etc..[1]

In present study, predictive models of chemical mutagenicity were achieved by a novel substructure pattern recognition method with the latest benchmark data set of Ames mutagenicity, containing 6512 compounds.[2] In our method, each molecule is represented as a substructure pattern fingerprint according to the MACCS substructure dictionary (http://openbabel.org), and then support vector machine algorithm is applied to build the prediction model. The major advantage is that the mutagenicity can be predicted directly from the 2D structure of a molecule, without requiring calculation of electronic properties or generation of 3D structures. In addition, the most important substructure patterns can be identified via the information gain method, which can help interpret the models from medicinal chemistry perspective. Models generated with our method were evaluated in a 5-fold cross-validation procedure. The average correctness of training sets and test sets were 93.1% and 78.3% respectively. And the overall correctness was 92.7%. To further investigate the reliability of our method, two quantitative models were built and evaluated using the Receiver Operating Characteristic (ROC). The area under the ROC curves (AUC) were 0.91 and 0.93 respectively (Figure 1), which were significantly improved compared with other methods.2 Moreover, some of the representative key substructure patterns which significantly correlated with the mutagenicity (alert and safe substructures) were also shown (Figure 2).

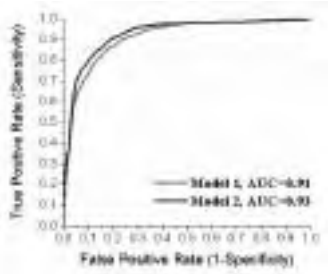


Figure 1. Roc curves of the quantitative models generated using our method.

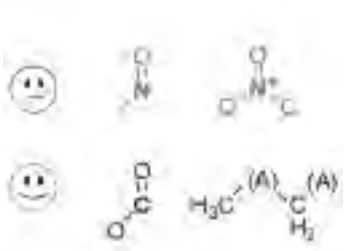


Figure 2. Representative key substructures of mutagenicity recognized in this study.

References:

(1) Benfenati, E., et al. *J. Environ. Sci. Health C-Environ. Carcinog. Ecotoxicol. Rev.* **2009**, 27, 57-90.
(2) Hansen, K., et al. *J. Chem. Inf. Model*. **2009**, 49, 2077-2081.

MAXIMUM COMMON SUBSTRUCTURE BASED ALGORITHM FOR TOXICITY PREDICTION

Hitesh Patel, Surajit Ray, Om Prakash, Indira Ghosh
School of Information Technology, Jawaharlal Nehru University, New Delhi – 110067, India

There is an urgent need to develop the efficient *in-silico* methods to predict the toxicity of chemicals that are going to be used in environmental or pharmaceutical industry. Computational predictive toxicology is one of important approach in toxicology. Here we are reporting a novel algorithm for generating computational predictive toxicology model(s) based only on 2-D substructure profile without using physico-chemical descriptors. This algorithm has applied earlier for analysis of non-congeneric compound having similar bioactivity e.g anti tuberculosis to derive bioactive motif [1]. Most often toxicity arises from different chemical substructures and a set of toxicophores have been listed in literature [2]. We have attempted here to determine the relationship between the toxicity of a molecule and the Maximum Common Substructure (MCSS) it shares with other toxic molecules. The present approach, mainly involves: (i) selecting a group of molecules based on common toxicity end-point, (ii) pair-wise comparison to generate a list of MCSS, (iii) removal of the redundant MCSS- i.e. formation of toxicity endpoint specific MCSS dictionary, (iv) generation of fingerprints of training set and test set molecules, and (v) building predictive model(s) using fingerprint as descriptors with the toxicity end-point as the dependent variable. For building classification based predictive model(s), we have used Random forest machine learning technique, which is based on decision tree. Using an internal set database(training set) dictionary, the test set fingerprints are developed and used to predict the toxicity through this model.

We have used two data-sets of aromatic amines tested for: (i) mutagenic activity in *Salmonella typhimurium* TA100, and (ii) carcinogenicity in rodents. The aforementioned data-sets have been used to check performance of our algorithm for comparative analysis with published classification based model(s) [3]. For comparative analysis, we have followed two approaches. In first approach, the classification statistics has been obtained from the model(s) built using MCSS dictionary approach and data-set comprising of exactly same training and test set reported in published study [3]- referred as "RomualdoDataPartition" henceforth. Whereas in second approach, the training and test-set has been partitioned by five replication of random selection based on equal probability of toxic and non-toxic classes. The five distinct training and test-set, thus obtained were used for model building using MCSS dictionary approach- referred as "NewDataPartition" henceforth. The MCSS dictionary has been generated using only training data-set and fingerprint based on this has been used for all the chemicals to build the classification model. Additionally, we have studied the effect of removal of less important MCSS from dictionary on performance of classification models. The effect has been measured on the basis of three classification statistics-namely selectivity, specificity and accuracy. The importance of MCSS based molecular scaffold has been calculated using Random Forest variable importance based on Mean Decrease in Accuracy.

Classification model built based on MCSS dictionary approach has following statistics for model trained on **RomualdoDataPartition:** for mutagenic data-set the prediction accuracy of 0.88 and 0.52 respectively with training and test set, while for carcinogenicity data-set, the prediction accuracy of 0.97 and 0.70 respectively or training and test set. Whereas, the classification model built based on MCSS dictionary approach has following statistics for model trained on **NewDataPartition** for mutagenic data-set the prediction accuracy of 0.90 and 0.71 respectively with training and test set, while for carcinogenicity data-set, the prediction accuracy of 1.0 and 0.69 respectively or training and test set. The published classification models obtained based on six physico-chemical descriptors using Linear discriminant analysis has

comparable accuracy with our model mostly. Removing less important MCSS leads to almost the same accuracy but increases specificity at the cost of sensitivity. The MCSS-dictionary based classification models of toxicity provides reasonable discriminatory power for Toxic and non-toxic selection. In addition one finds the small chemical motifs give rise to the toxicity in the set of data. While comparing them with the known 42 set of toxicophores [2] some of them were found to be present in our dictionary. The present approach not only provide a discriminatory power to decide toxicity amongst non-congeneric compounds but also could provide the medicinal chemist guideline for designing the non-toxic compound during the pre-clinical drug development cycle.

References:

1. Prakash et al. (2006) J. Chem. Inf. Model. 46, 17-23
2. K. Jeroen et al. (2005) J. Med. Chem., 48[1], 312-320
3. Romualdo et.al (2008), J. Chem. Inf. Model, 48 (5), 971-980.

INVESTIGATION OF THE POSSIBLE USE OF IN SILICO METHODS TO IMPROVE THE THRESHOLD OF TOXICOLOGICAL CONCERN APPROACH

Elena Fioravanzo[1], Sue Barlow[2], Arianna Bassan[1], Alan Boobis[3], Manuela Pavan[1], Andrew Worth[4], Chihae Yang[5]

[1]S-IN Soluzioni Informatiche, Vicenza, Italy
[2]Consultant, Brighton, UK
[3]Imperial College of London, London UK
[4]Joint Research Centre, Ispra, Italy
[5]Altamira, LLC, 1455 Candlewood Dr., Columbus, Ohio 43235, USA

The concept of Threshold of Toxicological Concern (TTC) has its roots in the concept that 'safe levels of exposure' can be identified for individual chemicals with known toxicological profiles. The TTC approach is a risk assessment tool based on the principle of establishing a human exposure threshold value for chemicals, below which no significant risk to human health and/or the environment is expected to exist. According to this approach, a safe level of exposure can be identified for many chemicals based on their chemical structure and the known toxicity of chemicals which share similar structural characteristics. Starting with the generic approach ('exposure threshold') used by the US FDA in the 80s, the TTC concept has evolved over the years to take into account extensive analysis of available data on mainly the oral toxicity data of substances, intake/exposures to the substances, and on the basis of a structure based decision tree to find applications mainly in the food area. The TTC approach has been used to evaluate flavouring substances (JECFA, EFSA)[1], food contact materials (US FDA)[2], genotoxic impurities in pharmaceuticals (EMA)[3] and for the risk assessment of chemicals (WHO IPCS)[4]. Recent publications have suggested that the TTC approach can also find uses in other categories of chemicals and more specifically on chemicals (or trace contaminants) in consumer products, food additives, pesticides and cosmetics. In 2010 the European Food Safety Authority (EFSA) funded the present research project aimed at investigating how the applicability of TTC schemes can be improved by incorporating physicochemical data (both experimental and predicted), as well as toxicity data generated by non testing methods such as Quantitative Structure-Activity Relationships (QSARs), experts systems and read-across within structurally related chemical groups. The concept of grouping has been used in the Cramer classification scheme[5], which is the most known approach for structuring chemical in order to make a TTC estimation. In the current study we explore in silico methods that can give further insight into the Cramer scheme. The study comprises three parts: a) retrieval of suitable datasets, b) analysis of the chemical covered by the datasets and c) investigation on how physicochemical properties and in silico methods can improve the Cramer scheme. The two major TTC datasets in relation to food are retrieved, verified and analyzed: the Carcinogen Potency Database (CPDB)6 including 657 carcinogens based on Kroes publication7 corrected to include TD50 values and the Munro dataset (Munro *et al*)[8] covering a variety of non-cancer toxicological endpoints and including 613 organic chemical substances. The chemical space is explored by means of a variety of chemoinformatics techniques using several sets of molecular descriptors, with the aim of obtaining information on physicochemical and structural similarities and dissimilarities on the TTC datasets by reducing the data dimensionality whilst retaining the information content. The analysis is also aimed at

providing a detailed characterisation of the datasets under investigation to explore a possible refinement of the subdivision in groups provided in the Cramer classification scheme. Finally, the last part of the study involves the exploration of the use of *in silico* methods to improve the TTC approach by means of different methodological strategies such as grouping, (Q)SAR models and the development of exclusion-inclusion rules based on structural/mechanistic rules. The work presented is in progress and the results are preliminary and do not represent an EFSA opinion.

References:

1. Schnabel, Regulatory and Scientific Affairs, Givaudan Flavours. The TTC Concept: Past Present and Future Developments Eurotox CEC 5,13 September 2009. **2.** Hirose et al., Toxicology Letters, Volume 172, Supplement 1, 7 October 2007, Page S195.
3. Müller et al., Regulatory Toxicology and Pharmacology, Volume 44, Issue 3, April 2006, Pages 198-211. **4.** Kroes et al., Toxicological Sciences 2005 86(2):226-230. **5.** Cramer et al, Food Cosmet Toxicol, 1978, 16, 255-276. **6.** Gold et al., Environmental Health Perspectives, 58, 9-319, 1984. **7.** Kroes et al., Food and Chemical Toxicology, 2004, 42, 65-83. 8**.** Munro et al., Food Chem Toxicol, 34, 829-867, 1996.

AN EXTENSIVE MULTI-LABEL ANALYSIS OF THE TOXCAST DATA SET

Jörg Wicker and Stefan Kramer
Institut für Informatik I12, Technische Universität München, Germany

The ToxCast™ data set was released by the EPA in beginning of 2009 [1]. Since August 2009, the data of phase one is publicly available. The data set contains 320 chemical structures and approximately 4000 endpoints. The chemicals are tested against the ToxRefDB, which contains 400 toxicological in vivo endpoints. The remaining features are mainly in vitro data. The main goal of the ToxCast™ data set is to break the in vitro / in vivo border and use the in vitro features to predict in vivo endpoints. Nevertheless, the data set is not easy to handle. Both the in vitro and in vivo data are of unknown and presumedly varying quality, many instances have missing values (in both in vitro and in vivo data), the structures are heterogeneous, and there is a slight skew in the class distributions. At the ToxCast™ Data Analysis Summit in May 2009, first analysis strategies were presented by partners in the project. First analyses of the data set showed that it is hard to find correlations between the in vitro data and individual in vivo endpoints.
In the presentation, we will discuss the results of a comprehensive multi-label classification analysis [3] of the data set. Multi label classifiers do not predict only one class but take into account interdependencies among several classes or labels to improve the prediction. In contrast to a previous multi-label approach presented by Jeliazkova et al. [2], we use all 320 structures and not just the 160 structures with all in vivo data available. We used the Mulan multi label library [3], as it provides several multi label classifiers and analysis methods. To handle missing values during testing, we still had to enhance the library to provide correct predictions in this case. With this modification, many classifiers (predicting in vivo endpoints from in vitro data) already gave acceptable results in a ten-fold cross-validation.
To further improve the results, we defined the applicability domain of a classifier in terms of the in vitro data. We optimized two parameters to obtain and predict instances only in the applicability domain of the model. One basic approach to defining the applicability domain is the bounding box method. It predicts every instance which is in the bounding box of the training set and discards the ones that have at least one attribute outside the boundaries of the bounding box. In the case of ToxCast™, this leads to a very limited model. Having approximately 3600 attributes and 320 instances, the probability is very high that an instance is outside the bounding box, i.e. has one attribute with a value outside the boundaries of the training set. We developed a method which is more robust to single extreme values in attributes. It takes into account how far the attributes are inside or outside of the training set limits. It sums up the normalized value of an attribute distance to the bounding box. Thus, an instance benefits if an attribute is in the bounding box and is penalized if an attribute is out of the bounding box. This gives a measure between -1 and 1, indicating to some extent how good the instance can be predicted. Thus, there is an upper and a lower threshold for the applicability of a model to an instance. Using this measure, we were able to further improve the performance of the multi label classifiers. Depending on the chosen thresholds, the labels can be predicted with an microAUC of up to 0.9 (microAUC is slightly over-optimistic while macroAUC is highly over-pessimistic). In the talk, we will present the result of an extensive analysis of the ToxCast data along those lines and discuss the sensitivity of the results against varying parameter values.

References

[1] Judson et al. (2010) "In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization - The ToxCast

Project", Environmental Health Perspectives, in press (published online in December 2009).
[2] Jeliazkova, N., Jeliazkov V. (2009) "Hierarchical Multi-Label Classification of ToxCast Datasets", ToxCast Data Analysis Summit, US EPA, Research Triangle Park, NC, May, 14-15, 2009.
[3] Tsoumakas, G., Katakis, I., Vlahavas, I. (2010) "Mining Multi-Label Data", Data Mining and Knowledge Discovery Handbook, O. Maimon, L. Rokach (Ed.), Springer, 2nd edition, 2010.

Acknowledgements:
This research was funded by the EU FP7 project OpenTox - An Open Source Predictive Toxicology Framework (Health-F5-2008-200787).

POSTER PRESENTATION I-101

CONSTRUCTION OF INTERPRETABLE TOXICOLOGY MODELS FOR PHOSPHOLIPIDOSIS AND AMES MUTAGENICITY

Axel Dietrich, Jens Karl, Andreas Evers, Sonja Franck, Friedemann Schmidt
Sanofi-Aventis Deutschland GmbH

Toxicity and clinical safety are major reasons for attrition of drug candidates. We approached a number of adverse effects by constructing a model. All models are biased towards a higher true positive rate to ensure no liable compound would pass through the filters in the drug discovery pipeline. To address the interpretability decision trees have been chosen as paradigm for the model.
Here we present the construction of two interpretable models for phospholipidosis and AMES mutagenicity prediction using KNIME as the primary data mining and integration tool. We use decision trees with low depth and few descriptors relevant for the underlying biological/chemical mechanism to model the underlying toxicity. Biasing the model towards high true positive rate (TPR / recall / sensitivity) resulted in high values for TPR and slightly lower true negative rate (TNR / Specificity) for both toxicities.

References:

[1] Kola I. et al, "Can the pharmaceutical industry reduce attrition rates?", NRDD, 2004, 3, 711-715
[2] Lüllmann H. et al., "Lipidosis induced by amphiphilic cationic drugs", Biochem. Pharmacol., 27, 1103-1108 (1978)
[3] Wang Q. et al., "Attenuation of amiodarone-induced lung fibrosis and phospholipidosis in hamsters by taurine and/or niacin treatment", J. Pharmacol. Exp. Ther., 262, 127-132, 1992
[4] N.L. Kruhlak, Set al. „Development of a phospholipidosis database and predictive quantitative structure-activity relationship (QSAR) models", Tox. Mech. Methods, 18, 217, 2008.
[5] Anderson N. et al., "Drug-induced phospholipidosis", FEBS Letters 580, 5533-5540, 2006
[6] Pospischil A. et al., "Phospholipidosis in healthy subjects participating in clinical studies: Ultrastructural findings in white blood cells", Exp. Tox. Path, article in press
[7] Pelletier D. J. et al., "Evaluation of a published in silico model and construction of a novel Bayesian model for predicting phospholipidosis inducing potential.", J. Chem. Inf. Model., 47, 1196-1205, 2007
[8] www.knime.org
[9] www.r-project.org
[10] Tetko, I. V., et al., "Can we estimate the accuracy of ADME–Tox predictions?", Drug Discovery Today, 11 15/16, 2006
[11] Jaworska, J., et al., "QSAR Applicability Domain Estimation by Projection of the Training Set in Descriptor Space: A Review", ATLA, 33, 445-459, 2005
[12] Ratcliffe, A. J., "Medicinal Chemistry Strategies to Minimize Phospholipidosis", Current Medicinal Chemistry, 16, 2816-2823, 2009

POSTER PRESENTATION I-102

ON THE CHARACTERIZATION OF VARIABILITY AND UNCERTAINTY FROM QSARS FOR PROBABILISTIC ENVIRONMENTAL RISK ASSESSMENT WITHIN REACH

Sahlin, U.*, Filipsson, M., and Öberg, T.
Environmental Science, Linneaus University, Kalmar, Sweden

When integrating predictions from QSAR models in probabilistic risk assessment, variability and uncertainty are to be quantified as probabilities. This study focuses on how to characterize uncertainties, including variability and (epistemic) uncertainty, related to predictions of QSAR models. First, we review existing methods to characterize variability and uncertainty in QSAR predictions of endpoints with a special attention to probabilistic approaches. This objective was addressed by providing an overview of current practice to characterize variability and uncertainty in endpoint predictions using QSARs for environmental fate and effect endpoints. We found several probabilistic approaches to quantify variability and uncertainty that could deserve more attention. Therefore, we continued discussing the current and potential extent of QSAR models that are probabilistic per se. Our aim was to initiate the empirical support for general characteristics of variability and/or uncertainty in predictions of environmental fate and effect endpoints. Both variability and/or uncertainty in predictions of endpoints are important for the development of probabilistic QSAR models.

DECIPHERING DISEASES AND BIOLOGICAL TARGETS FOR SMALL MOLECULES USING TOXICOGENOMCS NETWORKS

Audouze K.[1], Juncker A.S.[1], Roque F.J.[1], Baltyn K.K.[1], Weinhold N.[1], Taboureau O.[1], Jensen T.S.[1] & Brunak S.[1]

[1] Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Building 208, DK-2800 Lyngby, Denmark

Exposure to environmental chemicals and drugs may have a negative effect on human health. An essential step towards understanding the effect of small molecules on human health is to identify all possible molecular targets of a given chemical. Recently, various network-oriented chemical pharmacology approaches have been published. However, these methods limit the protein prediction to already known molecular drug targets. New findings can for example be made by using high-confidence protein-protein association databases.

We present a generic, computational systems biology model with the aim of understanding the underlying molecular mechanisms of chemicals and the biological pathways they perturb. We generated a novel and complementary approach to existing models by integrating toxicogenomics data, protein-protein interaction data, disease information and functional annotation of proteins. The core of our procedure is derived from the "target hopping" concept defined previously [1]. But instead of considering only binding activity, we extended the concept to gene expression. If two proteins are affected with two chemicals, then both proteins are deemed associating in chemical space. Our approach is not only a statistical model but mimics the true biological system by constructing a network of associations between human proteins defined as Protein-Protein Association Network (P-PAN) (Figure 1). We have validated our network by comparison with two high confidence protein-protein interaction (PPI) networks [2-3], and by assessing the functional enrichment of clusters in the network generated.

The high confidence P-PAN proposed reveals unexpected connections between chemicals and diseases or human proteins. This illustrates the usefulness of an approach that integrates toxicogenomics data with other diverse data types.

Figure 1: View of the high confidence human P-PAN.

1. Paolini GV et al (2006) Global mapping of pharmacological space. *Nat Biotechnol.* 24: 805-15

2. Lage K et al (2007) A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat. Biotechnol.* 25: 309-16

3. Rual JF et al (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437:1173-8

POSTER PRESENTATION I-104

COMBINATORIAL QSAR MODELING OF TOXICITY DATA USING 2D & 3D CHEMICAL DESCRIPTORS

Antreas Afantitis[a], Georgia Melagraki[b], Haralambos Sarimveis[b]
Liying Zhang[c], Hao Zhuc, Alex Tropsha[c]

[a] Department of ChemoInformatics, NovaMechanics Ltd, Nicosia, Cyprus
[b] School of Chemical Engineering, National Technical University of Athens, Athens, Greece
[c] Laboratory for Molecular Modeling, Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina at Chapel Hill, North Carolina, USA

The traditional approaches for in vivo animal chemical safety testing are costly, time consuming, and have a low throughput. Accurate prediction of the adverse effects of chemical substances on living systems, identification of possible toxic alerts, and compound prioritization for animal testing are the primary goals of computational toxicology. Rapid expansion of experimental data sets that combine data on chemical structure and various toxicity end points for numerous environmental agents (e.g Berkeley Carcinogenic Potency Database and Distributed Structure-Searchable Toxicity database (DSSTox) of U.S. Environmental Protection Agency) provides novel opportunities to explore the relationships between chemical structure and toxicity using cheminformatics approaches.

We have developed and validated predictive combinatorial toxicity QSAR models in order to determine what reliable predictors of potential toxic effects of chemicals can be derived using 2D and 3D chemical structure information. More specifically we have assembled approximately 2,000 Estrogen Receptor (ER) ligands in a single database compiled from public sources including data on relative binding affinity of compounds to ERα and/or ERβ. The database was curated (e.g., by removing unreliable data, duplicates and heavy metal compounds). The final dataset consists of 437 unique ERα and 113 unique ERβ chemicals respectively. Several QSAR modeling approaches including k-nearest neighbors (kNN), support vector machines (SVM), RBF Neural Networks, least squares support vector machines (LS-SVM) were combined with a pool of 2D and 3D descriptors (Dragon, MOPAC, TOPIX & Mold2) with the aim to develop ERα and ERβ binding affinity models. The predictive power of each of the produced models, assessed by various validation tests (external test set, cross-validation (LOO & L5O), Y-randomization etc), is discussed individually and in comparison. The successful models were interpreted in terms of the physical meaning of model-selected descriptors and their influence on ER binding in relation to the structural characteristics of the compounds. The domain of applicability defined the area of reliable predictions for each methodology applied. The internally validated and externally predictive models derived in this study could be used as alternative and reliable computational filters for the evaluation of chemical toxicity and prioritization of new compounds for experimental *in vivo* testing.

EUROPEAN FOOD SAFETY AUTHORITY DEVELOPMENTS IN THE USE OF IN SILICO TOOLS FOR THE TOXICOLOGICAL METABOLISM HAZARD AND RISK ASSESSMENT OF PLANT PROTECTION PRODUCTS

Miriam Jacobs, David Carlander, Daniele Court-Marques, Frederique Istace, Daniela Maurici, Luc Mohimont, Istvan Sebestyen, Hans Steinkellner, Manuela Tiramani
European Food Safety Authority, Parma, Italy

The metabolism of plant protection products (PPPs) is a key determinant of any related toxicological effects. An understanding of species differences in PPP metabolism and the ability to predict or to study (to compare and contrast) directly the metabolic fate of a PPP in animals and in humans will be of great value in the risk assessment of PPPs. Therefore, the elucidation of pesticide metabolism pathways plays a critical role in PPP risk assessment, particularly in identifying the residue definition in plants, livestock and drinking water. These metabolism pathways or 'maps' aid in the determination of the particular metabolites and/or breakdown products which should be included in a dietary exposure/risk assessment and/or MRL expression for foods and livestock feeds. In order to perform these tasks, regulatory authorities must document and review large volumes of data on metabolism and environmental transformation products (in ground water and surface water) for PPPs.

In practise, from the mixture (active substance, its metabolites, and degradates) to which the consumer is exposed, only the toxicological properties of the active substance are directly investigated through the range of toxicological studies required by Directive 91/414/EEC. In contrast, very limited information about the toxicological properties of metabolites and degradates is available in the majority of cases, while requests for further toxicological studies are restricted as far as possible to minimise the use of animals in toxicological testing.

In order to ensure consistency and robustness of expert judgement, EFSA has started a review of relevant hazard and risk assessment tools regarding their aptitude to be used optimally in evaluating the toxicological burden of metabolites and degradates of pesticides. This review will be used to develop and propose a conceptual evaluation framework which will combine these tools for setting the residue definition for risk assessment under the decision making process of plant protection products.

Additionally, to expedite the review of metabolism data, EFSA is also collaborating with regulatory authorities internationally and is supporting a project proposal to the OECD to develop the US EPA initiative of a new computational tool, MetaPath. Metapath is a relational and searchable database which also includes an embedded set of data evaluation tools. MetaPath will allow efficient and systematic comparison of metabolites across chemicals, species, and environmental matrices to identify and quantify exposure to metabolites of potential concern. Having a collectively populated database will also allow for data sharing nationally and internationally and will aid in the development of new in silico predictive methods for prospective and retrospective analyses.

Further areas of computational tool developments in relation to PPPs at EFSA include the review of expert systems for endocrine disrupting properties, as well as Thresholds of Toxicological Concern (TTC) concept tools. A TTC initiative from the EFSA Scientific Committee will be reported in complementary presentations.

CADASTER MODELS FOR BROMINATED FLAME RETARDANTS

Ester Papa, Simona Kovarich, Paola Gramatica
QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Dep. of Structural and Functional Biology, Univ. of Insubria, , Varese, Italy

The EU-REACH regulation encourages the use of alternative *in vitro* and *in silico* methods in order to minimize animal testing, costs and time. In this context the use of Quantitative Structure Activity Relationships (QSAR) becomes particularly useful to predict unknown activities/properties for existing or even not yet synthesized chemicals. The development and validation of QSAR models for four classes of emerging pollutants (brominated flame retardants, fragrances, perfluorinated compounds and (benzo)triazoles) is the central topic of Work Package 3 (WP3) within the FP7 European project CADASTER[1] (CAse studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment). The final goal of the project is to exemplify the integration of information, models and strategies for carrying out hazard and risk assessments for large numbers of substances, organized in the four representative chemical classes.

In this study are presented the QSAR models that were developed for Brominated Flame Retardants (BFRs) during the first year of the project. Briefly, QSPR models were developed for some SIDS physico-chemical properties[2] (i.e. Henry's low constant, vapour pressure, water solubility, $LogK_{OW}$, $LogK_{OA}$, photodegradation rate) and then compared with publicly available EPI Suite models.
In addition, endocrine disrupting activities of BFRs[3, 4] (i.e. interaction with Aryl hydrocarbon receptor, Estrogen receptor, Progesterone receptor, Androgen receptor, T4-TTR competition and E2SULT inhibition) were modelled by both regression (MLR)[5] and classification (K-NN) methods.
All the QSAR models were developed taking into account the OECD principles for validation, for regulatory purposes, of QSARs[6]. This implied internal and external validations, the analysis of the applicability domain and, when possible, a mechanistic interpretation of the models.

Financial support, given by European Union through the project CADASTER (FP7-ENV-2007-1-212668), is gratefully acknowledged.

References:

1. Cadaster FP7 Project:
http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_LANG=EN&PJ_RCN=10374303;
www.cadaster.eu
2. E. Papa, S. Kovarich, P. Gramatica. *Development, Validation and Inspection of the Applicability Domain of QSPR Models for Physicochemical Properties of Polybrominated Diphenyl Ethers*. QSAR Comb. Sci. 2009, 28 (8), 790-796.
3. T. Hamers, J. H. Kamstra, E. Sonneveld, A. J. Murk, M. H. A. Kester, P. L. Andersson, J. Legler, A. Brouwer. *In vitro profiling of the endocrine-disrupting potency of brominated flame retardants*. Toxicol. Sci. 2006, 92, 157-173.
4. T. Hamers, J. H. Kamstra, E. Sonneveld, A. J. Murk, T. J. Visser, M. J. M. Van Velzen, A. Brouwer, A. Bergman. *Biotransformation of brominated flame retardants into potentially endocrine-disrupting metabolites, with special attention to 2,2',4,4'-tetrabromodiphenyl ether (BDE-47)*. Mol. Nutr. Food Res. 2008, 52, 284-298.
5. E. Papa, S. Kovarich, P. Gramatica, QSAR *modeling and prediction of the endocrine disrupting potencies of brominated flame retardants*. Accepted for publication in Chemical Research in Toxicology (April 2010).
6. OECD Quantitative Structure-Activity Relationships [(Q)SARs] Project Available online at:
http://www.oecd.org/document/23/0,3343,fr_2649_34365_33957015_1_1_1_1,00.html

POSTER PRESENTATION I-107

EVALUATION OF THE FACTORS THAT INFLUENCE THE PRECISION ON THE ESTIMATION OF BIOLOGICAL PROPERTIES BY MEANS OF PHYSICOCHEMICAL SYSTEMS

Martí Rosés, Marta Hidalgo, Elisabet Fuguet, Clara Ràfols
Universitat de Barcelona, Departament de Química Analítica,
Martí i Franquès 1-11, 08028 Barcelona, Spain

Experimental determination of biological properties, including toxicological and environmental ones, usually involves very expensive, time consuming, and difficult procedures because of the complexity of biological samples. Moreover, some of these procedures are even ethically questionable.
Since experimental determination of physicochemical properties is normally cheaper, faster, and easier, there is a lot of interest in developing physicochemical systems that emulate biological processes. The aim of the present work is being able to predict biological properties of chemical compounds through physicochemical measurements.
The performance of physicochemical systems to mimic biological ones is evaluated in terms of the precision that can be achieved. The variance obtained when biological properties are correlated against physicochemical ones is mainly contribution of three factors: the precision of the biological data, the precision of the physicochemical data, and the dissimilarity between the two correlated systems.
The Abraham solvation parameter model[1] gives a good approach to characterize both biological and physicochemical processes, and thus the precision of the analyzed data and the similarity between different characterized systems. In order to compare systems characterized through this model, the d distance[2] is a good parameter to know their similarity.
In this work, retention factors (log k) of some physicochemical systems of micellar electrokinetic chromatography (MEKC) have been correlated against data of several interesting biopartitioning systems, previously characterized by means of the Abraham solvation parameter model. The contributions of the three factors to the overall variance of the correlations have been evaluated from their characterizations and the distances between them. The goodness of the considered MEKC systems to predict particular biological properties, mainly partition and adsorption constants and toxicities, has been tested according to this method.

References:

[1]Abraham, M. H. *Chem. Soc. Rev.* **1993**, 22, 73-83.
[2]Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. *J. Med. Chem.* **2006**, 49, 4861-70.

PERFORMANCE OF PHYSICOCHEMICAL SYSTEMS TO MODEL SOIL-WATER SORPTION

Clara Ràfols, Marta Hidalgo, Elisabet Fuguet, Martí Rosés
Universitat de Barcelona, Departament de Química Analítica,
Martí i Franquès 1-11, 08028 Barcelona, Spain

The uptake of organic compounds by soil from water is a process of great importance to understand the mobility and fate of organic contaminants in ground water systems. However, very difficult, expensive and time consuming tasks are required to determine experimentally the soil-water sorption of organic chemicals. For this reason, development of estimation methods to predict this environmental property is really interesting.

Since the determination of physicochemical parameters is fast, cheap, and not so complex, the estimation of soil-water sorption by means of physicochemical measurements is highly desirable. In order to achieve this aim, physicochemical systems that mimic the sorption of organic compounds by soil from water are necessary.

Abraham developed the solvation parameter model[1] to characterize systems based on the distribution of solutes between two phases. Among the processes that have been described through this model, both biological and physicochemical ones, soil-water sorption[2] is one of them. Since the solvation parameter model has also been applied to characterize many chromatographic systems, it is possible to compare them with the soil-water sorption system by means of their characterizations. In this respect, the d parameter[3] is very useful to measure the mathematical similarity between pairs of systems characterized through the Abraham model.

This work is focussed on the study of the similarity between the soil-water sorption system and several physicochemical systems of liquid chromatography (HPLC) and micellar electrokinetic chromatography (MEKC). The d distance has been calculated between each pair of soil-water sorption and chromatographic systems. Then, the pairs of systems with lower d have been selected, since the lower d is, the more similar the two compared systems are. After measuring retention factors (log *k*) in the selected chromatographic systems, these have been correlated against soil-water sorption data (log K_{oc}). In order to determine which particular chromatographic system best model the sorption of organic chemicals by soil from water, the correlations have been evaluated according to the precision of the original data and the dissimilarity between the correlated systems.

References:

[1]Abraham, M. H. *Chem. Soc. Rev.* **1993**, 22, 73-83.
[2]Poole, S. K., Poole, C. F., *J. Chromatogr. A* **1999**, 845, 381-400.
[3]Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. *J. Med. Chem*. **2006**, 49, 4861-70.

POSTER PRESENTATION I-109

MODELING FATHEAD MINNOW FISH TOXICITY BY MEANS OF PHYSICOCHEMICAL SYSTEMS

Marta Hidalgo, Elisabet Fuguet, Clara Ràfols, Martí Rosés
Universitat de Barcelona, Departament de Química Analítica,
Martí i Franquès 1-11, 08028 Barcelona, Spain

Determination of aquatic toxicity is useful to assess the adverse environmental effects caused by the release of chemicals to waterways. Among the aquatic organisms, the fathead minnow fish (*Pimephales promelas*) is one of the species outlined by EPA guidelines[1] as a biological model in aquatic toxicology studies.
Although several standard test methods and protocols have been established in order to measure aquatic toxicity, experimental determinations are often very expensive, time consuming, and involve difficult procedures. Because of these experimental complications, predictive models are highly desirable.
Since the measurement of physicochemical parameters is cheaper, faster, and easier, physicochemical systems are thought to provide a good approach to estimate aquatic toxicity.
The solvation parameter model[2], which was developed by Abraham, has successfully described many interesting biological processes as well as physicochemical ones, ruled by the transfer of solutes between two phases. Thus, the Abraham solvation parameter model has been used to characterize the aquatic toxicity of organic compounds to fathead minnow[3].
In this work, several physicochemical systems have been compared with the aquatic toxicity for fathead minnow system, by means of their Abraham characterizations. The d distance[4] has been the parameter used to measure the mathematical similarity between each pair of systems. The lower the d parameter is, the more similar the two compared systems are. Accordingly, we have selected the physicochemical systems more similar to the aquatic toxicity for fathead minnow system. Then, the corresponding physicochemical properties, mainly retention factors (log k) in liquid chromatography (HPLC) and micellar electrokinetic chromatography (MEKC) systems have been determined and correlated with the toxicity data for the fathead minnow fish (-log LC_{50}). The performance of the correlated physicochemical systems to model the aquatic toxicity for fathead minnow has been evaluated through the correlations and taking into account the precision of the biological data, the precision of the physicochemical data and the dissimilarity between the correlated systems.

References:

[1]http://www.epa.gov/ncct/dsstox/sdf_epafhm.html
[2]Abraham, M. H. *Chem. Soc. Rev.* **1993**, 22, 73-83.
[3]Hoover, K. R.; Acree, W. E., Jr.; Abraham, M. H. *Chem. Res. Toxicol*. **2005**, 18, 1497-1505.
[4]Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. *J. Med. Chem*. **2006**, 49, 4861-70.

QSAR MODELING OF RODENT ACUTE TOXICITY ON THE BASIS OF PASS PREDICTIONS

Lagunin A., Zakharov A., Filimonov D., Poroikov V.
Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences,
Pogodinskaya Str., 10, Moscow, 119121, Russia

Prediction of acute rodent toxicity is very important in drug design and environmental risk assessments. About 200,000 different chemicals are produced commercially and consumed every year, and 2000 - 3000 new chemicals are added to this list annually. Current technologies can not provide a possibility to study all synthesized compounds on toxic and side effects. Complete toxicological data are available for less than 10,000 of them. The experimental study of toxic and side effects is time-consuming and highly expensive. Along with the ethical aspects, such experiments cause criticism owing to the insufficient validity of extrapolation of obtained data on human. In silico prediction of toxic and site effects is considered an alternative of experimental animal tests. The new European Society initiative (REACH), that has been started at 2007, anticipates development of computer-aided methods for analysis of "structure-activity" relationships (e.g. IUCLID) and study of toxic effects for 30,000 chemical substances.

Following this trend we have developed a new method for QSAR modelling that is based on PASS (Prediction of Activity Spectra for Substances) prediction results [1] and self-consistent regression [2]. The idea of the method is the use of the probabilities of randomly selected activities from the predicted biological activity spectra of compounds (4150 activities in PASS 10.1) as independent input variables for self-consistent regression. The number of the selected activities depends on the number of compounds in the training set, which was equal to the half of the number of compounds in the training sets at the study. Twenty QSAR models are created on the basis of this algorithm for the training set. Thus, the consensus results of 20 QSAR models are used for prediction of activity for compounds from the test set. Moreover, the correction of prediction results on the basis of the values of the nearest neighborhoods was used for improving of the prediction accuracy. Predictions for the test set are performed taking into account the applicability domain for each model.

The proposed method was evaluated on the set of compounds tested on acute rat toxicity by oral route of administration (7286 compounds) used for the test of known QSAR methods in article of Zhu and co-authors [3] and several sets of compounds tested on acute rat/mouse toxicity by different routes of administration selected from SYMYX MDL Toxicity Database. The analysis of prediction results for the test sets showed that the proposed method displays the satisfactory accuracy of prediction ($R^2_{test}>0.6$) and high computation performance.

Acknowledgements:
This work was supported by the ISTC grant 3777.

References:

[1] Filimonov D.A., Poroikov V.V. Probabilistic approach in activity prediction. In: *Chemoinformatics Approaches to Virtual Screening.* Eds. Alexandre Varnek and Alexander Tropsha. 2008, Cambridge (UK): RSC Publishing, 182-216; http://www.ibmc.msk.ru/PASS.

[2] Lagunin A., Zakharov A., Filimonov D., Poroikov V. A new approach to QSAR modelling of acute toxicity, SAR and QSAR in Environmental Research, 2007, 18, 285-298.

[3] Zhu H, Martin TM, Ye L, Sedykh A, Young DM, Tropsha A. Quantitative structure-activity relationship modeling of rat acute toxicity by oral exposure. Chem Res Toxicol. 2009, 22(12),1913-1921.

A NEW CLASSIFICATION METHOD SUITABLE FOR TOXICITY SCREENING OF CHEMICAL COMPOUNDS

Kohtaro Yuta
In-silico data Ltd.

A new outstanding classification method which is specially designed for applying on the toxicity screening research is proposed and discussed on this paper. This new method realized to simplify calculation process and algorithm of previously proposed KY-method (K-step Yard sampling method) by Yuta [1].
With K-step Yard sampling method (KY-method), a set of Ames test samples including 6965 compounds has been classified into two classes (Positive / Negative) correctly (100%) [1]. Applying this KY-method, we can classify any sample set into two classes perfectly regardless of sample size and distribution.
There are big differences between this proposed new method and the KY-method on procedures of calculation. But basic concepts of those two methods are absolutely same. Therefore, no differences on classification and prediction power between the KY-method and this newly proposed method.
KY-method has big advantages for toxicity screening, but this approach need to take some complex operations. For example KY-method uses two discriminant functions at a time which have opposite classification feature each others.
The new proposed method on this paper make better and overcome those operational disadvantages of KY-method. For example, only one discriminant function is used on this new method. All samples are classified into three areas ((1)correctly classified into the class1, (2)correctly classified into the class2, (3)grey region (correctly classified and misclassified samples are mixed)), based on the value of discriminant score (fig.1 and fig.2). Samples classified into the grey region are reclassified by using new parameter set. These processes are repeated till no-samples are fall into the grey region.

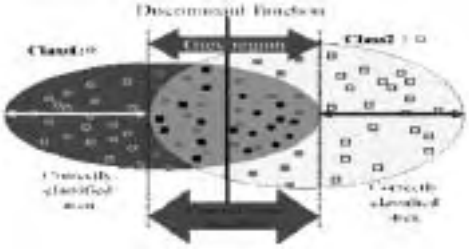


Figure1 Classification image

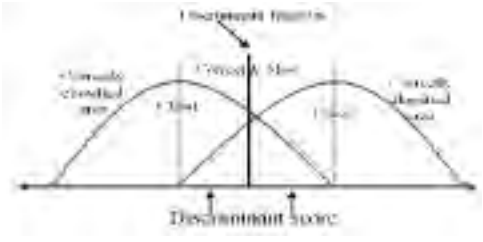


Figure2 Classification by discriminant score

Reference:

[1] Kohtaro Yuta, Development of the "K-step Yard sampling method" and Apply to the ADME-T In Silico Screening, The 34-th Symposium on Structure-Activity Relationships, Niigata, 2006.

POSTER PRESENTATION I-112

THE TOXTREE SOFTWARE FOR IMPLEMENTING THE CRAMER CLASSIFICATION SCHEME FOR CHEMICAL TOXICITY ASSESSMENT

Silvia Lapenna and Andrew Worth
Systems Toxicology Unit, Institute for Health and Consumer Protection, European Commission - Joint Research Centre, Ispra (Va), Italy

Every day man is exposed to thousands of low molecular weight, organic compounds, both naturally occurring and man-made, which, depending on the level of intake, may pose a risk to health. Currently, extensive animal testing is carried out to assess the safety of chemicals applied in food, consumer products and medicines, or their contaminants.
The Threshold of Toxicological Concern (TTC) is a concept that refers to the establishment of a generic oral exposure level for chemicals below which there would be no appreciable risk to human health and which can be identified for most chemicals based on their chemical structure [1]. The TTC principle may be useful to prioritise or even waive chemical testing based on estimation of human exposure, and has been incorporated into some risk assessment processes in the area of food additives and food contact materials. It has also been explored in relation to cosmetic ingredients, personal and household care products and pharmaceutical impurities. The most used toxicity classification according to structure is the decision tree by Cramer et al. [2], which has been encoded into the freely available, open-source software Toxtree [3], developed by Ideaconsult under the terms of a European Commission Joint research Centre contract.
The use of the Toxtree software to facilitate chemical classification according to the Cramer scheme is practical and is becoming widely used among scientists and regulators. In this work the current capabilities of the Toxtree software are presented and critically assessed in terms of correct implementation of the Cramer scheme. Furthermore, the possible strengths and limitations of the Cramer scheme for assessing the toxicity of chemicals are discussed, with an emphasis on possible ways to improve the current set of structural rules.

References:

[1] Munro IC, Ford RA, Kennepohl E & Sprenger JG (1996). Correlation of structural class with No-Observed-Effect Levels: A proposal for establishing a threshold of concern, Food Chem. Toxicol. 34, 829–867.
[2] Cramer GM, Ford RA & Hall RL (1978). Estimation of Toxic Hazard - A Decision Tree Approach. J. Cosmet. Toxicol. 16, 255-276.
[3] JRC Toxtree software: http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE

POSTER PRESENTATION I-113

STRUCTURE - ECOTOXICITY RELATIONSHIPS OF CHEMICAL SEDIMENT DATA USING QSAR AND CHEMOMETRIC APPROACHES

Ivanka Tsakovska [a], Ilza Pajeva [a], Stefan Tsakovski [b], Vasil Simeonov [b]
[a] Centre of Biomedical Engineering, Bulgarian Academy of Sciences, Sofia, Bulgaria
[b] Faculty of Chemistry, Sofia University "St. Kl. Ohridski", Sofia, Bulgaria

The assessment of the effects of pollution on water ecosystems requires bottom-sediment samples to be monitored. These effects are extremely complex since they result from the presence of a mixture of various pollutants. Further various factors as structural characteristics of the wide spectrum of compounds included in the mixtures, their concentration levels and aquatic species specificity should be considered. Thus, the search for relationships between ecotoxicity endpoints and particular structural parameters of the chemicals presented in the mixtures of pollutants may be a source of valuable information. However, analysis of such complex data is rather complicated. Chemometric and other methods of intelligent data analysis could be a suitable tool to achieve this goal [1].

In our previous study the relationship between ecotoxicity parameters (acute and chronic toxicity) and chemical components (polluting chemicals like polychlorinated biphenyls (PCBs), pesticides, polycyclic aromatic hydrocarbons (PAH), heavy metals) of lake sediments samples from Turawa Lake, Poland was analysed by application of self-organizing maps (SOMs). From the SOMs obtained, it was possible to select groups of similar ecotoxicity (either acute or chronic) and to analyse within each one of them the relationship between the chemical concentrations and the toxicity endpoints (EC50 and mortality). The study has shown, convincingly, that different regions from the Turawa Lake bottom indicate different patterns of ecotoxicity related to various chemical pollutants, such as the “heptachlor-B” pattern, “pesticide and PAH” pattern, “structural” pattern (heavy metals as a major component) or “PCB congeners” pattern [2].

The aim of the present study is to get a deeper insight into the relationships between the sediment toxicity and polluting chemicals in the Turava lake sediment dataset previously analysed [2]. In particular, QSAR and chemometric approaches are used to clarify the importance of the different groups of chemicals within a given cluster of sediment samples with similar toxicity for the toxicity profile of the cluster. Since the scope of application of the QSAR approaches is limited to organic compounds, from the initial chemicals dataset 28 compounds are selected and further investigated. First, a number of structural parameters is calculated using the MOE molecular modelling software [http://www.chemcomp.com/]. They describe the geometry, hydrophicity and electronic properties of the structures – features that could be significant for the aquatic toxicity of the chemicals. Further cluster analysis is performed based on the calculated structural properties. “PCB congeners” and “PAH” categories are selected. The third category comprises a number of organochlorine pesticides. For each category the "aquatic" toxicological profile is predicted using the ECOSAR models of ECOWIN program [http://www.epa.gov/oppt/exposure/pubs/episuite.htm] for acute and chronic toxicity to fish, daphnia and green algae. In order to identify the possible mechanisms of aquatic toxicity, the Verhaar classification scheme as implemented in TOXTREE program [http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE] is applied. Among the structures covered by the Verhaar rules (46% from the total) the most are classified as a class 1 (non-polar narcotics). Five of the chemicals (aldrine, dieldrine, endrine, heptachlor and its metabolite heptachlor epoxide isomer B), however, are classified as a class 3 (reactive chemicals). In the next step of the study the identified chemical categories and the toxicological QSAR predictions are related to the SOM patterns as proposed

within the sediment dataset [2]. For that purpose, the SOM clusters are re-developed after exclusion of the heavy metals from the initial dataset. Within the obtained new patterns the chemical categories as formed by the QSAR analysis are identified and their influence on the toxicity of the pattern is discussed based on the predicted toxicological profile of the particular category.
The proposed multistep approach is an example for intelligent data treatment of complex sediment data, going from the more general SOM based classification approach to the more detailed, based on the chemical structures QSAR approach. The general purpose is to help identification of major sources of threats caused by mixtures of numerous pollutants present in the sediments.

References:

1. V. Simeonov, L. Wolska, A. Kuczyńska, J. Gurwin, S. Tsakovski, M. Protasowicki, J. Namieśnik. Sediment-quality assessment by intelligent data analysis, *Trends in Analytical Chemistry,* 26(4), 323-331.
2. S. Tsakovski, B. Kudlak, V. Simeonov, L. Wolska, J. Namiesnik. Ecotoxicity and chemical sediment data classification by the use of self-organising maps, *Analytica Chimica Acta*, 631(2), 142-152.

Acknowledgement: The authors acknowledge the financial support from the National Science fund of Bulgaria (grant DTK02/58).

POSTER PRESENTATION I-114

IN SILICO STUDIES OF THE MULTIDRUG RESISTANCE TRANSPORTER P-GLYCOPROTEIN

Ilza Pajeva [a], Christoph Globisch [b]&, Michael Wiese [b]
[a] Centre of Biomedical Engineering, Bulgarian Academy of Sciences, Sofia, Bulgaria
[b] Pharmaceutical Institute, University of Bonn, 53121 Bonn, Germany

P-glycoprotein (P-gp) belongs to the family of ABC transporters that use energy from ATP hydrolysis to actively extrude xenobiotics out of cells. The protein is involved in multidrug resistance (MDR) in tumor cells and in the absorption, distribution, metabolism, and excretion of drugs. The investigations of P-gp by computational methods have developed over the years depending on the data available for modeling. Recently the 3D structure of mouse P-gp has been reported in the apo- and drug-bound conformations. These are the most appropriate data available so far for structure-based modeling of this transporter.
In this work results are presented obtained from a combined structure- and ligand-based approach to model P-gp and its interactions with ligands. A homology model of human P-gp is developed that corresponds to the inward-facing (open to the cytosole) conformation competent for drug binding. From the model the residues involved in the protein binding cavity are identified and compared to those in the outward-facing conformation of human P-gp generated previously. The analysis of the cavities in both models suggests that the ligands remain bound to the same residues during the transition from the inward- to outward-facing conformation. Next, a detailed analysis of the interactions of the cyclic peptides QZ59-RRR and QZ59-SSS is performed in both, the X-ray structures of mouse P-gp and the human P-gp model. The analysis of the ligand–protein interactions in the X-ray complexes shows differences in the residues involved as well as in the specific interactions performed by the same ligand and this observation has also been confirmed by docking of the QZ59 ligands into the human P-gp model. The same model has been further used to dock selective P-gp inhibitors (representatives of the 3rd generation MDR modulators) from the group of quinazolinones (Xenova Ltd.) into the binding cavity of the protein. The results reveal the interactions of particular functional groups and atoms in the compound structures with particular protein residues thus helping in elucidation of the pharmacophore pattern of the MDR modulators and the differences in their effects. Finally, the possible location of the Hoechst-binding site of the protein is investigated by pharmacophore modeling and docking of the P-gp substrate Hoechst 33342.
The results of the combined modeling approach complement each other and could help for better understanding of the protein-ligand interactions and development of highly selective and potent P-gp inhibitors.

Acknowledgement: I.P. and M.W. gratefully acknowledge the generous support by the Alexander von Humboldt Foundation, Germany and the financial support from the National Science Fund of Bulgaria (grant No. DTK02/58)

QSAR MODELS OF ACUTE INTRAVENOUS TOXICITY TO MICE ON THE BASIS OF INTERSPECIES CORRELATIONS, LIPOPHILICITY PARAMETERS AND PHYSICOCHEMICAL DESCRIPTORS

O.A. Raevsky [a] *, Y.V.Liplavskiy[a] , O.E.Raevskaya [a] J.Devillers [b] and A.Worth [c]

[a] Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds of Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia; [b]CTIS, 3 Chemin de la Graviere, 69140 Rillieux La Pape, France; [c] Institute for Health and Consumer Protection, TP 582, European Commission - Joint Research Centre,Via Enrico Fermi 2749, 21027 Ispra (Va), Italy

Five independent QSAR approaches were used for the calculation of acute intravenous toxicity to mice of 127 diverse chemicals with the aim of developing approaches that may serve to reduce the amount of vertebrate animal testing in the assessment of acute toxicity by the intravenous route:

Correlations between chemical acute toxicity to mice (intravenous injection; LD_{50}) and aquatic toxicity to *Tetrahymena pyriformis* (LC_{50}):

$\log(1/LD_{50})_{exper\_mi\_iv}=-0.11(\pm0.04)+0.47(\pm0.03)\log(1/LC_{50})_{exper\_ThP}$

$n=127$; $R^2=0.621$; $Q^2=0.606$; $SD=0.45$; $SD_{cv}=0.46$; $F=204.9$

$\log(1/LD_{50})_{exper\_mi\_iv}=0.05(\pm0.05)+0.46(\pm0.03)\log(1/LC_{50})_{exper\_ThP}-0.101(\pm0.021)[\log(1/LC_{50})_{exper\_ThP}]^2$

$n=127$; $R^2=0.682$; $Q^2=0.667$; $SD=0.42$; $SD_{cv}=0.43$; $F=132.8$

2. QSAR for acute toxicity to mice on the basis of lipophilicity parameters [1-3]:

$\log(1/LD_{50})_{exper\_mi\_iv}=-0.90(\pm0.09)+0.37(\pm0.04)\log P_{o/w(exper)}$

$n=85$; $R^2=0.500$; $Q^2=0.467$; $SD=0.53$; $SD_{cv}=0.54$; $F=82.9$

$\log(1/LD_{50})_{xper\_mi\_iv}=-0.92(\pm0.08)+0.59(\pm0.08)\log P_{o/w(exper)}-0.075(\pm0.025)[\log P_{o/w(exper)}]^2$

$n=85$; $R^2=0.550$; $Q^2=0.512$; $SD=0.50$; $SD_{cv}=0.52$; $F=50.1$

$\log(1/LD_{50})_{exper\_mi\_iv}=-0.97(\pm0.08)+0.46(\pm0.05)\log P_{o/w(exper}-1.34(\pm0.40)\log[0.000241P_{o/w(exper}+1]$

$n=85$; $R^2=0.559$; $Q^2=0.529$; $SD=0.50$; $SD_{cv}=0.51$; $F=52.0$

3. QSAR on the basis of HYBOT descriptors [4]:

$\log(1/LD_{50})_{exper\_mi\_iv}=-1.43(\pm0.19)+0.098(\pm0.011)\alpha-0.02(\pm0.06)\sum Ca(o)$

$n=127$; $R^2=0.395$; $Q^2=0.346$; $SD=0.57$; $SD_{cv}=0.60$; $F=40.4$

$\log(1/LD_{50})_{exper\_mi\_iv}=-2.37(\pm0.38)+0.343(\pm0.041)\alpha-0.31(\pm0.17)\sum Ca(o)-$

$0082(\pm0.0013)\alpha^2+0.06(\pm0.03)[\sum Ca(o)]^2$

$n=127$; $R^2=0.554$; $Q^2=0.541$; $SD=0.50$; $SD_{cv}=0.51$; $F=37.8$

4. Application of Arithmetic Mean Toxicity Model (AMT model) [5]:

$\log(1/LD_{50})_{exper\_mi\_iv}=-0.17(\pm0.05)+1.00(\pm0.10)\log(1/LD_{50})_{mi\_iv\_calc\_AMT}$

$n=116$; $R^2=0.466$; $Q^2=0.448$; $SD=0.50$; $SD_{cv}=0.50$; $F=99.4$

5.Application of Super overlapping clusterization/regression (SOCR) QSAR model:

$\log(1/LD_{50})_{exper\_mi\_iv}=-0.33(\pm0.05)+1,22(\pm0.13)\log(1/LD_{50})_{mi\_iv\_calc\_SOCR}$

$n=96$; $R^2=0.468$; $Q^2=0.440$; $SD=0.50$; $SD_{cv}=0.51$; $F=82.7$

A few combinations of QSAR models (parabolic interspecies correlation and AMT; parabolic interspecies correlation, parabolic model based on HYBOT descriptors and AMT) improved the results of toxicity prediction compared with the application of any one approach:

$\log(1/LD_{50})_{exper\_mi\_iv}=-0.07(\pm0.04)+1,17(\pm0.07)\log(1/LD_{50})_{calc\ ((eq.1)+AMT)/2}$

n=127; R^2=0.703; Q^2=0.693; SD=0.40; SD_{cv}=0.41; F=295.8

$\log(1/LD_{50})_{exper\_mi\_iv}$=-0.03(±0.04)+1,23(±0.07)$\log(1/LD_{50})_{calc\ ((eq.1)+eq(3)+AMT))/3\_}$

n=127; R^2=0.709; Q^2=0.700; SD=0.39; SD_{cv}=0.40; F=305.1

References

[1] J.W. McFarland, *On the Parabolic Relationship between Drug Potency and Hydrophobicity,* J. Med. Chem. 13 (1970), pp. 1192-1186;[2] H. Kubinyi, *Quantitative Structure-Activity Relationships, 7. The Bilinear Model, a New Model for Nonlinear Dependence of Biological Activity on Hydrophobic Character,* J. Med. Chem. 20 (1977), 625-629: [3] S.Bintein, J.Devillers, W.Karcher, Nonlinear dependence of fish bioconcentration on n-octanol/water partition coefficient, *SAR QSAR i Environ.Res.* 1993, 1, pp.29-39. [4] O.A.Raevsky, *Physicochemical Descriptors in Property-Based Drug Desig* , Mini-Reviews in Medicinal Chemistry, 4(2004), pp.1041-1052; [5] O.A. Raevsky, V.Yu.Grigor'ev, E.A. Modina and A.Worth, Prediction of Acute Toxicity to Mice by the Arithmetic Mean Toxicity (AMT) Modelling Approach, *SAR QSAR h Environ. Res.,* **2010**, 21, *№* 3/4.


18th EURO QSAR

POSTER SESSION II

DEVELOPMENT OF NEW SOFTWARE IN DRUG DESIGN AND APPLICATION OF ELECTRON CONFORMATIONAL-GENETIC ALGORITHM METHOD TO 1,4- DIHYDROPYRIDINE DERIVATIVES

Emin Sarıpınar*, Nazmiye Geçen, Kader Şahin, Ersin Yanmaz
Erciyes University, Science Faculty, Department of Chemistry, 38039 Kayseri, Turkey

The electron conformational-genetic algorithm method has been employed as a 4D-QSAR [1] approach to reveal the pharmacophore (Pha) and to predict biological activities in the 1,4- dihydropyridine derivatives [2]. Hence we present a comprehensive pharmacophore identification, molecular descriptor and activity calculation program package (EMRE V2.0 and ECSP 1.0) which runs on personal computers. These programmes have been developed mainly for computer-aided drug design using the methods of three-dimensional quantitative structure–activity relationships. We report here the results of pharmacophore (Pha) identification and quantitative bioactivity prediction for the class of 1,4- dihydropyridine derivatives using the EC-GA method. In the first part, the EC matrices of congruity (ECMCs) [3, 4] are constructed from data of conformational analysis and electronic structure calculation of each of the molecules in the compound series by EMRE V2.0 programme. Later, comparing ECMC of one of the most active compounds with other ECMCs by means of a special program of matrix comparisons (ECSP 1.0), we obtain features (pharmacophore) responsible for the activity as submatrices of the template that are called Electron Conformational Submatrices of Activity (ECSA). This is the EC submatrix of activity that represent the Pha, while the tolerances characterize the Pha flexibilities. The molecular descriptors of the compounds were performed with EMRE V2.0 which is prepared by us and these descriptors are used in Matlab 7.0 programme for calculating the activity. Over 250 molecular descriptors are available for predicting biological activity. To predict the theoretical activity of training and test set and to select important variables for describing the activities, genetic algorithm and nonlinear least square regression methods were applied. Statistical analysis was performed to calculate regression coefficient.

References:

1. Duca J.S., Hopfinger A.J., J Chem Inf Comput Sci, 41, 1367-1387, 2001.
2. Safarpour M.A., Hemmateenejad B., Miri R. and Jamali M., SAR Comb.Sci., 22, 997-1005, 2003.
3. Bersuker I.B., Current Pharmaceutical Design, 9, 1575-1606, 2003.
4. Bersuker I.B., J Comp Aid Mol Des, 22, 423-430, 2008.

QSAR PREDICTION OF D^2 RECEPTOR ANTAGONISTIC ACTIVITY OF 6-METHOXY BENZAMIDES

Mohammad. H. Fatemi*, Fereshteh Dorostkar
Laboratory of Chemometrics, Faculty of Chemistry,
University of Mazandaran, Babolsar, Iran

In the present work, quantitative structure–activity relationship (QSAR) method was used to predict the pIC_{50} value of 58 derivatives of 6-methoxy benzamides. In the first step, the data set was split to the training, internal and external test sets, that each of them has 46, six and six members, respectively, by y-ranking methods. The distributions of molecules in these sets were examined by molecular diversity test. After the calculation of molecular descriptors, the stepwise multiple linear regression method was employed to screening descriptors spaces. The selected parameters are: Geary autocorrelation-lag 5 / weighted by atomic Sanderson electronegativities (GATS5e), electrotoplogical state value of R-CH_2-R group ($SssCH_2$), electrotoplogical state value of R-NH-R (SssNH) non aromatic group, total molecular 2-center exchange energy/number of atoms (TEMM2), hydrophobicity of the substituent at R_3 position (π-R_3) and resonance effect at R5 position ($\Re$-R_5). Then artificial neural network (ANN) and multiple linear regressions (MLR) were used to construct the nonlinear and linear quantitative structure-activity relationship models, respectively. The standard errors in the prediction of pIC_{50} by MLR model are; 0.280, 0.446 and 0.382 and for ANN model are 0.175, 0.326 and 0.296 for training, internal and external test sets, respectively. Also these models were further examined by cross-validation methods which produce the statistics of Q^2 = 0.8340 and SPRESS= 0.322 for MLR model and Q^2 = 0.8055 and SPRESS= 0.219 for ANN model. Inspection to obtained results revealed the applicability of QSAR model using ANN techniques in prediction of antagonistic activities of 6-methoxy benzamides from their theoretical derived molecular descriptors.

S. Deswal, N. Roy, Eur. J. Med. Chem. 42 (2007) 463-470.
S. Samanta, B. Debnath, Sh. Gayen, G. Ghosh, A. Basu, K. Srikanth, T. Jha, Il Farmaco, 60 (2005) 818-825.

MPO INHIBITORS SELECTED BY VIRTUAL SCREENING

Alberto Malvezzi[1,2] Raphael F Queiroz[1] Leandro de Rezende[1,2] Ohara Augusto[1] Antonia T-do Amaral

[1]Departamento de Bioquímica and [2]Departamento de Química Fundamental, Instituto de Química, Universidade de São Paulo, Brazil

The heme peroxidase enzyme myeloperoxidase (MPO) participates in the innate immune response by generating microbicidal reactive oxidants such as hypochlorous acid, hypobromous acid and hypothiocyanous acid from halides and pseudohalide (SCN-) ions. These oxidants have been implicated as key mediators of tissue damage in many human inflammatory diseases including atherosclerosis, asthma, rheumatoid arthritis, cystic fibrosis and some cancers[1]. Until now there are no known MPO inhibitors available in the clinic and there is considerable interest in the development of MPO inhibitors that could be therapeutically useful[1,2]. Currently there is thorough information on the structure of MPO[3,4], as well as on its reaction mechanisms and redox intermediates[1], which enables the development of specific MPO inhibitors that could potentially prevent the pathophysiological burden caused by the enzyme[2].
In this work we have applied SBDD and LBDD strategies[5] for the selection of new, potent and selective MPO inhibitors. Initially, a pharmacophore model5 of this enzyme was developed based on the structure of the inhibitor salicylhydroxamic acid bound to the enzyme active site[4]. This pharmacophore model comprised of two hydrogen acceptor regions, one hydrogen donor region and one hydrophobic region. These regions matched with the positions of conserved water molecules crystallized in the active site[4,3]. Additionally, the pharmacophore model also included several exclusion volumes corresponding to the positions of all the heavy atoms of the residues comprising the active site. The final pharmacophore model was used to screen the ZINC database[6] for potential MPO ligands, using the Catalyst[7] program, which resulted in the selection of 1051 compounds. The selected compounds were further filtered based on their physical-chemical properties. Cut-off values of MW=300 and calculated logP = 5 were used to exclude large and lipophilic compounds. The 208 compounds that passed these filtering steps were further submitted to a docking inside the active site of MPO. This docking procedure was performed using the default parameters implemented in the program Gold[8]. Only 54 compounds were docked keeping their H-donor, H-acceptor and hydrophobic groups in the corresponding positions as predicted by the pharmacophore model. These top docked compounds were visually inspected considering the overall quality of the binding interactions and 9 compounds were finally selected and purchased. Surprisingly, all selected compounds were aromatic hydrazides, which are known to be MPO inhibitors[9]. The 9 purchased compounds were then submitted to the taurine chloramine assay which measures the inhibition of hypochlorous acid production by the MPO enzyme[10]. For screening purposes, this assay was, initially, performed on a plate. IC_{50} values, were determined for the top 4 most promising compounds and ranged from 0.97 to 2.78 μM.

Support: FAPESP; CAPES, INCT de Processos Redox em Biomedicina Redoxoma (CNPq/FAPESP/MCT).

1 Malle, E., Furtmuller, P.G., Sattler, W., e Obinger, C., *Brit. J. Pharmacol*., 152, 838, **2007**.
2 Rees, M.D., Bottle, S.E., Fairfull, K.E., Malle, E., Whitelock, J.M., e Davies, M.J., *Biochem J*., 421, 79, **2009**.
3 Fiedler, T.J., Davey, C.A., e Fenna, R.E., *J. Bio. Chem.*, 275, 11964, **2000**.
4 Davey, C.A. e Fenna, R.E., *Biochemistry*, 35, 10967, **1996**.
5 Malvezzi, A., de Rezende, L., e do Amaral, A.T., *Qsar & Comb. Sci.*, **2009**.
6 Irwin, J.J. e Shoichet, B.K., *J. Chem. Inf. Model.*, 45, 177, **2005**.
7 Catalyst, © Accelrys Software Inc. San Diego, CA, USA, **2005**.
8 Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W., Taylor, R.D., *Prot-Struc Func Gen*, 52, 609, **2003**.
9 Burner, U., Obinger, C., Paumann, M., Furtmuller, P.G., e Kettle, A.J., *J. Biol. Chem.*, 274, 9494, **1999**.
10 Kettle, A.J. e Winterbourn, C.C., *Oxy. Radi Bio. Syst., Pt C*, 233, 502, **1994**

IN SILICO BINDING STUDIES OF NOVEL QUINOLINONE 3-CARBOXAMIDES: AN EFFORT TO ELUCIDATE THEIR ANTI-INFLAMMATORY PROFILE

M. Zervou[1*], P. Zoumpoulakis[1*], C. Potamitis[1], K. C. Prousis[2], A. Detsi[2]

[1]Laboratory of Molecular Analysis, Institute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation, 48, Vas. Constantinou Ave., 11635, Athens, Greece.

[2]National Technical University of Athens, School of Chemical Engineering, Department of Chemical Sciences, Laboratory of Organic Chemistry, Zografou Campus, 15773, Athens, Greece

A chemical library consisting of 21 structurally diverse novel quinolinone-3- carboxamides was designed, synthesized and evaluated in vitro for their anti-inflammatory activity. Nine potent soybean lipoxygenase inhibitors were identified, exhibiting IC50 in the range of 10-83μM.

In silico docking of these compounds was attempted in an effort to elucidate the mechanism underlying their LOX inhibitory activity and more specifically the crucial interactions between enzyme amino acids and the quinolinone analogues. Several different binding modes were produced due to the conformational flexibility of the small ligands compared to the dimensions of the active site.

The ability of the studied molecules to inhibit LOX enzyme is mainly attributed to their coordination with Fe ion, responsible for the oxidation of fatty acids. Docking results revealed the participation of the C=O group in the metal coordination sphere, retaining a somewhat distorted trigonal bipyramid geometry. Furthermore, a cis conformation for the amide bond results in the formation of H-bonding with backbone carboxylate of Ile857 which may stabilize the binding complex.

Methyl and benzyl substitutions at the N-position of the heterocycle ring increase the lipophilicity of the ligand and orient the substituents towards Val571.

Para substitution of the aromatic ring, adjacent to the amide bond, with electronegative groups results in the formation of hydrogen bonds with different amino acids like Leu560 or Thr274, causing subsequent increase of the binding score.

Overall, the obtained results reveal the ability of the active compounds to efficiently mimic the bending pattern of the native fatty acid substrate (9Z,11E)-13(S)-hydroperoxy-9,11-octadecadienoic acid, 13-HPOD), perfectly superimposing its carboxamide C=O with the peroxy group of the native ligand.

Representative docking pose of N-phenyl quinolinone-3-carboxamide analog superimposed with the native fatty acid substrate in the binding site of Soy Lox-3 enzyme (pdb:1ik3). The iron coordination sphere is illustrated

MOLECULAR DOCKING STUDIES OF INHIBITORS CONTAINING ACTIVATED CARBONYL FUNCTIONALITIES IN THE ACTIVE SITE OF CYTOSOLIC PHOSPHOLIPASE A_2

V. Michopoulou, V. Mouchlis, T. Mavromoustakos, G. Kokotos
Laboratory of Organic Chemistry, Department of Chemistry, University of Athens, Panepistimiopolis, 15771 Athens, Greece

Group IVA cytosolic phospholipase A2 (GIVA cPLA$_2$) is a member of the superfamily of phospholipase A2 enzymes that selectively catalyses the hydrolysis at the *sn*-2 position of arachidonyl-glycerophospholips, releasing arachidonic acid and lysophospholipids. Arachidonic acid is metabolised to a variety of inflammatory mediators such as prostaglandins and leukotrienes. Since GIVA cPLA$_2$ plays an important role in several inflammatory diseases, the design of new anti-inflammatory drugs is considered to be a promising field of research. In the present study, we used the molecular model of the cPLA$_2$ complexed with inhibitor AX007 synthesized in our laboratory and recently derived by the combination of molecular dynamics simulations and deuterium exchange mass spectrometry[1].
Several known inhibitors of GIVA cPLA$_2$ possessing diverse biological activity and containing different activated carbonyl functionalities[2] were chosen in order to study their interactions with the active site of the enzyme using the Surflex-Dock of Sybyl 8.0 algorithm. Important stereoelectronic features responsible for anti-inflammatory activity have been revealed.

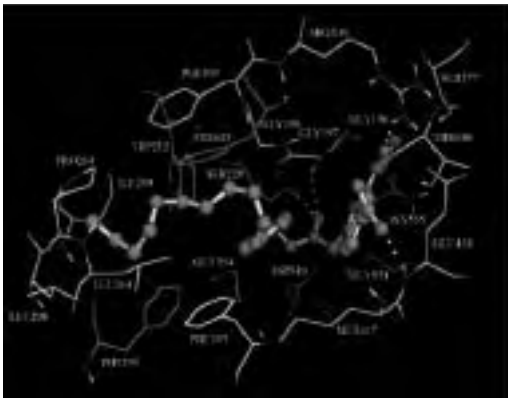


Figure 1: The high binding score of the most active in vitro inhibitor, a 1-indol-1-yl-propan-2-one derivative, is attributed to its orientation in the cleft of the active site. Its alkyl chain is surrounded by a cluster of lipophilic amino acids (TRP232, ILE299, LEU264, PHE683), its activated carbonyl group is hydrogen bonded with SER228 and GLY197, its carboxylic group with GLU418 and its methyl ester group with ARG200.

References:

1. Burke, J. E.; Babakhani, A.; Gorfe, A. A.; Kokotos, G.; Li, S.; Woods, V. L., Jr.; McCammon, J. A.; Dennis, E. A., *J Am Chem Soc* **2009**, 131, 8083-91.
2. Lehr M., *Anti-Inflamm. Anti-Allergy Agents Med. Chem* **2006**, 5, 149-161.

MOLECULAR DOCKING AND 3D-QSAR COMFA STUDIES ON INDOLE INHIBITORS OF GIIA SECRETED PHOSPHOLIPASE A_2

Varnavas D. Mouchlis, Thomas M. Mavromoustakos, George Kokotos
Laboratory of Organic Chemistry, Department of Chemistry, University of Athens, Panepistimiopolis, Athens 15771, Greece

Phospholipases A_2 (PLA_2) are a superfamily of enzymes which are characterised by the their ability to catalyse the hydrolysis of the *sn*-2 position of phospholipids molecules.[1] The products of PLA_2 activity include free fatty acids, predominantly arachidonic acid (AA), and lysophospholipids. The AA is further metabolised by downstream enzymes (COX-1, COX-2 and 5-LO) to form a variety of pro-inflammatory lipid mediators including prostaglandins, leukotrienes and thromboxanes. The bioactive lysophospholipids are converted to the platelet activating factor (PAF). The $sPLA_2$ group of enzymes plays an important role in a variety of physiological functions. Among the members of the mammalian $sPLA_2$ enzymes, the GIIA $sPLA_2$ is an interesting anti-inflammatory drug target because it was found in high levels in the synovial fluid from rheumatoid arthritis patients. In addition, GIIA $sPLA_2$ is crystallised with or without a ligand bound in the active site and as a result is an attractive target in drug design using computational approaches.[2] In this study, an automated docking allowing a protein-based alignment was performed on a set of indole inhibitors[3] using GLIDE 5.5.[4] A correlation between the binding score and the experimental inhibitory activity was observed ($r^2 = 0.666$, $N = 34$). Thereafter, best score pose of each inhibitor was used for a protein-based alignment of the compounds' set. A three-dimensional QSAR model was then established using the CoMFA method.[5] The set of 34 indole inhibitors was divided to two subsets. The training set, composed of 26 compounds and the test set with eight compounds. The robustness and the predictive ability of the generated CoMFA model were examined by using the test set. A good correlation between the predicted and the experimental activity ($R^2=0.997$) confirmed the validation of the CoMFA model. Finally, the generated CoMFA model was used for the design and evaluation of new compounds. The new designed compounds exerting higher predicted inhibitory activity were then docked in the GIIA $sPLA_2$ active site in order to see how the structural modifications affect the binding score.

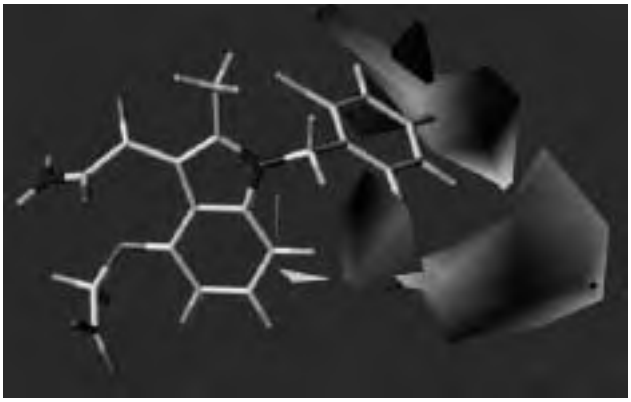


Figure 1: Generated CoMFA steric and electrostatic StDev*Coeff field contour map for the most active compound on the set of indole inhibitors. Bulky groups in the green region favor the inhibitory activity but bulky groups in the yellow region are not desirable for the inhibitory activity. Negative potentials in the red region favor the inhibitory activity but negative potential in the blue region are not desirable.

References:

1. Burke, J. E.; Dennis, E. A., *Cardiovasc Drugs Ther* **2009**, 23, 49-59.
2. Mouchlis, V. D.; Mavromoustakos, T. M.; Kokotos, G., *J Comput Aided Mol Des* **2010**, 24, 107-15.
3. Smart, B. P.; Pan, Y. H.; Weeks, A. K.; Bollinger, J. G.; Bahnson, B. J.; Gelb, M. H., *Bioorg Med Chem* **2004**, 12, 1737-49.
4. Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009.
5. SYBYL/QSAR and COMFA, molecular modeling software packages, version 8.0, 2007, Tripos Inc., 1699 South Hanley Rd., St. Louis, MO 63144-2917.

COMPARATIVE DOCKING STUDIES AT THE ACTIVE SITE OF LIPOXYGENASE-3, LIPOXYGENASE-5 AND CYCLOOXYGENASE-2

N. Neophytou[1], G. Michas[1], D. Hadjipavlou-Litina [2], M. Zheng[3], G. Kokotos[1], P. Moutevelis-Minakakis[1], T. Mavromoustakos[1]

[1]National Kapodistrian University of Athens, Department of Chemistry, Panepistimiopolis, Zographou 15784, Greece

[2]Aristotle University of Thessaloniki, School of Pharmacy, Department of Pharmaceutical Chemistry, Thessaloniki 54124, Greece

[3] Drug Discovery and Design Center (DDDC) Box 1201, Shanghai Institute of Materia Medica. No. 555 Rd. Zuchongzhi, Shanghai, China

Arachidonic acid is mainly metabolized by the two classes of enzymes lipoxygenases (LOX) and cyclooxygenases (COX). Lipoxygenase-5 is responsible for the formation of leukotrienes, lipid mediators involved in the symptoms of asthma and allergic disorders. Cyclooxygenase-2 is the inducible isoform of the enzyme during inflammation. Our laboratory is involved in producing molecules that equally block LOX-5 and COX-2.

MMK 16 (4-[(2S)-2-(1H-imidazol-1-ylmethyl)-5-oxotetrahydro-1H-pyrrol-1- yl]methylbenzenecarboxylic acid) is a synthetic pyrrolidinone derivative that was evaluated in soyabean enzyme LOX-3, and was found to have high inhibitory effect (0.08 mM versus 0.60 mM for caffeic acid). In vivo tests confirmed the anti-inflammatory activity of MMK 16 (41% Inhibition of induced carrageenin rat paw oedema against 47% for the standard Indomethacin) Since all the isoforms of the superfamily of lipoxygenases share analogous overall folding pattern, docking calculations have been performed for MMK16 at LOX-3 crystal structure [1] and compared to the results derived from the docking calculations at LOX-5 homology model [2] Further docking studies have been performed at the active site of COX-2 using two crystal structures and established the hypothesis that MMK 16 is a potential anti-inflammatory compound, that acts through the inhibition of LOX-5 and COX-2.

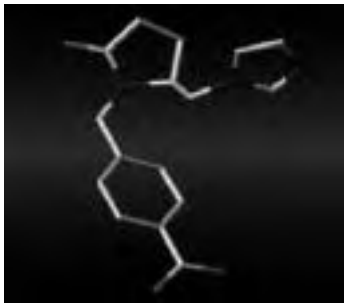


Figure 1:Low energy conformation of MMK 16

References

1. Skrzypezak-Jankun, E., Bross, A. R., Carroll, R.T., Dunham, W.R., Funk, O.M. J. Am. Chem. Soc. **2001**, 123, 10814-10820.

2. Du Li, Zhang Z., Luo, X., Chen, K., Shen, X. Hualiang, J. J. Biochem. **2006**, 139, 715-723

SIMPLE CLASSIFICATION APPROACH TO QSAR MODELLING OF PLANT POLYPHENOL ANTIOXIDANT ACTIVITY MEASURED IN STOICHIOMETRIC ASSAYS

Petko Alov, Ivanka Tsakovska, Ilza Pajeva

Centre of Biomedical Engineering, Bulgarian Academy of Science, 105 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria

Plant polyphenols, a group of ubiquitous natural compounds, are claimed to be responsible for many of the beneficial effects of fruits, vegetables, tea, vines, etc. Their biochemical and pharmacological effects are related to interactions with certain protein receptors, free radical scavenging in hydrophilic environments, mild prooxidant action leading to up-regulation of the enzymatic antioxidant defence, etc. It is shown that some of these polyphenols exert also cytostatic and multidrug resistance-reversal effects important for efficient anticancer therapy.

The antioxidant/antiradical effects of many natural and synthetic mono- and polyphenolic compounds are extensively studied by a number of assays, most of them characterized by stoichiometric endpoints. While the selection of molecular descriptors for modelling kinetic assays is quite straightforward, the proper selection of descriptors useful in predicting antioxidant/antiradical properties in stoichiometric assays is not clear yet. The main obstacle for obtaining reliable models in this case is related to the presence of multiple centres with similar reactivity in the molecules.

In this study antiradical capacities of a number of natural polyphenols in the stoichiometric assays FRAP (1), ABTS and DPPH (2) were studied by QSAR analyses. The descriptors used were the total number of OH-groups and the number of OH-groups classified as active in a particular assay by quantum-chemical parameters characterizing the O–H bond dissociation energy in the polyphenol molecules (difference between heats of formation of the molecule and its radicals, O–H bond length and partial charges in the phenolic group), and unpaired electron delocalization in the phenoxyl radicals (maximal spin density on a single carbon or oxygen atom). The quantum-chemical parameters were calculated semi-empirically using AM1, PM6 and RM1 parametrisations in MOPAC2009 (3) by both vertical and adiabatic geometry optimizations.

Classification of the OH-groups was done using a threshold calculated based on the sum of antioxidant phenolic equivalents (trolox equivalents, TEAC) in the particular group of compounds studied.

The number of OH-groups classified as active by the difference between heats of formation of the molecule and its radicals and by the maximal spin density on a single atom in the phenoxyl radicals correlated better with the antiradical capacity than the total number of OH-groups, especially for the subsample of compounds possessing a chromene moiety (flavonoids, isoflavonoids, coumarins, trolox).

The differences in the classifications with respect to the semi-empirical parametrisations and optimization approaches used, and the quality of the resulting univariate and multivariate QSARs is discussed.

Acknowledgement:

This work was partially supported by the National Science Fund of Bulgaria (Grant No DTK02/58).

References:

8) Firuzi, O.; Lacanna, A.; Petrucci, R.; Marrosu, G.; Saso, L. Biochim. Biophys. Acta 2005, 1721, 174-184.

9) Cai, Y.-Z.; Sun, M.; Xing, J.; Luo, Q.; Corke, H. Life Sci. 2006, 78, 2872-2888.

10) MOPAC2009, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, HTTP://OpenMOPAC.net (2008).

STRUCTURAL AND ELECTRONIC FEATURES RESPONSIBLE FOR ANTIAGGREGATORY EFFECT OF FLAVONOIDS

M. Medić-Šarić, M. Bojić, Ž. Debeljak

Faculty of Pharmacy and Biochemistry, University of Zagreb, A. Kovačića 1, 10000 Zagreb, Croatia

Leading causes of death in developed countries are cardiovascular diseases. Polyphenols, namely flavonoids and phenolic acids, have many pharmacological properties that can explain their protective role in the prevention of cardiovascular diseases: antioxidant capacity, vasorelaxation, anti-inflammatory and antiaggregatory effects. Benfitial effects of natural sources of polyphenolics on platelet aggregation souch as wine (Franch paradox), green tea (Asian paradox), dark chocolate, gingko, propolis are well documented. The goal of this work is to determine molecular descriptors responsible for antiaggregatory effect of flavonoids based on the set of thirty flavonoids.

The biological end point was defined as minimal concentration of flavonoid that can lower the platelet aggregation statistically compared to the negative control. Measurements were done on the whole blood using Multiplate® (multiple platelet functional analyzer) and adenosine diphosphate as weak agonist of aggregation (ADP test).

Molecular descriptors were generated using MOPAC 7 (Stewart Computational Chemistry, USA), ACD/ChemSketch Freeware (Advanced Chemistry Development, USA), ChemFileBrowser (Hyleos, USA) and TAM (Faculty of Pharmacy and Biochemistry, Croatia). These included: steric (molecular weight, volume, area), hydrophobic parameters (Hansch's parameter), molar refraction and electronic parameters (total energy, HOMO, LUMO, dipole moment, partial charges, bond orders). Statistical analysis was done in Statistica v7.1 while prediction, based on random forests and leave-one-out external validation, has been made within R v2.8.1 environment.

Preliminary results suggest Mulliken's partial atom charge and substituent charge as the most important molecular features for modeling of antiaggregatory effects of selected flavonoids. The determination of structural and electronic features responsible for antiaggregatory efficacy of flavonoids will provide basis for either synthesis of new flavonoids or selection of potent flavonoids from natural sources.

2D-QSAR STUDIES IN THE DESIGN OF NEW COMPOUNDS. THE EXAMPLE OF ANTITHROMBOTIC AGENTS

Kontogiorgis A. Christos[1,2] & Hadjipavlou – Litina Dimitra[1]

[1] Department of Pharmaceutical Chemistry, School of Pharmacy, Department of Pharmaceutical Chemistry, 54124, Thessaloniki, Greece,

[2] Present address: Department of Chemistry, Imperial College London, London SW7 2AZ

Cardiovascular diseases are one of the major causes of morbidity and mortality in the world. Platelets' activation and aggregation play a critical role in the pathophysiology of thrombotic disorders and this would be expected to modify the natural history of cardiovascular diseases [1,2]. Antiplatelet drugs have found critical application in the secondary prevention of vascular events, including acute myocardial infarction, stroke and cardiovascular death [3]. It has been observed that the currently used antiplatelet drugs are effective against certain but not all of the many endogenous platelet activators. A significant number of serious thromboembolic complications still occur, because of the limited efficacy of these drugs, thus new potent agents need to be designed.

In this study, Quantitative Structure Activity Relationships (QSARs) were developed for the following different sets of compounds: Coumarin derivatives (I) [4], Phenyl-ethyl-benzene derivatives (II) [5], Tetrahydronaphthalene derivatives (III) [6], Caffeic-acid amide analogues (IV) and Pyrrolo-benzylisoquinolinones (V) [8]. Lipophilicity and stereochemical parameters of substituents, such as Verloop's sterimol parameters, seem to highly influence the antiplatelet activity of these compounds. Molar refractivity seems also to play a crucial role. The above results would be quite useful in the design of new potent antiplatelet agents.

Acknowledgments.

Dr Kontogiorgis Christos is thankful to the Foundation for Education and European Culture, established in Athens for the financial support of this postdoctoral research. The authors are also thankful to Biobyte Corp. and Dr Hansch and Dr Leo for their support and free access to the C-QSAR program. This study has been performed using the above program via Internet, Biobyte Corp., 201 West 4th Street, Suite 204, Claremont, Pomona College, USA.

References:

[1]. Weston, C.; Rao, U. Int. J. Clin. Practice 2003, 57, 898

[2]. Clutton, P.; Folts, J.D.; Freedman, J.E. Pharmacol. Res.,2001, 44, 255.

[3]. Ault, K.A.; Cannon, C.P.; Mitchell, J.; McCahan, J.; Tracy, R.P.; Novothy, W.F.; Reinmann, J.D.; Braunwald E. J. Am. Coll. Cardiol. 1999,33,634-9.

[4]. Roma, G.; Di Braccio, M.; Grossi, G.; Piras, D.; Leoncini, G.; Bruzzese, D.; Signorello, M.G.; Fossa, P.; Mosti, L. J. Med Chem. 2007, 50, 2886-95.

[5]. Kikumoto, R.; Hara, H.; Ninomiya, K.; Osakabe, M.; Sugano, M.; Fukami, H.; Tamao, Y. J. Med. Chem. 1990, 33, 1818-23.

[6]. Cimetière, B.; Dubuffet, T.; Muller, O.; Descombes, J.J.; Simonet, S.; Laubie, M.; Verbeuren, T.J.; Lavielle, G. Bioorg. Med. Chem. Lett. 1998, 8, 1375-80.

[7]. Hung, C.C.; Tsai, W.J.; Kuo, L.M.; Kuo, Y.H. Bioorg. Med. Chem. 2005, 13, 1791-7

[8]. Kuo, R.Y.; Wu, C.C.; Chang, FR, Yeh JL, Chen IJ, Wu YC.Bioorg Med Chem Lett. 2003,13(5):821-3.

POSTER PRESENTATION II-11

QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIPS (QSARs) OF HISTAMINE H_4 RECEPTOR ANTAGONISTS

Pontiki Eleni[1, 2] and Hadjipavlou-Litina Dimitra[1]

[1] Department of Pharmaceutical Chemistry, School of Pharmacy, Aristotelian University of Thessaloniki, Thessaloniki, 54124, Greece

[2] Present address: UCL Department of Chemistry, University College London Gower Street London WC1E 6BT, UK

Histamine, a low molecular weight amine, is an important chemical mediator in physiological and pathological responses through interaction with four G-protein coupled receptors: H_1, H_2, H_3 and H_4. There is grown body of evidence that H_4 receptor plays a role in immune and inflammatory responses and modulates pruritus-itch responses as well. It has a distinct expression profile on immune cells including mast cells, eosinophils, dendritic cells, and T cells[1] and has modulatory effects on their function, as activation, migration, cytokine and chemokine production. Currently, the H_4 receptor is considered as a promising target for the treatment of various chronic inflammatory diseases such as inflammatory bowel disease, asthma and rheumatoid arthritis[2]. H_4 receptor shares low sequence homology with other histamine receptors, especially H_1 and H_2, so agonists or antagonists specific for H_1 or H_2 do not bind to H_4. The H_4 receptor has highest homology with the H_3, and accordingly, is bound by some H_3 ligands.

The development of potent and selective H_4 receptor ligands is of utmost importance as it will provide molecular tools to further characterize the H_4 receptor and to explore the therapeutic potential of H_4 related drugs. From Terzioglu et al.[3] a novel series of indole piperazines with enhanced potency and selectivity have been synthesized. The compounds have been studied for the inhibition of human histamine H_4 receptor. Also, Sander, K. et al.[4] have synthesized 2,4-diaminopyrimidines as H_4 receptor ligands and tested them for their affinity in histamine displacement assay with membrane prapatration of Sf9 cells expressing hH_4R, co-expressed with $G\alpha i_2$ and $G\beta_1\gamma_2$ subunits. Finally, Cowart, M. et al.,[5] have synthesized 2,4-diamino-5,6-disubstituted pyrimidines and biologically evaluated them to human H_4 binding potency.

In the present study, these three different groups of H_4 antagonists (indole piperazines, 2,4-diaminopyrimidines and 2,4-diamino-5,6-disubstituted pyrimidines) have been subjected to a QSAR study, in order to provide a simple description of the physicochemical parameters, implicated in the H4 antagonists using the C-QSAR suite of programs (Biobyte).

It seems that the lipophilicity of the whole molecule as clogP, steric parameters as molar refractivity of the substituents and B_5 of Verloop and indicator variables are significant for the biological response. The results of this study have been exploited in order to be used in the synthesis of new compounds with the proper characteristics for the antagonism of H_4 receptor.

Acknowledgements:

To Drs. C. Hansch and A. Leo for help and support, as well as Biobyte Corp., 201 West 4th St., Suite 204, Claremont CA 91711, USA, for free access to C-QSAR program.

Eleni Pontiki is grateful to the "Onassis Foundation" for scholarship for postdoctoral research at the Chemistry Department of University College of London.

References

[1] Zhang, M., Thurmond, R., Dunford, P., *Pharmacol Ther*., **2007,** 113, 594-606.

[2] Tiligada, E., Zampeli, E., Sander, K., & Stark, H., *Expert Opin. Investig. Drugs*, **2009**, 18(10), 1519-1531.

[3] Terzioglu, N., van Rijn, R.M., Bakker, R.A., De Esch, I.J,, Leurs, R. *Bioorg Med Chem Lett*. **2004**, 14(21), 5251-5256.

[4] Sander, K., Kottke, T., Tanrikulu, Y., Proschak, E., Weizel, L., Schneider, E.H., Seifert, R., Schneider, G., Stark, H. *Bioorg Med Chem,* **2009**, 17(20), 7186-7196.
[5] Cowart, M., Altenbach, R., Liu, H., Hsieh, G., Drizin, I., Milicic, I., Miller, T., Witte, G., Wishart, N., Fix-Stenzel, S., McPherson, M., Adair, R., Wetter, J., Bettencourt, B., Marsh, K., Sullivan, J., Honore, P., Esbenshade, T. and Brioni, J. *J. Med. Chem*. **2008**, 51, 6547–6557.

AN INVESTIGATION OF PROTEINS INVOLVED IN AUTOIMMUNE ARTHRITIS BY DESIGN AND BIOLOGICAL EVALUATION OF CHEMICAL PROBES

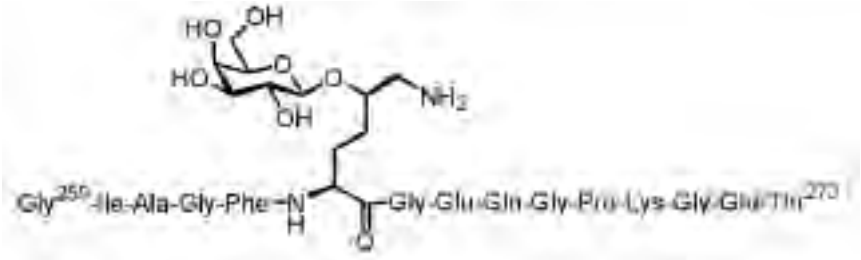
C. David Andersson,[1*] Ida E. Andersson,[1] Tsvetelina Batsalova,[2] Balik Dzhambazov,[2] Rikard Holmdahl,[2] Jan Kihlberg,[1,3] and Anna Linusson[1]

[1]Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden.

[2]Medical Inflammation Research, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

[3]AstraZeneca R&D Mölndal, SE-431 83 Mölndal, Sweden.

Here we present a methodology to design glycopeptide chemical probes that were used to investigate protein function associated with autoimmune arthritis in a chemical approach. Glycopeptide **1** is an inducer of collagen-induced arthritis (CIA) in mice. Nevertheless, vaccination of mice with glycopeptide **1** in complex with major histocompatibility complex (MHC) class II A^q protein can prevent the development of CIA and also relieve symptoms in mice who are suffering from chronic arthritis. This gives hope that human rheumatoid arthritis (RA) can be treated by immunization with glycopeptides. By combining structure-based design, ligand-based design and subsequent synthesis, a library of chemical probes was created including analogues of glycopeptide **1**, modified at amino acid positions crucial for anchoring to the A^q protein. The chemical probes were biologically evaluated for binding to two class II MHC proteins (A^q and human DR4) and for T-cell recognition of eight different hybridomas. The study revealed that the chemical probes acted differently on the two proteins depending on the amino acid patterns and the panel of T-cell hybridomas displayed diversity in responses to the probes. The creation of chemical probes presented here is one step towards the goal of developing a therapeutic vaccine for treatment of RA.



1

TARGETING ACETYLCHOLINESTERASE; THE EFFECT OF WATER AND PROTEIN FLEXIBILITY ON POSE PREDICTION AND AFFINITY ESTIMATION IN MOLECULAR DOCKING SIMULATIONS

Lotta Edvinsson,[1*] David Andersson,[1] Elisabet Artursson,[2] Fredrik Ekström,[2] Anna-Karin Tunemalm,[2] Anna Linusson[1]

[1]Department of Chemistry, Umeå University, Umeå, Sweden

[2]Swedish Defence Research Agency, CBRN Defence and Security, Umeå, Sweden

Acetylcholinesterase (AChE) is an essential enzyme involved in termination of neurotransmission by hydrolyzing acetylcholine. Nerve agents, a class of highly toxic organophosphorous compounds, irreversibly inhibit AChE by covalently binding to a catalytic serine residue in the active site. We are combining computational methods with organic synthesis to design novel molecules with improved potency as reactivators compared to the nerve agent antidotes available today. The attempted strategy is to design nucleophiles that are capable of removing the phosphorous conjugate from the serine residue, starting from molecules that inhibit AChE by binding to the peripheral anionic site. In an ongoing study, we have developed a docking protocol aimed at distinguishing AChE inhibitors from similar inactive compounds. In the study, a multivariate approach was used where multiple protein models were included to account for protein flexibility and the presence of water in the binding site. The protocol can be used to identify lead compounds as well as to further guide the design by predicting the binding mode of antidote candidates.

A MOLECULAR DYNAMICS STUDY OF A SUB-NANOMOLAR DUAL BINDING SITE HETERODIMERIC AChE INHIBITOR

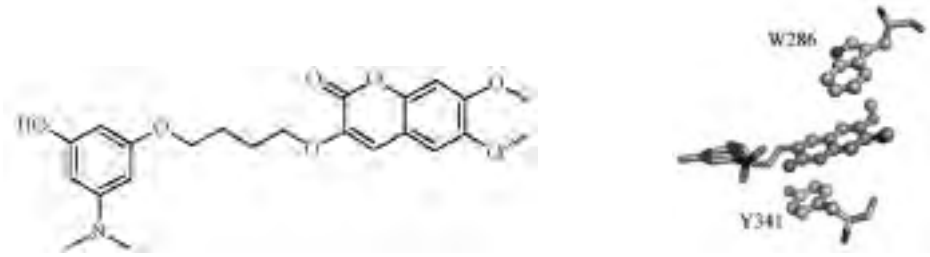
Nicolotti Orazio, Giangreco Ilenia, Pisani Leonardo, Catto Marco, Leonetti Francesco, Stefanachi Angela, Cellamare Saverio and Carotti Angelo

Dipartimento Farmaco-Chimico, Università di Bari

Via Orabona 4, 70125, Bari, Italy

Inhibition of acetylcholinesterase (AChE) represents a successful strategy for the symptomatic treatment of Alzheimer's disease. [1] In this regard, a brick-based approach was recently [2] developed to design and prepare a series of AChE inhibitors assembling via a proper sized linker two weak AChE bioactive fragments, i.e., an edrophonium-like and a coumarin moiety addressing the catalytic and peripheral AChE binding sites, respectively.

A small series of congeners, exhibiting an inhibitory potency in the low and sub-nanomolar range, was thus prepared [3]. Among them, the 6,7-dimethoxy-3-alkoxysubstituted coumarin derivative 1, shown in Figure 1A, exhibited both outstanding affinity (IC_{50} = 0.23 nM) and an excellent AChE/BChE selectivity (SI = 309000).



A B

Figure 1. A. 6,7-dimethoxy-3-alkoxysubstituted coumarin derivative 1. B. Representative sandwich-like binding conformation of 1 taken after 5 ns MD simulations.

The reason of such impressive potency and selectivity was nicely interpreted via molecular dynamics (MD) executed by using AMBER 10.0 suite with a 5 ns MD simulation performed through NAMD on human AChE (hAChE) isoform (code 1B41). Interestingly, the sampling of compound 1 demonstrated the occurrence of stable sandwich-like interaction between 6,7-dimethoxy-3-substituted coumarin ring (miming the 5,6-dimethoxy-1-indanone-moiety of donepezil) with the Trp286 and Tyr341 reported in Figure 1B. The removal of the two dimethoxy groups decreased the activity (IC_{50}=143 nM) for the disruption of the sandwich-like interaction in benefit of T-shaped orthogonal π-π interaction with Trp286 or, alternatively, in a parallel π-π interaction with Tyr341.

1. Abraham, D. J. Anticholinergic Drugs. Burger's Medicinal Chemistry, vol. 6, chap. 4.2. 6th Ed Wiley: New York, 2003.
2. Leonetti, F.; Catto, M.; Nicolotti, O.; Pisani, L.; Cappa, A.; Stefanachi, A.; Carotti, A. Bioorg. Med. Chem. 2008, 16, 7450–7456.
3. Pisani, L.; Catto, M.; Giangreco, I.; Leonetti, F.; Nicolotti, O.; Stefanachi, A.; Cellamare, S.;Carotti, A. Accepted for publication in *ChemMedChem*

4-ARYL-4-OXO-N-PHENYL-2-AMINYLBUTYRAMIDES AS NOVEL CLASS OF ACETYLCHOLINESTERASE INHIBITORS*

Maja D. Vitorovic-Todorović,[a,b*] Ivan O. Juranić,[b] Ljuba M. Mandić,[b] Branko J. Drakulić [c]

[a]Military-Technical Institute, Ratka Resanovića 1, Belgrade, Serbia; [b]Faculty of Chemistry, University of Belgrade, Studentski trg 12-16, Belgrade, Serbia; [c]Department of Chemistry-IChTM, University of Belgrade, Njegoševa 12, Belgrade, Serbia

The twenty 4-aryl-4-oxo-N-phenyl-2-aminylbutyramides (Fig. 1) were designed and synthesized. Their inhibition of acetylcholinesterase activity (AChE, EC 3.1.1.7) was tested using Ellman's spetrophotometric assay. Eight derivatives showed inhibition activity in low-micromolar range. The presence of voluminous alkyl substituents (4-*i*-Pr, 2,4-di-*i*-Pr and *β-tetralinoyl*) on the aroyl moiety of piperidino- and imidazolo-derivatives changes anti-AChE activity for two orders of magnitude in respect to unsubstituted congeners. Replacement of piperidino methylene group with oxygen in morpholino derivatives results with complete loss of anti-AChE activity. The most potent piperidino- and imidazolo- derivative showed mixed inhibition type, indicating their binding to free enzyme and enzyme-substrate complex. Alignment-independent 3D QSAR study, based on molecular interaction fields (MIF's) was performed on the set containing reported compounds and the ones taken from the literature. Totally 38 compounds were collected, including secolycorines [1], 1-[bis(4-fluorophenyl)-methyl]piperazines [2], litebamine derivatives [3], imidazolyl-isoxazolines [4], 3-aryl-N-methyl-tetrahydropyridine derivatives [5], spanning range of potencies 0.8 - 240 μM ($\sim$ 2.5 log(1/(IC_{50})) units). All compounds were submitted to Pentacle, and for the model building HBD (N1), HBA (O) and hydrophobic (DRY) probe were used. Obtained model has good statistics and predictivity (R_{acc}^2 = 0.91, Q_{acc}^2 (RG) = 0.54). The analysis confirmed that alkyl substitution on the aroyl moiety of molecules is requisite for inhibition activity. The presence of hydrophobic moiety at a close distance from hydrogen bond acceptor has favorable influence on inhibition potency. The presence of hydrogen bond acceptor and hydrogen bond donor on spatial distance of $\sim$ 12.5 Å yields less potent compounds. To investigate possible ligand-AChE interactions, docking studies were performed on electric eel AChE to generate binding model for the most active 2,4-di-*i*-Pr imidazolo derivative, using AutoDock 4.0.1. Docking studies showed that both enantiomers of the compound bind probably in the middle of the active site gorge of AChE, interacting with anionic site of the AChE (Trp 86), acyl pocket (Phe 295 and Phe 297) and peripheral anionic site (Tyr 72, Tyr 124).

Figure1. 4-Aryl-4-oxo-N-phenyl-2-aminylbutyramides, X = -CH2- or -O-;. R=H, 2,5-di-Me, 4-Cl, 4-MeO, 4-i-Pr, 2,4-di-i-Pr and β-tetralinoyl.

References:

[1] Lee, S-S.; Venkatesham, U.; Prasad Rao, C.P.; Lam, S-H.; Lin, J-H. *Bioorg. Med. Chem*. **2007**, 15, 1034; [2] Sadashiva, C. T.; Chandra, J. N. N. S.; Ponnappa, K. C.; Gowda, T. V.; Rangappa, K. S. *Bioorg. Med. Chem. Lett.* **2006**, 16, 3932; [3] Chiou, C-M.; Kang, J-J.; Lee, S-S. *J. Nat. Prod.* **1998**, 61, 46; [4] Rangappa, K. S. B. *J. Phys. Org. Chem.* **2005**, 18, 773; [5] Benaka Prasad, S. B.; Sunil Kumar, Y. C.; Ananda Kumar, C. S.; Sadashiva, C. T.; Vinaya, K.; Rangappa, K. S. *The Open Med. Chem.* J., **2007**, 1, 4.

* Reproduced in part from *Bioorg. Med. Chem.* **2010**, 18, 1181–1193.

MACHINE LEARNING APPROACHES TO THE SELECTIVITY PREDICTION FOR LIGANDS OF BENZODIAZEPINE BINDING SITE OF $GABA_A$ RECEPTORS

Vladimir I. Chupakhin[1], Maria N. Kurilo[2], Vladimir A. Palyulin[1,2], Nikolay S. Zefirov[1,2]

[1]Institute of Physiologically Active Compounds, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432, Russia

[2]Department of Chemistry, Moscow State University, Leninskie gory 1/3, Moscow, 119991, Russia

$GABA_A$ receptors are the main inhibitory receptors of the central nervous system. Five subunits form the ion channel selective to the chloride ions. There are 6 subtypes of α_{1-6} subunits, 3 subtypes of β_{1-3} and γ_{1-3} subunits, other subunits do not have subtypes (δ, ε, π, θ). The main subtype of $GABA_A$ receptor in the human brain is $\alpha_1\beta_2\gamma_2$ where α and β subunits are duplicated. The constitution of the receptor depends on the localization in brain regions, age, gender and other parameters. The ligands selective to the $GABA_A$ receptors containing α_1 subunit have sedative and hypnotic activity, but also cause amnesia and other side effects; the ligands selective to $\alpha_{2,3}$ subunits have anxiolytic activity without sedation; the ligands selective to α_5 subunit are cognition enhancers[1]. Current drug design of the modulators of $GABA_A$ receptors is targeted to discovery of highly selective positive and negative modulators of benzodiazepine binding site. We applied several QSAR approaches to the prediction of selectivity towards $\alpha_1\beta_3\gamma_2$, $\alpha_3\beta_3\gamma_2$, and $\alpha_5\beta_3\gamma_2$ $GABA_A$ receptor subtypes for the ligands of the benzodiazepine binding site (549 compounds) based on fragmental, pharmacophoric, and 3D descriptors using various machine learning techniques (ANN, PCA, SVM, Random forests). The results were compared with ligand-receptor interactions evaluated in the docking studies to the $GABA_A$ receptor homology models.[2] It was shown that ligand-receptor interactions alone do not allow one to find all structural features needed for the design of highly selective ligands while the QSAR models provide the additional necessary information.

References:

1. Qinmi Wang, Yifan Han, Hong Xue. Ligands of the GABAA Receptor Benzodiazepine Binding Site. *CNS Drug Reviews.* **1999,** 5, 125-144.

2. Chupakhin V.I., Palyulin V.A., Zefirov N.S. Modeling the Open and Closed Forms of GABAA Receptor: Analysis of Ligand-Receptor Interactions for the GABA-Binding Site. Dokl. *Biochem. Biophys*., **2006**, 408, 169-174.

MOLECULAR MODELLING OF INTERACTIONS OF FRIZZLED RECEPTORS EXTRACELLULAR DOMAINS WITH THEIR NATURAL LIGANDS

Andrey E. Voronkov, Vladimir A. Palyulin, Nikolay S. Zefirov
Department of Chemistry, M.V. Lomonosov Moscow State University, Moscow, 119991, Russia

The spatial models of CRD-domains of Frizzled receptors constructed by us were used for studies of molecular mechanisms of interactions between Frizzled receptors and their natural ligands. From the experimental studies performed by different scientific groups it is known, that CRD-domains of Frizzled receptors can interact with beta-amyloid protein and with lipid modifications of the Wnt proteins. The interactions of these ligands with constructed models of CRD-domains were modelled by molecular dynamics methods and analyzed. The potential small molecule binding sites on the surface of Frizzled receptors CRD-domains were identified and evaluated. The active and inactive forms of CRD-domains were studied. The goal of this research was to evaluate the druggability of Frizzled receptors extracellular domains and to suggest the rational in silico approaches to small molecules drug design for these biotargets.

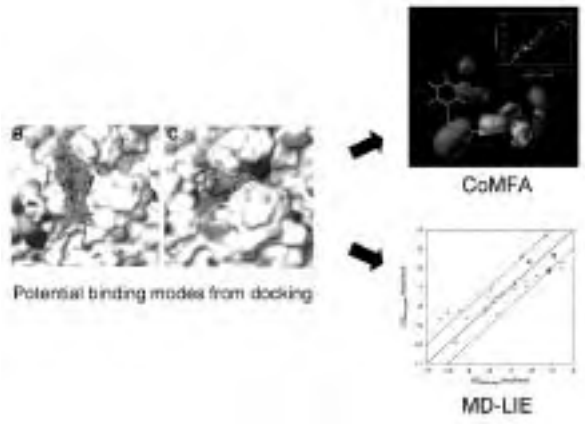
STRUCTURAL DETERMINANTS OF THE ACTIVITY OF A NEW FAMILY OF GLUCOCEREBROSIDASE INHIBITORS DERIVED FROM QSAR AND SIMULATION STUDIES

Lucía Díaz[1], Jordi Bujons[2], Hugo Gutierrez de Teran[3], Johan Aqvist[4], Ismael Zamora[5], and Antonio Delgado[1,6]

[1]Institute for Advanced Chemistry of Catalonia (IQAC-CSIC), Research Unit on Bioactive Molecules (RUBAM), Department of Biomedical Chemistry; [2]IQAC-CSIC, Department of Biological Chemistry and Molecular Modeling, Jordi Girona 18-26, 08034 Barcelona, Spain; [3]Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, Spain; [4]Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden; [5]Research Unit on Biomedical Informatics (GRIB), IMIM/Universitat Pompeu Fabra, Dr. Aiguader 88, 08003 Barcelona, Spain; [6]University of Barcelona, Faculty of Pharmacy, Unit of Medicinal Chemistry (Associated to CSIC), Avda. Joan XXIII s/n, 08028 Barcelona, Spain

Gaucher disease is one of the most prevalent lysosomal storage disorder characterized by the accumulation of the sphingolipid glucosylceramide (GlcCer) in the lysosomes. The disease is caused by the deficient activity elicited by several mutated forms of the enzyme glucocerebrosidase (GlcCerase), a β glucosidase that hydrolyzes GlcCer into glucose and ceramide (Cer) (1). Among the several therapeutic strategies for this disease,(2) the use of selective inhibitors as pharmacological chaperones, has become an active field of research.(3)

During the last years, our group has been actively working on the discovery of new GlcCerase inhibitors with chaperone activity (4). Here we report the results of modeling studies using docking, 3D-QSAR (CoMFA and GRID-GOLPE) as well as MD-LIE (molecular dynamics coupled to linear interaction energy approximation) methodologies, which were carried out to disclose the structural parameters involved in the interaction of a new class of potent inhibitors with the active centre of GlcCerase.



1. Butters, T. D. *Curr Opin Chem Biol* **2007**, 11, 412-418.
2. Sawkar, A. R.; D'Haeze, W.; Kelly, J. W. *Cell Mol Life Sci* **2006**, 63, 1179-1192.
3. Butters, T. D. Expert *Opin Pharmacother* **2007**, 8, 427-435.
4. Egido-Gabas, M.; Canals, D.; Casas, J.; Llebaria, A.; Delgado, A. *ChemMedChem* **2007**, 2, 992-994

AN AMYLOIDOGENIC DETERMINANT IN PRO-BRAIN NATRIURETIC PEPTIDE (pro-BNP)

Vassiliki A. Iconomidou, Danae Pheida and Stavros J. Hamodrakas
Department of Cell Biology and Biophysics, Faculty of Biology, University of Athens, Panepistimiopolis, Athens 157 01, Greece

Deposition of amyloid in the atria (isolated atrial amyloid) is fairly common in the ageing heart. It consists of amyloid fibrils, formed by both atrial natriuretic peptide (ANP) and brain natriuretic peptide (BNP) and their precursor molecules pro-ANP and pro-BNP. We have predicted a possible amyloidogenic determinant from the sequence of the pro-BNP peptide and we conclusively show that it forms amyloid-like fibrils in-vitro, utilizing transmission electron microscopy, X-ray diffraction, ATR FT-IR and polarizing microscopy. This peptide and its corresponding peptide in pro-ANP should be the targets of efforts towards inhibiting formation of isolated atrial amyloid.

LIGAND-BINDING PROPERTIES OF HUMAN TRANSTHYRETIN

Marta Pinto[a], Dani Blasi[b], Joan Nieto[c], Gemma Arsequell[d], Gregori Valencia[d], Antoni Planas[c], Jordi Quintana[b], Nuria B. Centeno[a]

[a] Computer-Assisted Drug Design Laboratory, Research Group on Biomedical Informatics, (GRIB) IMIM-Universitat Pompeu Fabra, Barcelona, Spain.

[b] Drug Discovery Platform (PDD), Parc Científic de Barcelona, Barcelona, Spain.

[c] Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, Barcelona, Spain.

[d] Unit of Glycoconjugate Chemistry, Institut de Química Avançada de Catalunya, IQAC-CSIC, Barcelona, Spain

Transthyretin (TTR) is one of the more than 20 known amyloidogenic proteins, which cause different amyloidogenic pathologies such as FAP, FAC or SSA. It is thought that the aggregation pathway of TTR into amyloid fibrils occurs via tetramer dissociation, a process which may be blocked by stabilisation of the tetramer via ligand binding. In this regard, several small molecules have been reported as tetramer stabilisers.

There are more than 25 crystallographic TTR-ligand complexes deposited in the Protein Data Bank. However, until now, no attempt to critically review the relevant information for ligand binding that is embedded on them has been performed. In this context, the aim of the present work was the computational analysis of these structures in order to deduce useful information for drug discovery and design.

The results obtained in this work showed that: (1) the binding site of TTR is a large and very flexible cavity, which may be described by three regions with different chemical features; (2) ligands can bind to the protein in *forward* or reverse modes depending on the conformation adopted by the serine and threonine residues located at the end of the cavity; (3) no relationship could be found between the binding mode of the ligands and their activity; (4) regardless of the structure, chemical properties or binding mode of the ligand to TTR, there is always a contribution of residues Lys^{15}, Leu^{17}, Ala^{108}, Leu^{110}, Ser^{117} and Thr^{119} to ligand binding; (5) the affinity of HBP1/HBP1′ and HBP3/HBP3′ pockets for halogens is greater than that of HBP2/HBP2′[1] and finally, (6) the most active compounds are characterised by the presence of at least one halogen atom in the HBP1/HBP1′ or HBP3/HBP3′ pockets.

[1]Mairal T, Nieto J, Pinto M, Almeida MR, Gales L, Ballesteros A, Barluenga J, Pérez JJ, Vázquez JT, Centeno NB, Saraiva MJ, Damas AM, Planas A, Arsequell G, Valencia G. (2009) Iodine Atoms: A New Molecular Feature for the Design of Potent Transthyretin Fibrillogenesis Inhibitors. PLoS ONE 4(1): e4124.

CARBOXYMETHYLATED PYRIDOINDOLES IN MULTITARGET APPROACH FOR THE TREATMENT OF DIABETIC COMPLICATIONS: STRUCTURAL ASPECTS OF ANTIOXIDANT/ALDOSE REDUCTASE INHIBITORY ACTIVITY AND BIOAVAILABILITY

M. Stefek[1], M. Majekova[1], N. Triantos[2], A. Tsantili-Kakoulidou[2], V.J. Demopoulos[3]

[1] Institute of Experimental Pharmacology and Toxicology, Slovak Academy of Sciences, Bratislava, Slovakia

[2] Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, Athens, Greece

[3] Department of Pharmaceutical Chemistry, School of Pharmacy, Aristotle University of Thessaloniki, Thessaloniki, Greece

Innovative strategies in the treatment of diseases of multifactorial origin are focused on rational design of chemical entities able to affect simultaneously multiple key mechanisms. This approach increases the chance of successful therapeutic intervention, decreases the risk of side effects and is economical. Novel carboxymethylated pyridoindoles, characterized by antioxidant activity combined with the ability to inhibit the aldose reductase enzyme, represent an example of a multitarget approach for the treatment of diabetic complications (DCs), which constitute severe health disorders of multifunctional nature. Starting from the efficient hexahydropyridoindole antioxidant stobadine, a series of carboxymethylated tetrahydro- and hexahydropyridoindole derivatives were synthesized and tested for inhibition of aldose reductase, an enzyme involved in the etiology of DCs. In vitro inhibition of rat lens aldose reductase of the most potent compounds in this series showed uncompetitive inhibition in micromolar range. On comparing tetrahydro- and hexahydropyridoindole congeners, inhibition efficacy was highly influenced by the steric conformation of the aromatic tricyclic skeleton. Selectivity with respect to the congeneric aldehyde reductase, an enzyme involved in the detoxification mechanism, was determined and related to structural variations. Antioxidant action of the novel compounds was documented in a DPPH test, in a liposomal membrane model, and in a cellular model of isolated erythrocytes oxidatively stressed by peroxyl radicals. Highly significant structure-activity correlation was obtained for DPPH scavenging.

Owing to ALR-2 pharmacophore requirements for an acidic proton, most aldose reductase inhibitors contain an acetic acid moiety or N-unsubstituted cyclic imides. Carboxylic acids are ionized at physiological pH, resulting in poor bioavailability. The presence of a basicity center at the tertiary nitrogen of the pyridoindoles, in addition to the acidic carboxylic function, predisposes these compounds to form double charged zwitterionic species, a characteristic which remarkably affects their pH-lipophilicity profile allowing for increased membrane penetration in the pH region around the isoelectric point. For the synthesized ampholytic pyridoindoles the maximal distribution ratio in the system of 1-octanol/phosphate buffer was recorded near the neutral physiological pH. The presence of zwitterionic species was experimentally proved by the concentration-dependent effect of the counter anion 1-sodium hexanesulfonate on the distribution profile. Molecular docking simulations into the ALR2 active site performed for the zwitterionic species provided an explanation for the observed structure-activity relationships and the calculated parameters were in agreement with characteristic differences in the stereoelectronic profiles of the tetrahydro- versus hexahydropyridoindoles.

VALIDATION RESULTS FOR A NEURAL NETWORK ENSEMBLE PREDICTOR OF ALDOSE REDUCTASE INHIBITORY ACTIVITY

N. J. Dimopoulos[1], R. Parra-Hernandez[1], E. M. Laxdal[1], P. Alexiou[2], V. J. Demopoulos[2]

[1]Department of Electrical and Computer Engineering, University of Victoria, B.C., Canada

[2]Department of Pharmaceutical Chemistry, Aristotle University of Thessaloniki, Greece

In this work, we present and validate a novel method of applying NN techniques in the area of QSAR. Our technique is applicable to cases where the set of exemplars is sparse; in such a situation, the classical back-propagation training methods do not result into NN that can adequately generalize.

Our technique, the guided selection mechanism [1], iteratively constructs a training subset from the available sparse set of exemplars, which will result in generalization-capable Neural Networks.

Using the thus constructed training subset, we train a number of neural networks. We further proceed to select, from this set of trained neural networks, the ones that exhibit the least sensitivity to variations of their inputs. The thus assembled set of NN constitutes the model, its response being the average of the responses of the individual NNs in the assembled set.

We have used this methodology, to construct a model of the Aldose Reductase Inhibitory activity. Our training method, as summarized in the previous paragraphs, utilized a set of 61 exemplars drawn from three chemical families, namely nitrophenyl derivatives, phenolic derivatives, and pyridazine derivatives.

Following, we validated the model by obtaining its response on three unknown chemical compounds that did not belong to the families the model was trained with.

The compounds used for this (blind) validation trial, together with the response of the model and the measured responses, are given in the Table below, and show a remarkable agreement between the model and the experimentally determined inhibitory activities of the compounds.

	Inhibitory activity (pIC_{50})			
Compound	Predicted	Measured	Difference	Normalized Difference
N-(3,5-difluoro-4-hydroxyphenyl) benzenesulfonamide	4.557	4.483	0.075	0.016
N-(3,5-difluoro-4-hydroxyphenyl) -4-methoxybenzenesulfonamide	4.884	4.81	0.074	0.015
N-(3,5-difluoro-4-hydroxyphenyl) -4-nitrobenzenesulfonamide	4.868	4.352	0.515	0.106

These results are significant not only because the errors are quite small, but more importantly because they are within a generally accepted absolute error of 1.

Further we developed our model based on exemplars that represented three different families of chemical compounds, and validated it with chemical compounds that did not belong to the classes our model was trained with. Thus, our technique was able to produce a NN model that is accurate, and to the extent of the compounds it was tested on, can be considered chemical family-independent.

Although the results we presented here very strongly indicate a NN model that can accurately predict the

Aldose Reductase Inhibitory activity of arbitrary chemical compounds, we feel that further validation trials, involving a larger number of compounds, are necessary to fully explore the predictive abilities and limitations of the developed model.
We further plan to study the significance of each of the descriptors used (based on the Sensitivity Heuristic) and use this information together with a Principal Component Analysis to perhaps curtail the number of descriptors used.

Reference:

[1] E. M. Laxdal, R. Parra-Hernandez, and N. J. Dimopoulos, "Guided Construction of Training Data Set for Neural Networks". IEEE International Conference on Systems, Man & Cybernetics, October 2004, The Hague, The Netherlands, pp. 5905-5910.

BINDING MODE PREDICTION AND ANALYSIS FOR SALACINOL DERIVATIVES AS ALPHA-GLUCOSIDASE INHIBITORS

Shinya Nakamura, Kazunori Takahira, Genzo Tanabe, Toshio Morikawa, Mika Sakano, Kiyofumi Ninomiya, Masayuki Yoshikawa, Osamu Muraoka, and Isao Nakanishi

Kinki University, 3-4-1 Kowakae, Higashi-osaka, Osaka, Japan

Salacinol is a potent α-glucosidase inhibitor isolated from *Salacia reticulata*, and a good lead compound for an antidiabetic drug. It is essential to clarify the binding state of salacinol to α-glucosidase for efficient optimization study using structure-based drug design. The binding mode of salacinol and its derivatives has been predicted using ASEDock implemented in MOE. To calibrate docking protocol, redocking simulations of two inhibitors, acarbose and casuarine whose complex structures are known, were performed. The RMSD values excluding hydrogen atoms between experimental and simulated structure were 0.7Å and 0.2Å for the acarbose and casuarine, respectively. Using this docking protocol, the salacinol binding state was explored. Salacinol bound to the protein with a similar binding mode as casuarine,[1] and the schematic interaction diagram was shown in Figure 1. Recently the experimental binding mode of salacinol has been reported.2 The predicted binding mode was quite similar (Figure 2) and RMSD value of 0.9Å.

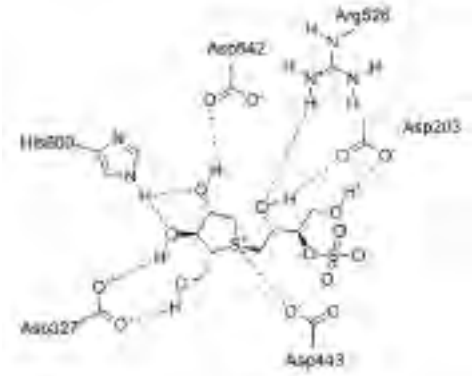


Figure 1. Schematic interaction diagram between salacinol and α-glucosidase

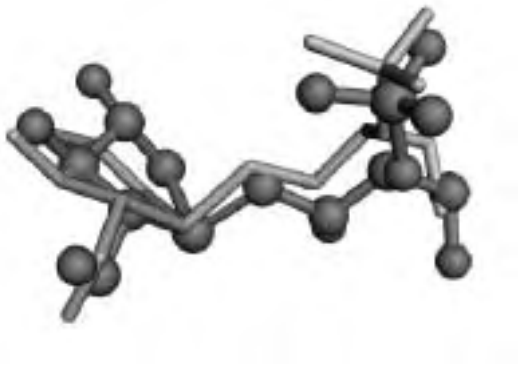


Figure 2. The comparison of binding mode of salacinol (cyan-stick: experimantal structure, magenta-ball&stick: predicted structure)

The binding mode was able to explain the differences among salacinol derivatives such as the weakened affinity by the lack of hydrogen bonds. Furthermore, as the sulfate group at the C3′ position on the methylene chain was located almost outside the binding pocket, it was structurally seemed that the lack of this group did not affect to the binding activity. However, our calculation showed that this group contributed to the binding affinity by the weak but sure interaction. These interactions were consistent to the structure-activity relationships of salacinol derivatives.

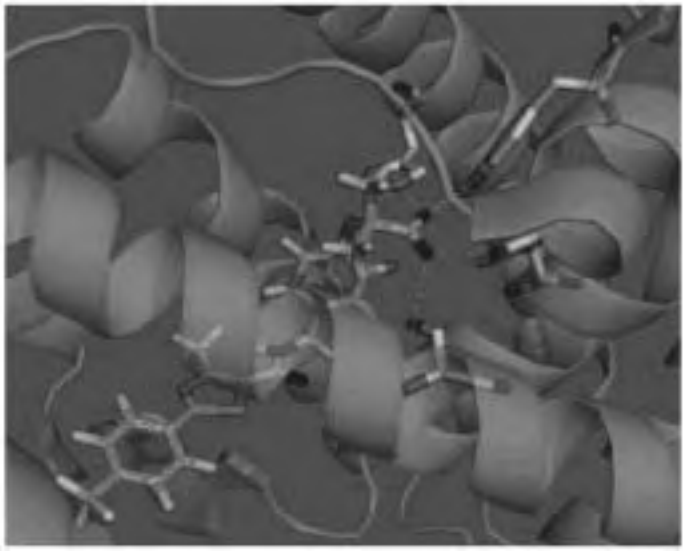
1) Nakamura S, Takahira K, Tanabe G, Morikawa T, Sakano M, Ninomiya K, Yoshikawa M, Muraoka O, Nakanishi Computational Simulation of salacinol binding to α-glucosidase. Bioorg. Med. Chem. Lett. submitted.

2) Sim L, Jayakanthan K, Mohan S, Nasi R, Johnston BD, Pinto BM, Rose DR, New glucosidase inhibitors from an ayurvedic herbal treatment for type 2 diabetes: structures and inhibition of human intestinal maltase-glucoamylase with compounds from Salacia reticulata. Biochemistry 2010;49:443-451

COMPUTATIONAL STUDY OF BINDING PROPERTIES OF FIBRATE AGONISTS TO PPARα RECEPTOR

F. Lannutti, A. Marrone, N. Re
Dipartimento di Scienze del Farmaco, Università "G. d'Annunzio", Chieti, Italy

Atherosclerosis is a progressive metabolic disease that involve perturbation of both lipoprotein metabolism and arterial. Several clinical studies have shown an improvement of the lipidic profile and reduction of cardiovascular risk upon administration of fibrates. These clinical effects are mainly mediate by Peroxisome proliferator-activated receptor, PPARα, a member of the nuclear receptor family that plays a key role in the regulation, storage and catabolism of fatty acids.



The binding properties of fibrate PPARα agonists have been evaluated through a combined approach of docking and molecular mechanics. In particular, we developed a computational protocol that permits us to: i) considering, although partially, the receptor flexibility, and so, better estimate the binding pose and energy; ii) estimate the main energetic contribute to free binding energy.
The calculated binding geometries were in good agreement with the available experimental poses and allowed to obtain several information on the binding fashion of this class of agonists. Moreover, predicting models of the binding to PPARα receptor were developed and validated through MLR (multi linear regression). The obtained models allowed us to estimate with fair accuracy the agonist activity of the considered fibrates, paving the way to a possible preliminary evaluation of new potential ligands of PPARα.

Bibliography

1 - S.J. Robins, *J. Cardiovasc. Risk* **2001**, 8, 195-201
2 - S. Verste, B. Desvergne, W. Wahli, *Nature* **2000**, 405, 421-424
3 - S. M. Rangwala et al. *Chirality* **1997**, 9, 37-47
4 - M. L. Sierra et al. *J. Med. Chem*. **2007**, 50, 685-695

DRUGLIKENESS NAVIGATION IN THE CHEMICAL SPACE OF PPAR-γ AND PPAR-α/γ AGONISTS

Costas Giaginis[1,2], Theodosia Vallianatou[1], Stamatios Theocharis[2], Anna Tsantili-Kakoulidou[1*]

[1]Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, Panepistimiopolis, Zografou, Athens 15771, Greece.

[2]Department of Forensic Medicine and Toxicology, Medical School, University of Athens, 75 Mikras Asias Street, Athens 11527, Greece

High-throughput screening of absorption, distribution, metabolism, elimination/toxicology (ADME/Tox) characteristics has been introduced in the early steps of Drug Design as an emergent demand to reduce attrition rate of drug candidates due to poor pharmacokinetic and pharmacodynamic profiles, as well as toxicity side effects. In this aspect, certain physicochemical and molecular properties have been considered as in silico predictors of oral bioavailability, while tolerant cut off values or ranges have been proposed to reflect druglike characteristics. In the present study, evaluation of druglikeness was performed in the case of PPAR-γ and dual PPAR-α/γ agonists, a category of chemical compounds with increasing interest as orally administrated drug candidates against diabetes mellitus and hyperlipidemia. For this purpose, a large data set of approximately 1200 relevant compounds available in literature was compiled and explored in respect to several physicochemical and molecular properties. Lipophilicity expressed as logP (or $logD_{7.4}$), the number of hydrogen bond acceptor/donor sites (HBA/HBA) and molecular weight (MW) implemented in 'Lipinski's rule of five' (ROF), as well as additional molecular properties, such as molar refractivity (CMR), topological polar surface area (TPSA), rotatable bonds (RB) and the number of rings (RNG) were calculated by familiar and widely recognized software packages. The distribution of the above properties was established and evaluated in respect to the proposed cut off values or ranges. Violations concerning excess logP, MW and RB were observed for a significant number of compounds, while TPSA, CMR and RNG were found to lie on the limit of the proposed cut off ranges. The same analysis was repeated discriminating the data set into two groups according to activity levels expressed by PPAR-α and PPAR-γ gene transactivation ($pEC50 < 7$ and $pEC50 \geq 7$) and the relevant physicochemical characteristics were assigned to each group. This information is important and can be incorporated in a more holistic approach for future design of dual PPAR-α/γ and PPAR-γ agonists with improved ADME/Tox characteristics and therefore enhanced chance to success within clinical development.

APPLICATION OF MULTIVARIATE DATA ANALYSIS FOR CONSENSUS MODELLING OF PPAR-α/γ GENE TRANSACTIVATION DATA

Vallianatou Theodosia, Giaginis Costas, Tsantili-Kakoulidou Anna
Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, Panepistimiopolis, Zografou, Athens 157 71, Greece

Peroxisome proliferator-activated receptors (PPARs) play a crucial role in the regulation of lipid and glucose metabolism. Among the three subtypes (PPAR-α, PPAR-γ, PPAR-β/δ), the most investigated is the γ-subtype and currently PPAR-γ agonists (glitazones) are marketed as anti-diabetic agents. To overcome certain side effects, common in this type of drugs, research interest has recently been shifted towards dual agonists acting on both γ and α subtypes. In the present work, Multivariate Data Analysis (MVDA) was used to analyze PPAR α and γ gene transactivation ($pEC_{50}\alpha$ and $pEC_{50}\gamma$) produced by several phenoxyacetic acids, meta-substituted phenyl propanoic acids and oxazole containing carboxylic acids, collected from ref. [1-4]. The pool of descriptors comprised physicochemical/molecular properties, 3-D descriptors, connectivity and electrotopological state indices. PLS models were initially derived for PPAR-α activity considering the data sets separately because of small differences in the experimental protocols. Local models although accompanied by satisfactory statistics showed poor predictivity and narrow applicability domain. As a next step a consensus 2 component PLS model for all four data sets was established with $R^2=0.878$, $Q^2=0.813$ RMSEE=0.42, denoting that reduced uniformity in experimental conditions within a data set can be tolerated by MVDA. Considering both PPAR α and PPAR-γ activities a consensus 4 component model for dual agonists was obtained with overall $R^2=0.827$ and $Q^2=0.627$ and analogous individual statistics for pEC_{50}-α and pEC_{50}-γ data. However significant differences in the contribution of the parameters in PPAR-α and PPAR- γ activity were revealed. Lipophilicity seems to contribute almost equally, while flexibility (expressed by rotatable bonds, RB) and bulk descriptors have more positive impact in PPAR-γ activity. Differences in the contribution of MOLCONNZ descriptors reflect the requirements to specific structural characteristics for each receptor subtype. Models were further validated using a blind test set with compounds taken from ref [5]. Activity predictions were satisfactory in most cases for both receptor subtypes.

References

1. Fracchiolla G. et al, *ChemMedChem*, **5**, 651-74, 2007
2. Pingali H. et al, *Bioorg.Med.Chem*, **16**, 7117-7127, 2008
3. Suh Y.G. et al, *J.Med.Chem*, **51**, 6318-6333, 2008
4. Pingali H. et al. Bioorg.*Med.Chem.Lett.* 18, 6471-6475, 2008
5. Benardeau A. *et al, Bioorg.Med.Chem Let*, **19** :2468-2473, 2009

IN SILICO CONFORMATIONAL DYNAMICS ON TOXIN COMPLEXES OF HUMAN A1 nAChR SUBUNIT: THE ROLE OF GLYCOSYLATION IN TOXIN BINDING

N. Dimitropoulos, G.A. Dalkas, A. Papakyriakou, C.T. Chassapis, G.A. Spyroulias
Department of Pharmacy, University of Patras, GR-26504, Patras, Greece

Nicotinic acetylcholine receptors (nAChRs) belong to the superfamily of the Cys-loop ligand-gated ion channels (LGICs), which also includes the GABA, glycine, and 5-HT3 receptors [1]. LGICs form homo- or heteropentamers of related subunits, each of them is divided into an N-terminal or ligand-binding domain (LBD), a transmembrane region and an intracellular region. In the last years, the knowledge regarding the nAChR structure has been dramatically increased by the determination of (a) the X-ray crystal structures of molluscan ACh-binding proteins (AChBPs) [2], (b) the electron microscopy (EM) structure of the Torpedo nAChR [3], (c) the X-ray crystal structure of the mouse nAChR a1-ECD (extracellular domain) [4] and (d) the X-ray crystal structures of two prokaryotic LGIC [5-6]. The crystal structure of the mouse nAChR a1-ECD bound to a-Bungarotoxin at 1.94 Å resolution reveals atomic details for several key functional elements of nAChR. The functional studies suggest that the carbohydrate is involved in a-Bungarotoxin binding and channel gating.

In this study we conducted molecular dynamics (MD) simulations of the mouse a1 nAChR ligand-binding domain bound to a-Bungarotoxin, as well as the human a1 nAChR ligand-binding domain, built through homology modelling, bound to three distinct toxins: a-Bungarotoxin, a-Cobratoxin and a-Conotoxin. Each MD simulation was conducted both in absence and presence of the carbohydrate chain linked to the Asn141 glycosylation site, in order to study the general role of glycosylation in channel function.

References

[1] Corringer, P.J et al. Nicotinic receptors at the amino acid level. Annu. Rev. Pharmacol. Toxicol. 40, 431–458 (2000).
[2] Brejc, K. et al. (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. Nature 411, 269–276.
[3] Unwin, N. (2005) Refined structure of the nicotinic acetylcholine receptor at 4 A ° resolution. J. Mol. Biol. 346, 967–989.
[4] Dellisanti, C.D. et al. (2007) Crystal structure of the extracellular domain of nAChR a1 bound to a-bungarotoxin at 1.94 Å resolution. Nat. Neurosci. 10, 953–962.
[5] Hilf, R.J. and Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. Nature 2009, 457; 115-8
[6] Bocquet, N. et al. J. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. Nature, 2009, 457: 111-4

CAPTURE COMPOUND MASS SPECTROMETRY – A NOVEL TOOL FOR TARGET DETECTION, DRUG REPURPOSING AND TOXICITY PROFILING AND PREDICTION

Jenny J. Fischer, Simon Michaelis, Anna K. Schrey, Olivia Graebner, Mirko Glinski, Mathias Dreger, Friedrich Kroll, Hubert Koester

Caprotec Bioanalytics GmbH, Volmerstrasse 5, 12489 Berlin, Germany

Assessment of protein binding profiles for small molecules has become a crucial issue in different fields of research. Extensive target identification for drugs or drug candidates gives insight in the mode of action, but also unravels new binding partners for an already-known structure, opening the door to new indications for the drug or drug candidate. Also, proteins correlated to toxic side effects such as hepatotoxicity can be detected. Such side effects are a major cause for failure of drugs in clinical trials, and the interactions between drug and off-targets underlying this toxicity are difficult to assess.

Capture Compound Mass Spectrometry (CCMS) is a novel experimental approach for the isolation and identification of proteome subsets on the basis of small molecule-protein interactions that complements and fosters theoretical approaches. Capture Compounds™ are tri-functional probes: a selectivity function (a small molecule) interacts with the target protein(s) in a biological sample, a reactivity function irreversibly forms a covalent bond, and a sorting function allows the captured protein(s) to be isolated for mass spectrometric analysis. The formation of a covalent bond between the Capture Compounds and the protein(s) of interest allows the isolation even of weakly interacting or low abundant proteins. In an in-house study, we investigated the hepatotoxicity profile of Tolcapone, a Parkinson drug which had temporarily been withdrawn from the market due to serious liver impairment.

A series of capture compounds based on Tolcapone and the less toxic, but also less efficient Parkinson drug Entacapone was designed aided by Molecular docking and MOLCAD protein surface investigations on the actual target, catechol-O-methyl transferase. With the Tolcapone capture compound, off-target binding occured with components of the mitochondrial respiratory chain and the fatty acid beta-oxidation pathway. This provides an explanation at the molecular level for physiological findings described in the literature.

In the theoretical work, these results can be exploited in different manners. Newly identified off-targets can be included into structure-based drug design either to reduce toxicity or to improve the structure towards the new target. Binding pockets can be mapped by detection of the crosslink positions in the protein. Further, it has to be investigated, if quantified CCMS protein lists can serve as a basis for – or can be included into – theoretical hepatotoxicity models.

OXIDIZED LIPIDS INHIBITS PARAOXONASE: A MECHANISTIC STUDY USING DOCKING TECHNIQUES AND LC-MS ANALYSIS

Soliman Khatib[1], Hagai Tavori[1, 2], Michael Aviram[2], Jacob Vaya[1]

[1]Oxidative Stress Research laboratory, MIGAL - Galilee Technology Center, P.O. Box 831, Kiryat Shmona 11016, and Department of Biotechnology, Tel-Hai Academic College, Israel. [2]The Lipid Research Laboratory, Rappaport Family Institute for Research in the Medical Sciences, Technion, Rambam Medical Center, Haifa 31096, Israel

Paraoxonase1 (PON1) is a HDL bound enzyme and many of the anti-atherogenic properties of HDL are attributed to PON1. PON1 hydrolyzes organophosphates, arylesters and lactones, whereas the lactones activity is assumed as the physio/pathological one. Recently we have published that the human carotid atherosclerotic plaque lipid extract (LE) accelerate atherosclerotic process by oxidizing LDL, macrophages and inhibit HDL mediated-cholesterol efflux and at the same time LE inhibit PON1 lactonase activity in a dose and time-dependent manner. The LE component responsible for PON1 inhibition was found to be oxidized lipid (OxLP). The aim of the present study is to explore the mechanism of OxLP inhibitory effect on PON1 using LC-MS analysis and by employing modeling techniques using AutoDock3 program. OxLP specifically acts to oxidize PON1 free thiol at position 284, this inactivation could be prevented and even reversed by a free thiol agents such as Cystein and propanethiol. The docking results showed that specific OxLP can interact with the thiol 284 site of the enzyme with high affinity while other OxLPs does not. This docking analysis was confirmed by LC-MS results showed that PON1 decreased these specific OxLP levels and not affected the other OxLPs.

CHARACTERIZATION OF PROTEIN LIGAND INTERACTION OF THE FACTOR VIII C2 DOMAIN WITH THE PHOSPHOLIPID MEMBRANE USING DEUTERIUM EXCHANGE MASS SPECTROMETRY

Dennis P. Pantazatos[1], Chris Gessner, Gary E. Gilbert[1] and Virgil L. Woods, Jr.[2]

From the [1]Department of Medicine, Veterans Affairs Boston Healthcare System, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02132 USA,

[2]Department of Medicine, University of California, San Diego, CA 92093 USA

Factor VIII (fVIII) functions as a membrane-bound cofactor in the Factor Xase complex in the intrinsic pathway of the blood coagulation cascade. Binding to a phospholipid membrane enhances activity of the factor VIIIa-factor IXa complex approx. 100,000-fold. While membrane binding increases the affinity of factor VIIIa for factor IXa and factor X, the major effect is upon the catalytic activity of the assembled complex. The mechanism by which activity is enhanced remains unknown. The C2 domain of factor VIII (fVIII-C2) contains the major membrane-binding function. The crystal structures of the C2 domain and of intact factor VIII have enabled hypotheses about the mechanism of membrane binding and enhanced activity. Functional motifs of the C2 domain include two pairs of hydrophobic, membrane-interactive amino acids at the tips of "spikes," on the lower end of the C2 domain and factor IXa-binding site on the upper surface. In this study we used deuterium exchange mass spectrometry (DXMS) to identify membrane interacting structural motifs and regions of allosteric change and altered dynamics induced by ligand binding. Our results not only confirmed known membrane interactive regions identified using site directed mutagenesis[1] but also identified an additional membrane interactive motif as well as regions of conformational change and altered stability. (DXMS) is a powerful, and sensitive method for QSAR studies of protein ligand interactions and characterization of protein structure, and dynamics. It is has proven effective for mapping ligand interfaces[2], localizing regions of disorder and solvent exposure[3], and studying protein small molecule interactions[4]. DXMS relies on the hydrogen exchange chemical reaction in which backbone amide hydrogens exchange with those in the solvent, contingent upon physical contact. The exchange rate is dependent on solvent accessibility and pH. DXMS is performed by isotopic labeling using deuterium oxide (D_2O). The labeling is quenched at varying time points followed by proteolysis to produce overlapping peptides spanning the length of the protein sequence. The degree of labeling is assessed using LC/MS analysis. Changes in percent deuterium incorporation ($\%D_{inc}$) can be determined and represent the degree of solvent accessibility and structural stability. We utilized DXMS to characterize the fVIII-C2 phospholipid interaction. Serial digestion using pepsin and fungal protease XIII produced overlapping peptides spanning the entire fVIII-C2 structure,. In the absence of pospholipid vesicles (PLVs) the extent of $\%D_{inc}$ upon a 10s labeling at the membrane-interactive spikes (M2199/F2200) and (L2251/L2252), was 77.5% and 100% respectively, confirming their flexibility. Upon phospholipid binding a decrease in $\%D_{inc}$ at spikes 1 and 3 was observed to 49% and 83.5% respectively. The $\%D_{inc}$ also decreased from 30% →11.6% in a 5 residue segment Q2311-Q2316 in the presence of PLVs, identifying it as an additional membrane interactive motif. The reduction in %Dinc in these regions implied increased structural rigidity in addition to any solvent protection occurring upon membrane immersion or interaction of the hydrophobic side chains. A substantial four fold decrease in $\%D_{inc}$ (84.9% →20.6%) was also observed in the distal putative FIXa binding site E2228-V2240 indicating significant conformational change, while minimal reductions are observed in $\%D_{inc}$ in most peptides beta-barrel core . These observations indicate that membrane engagement alters the mobility of membrane interactive spikes and induces a conformational change and

stabilization of the factor IXa binding site. Other regions of stabilization include residues at the between the A1-C1 and C1-C2 domain interfaces. This study yields informative insight in structural factors that may play a key role in the mechanism of Factor VIII activation. It identifies key regions of ligand binding and induced allosteric changes that may contribute to the enhancement of factor fVIIIa-factor IXa activity.

References:

1. Gilbert, G.E., et al., *Four hydrophobic amino acids of the factor VIII C2 domain are constituents of both the membrane-binding and von Willebrand factor-binding motifs*. J Biol Chem, 2002. **277**(8): p. 6374-81.
2. Garcia, R.A., D. Pantazatos, and F.J. Villarreal, *Hydrogen/deuterium exchange mass spectrometry for investigating protein-ligand interactions*. Assay Drug Dev Technol, 2004. **2**(1): p. 81-91.
3. Pantazatos, D., et al., *Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS*. Proc Natl Acad Sci U S A, 2004. **101**(3): p. 751-6.
4. Garcia, R.A., et al., *Molecular interactions between matrilysin and the matrix metalloproteinase inhibitor doxycycline investigated by deuterium exchange mass spectrometry*. Mol Pharmacol, 2005. **67**(4): p. 1128-36.

RECEPTOR DEPENDENT 3D-LQTA-QSAR OF A SERIES OF SUBSTITUTED AMPHETAMINES

M.A. Carvalho-Fresqui, E.G. Barbosa and M. M. C. Ferreira
Instituto de Química, Universidade Estadual de Campinas, 13083-977, Campinas, SP, Brazil

The amphetamine family of drugs is the most common group of psycho-stimulant drugs. Amphetamines inhibit the monoamine oxidase enzyme (MAO, isoforms A and B) which catalyzes the oxidation of neurotransmitters. Two different amphetamines derivatives, selegiline and rasagiline are of particular interest. They are selective irreversible inhibitors of type B monoamine oxidase (MAO-B), that is used primarily in the form of a covalent adduct with the isolloxazone moiety of the FAD cofactor in the treatment of Parkinson's and Alzheimer's diseases and depression.

Docking studies were performed with GOLD program [1] on 30 selegiline and rasagiline derivatives [2] and according to the results the most potent compounds had the propargyl group in an orientation suitable for their reaction with the FAD cofactor placed at the receptor binding site. However, this orientation was not obtained for some compounds, so, a manual docking was done to obtain the best orientation to form such FAD-ligands adducts as suggest by Potashman [3].

Docked reactive poses from the superimposed ligands at the binding site provided aligned conformations for receptor dependent 3D-LQTA-QSAR [4], where electrostatics and Lennard-Jones interaction energies were calculated using (NH_3^+) as probe in a grid box of 1 Å and used as descriptors for modelling.

Regression models were built employing the ordered predictor selection [5] algorithm for variable selection and multiple linear regression (MLR). The y–randomization and leave–N–out cross-validation procedures were carried out in addition to the external validation. MLR models provided the following statistics: $Q^2 = 0.64$, $R^2 = 0.77$ for 6 variables selected. The selected descriptors illustrated in Figure 1 provide information regarding the interaction of the derivatives with FAD and two isoleucines (172,199). These preliminary results are promising and useful for further receptor dependent studies with adduct formed from these selected selegiline and rasagiline derivatives.

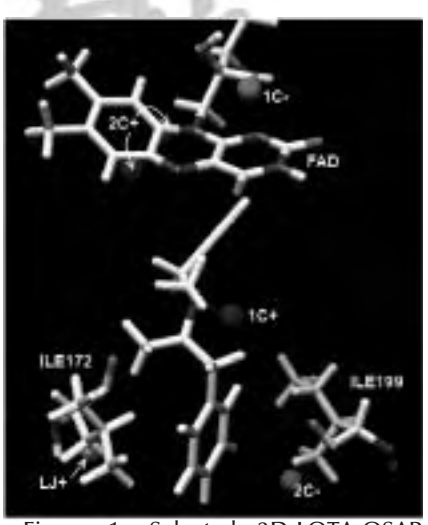


Figure 1: Selected 3D-LQTA-QSAR descriptors (ball) and the interaction of the derivatives with FAD and isoleucines 172 and 199.

Acknowledgments:
Dr. Phil Biggin for collaboration in docking studies.
CNPq, FAPESP and CAPES (Brazilian Agencies).

[1] GOLD 3.1: CCDC Software Ltd. Cambridge, UK, 2004. CD-ROM.
[2] Sterling J., et al. *J. Med. Chem.*; **45**, 5260 (2002)
[3] Potashman M.H., Duggan M.E., *J. Med. Chem.;* **52**, 1231 (2009)
[4] Martins J.P.A., Barbosa E.G.,Pasqualoto K.F.M., Ferreira M.M.C., *J Chem Inf Model*,;**49**, 1428, (2009)
[5] Teófilo R.F., Martins J.P.A., Ferreira M.M.C., *J. Chemom.;* **23**, 32 (2009)

A COMPARATIVE QSAR APPROACH FOR UNDERSTANDING ALLOSTERIC EFFECTS IN GLYCINE/NMDA RECEPTOR ANTAGONISM

Manish Sharma and V. B. Gupta

Centre for QSAR and Molecular Modeling, Department of Medicinal and Pharmaceutical Chemistry, TIFAC-CORE, Innovation Square, B. R. Nahata College of Pharmacy, MP, India

A comparative Hansch type QSAR study was conducted using multiple regression analysis on various sets[1-5] of quinoxalines, quinoxalin-4-ones, quinazoline-2-carboxylates, 4-hydroxyquinolin-2(1*H*)-ones, 2-carboxy tetrahydro-quinolines, phenyl hydroxy quinolones, nitro quinolones and 4-substituted-3-phenyl quinolin-2(1*H*)-ones as selective glycine/NMDA site antagonists. They produce 50% inhibition of [^{3}H] glycine or [^{3}H]-L-689,560 binding to rat brain cortical membranes. Ten statistically validated equations were developed which indicated importance of calculated molar refractivity (CMR), verloop's sterimol length (L1) and calculated logP (Clog*P*) parameters. CMR was present in eight equations indicating involvement of charge transfer reaction in ligand receptor interaction. Three indicator variables (I_a, I_1 and I_2) were also utilized in equations. Normal and inverse non-linear parabolic relationships were found with CMR in different series indicating dual allosteric binding mode in glycine/NMDA antagonism. Equations (4,8,10) revealed normal parabolic relationship (optimum CMR:10±10%) and equations (6,7) indicated inverse parabolic relationship (inversion point:10.8). Equations (2,3) indicated importance of ClogP parameter with an inverse parabolic relationship in equation (2). Good range of data points on both sides of parabola and 95% confidence at inversion point confirm quality of our conclusion. L1 negatively contributed in equation (4). Equations (7,8,9,10) indicated presence of anionic functionality at 4-position in all ring systems studied is not absolutely required for effective binding. CMR values of some molecules undergone clinical trials were close to inversion point discussed for equations (4,8,10). All equations were laterally validated and some tested for their external predictivity. Applicability domains of the equations were also estimated.

References:

1. Rowley, M.; Leeson, P. D.; Stevenson, G. I.; Mosley, A. M.; Stansfield, I.; Sanderson, I.; Robinson, L.; Baker, R.; Kemp, J. A.; Marshall, G. R.; Foster, A. C.; Grimwood, S.; Tricklebank, M. D.; Saywell, K. L. J. *Med. Chem*. **1993**,36,3386.
2. Leeson, P. D.; Carling, R. W.; Moore, K.W.; Moseley, A. M.; Smith, J. D.; Stevenson, T. C.; Baker, R.; Foster, A. C.; Grimwood, S.; Kemp, J. A.; Marshall, G. R.; Hoogsteen, K. J. *Med. Chem*. **1992**,35,1954.
3. Kulagowski, J. J.; Baker, R.; Curtis, N. R.; Leeson, P. D.; Mawer, I. M.; Moseley, A. M.; Ridgill, M. P.; Rowley, M.; Stansfield, I.; Foster, A. C.; Grimwood, S.; Hill, R. G.; Saywell, K. L.; Tricklebank, M. D. J. *Med. Chem*.**1994**,37,1402.
4. Carling, R. W.; Leeson, P. D.; Moore, K.; Smith, J. D.; Moyes, C. R.; Mawer, I. M.; Thomas, S.; Chan, T.; Baker, R.; Foster, A. C.; Grimwood, S.; Kemp, J. A.; Marshall, G. R.; Tricklebank, M. D.; Saywell, K. L. J. *Med. Chem*.**1993**,36,3397.
5. Carling, R.W.; Leeson, P. D.; Moore, K.; Smith J. D.; Moyes, C. R.; Dunction, M.; Hudson, M. L.; Baker, R.; Foster, A.C.; Grimwood, S.; Kemp, J. A.; Marshall, G. R.; Tricklebank, M. D.; Saywell, K. L. J. *Med. Chem*.**1997**,40,754.

HUMAN METABOTROPIC GLUTAMATE RECEPTOR-5 (mGluR5) HOMOLOGY MODELLING: AN INNOVATIVE STRATEGY TO MODEL RECEPTOR FLEXIBILITY

Bettinelli I. [1], De Toma C. [1], Graziani D. [1], Leonardi A. [1], Marconi C. [*], Motta, G. [1], Pedretti A[*], Riva, C. [1], and Vistoli, G. [*]

[1] Recordati S. p. A., Drug Discovery, via Civitali 1, 20148 Milano, Italy

[*] Dipartimento di Scienze Farmaceutiche "Pietro Pratesi", Via Mangiagalli 25, Università degli Studi di Milano, 20133 Milan, Italy

During the last years, a big effort has been focused on the study of modulators at GPCR allosteric sites. Recordati claimed[1] the possibility to treat Overactive Bladder (OAB) dysfunctions using selective metabotropic glutamate receptor subtype 5 (mGluR5) negative allosteric modulators (NAM). From this patent application, a number of close analogues of MPEP and other hit compounds endowed with good affinity for the receptor became available for our studies, all of them sharing a common phenylethynyl portion.

Using these newly synthesised molecules as training set and available binding pocket mutagenesis data we worked on the validation of a mGluR5 homology model, which was generated using the fragmental approach as recently proposed by some of us.[2] In agreement with previous models and mutagenesis data,[3] the allosteric binding pocket was placed among TM3, TM5, TM6 and TM7 helices.

As recently described, [4] conserved proline residues may play a crucial role in conformational changes of receptor since they can assume two well defined conformations (straight or kinked) switching the receptor between the active and inactive state. An innovative strategy to model the receptor flexibility, considering the conformational effect of Pro-containing transmembrane helices, was defined in this study. This approach allowed the generation of an exhaustive set of mGluR5 homology models (also termed chimeras), which differ for the bending of Pro-containing transmembrane helices and thus for the wideness of the TM bundle.

With such constructed chimeras, we decided to focus the docking studies on three chimeras, which were supposed to be representative of the close, intermediate and open receptor structure. The chimera which afforded the best predictive results was characterized by an intermediate opening of TM bundle, confirming that these ligands, even being allosteric modulators, act with an inverse agonism mechanism based on their ability to prevent the full closure and thus the activation of the receptor.

Specifically, Pro-655 (TM3), Pro-743 (TM5) and Pro-790 (TM6) seem to play a main role in the allosteric binding site. The first one may be directly involved in the binding and all together work as a filtering "proline sieve", discriminating among active and inactive compounds and conferring to the receptor a particular rigidity, that can explain the dramatic loss of affinity obtained with minor molecular changes. In agreement with a pharmacophore model developed in house, the "proline sieve effect" is defined by two main ligand aromatic portions, characterized by a different lipophilicity, and rigidly spaced at a distance of about 10 Å with an optimal angle of about 120°.

In summary, this study emphasized the promising potentialities of chimera modelling, confirming the key role of proline residues, within the TM bundle, and suggesting that improved results can be pursued by considering the often disregarded receptor flexibility. This approach could find fertile applications in GPCR modelling affording relevant results both in rational ligand optimization and in virtual screening.

[1]Selective mGlu5 Antagonists for Treatment of Neuromuscular Dysfunction of the Lower Urinary Tract, *Leonardi,*

Amedeo; Testa, Rodolfo; Poggesi, Elena; ***European Patent Application EP1599204***
[2]Construction of Human Ghrelin Receptor (hGHS-R1a) Model Using a Fragmental Prediction Approach and Validation through Docking Analysis; *Alessandro Pedretti, Marco Villa, Marco Pallavicini, Ermanno Valoti, and Giulio Vistoli*, ***J. Med. Chem. 2006, 49, 3077-3085***.
[3]Preferential binding of allosteric modulators to active and inactive conformational states of metabotropic glutamate receptors; *Naveena Yanamala, Kalyan C Tirupula and Judith Klein-Seetharaman;* ***BMC Bioinformatics 2008, 9 (Suppl 1):S16***
[4]The Role of a Conserved Proline Residue in Mediating Conformational Changes Associated with Voltage Gating of Cx32 Gap Junctions; *Yi Ri, Juan A. Ballesteros, Charles K. Abrams, Seunghoon Oh, Vytas K. Verselis, Harel Weinstein, and Thaddeus A. Bargiello*; ***Biophysical Journal Volume 76 June 1999 2887–2898.***

HOMOLOGY MODELING AND MOLECULAR DYNAMICS SIMULATION OF THE ANGIOTENSIN II (AT1) RECEPTOR: INSIGHT INTO THE LIGAND-RECEPTOR INTERACTIONS

Minos-Timotheos Matsoukas[1], Arnau Cordomí[2], Leonardo Pardo[2] and Theodore Tselios[1]

[1]Department of Chemistry, Section of Organic Chemistry, Biochemistry and Natural Products, University of Patras, 26500 Patras, Greece

[2]Laboratori de Medicina Computational, Unitat de Bioestadistica, Facultat de Medicina, Universitat Autonoma de Barcelona, 08193 Bellatera, Spain

The Renin-Angiotensin System (RAS) plays a major role in blood pressure regulation. A sequence of enzyme reactions leads to the release of angiotensin II which interacts principally with the type-1 angiotensin II receptor (AT1), a 359-residue, which is part of the G protein-coupled receptor family. Recent advances in determining the crystal structures of other GPCRs have largely contributed to a better understanding of the structural features of these proteins[1]. In the present study, a homology model of the human AT1 receptor was constructed on the basis of the crystal structure of the β_2-adrenergic receptor (PDB code 2RH1)[2]. Molecular dynamics simulations on multiple receptor - ligand complexes embedded in an explicit membrane model were performed in order to gain a better understanding on the binding mode of non-peptide antagonists (Sartans). The positioning of critical amino acids S3.33(109), K5.42(199), H6.51(256), N7.46(295), known to effect antagonist binding affinity were thoroughly examined in the presence of known active compounds. The structural requirements, as well as interaction details, were obtained by the above mentioned molecular dynamics simulations for developing more potent AT1 inhibitors.

References:

1. Weis, W.I. and Kobilka, B.K. (2008) Structural insights into G-protein coupled receptor activation. Curr. Opin. Struct. Biol. 18, 734–740

2. Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Kuhn, P., Weis, W.I., Kobilka, B.K., Stevens, R.C. (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318, 1258–1265

THE SIGNIFICANCE OF STRUCTURAL MODIFICATIONS OF UPF PEPTIDES, GLUTATHIONE ANALOGUES, ON CuZnSOD ACTIVITY OF K562 CELLS

Ceslava Kairane, R. Mahlapuu, K. Ehrlich, U. Soomets and M. Zilmer
Department of Biochemistry, Medical Faculty, University of Tartu, The Centre of Excellence for Translational Medicine,19 Ravila St, 50411 Tartu, Estonia

In this research glutathione (GSH, L-γ-glutamyl-L-cysteinyl-L-glycine) analogues (UPFs) has been synthesized. These compounds are more stable and effective antioxidants than GSH. Glutathione analogue UPF1 is a tetrapeptide (o-methyl-L-tyrosine to the N-terminus of GSH). Investigation of the structure-activity relationships in radical scavenging *in vitro* showed that this change increased the hydroxyl radical scavenging ability 60-fold compared to GSH itself [1]. The most effective structural modification was the substitution of the γ-peptide bond with the α-peptide bond in GSH backbone [2]. UPF17 was synthesized so that original γ-glutamate was changed to the α-glutamate. The change of γ-glutamate bond to the α-glutamate bond improved remarkably the hydroxyl radical scavenging properties.
In the present study we have examined the influence of UPF1 and UPF17 on CuZnSOD activity in the K562 cells. The aim of this study was to get information about if and how the replacement of γ-peptide bond with α-peptide bond affects on behavior of UPF peptides. The results demonstrated that UPF1 increased the activity of CuZnSOD at investigated (1.0; 5.0 and 10 μM) concentration, but lower concentration showed an inhibition of the enzyme activity. The rate of the activation was concentration dependent. Contrary to UPF1, the UPF17 had an inhibitory effect and the inhibition was not concentration depend. Briefly, UPF1, which includes γ-peptide bond, activated CuZnSOD activity, but UPF17, which includes α-peptide bond, decreased the activity of CuZnSOD. Our previous research showed that GSH itself has a concentration dependent activating effect on the CuZnSOD activity of K562 cells; α-GSH has an inhibiting effect on the enzyme activity but only at the highest used (10 μM) concentrations.
The substitution of the γ-peptide bond with the α-peptide bond in GSH and its analogues backbone resulted in opposite influence on CuZnSOD activity of the K562 cells.

This work was supported by European Union through the European Regional development Fund and by targeted financing from Ministry of Education and Science of Estonia (SF0180105s08) and EFS Grant No 7856.

References:

1. K. Ehrlich, K. Ida, R. Mahlapuu, I. Oit, C. Kairane, M. Zilmer and U. Soomets. Characterization of UPF peptides, members of the glutathione analogues library, on their effects on oxidative stress-related enzymes, Free Radical Research, 2009, 43:6,572-580.
2. S. Viirlaid, R. Mahlapuu, K. Kilk, A. Kuznezov, U. Soomets and J. Järv. Mechanism and stoichiometry of 2,2-diphenyl-1-picrylhydrazyl radical scavenging by glutathione and its novel α-glutamyl derivate, Bioorganic Chemistry, 2009, 37, 126-132.

IN SILICO SUPPORT TO THE DISCOVERY AND OPTIMIZATION OF POTENT INHIBITORS OF 11β-HYDROXYSTEROID DEHYDROGENASE TYPE I

Cele Abad-Zapatero,[a] Denis Carniato,[c,‡] Sophie Bozec,[c,¤] Christine Charon,[c,¦] Serge Christmann-Franck,[1,b] Liliane Doare, [c] Per Hillertz ,[d] Franck Lepifre,[c,¦] Djordje Musil,[d] Didier Roche, [c,¶] Fabien Schmidlin, [c,§] Eric Valeur.[c,†]

[a] Center for Pharmaceutical Biotechnology, University of Illinois at Chicago
[b] Merck Serono S.A., 9, Chemin des Mines, 1202 Geneva, Switzerland
[c] Merck-Serono, 4 Avenue du Président Mitterrand, 91380 Chilly-Mazarin, France
[d] Merck-Serono, Frankfurter Strasse 250, 64293 Darmstadt, Germany.

[¶] Currently at Novartis Pharma AG, WSJ 152-4.06.6 CH-4002 Basel, Switzerland
[¦] Currently at MetaBrain Research, 4 Avenue du Président Mitterrand, 91380 Chilly-Mazarin, France
[‡] Currently at Cytomics Pharmaceuticals, 86 rue de Paris, Bâtiment Mélèze, 91400 Orsay, France
[¤] Currently at Poxel Pharma, 200 avenue Jean Jaures, 69007 Lyon, France
[§] Currently at Ipsen, 5 avenue du Canada, 91966 Les Ulis Cedex, France
[¶] Currently at Edelris, 115 avenue Lacassagne, 69003 Lyon, France

Hydroxysteroid dehydrogenases regulate ligand-binding to diverse steroid hormone receptors at the pre-receptor level, like a switch, converting steroid hormones to inactive or precursor forms and vice-versa.[1] Among these, 11β-hydroxysteroid dehydrogenases 1 and 2 regulate glucocorticoid chemistry, oxidizing the 11β-hydroxyl position or reducing the 11-oxo group, which determines whether or not steroids are able to bind to glucocorticoid (and mineralocorticoid) receptors.[2]

The type 1 11β-hydroxysteroid dehydrogenase isoform (11β-HSD1) catalyses intracellular regeneration of active glucocorticoids from inert precursors (cortisone, 11-dehydrocorticosterone) by 11-oxoreductase activity in liver, adipose tissue, brain, skeletal muscle, vascular smooth muscle cells and other organs.[3] Several transgenic and in vivo models performed in rodents have documented the physiological importance of this glucocorticoid converting enzyme in insulin action. In particular, 11β-HSD1 knockout mice were shown to resist hyperglycaemia and to have increased hepatic insulin sensitivity.[4] These promising studies, as well as others,[5] strongly suggest that 11β-HSD1 inhibition is a potential therapeutic target for a broad range of disorders that could be improved by decreased intracellular glucocorticoid levels (i.e., cortisol), including the obesity-induced metabolic syndrome. In view of this, many efforts towards the discovery and development of selective 11β-HSD1 inhibitors have been carried out.[6]

We present here the in silico support to hits characterisation after a high-throughput screening campaign, from series selection to hits potency optimization. That includes:

- the use of a new framework to analyze SAR data by combining Ligand Efficiency Indices in terms of potency per size and potency per polarity; combination of both indices in an efficiency plane provides a graphical representation that can aid in comparing series and developing medicinal chemistry optimization strategies
- the consistent support of medicinal chemistry by ways of structure-based drug design.

1. Nobel, S.; Abrahmsen, L.; Oppermann, U. Eur. J. Biochem. 2001, 268, 4113.
2. Tomlinson, J. W.; Walker, E. A.; Bujalska, I. J.; Draper, N.; Lavery, G. G.; Cooper, M. S.; Hewison, M.; Stewart, P. M.

Endocrinol. Rev. 2004, 25, 831.
3. (a) Brereton, P. S.; van Driel, R. R.; Suhaimi, F.; Koyama, K.; Dilley, R.; Krozowski, Z. Endocrinology 2001, 142, 1644; (b) Bujalska, I. J.; Kumar, S.;
Stewart, P. M. Lancet 1997, 349, 1210; (c) Jellinck, P. H.; Pavlides, C.; Sakai, R. R.; McEwen, B. S. J. Steroid. Biochem. Mol. Biol. 1999, 71, 139; (d) Moisan, M. M.; Seckl, J. R.; Edwards, C. R. W. Endocrinology 1990, 127, 1450; (e) Ricketts, M. L.; Verhaeg, J. M.; Bujalska, I. J.; Howie, A. J.; Rainey, W. E.; Stewart, P. M. J. Clin. Endocrinol. Metab. 1998, 1325; (f) Whorwood, C. B.; Mason, C. B.; Ricketts, M. L.; Howie, A. J.; Stewart, P. M. Mol. Cell. Endocrinol. 1995, 110, R7.
4. Morton, N. M.; Holmes, M. C.; Fiévet, C.; Staels, B.; Tailleux, A.; Mullins, J. J.; Seckl, J. R. J. Biol. Chem. 2001, 276, 41293.
5. (a) Schweizer, R. A. S.; Zürcher, M.; Balazs, Z.; Dick, B.; Odermatt, A. J. Biol. Chem. 2004, 279, 18415; (b) Small, G. R.; Hadoke, P. W. F.; Sharif, I.; Dover, A. R.; Armour, D.; Kenyon, C. J.; Gray, G. A.; Walker, B. R. Proc. Natl. Acad. Sci. U.S.A. 2005, 102, 12165; (c) Sandeep, T. C.; Yau, J. L. W.; MacLullich, A. M. J.; Noble, J.; Deary, I. J.; Walker, B. R.; Seckl, J. R. Proc. Natl. Acad. Sci. U.S.A. 2004, 101, 6734; (d) Yau, J. L. W.; Noble, J.; Kenyon, C. J.; Hibberd, C.; Kotelevtsev, Y.; Mullins, J. J.; Seckl, J. R. Proc. Natl. Acad. Sci. U.S.A. 2001, 98, 4716.
6. For a review see Webster, S. P.; Pallin, T. D. Expert Opin. Ther. Patents 2007, 17, 1407.

PREDICTION OF THE ANOMERIC CONFIGURATION, TYPE OF LINKAGE AND RESIDUES IN DISACCHARIDES FROM 1D ^{13}C NMR DATA

Florbela Pereira

REQUIMTE/CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

Carbohydrates play key roles in many biological processes, such as cell-cell recognition (adhesion and immunological reactions), regulatory and disease processes. Their biological activity is mainly due to its surface properties, which depend on their structure and conformation.

In the structural analysis of carbohydrates, NMR spectroscopy is the most commonly used technique. The complete analysis carbohydrates is a complex, time-consuming process that usually makes use of a variety of 2D techniques, such as ^{1}H-^{1}H TOCSY and DQF-COSY, ^{1}H-^{1}H NOESY, ^{1}H-^{1}H ROESY, ^{13}C-^{1}H HSQC and ^{13}C-^{1}H HMQC or HMBC.

We have explored a machine learning approach for the prediction of the anomeric configuration, residues and type of linkages in disaccharides using ^{13}C NMR chemical shifts. For these studies, 154 pyranosyl disaccharides were used that are dimers of the α- or β-anomers of D-glucose, D-galactose or D-mannose residues bonded through α or βglycosidic linkage of types 1→2, 1→3, 1→4 or 1→6, as well as methoxylated disaccharides.

The ^{13}C NMR chemical shifts of the training set (112 disaccharides) were calculated using the program CASPER[1] and chemical shifts of the test set (42 disaccharides) were experimental values obtained from the literature. Two basic approaches were used for the encoding of the chemical shifts (independent variables) – as a sequence of chemical shifts sorted in ascending order, and as a sequence of chemical shifts sorted according to their assignment.

The two anomeric configurations of the 154 disaccharides result in four classes, (α,α), (α,β), (β,α) and (β,β) corresponding to the stereochemistry of the glycosidic linkage and the free end of disaccharide, respectively. The three monomers generate eight classes, Glc-Glc, Man-Man, Gal-Gal, Glc-Man, Man-Glc, Glc-Gal, Gal-Glc, and Man-Gal.

Experiments were performed for 1) the classification of anomeric configuration, 2) classification of type of linkage, and 3) classification of residues. Classification trees could correctly classify 67%, 74%, and 48% of the test set for the three tasks, respectively, on the basis of unassigned chemical shifts. The results for the same experiments using Random Forests were 93%, 90% and 68%.

The results indicate that machine learning techniques can be trained to predict the anomeric configuration and type of linkage for disaccharides from the unassigned list of ^{13}C chemical shifts with high accuracy.

1 P.E. Jansson, R. Stenutz and G. Widmalm, *Carbohydr. Res.,* **2006**, 341, 1003-1010.

PRELIMINARY NMR STUDIES OF A PROKARYOTIC LIGAND-GATED ION CHANNEL (LGIC) MONOMER

C.T. Chasapis[1], A. Argyriou[1], D. Bentrop[2], N. Bocquet[3], P.-J. Corringer[3], G.A. Spyroulias[1]

[1] Department of Pharmacy, University of Patras, GR-26504, Patras, Greece.
[2] Institute of Physiology II, University of Freiburg, D-79104 Freiburg, Germany.
[3] Pasteur Institute,G5 Group of Channel-Receptor, CNRS URA 2182, France

Pentameric ligand-gated ion channels from the Cys-loop family are of special importance for the rapid chemo-electrical transduction [1–3], but the mechanisms of ion permeation and gating of these membrane proteins remain elusive. Recently the X-ray structures of two prokaryotic homologues of the LGIC family most studied member, the nicotinic acetylcholine receptor (nAChR) have been determined. The first is the bacterial Gloeobacter violaceus pentameric ligand-gated ion channel homologue4 (GLIC) studied at 2.9 Å resolution in an apparently open conformation [2] and the second is the bacterium Erwinia chrysanthemi (ELIC) pentamer, studied at 3.3 Å resolution defining a closed conformation of the channel [3]. Interestingly, the extracellular soluble domain of GLIC is found to be in monomeric state in solution.

The 200-residues extra-cellular polypeptide fragment of GLIC, was cloned and expressed in high yields in E. Coli. The ^{1}H-^{15}N HSQC exhibits signal dispersion typical for polypeptides that adopt mainly beta structure. The uniformly labeled ^{13}C and ^{15}N GLIC is now studied using multi-nuclear and multi-dimensional NMR Spectroscopy and the 20% of the backbone $^{13}C/^{15}N$ nuclei have been assigned [4]. However, a complete assignment of the backbone and side-chain nuclei is difficult to be performed for a polypeptide of this size ($\sim$25 kDa) and deuteration is usually required. Various samples of recombinant proteins have been prepared where replacement of the ^{1}H nuclei by deuterium (^{2}H) varies from 60 to $\sim$100% leads to the attenuation of the relaxation rates of NMR active nuclei and concomitant sensitivity and resolution gains in spectra. Production of deuterated, triple labeled ^{13}C, ^{15}N and ^{2}H, samples suitable for NMR studies has been used for the acquisition of triple-resonance NMR spectra and the so-far analysis of the data have allowed the identification of 30-40% of the backbone resonance assignment. Molecular biology techniques are currently in progress for achieving selective labeling of certain protein residues increasing thus the number of identified nuclei and collecting high-resolution NMR data for structure determination.

References:

1. Corringer, PJ & Changeux, JP Scholarpedia 2008, 3,3468..
2. Bocquet N, Nury H, Baaden M, Le Poupon C, Changeux JP, Delarue M, Corringer PJ. *Nature* **2009** 457, 111-4.
3. Hilf RJ, Dutzler R. *Nature* **2009** 457, 115-8
4. Chasapis CT, Bentrop D., N. Bocquet, P.-J. Corringer, Spyroulias GA. *Manuscript in preparation.*

OVEREXPRESSION, CHARACTERIZATION AND NMR STUDIES OF ANTHRAX LETHAL FACTOR CATALYTIC SITE

P.V. Gkazonis[1], G.A. Dalkas[1], C.T. Chasapis[1], A Vlamis Gardikas[2], D. Bentrop[3], G.A. Spyroulias[1]

[1]Department of Pharmacy, University of Patras, GR-26504, Patras, Greece.
[2]Biomedical Research Foundation (BRFAA), Soranou Effesiou 4, GR-11527, Athens,Greece
[3]Institute of Physiology II, University of Freiburg, D-79104 Freiburg, Germany

The most prominent virulence factor of the disease anthrax is the bacterium's lethal toxin (LeTx) and in particular a 90 kDa Zn-dependent highly specific metalloprotease called Anthrax Lethal Factor (LF)[1]. LF exhibits high proteolytic specificity towards vital cellular signal transducers, the family of mitogen-activated protein kinase kinases (MAPKKs) cleaving them close to their N-termini, thus disrupting their ability to interact with and phosphorylate downstream substrates. The overall effect is alteration of signalling pathways and ultimately apoptosis[2]. Moreover, the high cleavage specification of LF against these kinases, often found overexpressed in tumor cells and thus associated with cancer tumorgenesis[3], might spearhead to the development of an innovatory therapeutic treatment of nascent tumour[4] contrasting its potential use as a biological weapon.

Aiming to understand the structural-functional activity of the catalytic site of LF towards its kinase substrates, we studied the interaction in silico, performing Molecular Dynamics Simulations in eight LF-MEK/MKKs complexes[5]. Likewise, NMR spectroscopy can provide experimental evidence for atomic level insights of the structure and the dynamics of proteins and enzymes, hence two polypeptide constructs of the LF catalytic domain (106a.a. and 170a.a.) were overxpessed as recombinant proteins for NMR structural studies. The recombinant LF catalytic domains were purified and reconstituted with the Zn^{2+} ion in order to be subjected to biochemical, biophysical and structural characterization in vitro.

References

1. Pannifer, A.D.; Wong, T.Y.; Schwarzenbacher, R.; Renatus, M.; Petosa, C.; Bienkowska, J.; Lacy, D.B.; Collier, R.J.; Park, S.; Leppla, S.H.; Hanna, P.; Liddington, R.C. *Nature*, **2001**, 414, 229.
2. Duesbery, N.S.; Webb, C.P.; Leppla, S.H.; Gordon, V.M.; Klimpel, K.R.; Copeland, T.D.; Ahn, N.G.; Oskarsson, M.K.; Fukasawa, K.; Paull, K.D.; Vande Woude, G.F. *Science* **1998**, 280, 734-737.
3. Davies H, Bignell GR, Cox C. *Nature*, 417, **2002**, 949
4. Huang D, Ding Y, Luo WM, Bender S, Qian CN, Kort E, Zhang ZF, VandenBeldt K, Duesbery NS, Resau JH, Teh BT. *Cancer Res*. **2008** 1;68(1):81-8
5. Dalkas G.A., Papakyriakou A., Vlamis-Gardikas A., Spyroulias G.A., *Protein Sci*. **2009** 18, 1774-85

3-SUBSTITUTED PHENYLALANINES, SELECTIVE LIGANDS FOR AMPA- AND KAINATE RECEPTOR SUBUNITS

Serafimoska Zorica[1], Kapedanovska Nestororvska Aleksandra[1], Suturkova Ljubica[1], Szymańska Ewa[2], Pickering Darryl[3], Tommy N. Johansen[4], Karla Frydenvang[4]

[1]Institute for Pharmaceutical chemistry, faculty of Pharmacy, "St Cyril and Methodius" University Skopje, [2] Department of Chemical Technology of Drugs, Faculty of Pharmacy, Jagiellonian University Medical College, Kraków, Poland, Department of Pharmacology and Pharmacotherapy[3] and Medical chemistry[4], Faculty of Pharmaceutical Sciences, University of Copenhagen, Denmark

Ionotropic glutamate receptors (iGluRs) constitute a family of ligand gated ion channels subdivided in three classes, NMDA, AMPA (iGluR1-4) and KA (iGluR5-7 and KA1,2) according to the agonists that selectively activate them. iGluRs are tetrameric assemblies of highly homologous receptor subunits. They are critically important for normal brain function and are considered to be involved on neurological disorders and degenerative diseases such as schizophrenia, Alzheimer's disease, brain damage following stroke and epilepsy.

AMPA receptor antagonists are considered to have clinical potential as neuroprotective drug candidates. None of the competitive AMPA receptor antagonists known today is able to discriminate between individual AMPA and KA subunits. In order to identify the structural determinants for receptor selectivity between homomeric AMPA and KA receptors a serial of rigid as well as flexible biaromatic alanine derivatives carrying selected hydrogen bond acceptors and donors has been synthesized based on the published X-ray structure of competitive antagonist (S)-ATPO co-crystallized with the ligand-binding domain of iGluR2 (S1S2J). The compounds were tested in radioligand binding studies on recombinant iGluR1-7 receptors. Based on these results as well as on the results obtained from molecular modeling studies, important structure-activity relationships at AMPA and GluR5 were established. A group of compounds selective for either GluR5 or AMPA receptors were identified. One particular substituent position appeared to be of special importance for control of ligand selectivity.

In the current project, a crystal structure of a new competitive antagonist (S)-EL-7 in complex with iGluR2 (S1S2J) has been obtained. This X-ray structure provides interesting structural information on the water-mediated interactions between the ligand and particularly one amino acid residue, Tyr702. Tyr702 is the only non-conserved amino acid residue within the ligand binding pocket among the four AMPA receptor subunits and has been identified as an important determinant for AMPA receptor agonist subunit selectivity. Furthermore, the domain closure obtained by (S)-EL-7 is within the range 6.9°-9.2°. This range of domain closure is in between what has been observed previously for antagonists (ATPO: 2.5°-5.1°) and partial agonists (KA:13°).

References:

Hogner, A. et al., Competitive antagonism of AMPA receptors by ligands of different classes: crystal structure of ATPO bound to GluR2 ligand-binding core, in comparison with DNQX. J.Med Chem.46, 214-221 (2003);

Szymańska E, Pickering DS, Nielsen B, Johansen TN. 3-Substituted phenylalanines as selective AMPA- and kainate receptor ligands. Bioorg Med Chem.;17(17):6390-401 (2009).

DOCKING, 3D-QSAR, AND MD SIMULATIONS FOR THE DESIGN OF "CLICKED" MMP-2 INHIBITORS

Anna Puckowska[1,2], Marta Bruczko[1,3], Jose Maria Zapico[1], Pilar Serra[1], Sonsoles Martin-Santamaria[1], Beatriz de Pascual-Teresa[1], Ana Ramos[1]

[1]Department of Organic Chemistry, Faculty of Pharmacy, University of San Pablo CEU, Madrid, Spain,
[2]Department of Organic Chemistry, Faculty of Pharmacy, Medical University of Bialystok, Poland,
[3]Department of Medicinal Chemistry, Faculty of Pharmacy, Medical University of Bialystok, Poland

Matrix metalloproteinases are a family of zinc-dependent endopeptidases participating in a variety of physiological and pathological processes, involved in inflammatory, malignant and degenerative diseases. Since MMP-2 has been reported as one of the MMPs with a major role in cancer, there is a great need for the design of potent and highly selective inhibitors of this enzyme [1].

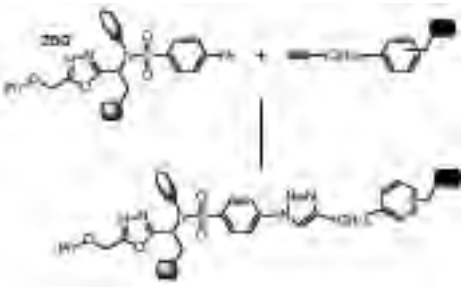
Effective MMP inhibitors (MMPIs) should be characterized by: 1) a functional group chelating catalytic zinc ion (zinc binding group, ZBG), 2) one or more side chains undergoing strong van der Waals interactions with the enzyme subsites, mainly with S1', responsible for MMP selectivity, 3) functional groups providing additional interactions with enzyme backbone [2].

Among the reported MMPIs, hydroxamic acid derivatives constitute the main family of selective MMP-2 inhibitors. In order to understand ligand-protein binding mode, and to define the structural requirements of hydroxamate bound to MMP-2, we have used different computational techniques like docking, 3D-QSAR, and molecular dynamics simulations. These studies allowed us to select the most appropriate side chains in terms of MMP-2 selectivity. Moreover, a new heterocyclic ZBG derived from an oxadiazole system was designed. Selected side chains and oxadiazole ZBGs have been connected through triazole linkers, by means of click reactions [3], leading to new MMP-2 inhibitors.

Figure1.

Docked structure of a newly designed compound

Scheme 1.



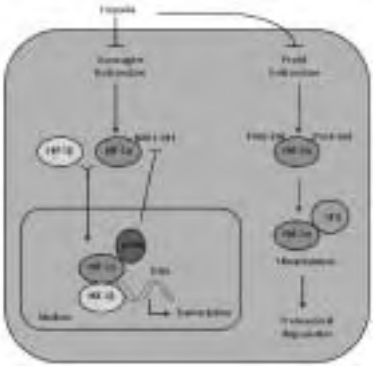
Click chemistry approach

[1] Murphy, G.; Nagase, H. *Molec. Asp. Med.* **2008**, 29, 290-308.
[2] Tingjun, H.; Wei, Z.; Xiaojie Xu. *Journal of Computer-Aided Molecular Design* **2002**, 16, 27-41.
[3] Kolb, C. K.; Finn, M. G.; Sharpless, K. B.; *Angew. Chem. Int. Ed*. **2001**, 40, 2004-2021.

THE DISCOVERY OF NEW HIF-1 INHIBITORS THROUGH MOLECULAR MODELING STUDIES

Antonino Lauria, Licia Pantano, Marco Tutone, Anna Maria Almerico
Dipartimento Farmacochimico, Tossicologico e Biologico
Università di Palermo, Via Archirafi 32, 90123 Palermo, Italy

Hypoxia, a frequent effect of solid tumor growth in head and neck cancer and in other types, serves to generate a cascade of molecular pathways which include angiogenesis, glycolysis, and various cell-cycle control proteins. These cell-salvaging mechanisms can be carried out rapidly by a transcription factor that reacts to hypoxic conditions, the hypoxia-inducible factor-1 (HIF-1) [1].
HIF-1 is a heterodimer consisting of an α and a β subunit, which dimerizes and binds to DNA. Whereas HIF-1β is constitutively present, HIF-1α is highly unstable, except under low-oxygen conditions. Under hypoxic conditions, the HIF-1α subunit accumulates, and is translocated to the nucleus. Here, upon dimerization with HIF-1β and interaction with cofactors, such as p300/CBP, the HIF-1 complex binds to hypoxia-responsive elements (HRE) and activates transcription. HIF-1α contains two transcriptional activation domains: the N-terminal transactivation domain (NTAD) and the C-terminal transactivation domain (CTAD) [2]. The primary function of the CTAD is to recruit widely employed transcriptional coactivators p300/CBP. Under normoxic conditions, this interaction is blocked when an asparagine residue (N803) within the CTAD is hydroxylated by factor inhibiting HIF-1 (FIH1) [3]. Under hypoxic conditions, the hydroxylation is abrogated, allowing CTAD to bind to a domain in p300/CBP that is known as the cysteine/histidine-rich1 (CH1). A recent study demonstrated that peptide inhibitors of the HIF-1α/p300 interaction caused suppression of tumor growth in vivo [4]. Thus, preventing or terminating HIF-1α transactivation has the potential to interfere with tumor growth.



In this work we performed docking studies with the GLIDE [5] software to evaluate selective inhibitors of HIF-1α p300 interaction. HIF-1α protein (PDB code: 1L3E) was preliminary pre-optimizated by PROTEIN PREPERATION WIZARD [6] of Schrodinger Inc.The p300 key amino acid residues were used as template to create a molecular database with ZINC. The more than 99000 obtained molecular structures were used in the following docking step, using the VIRTUAL SCREENING WORKFLOW of GLIDE. This one combines into one job all the steps of screening a library using Glide-ligand preparation, filtering, and docking in increasing accuracy, from HTVS through SP to XP.

References

[1] Semenza G.L., Wang G.L. (1992) Mol. Cell. Biol. 12: 5447-5454.
[2] Ruas J.L., Poellinger L., Pereira T. (2002) J. Biol. Chem. 277: 38723-38730.
[3] Lando D., Peet D.J., Whelan D.A., Gorman J.J., Whitelaw M. L. (2002) Science 295: 858-861.
[4] Kung A.L., Wang S., Klco J.M., Kaelin W.G., Livingston D.M. (2000) Nat. Med. 6: 1335–1340.
[5] Glide , Schrodinger, LLC, New York, NY, 2005.
[6] Schrodinger, LLC, New York, NY, 2005.

PHARMACOPHORIC MODEL AND 3D-QSAR OF 5-RESORCINOL-ISOXAZOLE DERIVATIVES AS HSP90 INHIBITORS

Andrea Ciacci[1], Giuseppe Giannini[1], Walter Cabri[1], Tiziana Brunetti[1], Domenico Alloatti[1], Daniele Simoni[2], Riccardo Baruchello[2], Giuseppina Grisolia[2], Cinzia Rullo[2], Loredana Vesci[1], Massimo Castorina[1], Paolo Carminati[3], Claudio Pisano[1]

[1]Sigma-Tau, S.p.A., Pomezia, Italy; [2]Università di Ferrara, Ferrara, Italy; [3]sigma-tau Research Switzerland S.A., Mendrisio, Switzerland

Molecular chaperones are a class of proteins that are responsible for the correct folding and maturation of several cellular client proteins.[1] Heat shock protein 90 (Hsp90) is the most abundant and ubiquitous molecular chaperone expressed both in eukaryotic and prokaryotic cells.

The prevalence of a high-affinity form of Hsp90 in tumor cells, the "addiction" of cancer cells to oncogenic client proteins and their greater dependence on Hsp90 has been proposed as rationales for selectivity of Hsp90 inhibitors for cancer versus normal cells. The interest of Hsp90 as validated target involves several diseases besides cancer, like neurodegenerative alteration, viral, fungal, and microbial infection.[2,3]

After pioneering studies with natural products, where geldanamycin and radicicol were found to bind to the ATP binding site of the N-terminal domain, many selective Hsp90 inhibitors from various institutions have entered in clinical trials. One of these classes of inhibitors is the 4,5-diaryl-pyrazole/isoxazole, originally discovered by high-throughput screening at the Centre for Cancer Therapeutics, which subsequently entered into a collaborative discovery program with Vernalis, culminating in the derivative NVP-AUY922 currently in clinical trials. [4]

We have mainly focused our research on the 4,5-diaryl-pyrazole/isoxazole scaffold, until now considered a prerequisite for the activity on Hsp90. We have found that compounds possessing a nitrogen atom, or better an amide, on the 4 position of isoxazole ring, are endowed with potent Hsp90 inhibitor properties similar or even better than the diaryl- analogues, probably due to a favorable interaction in the active site binding protein.

This series of compounds have pIC50 values ranging from 4.5 to 7.7 against Hsp90 as measured in a fluorescence polarization (FP) competitive binding assay, in concert with the inhibition proliferation activity against human cancer cell line (NCI-H460) and associated to a characteristic profile of depletion of oncogenic proteins and concomitant elevation of Hsp70.[5,6]

In order to rationalize structure-activity data, a 3D-QSAR analysis was performed by using the PHASE module of Schrodinger suite.[7] A hypothesis generation step came after, with a grid-based 3D QSAR method, in which the grid positions of atoms in molecules overlaid to the hypotheses are correlated to their activities using a partial-least-squares (PLS) fitting approach. The pharmacophore contains a H-bond donor-acceptor region, a lipophilic moiety and an aromatic portion. The generated 3D-QSAR model showed that the compounds could be aligned into a common orientation to produce a QSAR model with good explanatory and predictive power. Different alignment protocols (e.g. template alignment and docking) were also explored.

1) Whitesell, L. *et al. Nat. Rev. Cancer* **2005**, 5, 761.
2) Solit, D. B. *et al. Drug Discov. Today* **2008**, 13, 38.
3) Neckers, L. J. *et al. Bioscience* **2007**, 32, 517.
4) Eccles, S. A. *et al. Cancer Research* **2008**, 68, 2850.
5) Aryl isoxazole compounds with antitumoral activities: WO2010/000748A1
6) G.Giannini, *et al. AACR 101st Annual Meeting* **2010** - Abstract #731
7) Dixon, S. L. *et al. J. Comput. Aided Mol. Des.* **2006**, 20, 647. www.schrodinger.com

INVESTIGATION OF CYTOTOXIC AND APOPTOTIC EFFECTS OF BORON COMPOUNDS ON LUNG ADENOCARCINOMA (A549) CELLS

[1]Tunçay Tuğba, [1]Arslan A. Selda, [1]Ergene Emel

[1]Anadolu University, Faculty of Science, Dept. of Biology, Eskisehir, Turkiye

Recent studies indicated that boron compounds has a therapeutic role in various cancer types. Boron compounds have been widely used as biologically active agent and drug. The effects of the antitumoral properties of the boron compounds on the different cancer cell lines were shown by the various investigations. In this study, we investigated the effects of the certain boron compounds (boric acid, sodium perborate and diammonium tetraborate) both on the activation of the apoptotic pathway and cytotoxicity in lung adenocarcinoma (A549) cells by the using MTT and Neutral Red, Mitochondrial Membran Permeability (Mitocapture), Caspase 3 and 8 Enzyme Activation and DNA Fragmentation methods. We incubated and tested the A549 cells in the ranges of the 50-2,5 mM concentrations of boric acid and sodium perborate and diammonium tetraborate. The cell proliferation was measured after 24, 48 and 72h, using by the MTT and NR Assays. According to these results, the cytotoxic effect was found for boric acid and sodium perborate at a concentration at 5 mM, for diammonium tetraborate at 10 mM. According to Mitochondrial Membran Permeability (Mitocapture) assay, apoptotic effects were found for boric acid 80%, for sodium perborate 65% and for diammonium tetraborate 58% of the total cells. We reached to same results by the using, Caspase Enzyme Activation and DNA Fragmentation methods.

References:

Scorei, R., Ciubar, R., Ciofrangeanu, C. M., Mitran, V., Cimpean, A., and Iordachescu, D. Biological Trace Element Research, V.122, Number 3, June, 2008.

Dibas A., Howard J., Anwar S., Stewart D., Khan A. Biochem. Biophys. Res. Commun.270:383-386;2000

RANGE AND SENSITIVITIES OF 2-[(CARBOXYMETHYL)SULFANYL]-4-OXO-4-ARYLBUTANOIC ACIDS PROPERTY SPACES. PART 2. MULTIDIMENSIONAL FREE ENERGY LANDSCAPES

Branko J. Drakulić,[a*] Alessandro Pedrretti,[b] Mire Zloh,[c] Vladimir Slavnić,[d] Ivan O. Juranić,[e] Milan M. Dabović [a]

[a] Department of Chemistry - Institute of Chemistry, Technology and Metallurgy, University of Belgrade, Njegoševa 12, Belgrade, Serbia; [b] Dipartimento di Scienze Farmaceutiche "Pietro Pratesi", Universita degli Studi di Milano, Via Mangiagalli, 25, I-20133, Italy; [c] The School of Pharmacy, University of London, 29/39 Brunswick Square, London WC1N 1AX, U.K.; [d] Scientific Computing Laboratory, Institute of Physics Belgrade, Pregrevica 118, 11080 Belgrade, Serbia;
[e] Faculty of Chemistry, University of Belgrade, Studentski Trg 12-16, Belgrade, Serbia

The 2-[(carboxymethyl)sulfanyl]-4-oxo-4-arylbutanoic acids (CSAB, Figure 1) exert antiproliferative potency toward all NCI sixty human tumor cell lines[1 a)] in submicromolar to low micromolar concentrations.[1 b)] In this communication we are focused on the observed selectivity of compounds, comparing potency toward human cervix carcinoma cells (HeLa) vs. healthy human cells.[2] Previously, selectivity of selected compounds was correlated with properties derived from their conformational assemblies. Using the concept of ranges and sensitivities of property spaces[3] we demonstrated that the range of apolar surface areas, obtained by conformational search in vacuum (OPLS2005 FF), was well correlated with the selectivity giving bilinear correlation.[4] Here, we have extended our findings on the whole prepared set. In this way, along with statistical significance of correlation obtained, compounds can be classified from highly potent/highly selective via highly potent/moderately selective to those that exert less significant potency (IC_{50} in low μM concentrations) and relatively low selectivity. To put results on a solid theoretical background, 20 ns molecular dynamics (MD) simulations (CHARMM FF) were performed for all compounds in the explicit
n-octanol/water (4: 1) and in the explicit water. During simulations, a biasing force was applied over three collective variables (geometric parameter that measure progression of conformational change) to obtain free energy landscape of compounds.[5 a)] The enhanced sampling of MD trajectories was obtained, as well as the estimate of free energy gradient for the whole conformational space of the each compound in a given medium. The good correlation obtained between selectivity of compounds and the time needed for conformational change from the 'bent' to the 'extended' in n-octanol/water should be considered as significant; because the time of the whole simulation enables uniform sampling along each applied reaction coordinate. Next, few local minima appear on 3D maps of the free energy landscape of the compounds having both high potency and high selectivity, opposite to those exerting lower potency and selectivity. This sheds light on, at the first sight somewhat paradoxical, observation that the more flexible compounds exerts both higher selectivity and better potency. It should be noted that differences in

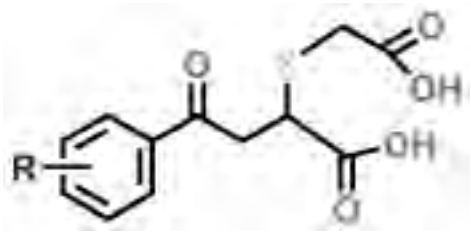


Figure 1. Structure of the title compounds.

flexibilities are due only to different substitution patterns on the phenyl ring.
To the best of our knowledge, this is the first, and the first successful application of the concept of the ranges and sensitivities of the property spaces on the selectivity data; as well as their first corroboration by physically well-grounded concept of the adaptive biasing force calculations,[5a, b] that give good distinction between the metastable and stable conformational states. On the other hand, derived results are very applicable for the further design of the title class of compounds.

Acknowledgement: Corresponding author gratefully acknowledge authors of the reference 5a and Axel Kohlmeyer for many useful suggestions on ABF implementation in NAMD 2.7; and is in debt to High performance computing facilities of the Institute of Physics, Belgrade. Ministry of Science and Technological Development of Serbia support this work. Grant 142010.

References:

[1] a) R. H. Shoemaker, *Nat. Rev. Cancer* **6** (2006) 813, b) Unpublished results; [2] B. J. Drakulić, Z. D. Juranić, T. P. Stanojković, I. O. Juranić, *J. Med. Chem.* **48** (2005) 5600; [3] a) G. Vistoli, A. Pedretti, L. Villa, B. Testa, J. *Am. Chem. Soc.* **124** (2002) 7472, b) G. Vistoli, A. Pedretti, B. Testa, *Drug Discov. Today* **13** (2008) 285; [4] B. J. Drakulić, Z. D. Juranić, T. P. Stanojković, M. F. Zloh, I. O. Juranić, The XVI European Symposium on Quantitative Structure-Activity Relationships and Molecular Modelling, *Book of Abstracts* (2006) pp. 274-275; [5] a) J. Hénin, G. Fiorin, C. Chipot, M. L. Klein, *J. Chem. Theory Comput.* **6** (2010) 35, and references cited therein, b) NAMD 2.7b2 User guide, http://www.ks.uiuc.edu/Research/namd/2.7b2/ug/.

INTERACTION ENERGY BETWEEN THYMIDINE PHOSPHORYLASE AND ITS INHIBITOR BY THE FRAGMENT MOLECULAR ORBITAL (FMO) METHOD

Yukio Tada[1], Masami Uebayasi[2], Dmitri G. Fedorov[3], Kazuo Kitaura[3,4]

[1]The University of Tokyo, [2]Tokyo Institute of Technology, [3]National Institute of Advanced Industrial Science and Technology, [4]Kyoto University, Japan

2'-Deoxy-5-trifluoromethyl-uridine(F_3dThd) is reported to have strong antitumor activities. However, F3dThd is metabolized to an inactive compound 5-trifluoromethyl-uracil by human thymidine phosphorylase(hTP). In order to prevent such inactivation we have developed a thymidine phosphorylase inhibitor(TPI) having K_i=20nM, which is now on phase II trial as TAS-102, based on homology model of hTP and by classical QSAR method.[1] As hTP is known to undergo conformational change from an open form to a closed active form, we have modeled the complex TPI cation with TP two active closed forms accompanied with and without phosphate anion(PO4-)(Fig. 1) based on crystal structure of pyrimidine nucleoside phosphorylase(1BRW).

Fig. 1. TPI cation with PO4- model, TPI cation without PO4- model, TPI zwitterion without PO4- (1UOU)

The reported crystal structure of hTP in complex with TPI(1UOU) is a closed form without PO4-, and the ionic state of TPI has been estimated to be the zwitterion. We have calculated the complex of TPI with hTP by the fragment molecular orbital (FMO) method [2,3] to investigate the state of TPI. There are 6 histidine residues in hTP. When all 6 histidine are not protonated, the cation model A has a lower interaction energy than zwitterion model A. On the other hand, the interaction energies of cation model B and zwitterions model B are almost same when outer 5 histidine residues are protonated except His116 in the active site of hTP. Then we are going to calculate in the case of TPI cation with PO4-.

Table 1. Caluculation system and interaction energy with TPI in gas phase

Ionic state	TPI charge	protein charge	6-31(+)G* (kcal/mol)
cation model A	1	-5	-165.26
cation model B	1	0	-70.35
zwitterion model A	0	-5	-68.70
zwitterion model B	0	0	-72.62

1. Fukushima, M., Suzuki, N, Emura, T. Yano S., Kazuno H., Tada, Y., Yamada Y., Asao T, Biochem Pharmacol 2000, 59; 1227-1236.
2. General FMO review: D. G. Fedorov, K. Kitaura, J. Phys. Chem. A 111 (2007) 6904.
3. FMO-QSAR review: T.Yoshida, T.Fujita, H.Chuman, Curr.Comp.-Aided Drug Des. 5 (2009) 38

A GENERAL QSAR METHODOLOGY TO PREDICT TELOMERASE INHIBITION BY G-QUADRUPLEX STABILIZATION

Daimel Castillo-González,[ab] Gisselle Pérez-Machado,[b] Guillermin Agüero-Chapín,[b] Aliuska Morales-Helguera,[ab] Marta Teijeira Bautista[c] and Miguel Ángel Cabrera-Pérez[b]

[a]Faculty of Chemistry and Pharmacy. Universidad Central de Las Villas.Santa Clara. 54830. Cuba

[b]Molecular Simulation & Drugs Design Group. Centro de Bioactivos Químicos (CBQ) Universidad Central de Las Villas. Santa Clara. 54830. Cuba

[c]Department of Organic Chemistry. Faculty of Chemistry. University of Vigo. Vigo 36310. Spain

The inhibition of telomerase is considered one of the most important targets in the cancer therapy. Among the different mechanisms of telomerase inhibition, the stabilization of telomere by the G-quadruplex formation is considered a promising alternative to avoid the elongation action of this enzyme. In order to predict the telomerase inhibitory potency a computational approach was applied to a non-congeneric 562 telomerase inhibitors by the G-quadruplex stabilization. Eighteen classification models were obtained by Lineal Discriminant Analysis (LDA) with a global percentage of good classification close to 80%. Four quantitative models achieved by Multiple Linear Regression (MLR) showed correlation coefficients greater than 80%. All models were widely validated using several statistic techniques. A virtual screening was performed on two different datasets: (1) 2500 compounds reported with different pharmacological properties and (2) "in house" dataset of 36400 compounds. The first screening predicted 17 compounds as active, 10 out of these had been already reported with antitumoral activity. The second one, identified 85 compounds as G-quadruplex ligands. These results demonstrated the quality of the QSAR models supporting their relevance in the antitumoral drug discovery process. The predictive models are being used to detect G-quadruplex ligands among commercial drugs. As preliminary screening, a fast and inexpensive experimental spectrophotometric method is being used to filter the better candidates.

PREDICTION OF TELOMERASE INHIBITORY ACTIVITY FOR TRIAZINE DERIVATIVES BASED ON CHEMICAL STRUCTURE

Daimel Castillo González[a], Gisselle Pérez Machado[b], Marta Teijeira Bautista[c] and Miguel Angel Cabrera Pérez[b]

[a]Faculty of Chemistry and Pharmacy. Universidad Central de Las Villas.Santa Clara. 54830. Cuba
[b]Molecular Simulation & Drugs Design Group. Centro de Bioactivos Químicos (CBQ) Universidad Central de Las Villas. Santa Clara. 54830. Cuba
[c]Department of Organic Chemistry. Faculty of Chemistry. University of Vigo. Vigo 36310. Spain

The reverse transcriptase enzyme telomerase is responsible for maintaining telomeric DNA length in over 85% of cancer cells by catalyzing the synthesis of further telomeric repeats. However, their expression is switched off after embryonic differentiation in most normal cells. Thus, telomerase is regarded as a promising novel molecular target for cancer treatment and different strategies to inhibit telomerase have been developed. One of them is to stabilize a quadruplex structure for the telomere end. Triazines, small planar aromatic molecules, have been widely described as telomerase inhibitors by stabilization of G-quadruplex.

Quantitative structure-activity relationships (QSARs) provide a useful tool for defining a mathematical relationship between chemical structure and biological activity, and for applying such statistically derived models for predicting the activity of untested chemicals.

In the present work was developed a combined QSAR strategy to predict the activity against telomerase (IC_{50}), by G-quadruplex stabilization, and cytotoxicity in A549 cells. A dataset of 219 triazines was used to predict IC_{50} values applying LDA and piecewise as statistical techniques. On the other hand, 63 triazines were employed to predict cytotoxicity on A549 cell line and MLR was used as statistical methods. With the LDA model the percentage of good classification was close to 80%. The sensibility, specificity and accuracy of the model were 78.82%, 78.12%, 79.12%, respectively. These results were complemented with a piecewise regression (R^2= 76.3). For cytotoxicity on A459 cells were achieved good results of fitness and prediction of the MLR model (R^2=79.5 and Q^2=76.47). Finally, all models were used to evaluate a dataset of patented triazines with action against telomerase by G-quadruplex stabilization mechanism. The models were able to explain the relationship between descriptors and activity, evidencing the powerful of QSAR methods in the design of novel compounds with inhibitory activity against telomerase.

QSAR APPROACH FOR THE MODELING AND DESIGN OF ANTITUMORAL Pt(IV) COMPLEXES

Ester Papa[1], Paola Gramatica[1], Mara Luini[1], Elena Monti[2], Marzia B. Gariboldi[2], Mauro Ravera[3], Elisabetta Gabano[3], Luca Gaviglio[3], Domenico Osella[3]

[1] Department of Structural and Functional Biology, University of Insubria, QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Via Dunant 3, 21100 Varese, Italy, [2]Department of Structural and Functional Biology, University of Insubria, Section of Pharmacology, Via A. da Giussano 10, 21052 Busto Arsizio (VA), Italy, and [3]Department of Environmental and Life Sciences, University of Piemonte Orientale, Viale Michel 11, 15121, Alessandria, Italy

A large series of Pt(IV) complexes, used as antitumor for ovarian and colon human carcinoma, has rationally designed[1,2] by the quantitative structure-activity relationship (QSAR) modeling approach and synthesized to try to develop a predictive QSAR model. In this study Multiple Linear Regression (MLR) models, based on experimental (i.e. reduction peak potential E_p and partition coefficient log $P_{o/w}$) and theoretical descriptors[3], were developed for a drug design approach. This set of descriptors was used as an input set for modelling[4], in order to identify different structural features of Pt(IV) complexes related to the in vitro cytotoxicity (i.e. half maximal inhibition concentration, IC_{50}, on A2780 ovarian and HCT116 colon carcinoma cell lines) and the above-said features of Pt(IV) complexes.

In the resulting best models, a lipophilic descriptor (i.e log $P_{o/w}$ or number of secondary sp^3 carbon atoms, nCs) plus an electronic descriptor (E_p, number of oxygen atoms, nO, or total polar surface area, TPSA(NO)) are necessary for the modeling. This result supports the general findings that the biological behavior of Pt(IV) complexes is related to their uptake, reduction, and to the structure of the corresponding Pt(II) metabolites.

Finally, a QSAR model based only on theoretical molecular descriptors is proposed. This model can be applied for the prediction of the studied endpoints also when the experimental descriptors, i.e. log $P_{o/w}$, and E_p, are not available.

References:

1. Hall, M.D.; Mellor, H.R.; Callaghan, R.; Hambley, T.W. Basis for disign and development of Platinum (IV) anticancer complexes, J.Med.Chem. 2007,50,3403-3411

2. Montaña, A.M.; Batalla, C. The rationa design of anticancer platinum complexes: the importance of the structure-activity relationship, Curr. Med. Chem., 2009, 16, 2235-2260.

3. Todeschini,R.; Consonni, V.; Mauri, A.; Pavan, M. DRAGON Software for the Calculation of Molecular Descriptors, version 5.5 for Windows; Talete S.r.l.; Milan, Italy, 2008; software available online at: http://www.talete.mi.it/.

4. Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M. MOBYDIGS: software for regression and classification models by genetic algorithms. In *Nature-inspired methods in chemometrics: genetic algorithms and artificial neural networks; Leardi, R., Ed.; Elsevier*: Amsterdan, 2003; pp 141-167.

MOLECULAR MODELING OF COMPOUNDS STIMULATING THE ACTIVITY OF PSA

Henna H. Härkönen[1], Erik Wallén[2], Antti Poso[1], Maija Lahtela-Kakkonen[1]

[1] School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland

[2] Division of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Helsinki, Finland

Prostate cancer (PCa) is the most common cancer of males in industrialized countries. Its incidence has increased significantly due to increased PSA-screenings and aging of the population. Common drug therapies of PCa have significant side effects affecting the patient's quality of life and eventually some of the drugs become inefficient. In all, novel drugs for the treatment of PCa are awaited.

Prostate-specific antigen (PSA) is a widely used marker for PCa screening and it has also proven to be useful in the diagnosis and monitoring of PCa. However, the exact physiological role of PSA in cancer remains unclear. Experimental studies have shown PSA to possess cancer promoting abilities as well as to prevent prostate cancer progression. This makes PSA an intriguing target for drug development.

Cyclic peptides consisting of 10-13 amino acids that bind specifically to PSA and stimulate its activity have been developed (Wu et al 2000, Koistinen et al 2008). The peptides have been shown to inhibit tube formation in angiogenesis (Mattson et al 2008). Compounds which stimulate PSA activity and in turn prevent angiogenesis in tumor cells have been suggested to be beneficial in fighting cancer progression. However, poor physicochemical properties, low bioavailability and degradation by proteolytic enzymes, typical for peptides present limitations for their usage as drugs. To overcome these problems we aim to develop non-peptidic compounds, which mimic the binding and biological effects of the peptides.

Structural studies by NMR spectroscopy and mutagenesis studies of the peptides have shed light on the structure-activity relationships of the peptides (Pakkala et al 2004). Our results from conformational analysis of the peptides, molecular dynamics simulations of PSA and molecular docking of the peptides with PSA were combined with the experimental data. This enabled us to generate 3D pharmacophore models capturing the essential features of the peptides for biological activity. The pharmacophores have been applied in virtual screening of compound databases to find potential lead compounds. The hit molecules from database screening were filtered based on their drug-likeness and examined using alignment-free molecular descriptors and principal components analysis. A representative set of distinct molecules were selected for biological screenings in vitro. The most potent compounds stimulating the activity of PSA are selected as lead compounds for further development of novel, potential compounds for the treatment of prostate cancer.

References:

1. Wu P et al. Eur. J. Biochem. 267:6212-20, 2000
2. Koistinen et al. Biol. Chem. 389:633-42, 2008
3. Mattson JM et al. Prostate. 68:945-54, 2008
4. Pakkala et al. J. Peptide Sci. 10:439-47, 2004

MODULATION OF AFFINITIES FOR BETA-TUBULIN IN A LARGE SERIES OF TAXANE DERIVATIVES STUDIED BY COMPARATIVE BINDING ENERGY (COMBINE) ANALYSIS

Claire Coderch[a], Javier Klett[b], Antonio Morreale[b], Angeles Canales[c], Fernando Dvaz[c], Wei-Shuo Fang[d], Federico Gago[a]

[a]Departmento de Farmacologva, Universidad de Alcalα, 28871 Alcalα de Henares, Spain.
[b] Unidad de Bioinformαtica, Centro de Biologia Molecular S.O. (CSIC-UAM), 28049 Madrid, Spain.
[c]Centro de Investigaciones Biologicas, CSIC, 28040 Madrid, Spain.
[d]Institute of Materia Medica, Chinese Academy of Medical Sciences, Beijing 100050, China.

The first atomic- detail structure of paclitaxel (PXL, Taxol) bound to beta- tubulin was deposited with the Protein Data Bank (PDB) in 1998[1] but the precise binding conformation of this antitumor drug is still not fully clear due to limited resolution of the experimental data. Refinement of this structure using molecular dynamics and energy minimization provided us with an updated model[2] that we used as a template for the modeling of the complexes of tubulin with cephalomannine (CPH), Docetaxel (DXL, *Taxotere*), and a large series of 44 synthetic taxanes (CTX). To assess the validity of the proposed binding modes and establish Quantitative Structure-Activity Relationships (QSAR) we have used the COMparative BINding Energy (COMBINE) approach, as recently implemented in the gCOMBINE program[3]. To this end we decomposed the calculated binding energy for each complex, using the AMBER[4] Force field (parm03), into a summatory of residue- based electrostatic and van der Waals contributions plus two additional variables describing the electrostatic Contributions to the desolvation energy of both the ligand and the receptor, using the continuum method implemented in DelPhi[5]. Protonation of the side chain of His229, as well as inclusion of two water molecules in the complexes, proved crucial for deriving a robust SAR model: a correlation coefficient (r^2) of 0.94 and a correlation coefficient after cross- validation (q^2, using random groups of 5 elements per group) of 0.86 for a model with only 5 latent variables. The only three outliers had in common the presence of an amide linker between the baccatin core and the C2 substituent. Interestingly, these compounds showed some of the largest entropic contributions to the free energy of binding. A hypothesis supported by NMR data will be provided to account for this behavior and to highlight the possible limitations of enthalpy- based QSAR methods.

References

[1] Nogales, E. et *al*. *Nature* 1998, 391, 199-‑-203.
[2] Matesanz, R. et al. *Chem Biol*. 2008, 15, 573-‑-585.
[3] Gil-Redondo, R. et al. *Proteins* 2010, 78,162-‑-72.
[4] URL: http://ambermd.org/
[5] Nicholls, A., et al. J. *Comput. Chem*, 1991, 12, 435-445.

2D-QSAR COMPARATIVE STUDIES ON CHALCONE DERIVATIVES WITH ANTICANCER ACTIVITY

Repanas A., Katsori A.-M. & Hadjipavlou-Litina D.
Department of Pharmaceutical Chemistry, School of Pharmacy, Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece

Chalcones or 1,3 diaryl-2-propen-1-one are open analogues of flavonoids in which the two aromatic rings are connected by a three carbon α, β unsaturated carbonyl system. In nature are precursors in flavonoids biosynthesis and have been isolated from various parts of plants **[1]**. Also they can be obtained by several synthetic procedures, among them and the most common by base-catalyzed Claisen-Schmidt condensation **[2]**. One of the most important characteristic of these compounds is that they exhibit a number of biological activities, such as antiviral, antiangiogenic, anti-inflammatory. Furthermore in the literature have been reported a variety of this type of derivatives that posses anticancer and anti-proliferative activity, suggesting also many different mechanisms of action **[3]**.
Herein we perform a 2D-QSAR study from bibliographic anticancer data from different chalcones analogues, using the C-QSAR program of Biobyte. The QSAR analysis presented here is an attempt to organize the knowledge on the chalcones anticancer and anti-proliferative activity with the purpose of designing new chemical entities with enhanced inhibitory potencies and to study the mechanism of action of the compounds. This study revealed that lipophilicity is one of the most important determinants of activity. Additionally, steric factors such as the overall molar refractivity (CMR), molar volume (MgVol), the substituent's molar refractivity (MR) are important. Electronic parameters are not found to be present.

1. Dimmock, J. R., Elias, D. W. Curr. Med. Chem., **1999**, 6, 1125-49.
2. Liu, M., Wilairat, P., Go. M. L. J. Med. Chem., **2001**, 44, 4443-52.
3. Katsori, A-M., Hadjipavlou-Litina, D. Curr. Med. Chem., **2009**, 16, 1062-81.

*We would like to thank **Dr. C. Hansch and Biobyte Corp.** 201 West 4th Street, Suite 204, Claremont, CA 91711, USA for free access to the C-QSAR program.*
*Katsori A-M is thankful to **"Bodossakis Foundation"** for PhD scholarship.*

POSTER PRESENTATION II-53

INSIGHT INTO IMPORTANT INHIBITOR-ENZYME INTERACTIONS AND STRUCTURAL BASIS FOR IMPROVING POTENCY OF ARYLAMIDE DERIVATIVES AS NOVEL DIRECT INHIBITOR OF InhA, BASED ON COMPUTER–AIDED MOLECULAR DESIGN APPROACHES

Auradee Punkvang[1], Patchreenart Saparpakorn[2], Supa Hannongbua[2], Peter Wolschann[3] Anton Beyer[3] and Pornpan Pungpo[1,*]

[1]Department of Chemistry, Ubon Ratchathani University, 85 Sthollmark Rd., Warinchamrap, Ubonratchathani, 34190, Thailand

[2]Department of Chemistry, Kasetsart University, Chatuchak, Bangkok, 10900, Thailand

[3]Institute for Theoretical Chemistry, University of Vienna, A-1090 Vienna, Austria

Arylamides is the novel inhibitors of the enoyl ACP reductase enzyme (InhA), involving in the type II fatty acid biosynthesis pathway of *M. tuberculosis*. Their remarkable property, inhibiting directly the InhA enzyme without requiring any coenzyme, makes them highly appropriate for the design of new antibacterials. In order to identify the specific interactions of arylamide derivative in the InhA binding pocket, molecular docking calculations and were performed on a set of arylamides. The docked results show a good ability to reproduce the X-ray bound conformation with rmsd of less than 1.0 Å. The analysis of arylamide binding modes could provide significant insight into inhibitor-enzyme interaction of arlylamide derivatives in the InhA binding pocket. Three QSAR approaches were performed using HQRAS, CoMFA and CoMSIA techniques. With statistically satisfied models, the obtained QSAR results show high correlation of the molecular structure properties of arlylamide derivatives with their biological activities. Molecular dynamics (MD) simulations were performed to get better insight into the binding of arylamide derivatives in the InhA binding pocket. Moreover, the estimate binding free energies of arylmides in InhA were also calculated using linear interaction energy (LIE) method. The obtained results are successful to simulate the binding modes of arylamides in InhA binding pocket.These results indicate that arylamides are tightly held in InhA binding pocket by these hydrogen bonds. The estimate free binding energy of each arylamide shows corresponding well with its inhibitory activity. Therefore, the integrated results from structure-based, ligand-based design approaches and MD simulations provide insight into drug-enzyme interactions and are helpful for better understanding the binding mechanism of arylamide derivatives in InhA binding pocket. Consequently, the derived structural information can be fruitful guideline for the design of new and more potentially effective antitubercular agents.

STUDY ON THE BINDING MODE OF AZANAPHTHOQUINONE ANNELATED PYRROLE DERIVATIVES AS ANTI-CANCER AGENT IN DNA STRADS USING MOLECULAR MODELING

P. Pungpo[1*], A. Punkvang[1], C. Kheawsod[1], N. Kuntiya[1], A. Kumkong[1], P. Saparpakorn[2], N. Pongprom[1], P. Wolschann[3] and S. Hannongbua[2]

[1]Department of Chemistry, Faculty of Science, Ubonratchathai University, Ubonratchathani, Thailand
[2]Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok, Thailand
[3]Institute for Theoretical Chemistry, University of Vienna, A-1090 Vienna, Austria

Azanaphthoquinone annelated pyrrole derivatives act as the synthetic DNA intercalating agent. The binding modes of these anti-cancer agents have not been established. In order to elucidate the potential binding modes of these compounds in DNA binding site, molecular docking calculations with Glide program were employed. Moreover, quantum chemical calculations based on MP/6-31G(d) method were also applied to explore the key residues of the molecular interactions between these anti-cancer agents and DNA chain. The results show that molecular docking calculations enable to model the potential binding modes of azanaphthoquinone annelated pyrrole derivatives with RMSD values less than 1 Å. Based on molecular docking results, the azanaphthoquinone scaffolds are inserted between hydrophobic sides of two base pairs of DNA chain. The pi-pi interactions between aromatic systems of the azanaphthoquinone scaffolds and two base pairs are crucial for binding of these anti-cancer agents. In addition, strong hydrogen bond interactions between oxygen of deoxyribose and the substituent in azanaphthoquinone scaffold of anti-cancer agents are also observed. The quantum chemical calculation results are consistent with the obtained results from docking calculations. The interactions between azanaphthoquinone annelated pyrrole derivatives and two base pairs of DNA chain show high attraction energy. Therefore, the integrated results from both approaches should be helpful for understanding the binding modes and the crucial interactions of azanaphthoquinone annelated pyrrole derivatives in DNA binding site.

References

[1] N. Pongprom, G.M ller, P. Schmidt, W. Holzer, H. Spreitzer, *Monatsh Chem.* 140 (2009), pp.309–313.
[2] K. Shanab, N. Pongprom, E. Wulz, W. Holzer, H. Spreitzer, P. Schmidt, B. Aicher, G. Muller, E. Guntherb, *Bio. Med. Chem. Lett.,* 17 (2007), pp.6091–6095

THE MOLECULAR BASIS OF THE RECOGNITION OF NATURAL PRODUCTS DERIVED FROM Stypopodium zonale BY PROTEIN TYROSINE KINASE p56^{Lck}

Nelilma Correia Romeiro, Juliana Stróligo Pecly, Angélica Ribeiro Soares
Universidade Federal do Rio de Janeiro-UFRJ/NUPEM, Brazil

Drugs isolated from natural products represent important and innovative molecular skeletons in the search of new drugs. A few examples of drugs derived from marine natural products have been described, some of which have been launched or are in clinical tests[1-5]. Atomaric and stipoquinonic acid are terpenes isolated from brown algae *Stypopodium zonale (Dictyotaceae)* that have been described as inhibitors of the tyrosine kinase p56^{lck}, a potential target in lymphoid malignancies[6], representing prototypes of antitumor agents, with similar IC_{50} values[7]. This data has encouraged the investigation of the putative binding modes of these marine natural products and other metabolites, potential tyrosine kinase inhibitors isolated from *Stypopodium zonale* by our group, with p56^{lck} [8] . Figure 1 A-D shows the top scored complexes of atomaric acid and one of the most promising metabolites with p56^{lck} as obtained by docking with FlexX software[9], and superimposition with imatinib, co-crystallized with p56^{lck}.

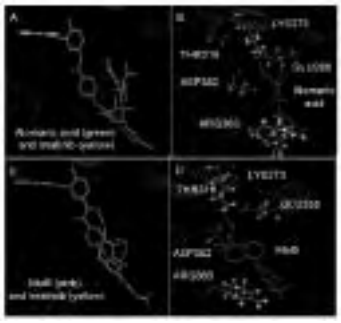


Figure 1. A, C - Molecular superimposition of atomaric acid and molecule 5, respectively, with imatinib, in the binding site of p56^{Lck} as obtained by docking with FlexX; B, D – enlarged vision of the top scored complexes of atomaric acid and Molecule 5 with p56^{Lck}.

The docking results have shown that molecule 5, one of the metabolites isolated by our group, fits the binding site better than atomaric acid due to its smaller volume and has a better theoretical free energy of binding than this substance (-19.40 vs. -14.32 kJ/mol) and stipoquinonic acid (-14.32 kJ/mol), supporting molecular modifications by derivatization, aiming at improving the binding energy and selection for biological assays.

1)- Terlau, H. & Olivera, B. M. *Physiol. Rev*. **2004**, 84, 41; 2)- Mayer, A. M. S. & Gustafson, K. R. *Eur. J. Cancer* **2008**, 44, 2357; 3)- Newman D. J. e Cragg, G. M. *J. Nat. Prod.* **2007**, 70, 461; 4)- Simone, M.; Erba, E.; Damia, G.; Vikhanskaya, F.; Di Francesco, A. M.; Riccardi, R. *et al. Eur. J. Cancer* **2005**, 41, 2366; 5)- Baunbek, D.; Trinkler, N.; Ferandin, Y.; Lozach, O.; Ploypradith, S.; e Rucirawat, F. *Marine Drugs* **2008**, 6, 514; 6)- Harr, M. W.; Caimi, P. F.; McColl, K. S.; Zhong, F.; Patel, S. N.; Barr, P. M.; Distelhors, C. W. Cell Death Differ. **2010** Mar 19. [Epub ahead of print]; 7)- Wessels, M.; Koniig, G. M. e Wright, A. D. J. *Nat. Prod*. **1999**, 62, 927; 8)- Jacobs, M. D.; Caron, P. R.; Hare, B. J. *Proteins* **2008**, 70, 1451; 9)- Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A. *J. Mol. Biol.* **1996**, 261, 470.

SCHARACTERIZING INTERACTIONS BETWEEN THE PI3 KINASE, P110α, ATP BINDING SITE AND A SERIES OF IMIDAZO [1,2-a]PYRIDINE-BASED INHIBITORS

Jack U. Flanagan, Andrew J. Marshall and William A Denny

Auckland Cancer Society Research Centre, Faculty of Medical and Health Sciences, The University of Auckland, Grafton, Auckland, 1000, New Zealand

Phosphoinositide-3-kinases phosphorylate the 3-OH group of phosphoinositides, generating lipid mediators of a diverse range of physiological processes including glucose homeostasis, cell growth, differentiation, and motility. Vertebrates possess eight unique PI3Ks, divided into three classes. Characterisation of the class IA p110α isoform gene in cancer has demonstrated that it undergoes amplification and mutation leading to gain of function and oncogenicity. A diverse range of chemotypes are known to inhibit p110α, with several in clinical trial as new anti-cancer agents.[1]. Most of the information on interactions between the class IA PI3K active site and these inhibitors is derived from crystallographic studies using the related p110α isoform.[2]. The data illustrated that the inhibitors work by targeting the ATP binding site, blocking nucleotide binding. Here, we explore the ability of different docking protocols to sample and score poses that reflect this experimental data within the recently determined atomic model of the p110α ATP binding site.[3,4]. Currently, there is no experimentally derived structural data that informs on the binding mode for imidazo[1,2-a]pyridine type compounds exhibiting p110α selectivity, with insights to date obtained from simulation only.[5,6]. We characterize binding modes for these imidazo[1,2-a]pyridines within the p110α crystal structure sampled by docking, using CoMFA to identify the mode that best describes the biochemical data, and gain insight into the structural and chemical features that influence binding affinity. The effect on these 3D QSAR models of different molecular mechanic energy minimization protocols, with and without a Poisson-Boltzmann implicit solvent model, is also followed.

1. Marone, R., Cmiljanovic, V., Giese, B. and Wymann, MP. (2008) Biochim. Biophys. Acta 1784:159-185.
2. Knight, ZA., Gonzalez, B., Feldman, ME., Zunder, ER., Goldenberg, DD., Williams, O., Loewith, R., Stokoe, D., Balla, A., Toth, B., Balla, T., Weiss, W., Williams, RL. and Shokat, KM. (2006) Cell 125:1-15.
3. Huang, C-H., Mandelker, D., Schmidt-Kittler, O., Samuels, Y., Velculescu, VE., Kinzler, KW., Vogelstein, B., Gabelli, SB. And Amzel, LM. (2007) Science 318:1744-1748.
4. Mandelker, D., Gabelli, SB., Schmidt-Kittler, O., Zhu, J., Cheong, I., Huang, C-H., Kinzler, KW., Vogelstein, B., Amzel, LM. (2009) PNAS 106:16996-17001.
5. Frederick, R. and Denny, WA. (2007) J. Chem. Inf. Model. 48:629-638.
6. Han, M. and Zhang, JZH. (2010) J. Chem. Inf.Model. 50:136-145.

CONSTRUCTION OF THE TRK-1 TYROSINE KINASE BY HOMOLOGY MODELING AND COMPARISON OF ATP-BINDING SITES OF IR, IGF-1R AND TRK-1

Jan H. Löhr and Thomas Lemcke
University of Hamburg, Institute of Pharmacy, Bundesstr. 45, 20146 Hamburg, Germany

Receptor tyrosine kinases (RTKs) have emerged as promising targets in cancer therapy. They play an important role in transforming extracellular signals from growth hormones into cellular responses - regulating growth, differentiation and cell survival. Increased expression of certain RTKs has been associated with various human tumor conditions.[1]

The search for additional antitumor targets has led to the insulin-like growth-factor receptor 1 (IGF-1R) and the neurotrophic tyrosine kinase receptor type 1 (TrK-1).[2,3] IGF-1R plays a critical role during normal growth and development regulating a variety of cellular processes such as proliferation, survival and angiogenesis.[4] However, increased levels of expression have been linked to various malignancies and its inhibition proved to be an effective anti-cancer intervention.[2,4] Similarly, the TrK-1 receptor contributes to normal neural development controlling differentiation and apoptosis, and its overexpression has been associated with tumors of the nervous system as well as pancreatic and prostate carcinomas.[5,6]

Their close relationship to the insulin-receptor (IR) with its central metabolic functions emphasizes the need for selectivity. Published inhibitors are not quite selective enough and their development has been hindered because more insight into the differences of the ATP-binding sites is required.

In order to compare the ATP-binding sites of the IR, IGF-1R and TrK-1 kinases, a three-dimensional homology model for TrK-1 had to be constructed, since experimental structure data is missing to date. Attempting to reflect the different impact of receptor tyrosine kinase conformation on inhibitor selectivity, a homology model representing both the active and inactive conformation was calculated. On the basis of high sequence identity and close phylogenetic relationship[7], IR[5], IGF-1R and MuSK were selected as suitable templates. We use molecular interaction fields (MIFs) to characterize potential binding interactions that could be exploited for the design of selective inhibitors and that would help to find the correct binding mode of already known inhibitors. We investigate the structural differences between the ATP-binding sites by calculating molecular interaction fields using selected GRID-probes.[8] The process of building a TrK-1 homology model and the results of the comparison will be presented.

References:

1. A. Arora, E. M. Scholar, *J. Pharmacol. Exp. Ther*. **2005**, 315(3), 971-979.
2. R. Li, A. Pourpak, S. W. Morris, *J. Med. Chem*. **2009**, 52(16), 4981-5004.
3. A. Nakagawara, *Cancer Lett.* **2001**, 169(2), 107-114.
4. R. D. Hubbard, J. L. Wilsbacher, *ChemMedChem* **2007**, 2(1), 41-46.
5. B. Lippa, J. Morris, M. Corbett, T. A. Kwan, M. C. Noe, S. L. Snow, T. G. Gant, M. Mangiaracina, H. A. Coffey, B. Foster, E. A. Knauth, M. D. Wessel, *Bioorg. Med. Chem. Lett*. **2006**, 16(13), 3444-3448.
6. A. M. Camoratto, J. P. Jani, T. S. Angeles, A. C. Maroney, C. Y. Sanders, C. Murakata, N. T. Neff, J. L. Vaught, J. T. Isaacs, C. A. Dionne, I*nt. J. Cancer* **1997**, 72(4), 673-679.
7. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, *Science* **2002**, 298, 1912-1934.
8. P. J. Goodford, *J. Med. Chem.* **1985**, 1985, 849-857

IDENTIFICATION OF INHIBITORS OF THE TYROSINE KINASE c-Met BY STRUCTURE-BASED VIRTUAL SCREENING

Thomas Lemcke,[1] Matthias Rarey,[2] Jan Dreher,[2,4] Frank Totzke,[3] Christoph Schächtele,[3] Michael H. G. Kubbutat,[3] and Conrad Kunick [4]

[1]University of Hamburg, Institute of Pharmacy, Bundesstr. 45, 20146 Hamburg, Germany

[2]University of Hamburg, Center of Bioinformatics, Bundesstr. 43, 20146 Hamburg, Germany

[3]ProQinase GmbH, Breisacher Str. 117, 79106 Freiburg, Germany

[4]Technische Universität Braunschweig, Institut für Pharmazeutische Chemie, Beethovenstr. 55, 38106 Braunschweig, Germany

The hepatocyte growth factor receptor tyrosine kinase, commonly called c-Met, is associated with cell proliferation, cell survival, motility, invasion and morphogenesis. There are reports about the essential role of c-Met in normal embryonic development and organogenesis. In adults, however, it is mainly involved in tissue damage repair and regeneration. Triggering the c-Met pathway results in the activation of several downstream signaling pathways, like the RAS/MAPK, the PI3K and AKT.

In tumor cells, c-Met signaling is strongly dysregulated, due to mutations, overexpression and amplification of the receptor. Thus activated c-Met promotes tumor cell growth, migration and invasion, tumor angiogenesis and protection from apoptosis.[1,2] Abnormal c-Met expressions have been observed in biopsies of most solid tumors and this documents the correlation between c-Met signaling and human malignancies like bladder, breast, cervical, colorectal, gastric, liver, lung, ovarian, pancreatic, prostrate, renal and thyroid cancers.[3] Targeting the c-Met kinase domain with selective inhibitors has emerged as a promising way to correct for this overactivated c-Met signaling and thereby to treat special human tumors.

In order to identify new and selective small molecule inhibitors of the c-Met tyrosine kinase, which bind to the ATP-binding site of the kinase, we initialized a structure-based virtual screening. The starting point was a subset of more than 2 million purchasable compounds taken from the ZINC database.[4] These compounds were prefiltered using a modified "Rule of Five" criterion[5] and subsequently docked into the ATP-binding site of a c-Met structure of 1R0P6 using the FlexX[7,8] docking program. Two compound sets were selected from the resulting docking poses and visually inspected. After elimination of chemically improbable or identical poses, 57 compounds were selected and subsequently tested in an IC50 profiling against 8 different protein kinases, including c-Met. Two of the tested compounds showed low µM inhibitory activity against c-Met and one additional kinase and thus are promising candidates for the development of lead structures for the design of selective c-Met inhibitors.

In this poster, the virtual screening process and the resulting ATP-binding site inhibitors of the c-Met will be presented.

References:

1. Liu, X.; Newton, R. C.; Scherle, P. A. *Trends Mol. Med*. **2009**, 16, 37.
2. Eder, J. P.; Vande Woude, G. F.; Boerner, S. A.; LoRusso, P. M. *Clin. Cancer Res.* **2009**, 15, 2207.
3. Peruzzi, B.; Bottaro, D. P. *Clin. Cancer Res.* **2006**, 12, 3657.
4. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Mod.* **2005**, 45, 177.
5. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, 23, 3.
6. Schiering, N.; Knapp, S.; Marconi, M.; Flocco, M. M.; Cui, J.; Perego, R.; Rusconi, L.; Cristiani, C. *Proc. Natl. Acad. of Sci. USA* **2003**, 100, 12654.
7. Hindle, S. A.; Rarey, M.; Buning, C.; Lengauer, T. *J. Comput.-Aided Mol. Des*. **2002**, 16, 129.
8. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, 261, 470.

POSTER PRESENTATION II-59

IMPLEMENTATION OF HIGH-LEVEL OF THEORY COMPUTATIONAL APPROACHES FOR THE RATIONALIZATION OF INTERACTIONS BETWEEN KINASES AND INHIBITORS. A PRESSING NECESSITY?

Vassilios Myrianthopoulos and Emmanuel Mikros*
Division of Pharmaceutical chemistry, School of Pharmacy, University of Athens, Panepistimiopolis Zografou, 15771, Hellas

Kinase function is a collective biological determinant of normal cell life. The correlation between aberrant kinase activity and disease is now well established and abnormal kinase function is recognized as a major cause of serious pathological states like neurodegeneration, inflammatory responses, as well as various forms of cancer. As a result, kinases have emerged as attractive and highly prioritized drug targets. Research and development programs focusing on kinase inhibitors currently count up to 25% of total projects worldwide. To this end, efforts employing structure-based methods in the design of novel inhibitory compounds have been generously favored by SAR studies based on experiment (X-ray crystallography) or theory (Molecular simulations, QSAR, chemoinformatics). Undoubfully, structural insight offered by such theoretical studies is of outstanding importance for the rationalization and subsequent development of specific, highly potent kinase inhibitors. However, there are several cases where a notable lack of agreement between biological results and theory exists. Theory fails to reproduce experimental results, implying that several of the factors underlying the physical systems are being neglected by theory, at least at the particular level implemented in the computational study.

Three representative cases demonstrating the efficacy of sophisticated simulation approaches as opposed to the inadequacy of lower level calculations are presented in this study. The first concerns the specificity reversal within a series of highly similar indirubin analogues towards the pair of protozoal kinases L*m*GSK3 and CRK3 of *Leishmania donovanii*. The addition of a methyl group on the indirubin scaffold while enhancing affinity for the former, it showed the reverse effect on the latter, which is a highly homologous kinase. The specificity reversal was not reproducible by empirical models correlating experimental affinities and binding energies. On the contrary, the impementation of Free Energy Perturbation simulations did result in a highly satisfactory agreement with biological data.

In the second case, the conversion of a benzyl amine moiety to the corresponding phenyl was found to improve affinity of a series of roscovitine analogues towards Casein kinase-1. In a similar fashion, docking-scoring calculations and the corresponding empirical models failed to reproduce the 10-fold increase in the binding affinity of the phenyl analogues. Elucidation of the observed difference was only achieved by the utilization of high level ab-initio calculations. A second order Møller-Plesset perturbation theory (MP2) interaction energy study afforded a remarkably accurate interaction energy prediction and an excellent agreement with biological results concerning the aforementioned system of closely related roscovitine analogues.

Finally, in the third case, a small variation on its substitution pattern converted a highly potent GSK3b kinase inhibitor based on the indirubin scaffold into a totally inactive compound, which demonstrated an affinity decrease of no less than 5 orders of magnitude. The variation concerned the conversion of the 6-bromoindirubin-3'oxime to its 7-bromo analogue. A qualitative explanation was rather simple to conclude; diminished potency was expected to originate from a steric clash formed between the bulky bromine at position -7 and the kinase backbone. Still, the attempts to quantify the observed difference were

unsuccessful, as the 10000-fold difference could not be reproduced. The use of MP2 ab initio calculations strongly indicated that the affinity decrease could not be solely attributed to the differentiated ligand orientation. Thus, an energetic contribution possibly resulting from the displacement of a structural water molecule was calculated on the basis of MD simulations and the double-decoupling method and consideration of this extra term resulted in a fairly good reproduction of the observed biological results.

POSTER PRESENTATION II-60

VALIDATION OF A MULTI-RANKING APPROACH OF VIRTUAL SCREENING AS A TOOL FOR CHARACTERIZING NOVEL KINASE INHIBITORS. THE CASE OF KINASE DYRK1α

Vassilios Myrianthopoulos[1], Evangelos Gikas[1], Laurent Meijer[2] and Emmanuel Mikros*[1]
[1]Division of Pharmaceutical chemistry, School of Pharmacy, University of Athens, Panepistimiopolis Zografou, 15771, Hellas [2]Cell Cycle Lab, CNRS, Roscoff cedex, BP 74, 26982, Bretagne, France

The dual-specificity tyrosine phosphorylation-regulated (DYRK) kinases belong to a family comprising seven isoforms. Several studies have revealed that DYRK kinases play key roles in cell proliferation and apoptosis induction. In addition, they have been implicated in many key cellular processes such as regulation of protein synthesis and differentiation. DYRK1a is implicated in the development of the nervous system. Its overexpression in Down syndrome (DS) fetal brains initiated several studies, the outcome of which indicated strong correlations between DYRK1a aberrant function and DS pathology. As DYRK kinases and particularly the -1a isoform are emerging as novel targets for chemical biology and possibly for therapeutics as well, the characterization of selective inhibitors can be of particular interest.
A compound library was screened for DYRK1a inhibitors. The library consisted of the NCI Diversity set-II compound collection (1346 compounds), which was enriched with natural products (~350 compounds) originating from the inventory of the Pharmacognosy lab (University of Athens). This afforded a collection of ~1700 molecules demonstrating a high degree of chemodiversity and drug-like properties. Virtual screening (VS) experiments were performed using a fast grid-based algorithm, while two different ranking approaches were used. In the first instance, ranking was based on scores obtained from the targeted protein. A small ensemble of homologous proteins was used in the second instance. Subsequent *in-vitro* biological evaluation of the total collection was performed in order to validate theoretical results and quantify the potency of the active molecules recovered. Screening afforded in total 21 novel, highly potent and selective DYRK1a inhibitors that demonstrated a >80% reduction of enzyme activity, while the IC_{50} values of the most potent hits were at the low μM range.
Several hits demonstrated a high degree of structural similarity with known kinase inhibitors, while others constitute novel chemotypes that could serve as leads for the development of highly original kinase inhibitors. A subsequent QSAR analysis of the studied compounds on the basis of a number of 2D, 3D and physicochemical descriptors coupled with sophisticated statistical methods was undertaken in order to identify common structural patterns and shed light on the pharmacophoric determinants of the active compounds. Concerning methodology, results showed that VS coupled with the multi-ranking scheme can effectively increase enrichment by a factor of 10 with respect to random sampling. Furthermore, the multi-ranking scheme can increase hit rate by ~65% compared to the single-target approach, thus representing a computationally feasible method of enhancing sampling within the initial docking step of VS.

STRUCTURAL BIOINFORMATICS APROACH OF CYCLIN - DEPENDENT KINASES 1 AND 3 COMPLEXED WITH INHIBITORS

L. A. Saraiva[1,2]; M. P. Veloso[2]; I. C. Rodrigues[1]; N. J. F. da Silveira[1]

[1.] Departamento de Ciências Exatas - Universidade Federal de Alfenas (Unifal-MG) Rua Gabriel Monteiro da Silva, 700 - Centro - Alfenas/MG, Brazil.

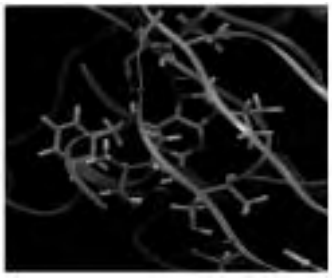
[2.] Departamento de Farmácia - Unifal-MG

The cyclin-dependent protein kinases, or CDKs, participate in regulation of both the cell proliferation cycle and the RNA polymerase-II (RNAP-II) transcription cycle.[1] Activities of CDKs are controlled by association with subunits known as cyclins and reversible phosphorylation reactions.[2] In several human tumors, deregulations of CDK-related mechanisms have been detected, e.g., overexpression of cyclins or deletion of genes encoding for CKIs.[3] Regarding these observations, CDKs came up to be interesting targets for elaboration of novel antitumor drugs.[4] Based on the importance of the CDKs, this research aims to describe, to characterize and to compare the molecular models of CDK3 and CDK1. The understanding of its structural features mainly the biding sites would lead to discovery and rationalization of drug design process.

Since the structures of human CDK1 and CDK3 are unavailable in the Protein Data Bank – PDB, models were created. A BLAST search in the PDB with human CDK1 sequence [Swissprot sequence CDC2_HUMAN (P06493)] and CDK3 [Swissprot sequence CDK3_HUMAN (P00526)] revealed 64,83% and 75,86% identity respectively with CDK2 (PDB 2r3i) whose resolution is 1,28Å. Model building of CDK1 and CDK3 was carried out using the program MODELLER[5] and the overall stereochemical quality of the final model for CDK1 and CDK3 was assessed by the program PROCHECK[6]. The structural studies of the CDK1 and CDK3 biding sites were conducted by molecular docking with 15 different CDKs inhibitors previously identified for CDK2. For the docking methodology was used the software Induced Fit Docking from the suite Schrödinger[7].

Table 1: Hydrogen bonds between the molecular forks of CDK1 and CDK3 with inhibitors
NO: Not observed

Inhibitor	Distance between inhibitor and C=O on Leu83 (Å)	Distance between inhibitor and N-H on Leu83 (Å)	Distance between inhibitor and C=O on Glu12 (Å)	Distance between inhibitor and C=O on Ile10 (Å)	Distance between inhibitor and N-H on Lys89 (Å)	Distance between inhibitor and N-H on Gln301 (Å)	GlideScore Kcal.mol^{-1}
Roscovitine CDK1	1,837	2,214	1,891	2,181	NO	NO	-10,318
Roscovitine CDK3	1,890	2,220	NO	1,719	2,007	NO	-10,521
Purvalanol B CDK1	1,977	2,358	2,047	2,197	1,889	NO	-10,328
Purvalanol B CDK3	1,794	NO	1,890	1,977	1,888	1,729	-11,933



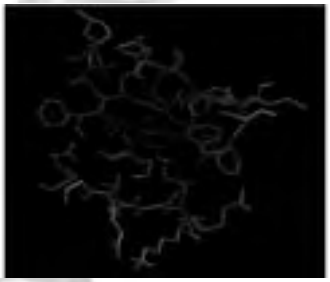


Fig 1a Fig 1b Fig 1c

Fig 1a: Inhibitor Roscovitine bound in human CDK1 (ribbons representation) showing hydrogen interactions (yellow dashed line) of the probe with CDK1. Fig 1b: Inhibitor Roscovitine bound in human CDK3 (ribbons representation) showing hydrogen interactions (yellow dashed line) of the probe with CDK3. Fig 1c: superimposed biding pockets of CDK1 - Roscovitine complex (thick representation) and CDK3 - Roscovitine complex (thin representation) (RSMD: 0,946 Å).

In conclusion, the overall structure of all complexes indicates that inhibitors Roscovitine, Purvalanol B are tightly bound to the ATP-binding pocket in both CDK1 and CDK3. Moreover, no further binding sites were identified in the CDK structures. Considering the comparison of the structures of the binary complexes with multiples inhibitors, previously identified for CDK2, with the complexes containing CDK1 and CDK3, strongly indicates that those inhibitors should inhibit CDK1 and CDK3 as well. Thus, studies in silico of different enzymes with high sequential identity with the same inhibitor may be used as preliminary studies before experimental approaches.

We would like to thank to FAPEMIG, CNPq and FINEP for financial support.

References:

1. Sausville, E. A. (2002). Complexities in the development of cyclin-dependent kinase inhibitor drugs. Trends Mol. Med. 8, 32 - 37.
2. Malumbres, M & Barbacid, M. (2005). Mammalian cyclin-dependent Kinases. Trends Biochem. Sci. 30, 630 - 641.
3. Sherr, C. J. (1966). Cancer cell cycles. Science, 274, 1672 - 1677.
4. Meijer, L. (1996) Chemical inhibitors of cyclin-dependent kinases. Trends Cell Biol. 6, 393 - 397.
5. Sali, A., Blundell, T. L. (1993). Comparative protein modeling by satisfaction of spatial restraints, J. Mol. Biol. 234, 779 - 815.
6. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structure. J. Appl. Cryst. 26, 283 - 291.
7. Schrödinger Suite (2008) Induced Fit Dockink protocol; Glide version 5.0, Schödinger, LLC, New York, NY, 2005; Prime version 1.7, Schödinger, LLC, New York, NY, 2005.

DESIGN OF HIGHLY SELECTIVE COVALENT KINASE INHIBITORS

Emeline Leproult, Dino Moras, Jean-Marie Wurtz
Department of Structural Biology and Genomics, IGBMC, Strasbourg, FRANCE

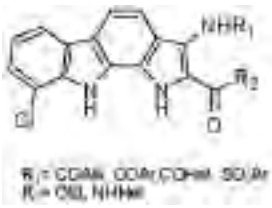
Most kinase inhibitors are rejected from clinical development phases, either because of a lack of selectivity, or appearance of drug resistance of their target. Emerging success of covalent inhibitors could solve both of these problems at the same time. In order to design such inhibitors, our main focus consists in automatically identifying all human kinases bearing a cysteine in their ATP binding site. Since the shape of the ATP binding site depends on the kinase domain conformation, we performed an important preliminary study to distinguish the different conformations among all the available human kinase structures. As a result, we designed a covalent inhibitor, derived from the non covalent drug Imatinib, that covalently binds Kit kinase by reacting with a non conserved cysteine accessible in the inactive kinase DFG-out conformation. Inhibition assays have confirmed the selectivity of this inhibitor. As a conclusion, this work gives insights on how to design a new generation of highly selective covalent kinase inhibitors.

OPTIMIZATION OF PYRROLO[2,3-a]CARBAZOLES AS CDK1 INHIBITORS. SYNTHESIS, KINASE INHIBITORY ACTIVITY, AND BINDING MODE THROUGH DOCKING SIMULATIONS

Efstathios Spyropoulos[1], Andreas Montzolis, Athanasia Chatzianastasiou[2], Athanasios Papakyriakou[3], Manolis Fousteris[1], Georgios Spyroulias[3], Andreas Papapetropoulos[2] and Sotiris Nikolaropoulos[1]

[1]Laboratory of Medicinal Chemistry, [2]Laboratory of Molecular Pharmacology, [3]Laboratory of Pharmacognosy and Chemistry of Natural Products, Department of Pharmacy, University of Patras, GR-26500, Patras, Greece

Cyclin dependent kinases (CDKs) constitute a family of serine/threonine protein kinases which are involved in the control of transcription and cell-cycle progression.[1] Deregulation of their activity is a common feature of many human tumors, various proliferative and neurodegenerative disorders.[2] Recent genetic studies reconsider the classical model of the cell cycle regulation by CDKs and reveal that CDK1 is sufficient and probably essential to drive the cell cycle in the absence of other interphase CDKs (CDK2, CDK4 and CDK6)[3], comprising an attractive target for cancer therapy.



We have recently reported on the CDK1 inhibitory activity of the pyrrolo[2,3-a]carbazole scaffold.[4] Molecular modeling and docking simulation studies explored the CDK1 binding mode of this core and especially its positioning into the ATP binding cleft.[4] CDK1 inhibitory potency or/and selectivity improvement efforts led to the synthesis of new selectively substituted pyrrolo[2,3-a]carbazoles. Preliminary biological evaluation showed that derivatives bearing 3-heteroaryl/aryl/alkyl/arylsulfon-amido substituents presented diverse CDK1 inhibitory activity. Furthermore, the conducted docking studies elucidated the possible alternative binding mode of the ligands into the ATP binding cleft and assigned their physicochemical and structural features which promote effective binding in the kinase.

References

1. Malumbres, M.; Barbacid, M. *Nat. Rev. Cancer* **2009**, 9, 153-166.
2. Lapenna, S.; Giordano, A. *Nat. Rev. Drug Discov*. **2009**, 8, 547-566.
3. Santamarıa, D.; Barriere, C.; Cerqueira, A.; Hunt, S.; Tardy, C.; Newton, K.; Caceres, J. F.; Dubus, P.; Malumbres, M.; Barbacid, M. *Nature* **2007**, 448, 811-815.
4. Fousteris, M. A.; Papakyriakou, A.; Koutsourea, A.; Manioudaki, M.; Lampropoulou, E.; Papadimitriou, E.; Spyroulias, G. A.; Nikolaropoulos, S. S. *J. Med. Chem*. **2008**, 51, 1048-1052.

4D QSAR STUDY OF p56^{lck} PROTEIN TYROSINE KINASE INHIBITORY ACTIVITY OF FLAVONOID DERIVATIVES USING MCET METHOD

Guzel, Yahya; Yilmaz, Hayriye; Boz, Mehmet; Altiparmak, Gokce
Department of Chemistry, Erciyes University, 38039 Kayseri, Turkey

The protein kinase family has been emerging as an exciting class of targets for drug discovery. They share a conserved structural similarity in the region of the ATP binding site, where most inhibitors interact. Quantitative relationships of the interaction between molecular structure and p56^{lck} protein tyrosine kinase for the inhibitory activity of 50 flavonoid derivatives were studied as 4D-QSAR by MCET method. Our main research includes choice and type of descriptors responsible from the activity with 3D-space of the conformers in the data set. The auxilary group (AG) and the anti-pharmacophore shielding (APS) besides pharmacophere (Pha) were investigated as effecting the activity. The "pharmacophere map" resulting from these groups shows the interactions made by all ligands with their receptors simultaneously. The choice of Pha map was based upon comparing and matching "electron topological matrix" (ETM) of all conformers of the understudied molecules with ones of references molecule. MCET attempts to find consistent relationships between the variations in the values of molecular properties and the biological activity. The structure-activity relationship could be described much more accurately via conformers for a series of compounds. We suggested that the outside atoms in related to Pha-centred might be probably interacted with receptor and responsible for AG and APS. These outside atoms consist of AG and APS were selected automatically besides atoms of Pha for the biological active structure. Only one atom with same position and orientation within certain tolerances for each accepted conformer was taken into account. An algorithm was described that Pha map for goodness of fit to the binding site description were obtained from the calculation. The resulting QSAR model revealed that the parameters of the pharmacophere map had significant impact on protein tyrosine kinase inhibitory activity of the compounds. All the electron topological parameters were autoloaded from ETM database, and the adjustable constants were calculated. Executing the regression analysis with the MCET program, the parameters used and the adjustable constant calculated in this report were discussed in detail along with their applications. It was shown that this model had a high statistical quality for predicting the activity of the inhibitors. To display the rightness of the model equation with 4D QSAR, r^2, the square of the correlation coefficient, q^2,the measure of quality of fit, and s,the standard deviation were used. The model based on the molecular descriptors of flavonoid derivatives was developed with high prediction accuracy for the training and test.

4D-QSAR STUDY OF ANALGESIC ACTIVITY OF 3(2H)-PYRIDAZINONE DERIVATIVES WITH MCET METHOD

Yahya Guzel*[1], Mehtap Gokce[2] and Hayriye Yilmaz[3]

[1]Erciyes University Faculty of Science Department of Chemistry Kayseri, Turkey

[2] Erciyes University Faculty of Pharmacy Kayseri, Turkey

[3] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Gazi University, Ankara, Turkey

Molecular Conformer Electron Topology (MCET) method has been used to build a fourth dimensional quantitative structure-activity relationship (4D-QSAR) for a series of 30 compounds of the 3(2H)-pyridazinone derivatives. 4D-QSAR analysis in MCET method incorporates conformational and alignment freedom into the development of 3D-QSAR models for training sets of structure-activity data by performing molecular state ensemble averaging, the fourth "dimension". In order to understand structural requirements and to predict activity, we need to develop a new computational method with 4D-QSAR.

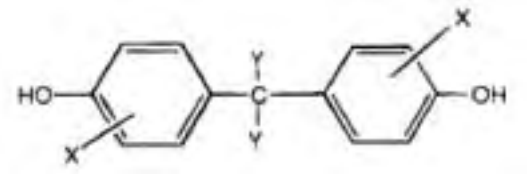
In this study, the results obtained from the MCET method have also been compared to information from the 3(2H)-pyridazinone derivatives. This is in an effort to determine structural features (steric and electrostatic) of compounds by the MCET for QSAR analysis. MCET method takes simultaneously into account for Pharmacophore (Pha), Auxiliary Group (AG) and Anti-Pharmacophore Shielding (APS) in all accepted conformers. Pha, AG and APS groups as descriptor in biological activity may correspond to structure of the 3(2H)-pyridazinone analogues and to further clarify the mode of binding of ligands. Using energetically accessible all conformations for each compound the identification of Pha, AG and APS groups were resulted in sensitive by means of the geometric parameters such as distance, angle and dihedral. In this model, κ, adjustable constant, was based upon the magnitude of the corresponding PLS regression coefficients. For a procedure which involves iterative selection of Pha poses and structures, extraction of AG and APS and construction of a statistical model, which explains the observed biological activity can be repeated until R^2 was reached to big value. Here AG or APS related to conformers of the molecules in which Pha especially exist. Three properties as descriptor were then subjected to quantitative analyses via partial least-squares (PLS) using the model. The results in the dataset were compared with corresponding MCET computed by accounting to the combine electrostatic and shape fields of Pha, AG and APS in model. For compounds of both active and inactive analogues, 4D QSAR model were predicted with high accuracy through MCET. It can be used to predict the activities and bioactive poses of new molecules. By defining the proper arrangement of atoms as the best Pha, AG and APS, the correlation coefficient (R^2) and root mean square error (RMS) between the estimated and the observed activities (0.84 and 2.48, respectively) were obtained. The result indicates that the model has good predictability.

4D-QSAR STUDY WITH MCET METHOD ON ESTROGENIC ACTIVITY OF 4,4-DIHYDROXYDIFENYLMETHANE AS BİSFENOLA (BSA) DERİVATES

Hayriye Yilmaz, Yahya Guzel, Nazmiye Geçen, Kader Şahin and M. Betul Aycan
Department of Chemitry, Erciyes University, 38039 Kayseri Turkey

Four dimensional quantitative structure activity relationship (4D QSAR) was applied to a series of 37 compounds of 4,4-dihydroxy difenylmethane and its derivatives by Molecular Conformer Electron Topological (MCET) method. In our current work QSAR analysis was done with Electron Topological Matrix (ETM), included both the electronic parametres (atomic charges, bond energy HOMO-LUMO participation) and geometrical parameters (inter atomic distance, bond length). The MCET method used for identification of the pharmacophore (Pha) is a different procedure from other QSAR approches. For a set of the electronic and geometric parametres in this method were simultaneously used to obtain and characterize Pha, together with auxilary group (AG) and the anti-pharmacophere shielding (APS). The Pha, AG and APS, which were responsible from relative binding rate (RBR), were defined by comparing the matrix of reference conformer with those of all conformers. The 4D-QSAR analysis prepared by MCET in this study was firstly used to construct the essential model. In this way the follows were aimed

- To predict affinities,
- To illustrate significant regions,
- To provide insight into possible interactions.



4,4-Dihydroxydiphenylmethanes

In order to find the more convenient set of descriptors as Pha, AG and APS in bio-structure, the partial least-squares (PLS) modeling and genetic algorithm in MCET method were used. After PLS analysis for 30 ligands in the training set the 4D-QSAR model validation was predicted on 7 samples in test set. The resulted MCET method possessed a high statistical quality (R^2=0.71 and se= 0.304) for predicting RBR of the molecules under consideration. This shows that MCET method can be useful for computer aided drug design.

STRUCTURE-ANTIPROLIFERATIVE ACTIVITY RELATIONSHIP OF ESTRONE BASED COMPOUNDS

Dragana Milić[a], Tatjana Kop[b], Bogdan Šolaja[b,c] and Miroslav J. Gašić[b,c]

[a] Faculty of Chemistry University of Belgrade, Studentski trg 12-16, 11000 Belgrade, Serbia

[b] ICTM - Center of Chemistry, Njegoševa 12, 11000 Belgrade, Serbia

[c] Serbian Academy of Sciences and Arts, Knez Mihailova 35, 11000 Belgrade, Serbia

Our previous research indicated that stuctural modifications of estrone afforded compounds with pronounced antiproliferative activity toward different cancer cell lines.

In the first, key-step, a light assisted oxidation of estrone with a non-metal oxidizing system consisting of m-chloroperoxybenzoic acid and benzoyl-peroxide provided corresponding quinol and epoxyquinol1. Obtained products appeared to be useful synthetic intermediates for further synthetic modifications into various A-ring polyoxygenated compounds2. In addition, A-ring re-aromatization and long range double bond introduction were also achieved.

Antineoplastic activity of obtained compounds on three cancer cell lines was examined, while some of structures were subjected to the extensive tests of in vitro anticancer action using panel of 60 cell lines, and also to acute citotoxicity and in vivo activity determination3.

Obtained results showed that A-ring oxidation led to compounds with significant anticancer activity. Also, the presence of the additional double bond in Δ9,11A-ring aromatized compounds induces strong antineoplastic activity, converting tumor-promoting catehols to antitumor ones.

Detailed analysis of structure-activity relationship will be presented.

References:

1 a) B. A. Solaja, D. R. Milic, M. J. Gasic, Tetrahedron Lett. 1996, 37, 3765-3768. b) D. Milic, M. Gasic, W. Muster, J. Csanadi, B. Solaja, Tetrahedron, 1997, 53, 14073-14084.

2 a) D. Milic, T. Kop, Z. Juranic, M. J. Gasic, B. Solaja, Bioorg. Chem. Lett. 2001, 11, 2197-2200. b) D. Milic, T. Kop, Z. Juranic, M. J. Gasic, B. Tinant, G. Pocsfalvi, B. A. Solaja, Steroids, 2005, 70, 922-932.

3 D. Milic, T. Kop, J. Csanadi, Z. Juranic, Z. Zizak, M. J. Gasic, B. Solaja, Steroids, 2009, 74, 890-895.

MOLECULAR MODELING OF ESTROGEN RECEPTOR: UNRAVELING BEHAVIOUR AND REGULATION

Masłyk, Maciej[ab]; Jakub, Ciepielski[ac]; Filipiak, Kamila[ab]; de Pascual-Teresa, Sonia[d], Martin-Santamaria, Sonsoles[a]; de Pascual-Teresa, Beatriz[a];. Ramos, Ana[a]

[a] Department of Chemistry, Faculty of Pharmacy, Universidad San Pablo CEU, Urbanizacion Monteprincipe, 28668- Madrid, Spain

[b] Department of Molecular Biology, The John Paul II Catholic University of Lublin, 20-718 Lublin, Poland

[c] Department of Biochemistry and Environmental Chemistry, The John Paul II Catholic University of Lublin, 20-718 Lublin, Poland

[d] Departamento de Metabolismo y Nutrición, Instituto del Frío, C.S.I.C. C/José Antonio Novais, 10, 28040-Madrid, Spain

Estrogen receptors (ER) play pivotal roles in human health and diseases, such as cancer, inflammation, and depression. ERs subtypes, alpha and beta, mediate their actions almost entirely by binding to 17□-estradiol and further function as transcription factors. Selective estrogen receptor modulators (SERMs) are synthetic molecules exhibiting agonistic or antagonistic biocharacter through the binding to ER, and can modulate its transcriptional capabilities in different ways and in diverse estrogen target tissues.[1] On the other hand, there is multitude of corregulator proteins that are involved in ER complex association, recruiting many enzymatic and structural activities that allow modulation of chromatin structure to facilitate stimulation or repression of gene expression. The activity of SERMs, as well as coregulator proteins, is not fully explained. A better understanding of the mechanisms involved in estrogen signaling could help to identify novel therapeutic candidates.

Here we report docking analysis of newly designed and synthesized SERMs (see Figure 1 for a representative compound), by means of GLIDE program. Docking studies reveal that these compounds are able to adopt antagonist-like conformation through the interaction with Asp351 and Asp303 in both, ERα and ERβ. Additionally, energetically more favored poses corresponded to complexes with ERβ, in agreement with our biological data.

We also report a computer-based insight into the ER-coregulator binding. In this case we performed molecular dynamics simulations using AMBER10, in conjunction with MM-PBSA/GBSA and NMODE analysis. Docking of reported coregulator binding inhibitors (CBIs), based on pyrimidine scaffolds,[2] has allowed us to compare binding modes of coregulator peptide and these CBIs. Selective binding to ERα/ERβ was also analyzed. These approaches will allow us to extract key information about the molecular recognition processes involved in ER behaviour, which will be useful for future structure based drug design.

Figure 1. Representative compound in complex with ERβ.

[1] V. C. Jordan, Nature Rev. Cancer 2007, 7, 46-53.

[2] Parent, A.A., J.R. Gunther, and J.A. Katzenellenbogen. J. Med. Chem., 2008, 51, 6512-30.

QSAR CLASSIFICATION ANALYSIS OF ESTROGEN RECEPTOR BINDERS

Jiazhong Li[a,b], Paola Gramatica [a] , Ester Papa [a], Mara Luini [a], Andrea Mauric, Marco Tasca[c], Viviana Consonnic and Roberto Todeschini[c]

[a] QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Dep. of Structural and Functional Biology, Univ. of Insubria, , Varerse, Italy, e-mail: paola.gramatica@uninsubria.it

[b] Dep. of Chemistry, Lanzhou University, Lanzhou, China

[c] Milano Chemometrics & QSAR, Dep. of Environmental Sciences, Univ. of Milano-Bicocca, Milan, Italy

Many environmental chemicals, both man-made and naturally occurring, termed "endocrine disrupting chemicals" (EDCs), are capable of mimicking or antagonizing natural hormones and disrupting the normal endocrine functions[1]. Exposure to EDCs can pose serious threats to the health of human and wildlife, resulting in abnormal physiological states that lead to adverse effects including reproductive disruption, hormonal imbalance, neurological and immune effects, or deformity in organisms and their offspring.

EDCs interact with the endocrine system through a variety of mechanisms, mainly receptor-mediated mechanisms of toxicity, such as estrogen receptor (ER). ER is one of the first identified targets for the disrupting activities of chemicals such as dioxins and dioxin-like compounds, polychlorinated biphenyls (PCB), DDT, bisphenol A and pesticides etc. Many experimental data evaluating estrogenic activity have been released. EDCs are recognized as substances of very high concern (SVHCs), requiring the most demanding step of authorization and plan for safer alternatives in the new European regulation of REACH[2]. Accordingly, computational tools such as quantitative structure-activity relationships (QSAR) have been recommended as a screening support. QSAR approach is particularly useful not only for identifying SVHC as EDCs, but also for the environmentally benign design of safer replacement solutions for recognized endocrine disruptors.

In this research, two famous ER binder databases, Japanese METI[3] and US FDA EDKB[4], were combined, resulting in so far the largest ER binder data set of 838 heterogeneous compounds. These two databases have experimental data measured on human ER (EDKB) or assignments as active/inactive (METI). Classification QSAR studies on this combined data set are investigated. All the structures were optimized to the minimal energy conformations by using semi-empirical AM1 method. DRAGON program was used to generate theoretical molecular descriptors and recursive feature elimination (RFE)[5], Genetic Algorithms (GAs), and Wilk's lambda were employed to rank and select the descriptors according to their abilities for distinguishing ER binders from non-binders. Then, several modeling methods (k nearest neighbor, random forest, local lazy method and PLS-DA) and consensus approach were used to build different models based on the top descriptors in the rank lists.

New predictive classifiers were derived with comparable predictive abilities and wide chemical applicability. After strict validation, these models were used to virtually screen a set of AR binders for possible pleiotropic EDCs.

Financial supports from PRIN 07 MIUR Program (2007R57KT7) and China Scholarship Council (CSC) Post-graduate Study Abroad Program (2008618039).

References:

1. R.J. Kavlock, et al. *Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. EPA-sponsored workshop, Environ.* Health Perspect. 104 (1996), pp. 715-740.

2. http://europa.eu.int/comm/environment/chemicals/reach.htm
3. METI, ministry of economy trade and industry, Japan. *Current status of testing methods development for endocrine disrupters. 6th meeting of the task force on endocrine disrupters testing and assessment (EDTA)*, 24–25 June 2002, Tokyo, Japan.
4. http://edkb.fda.gov/databasedoor.html, accessed March, 2009.
5. I. Guyon, et al., *Gene Selection for Cancer Classification using Support Vector Machines*, Mach. Learn. 46 (2002), pp. 389-422.

QUANTITATIVE STRUCTURE – ACTIVITY RELATIONSHIPS (QSAR) OF PHYTOESTROGENS, MYCOESTROGENS, AND DIETHYLSTILBESTROL DERIVATIVES

K Y B Williams[a,c], John A. McLachlan[b,c] , and Thomas Wiese[c,d]

[a] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL 35203

[b]Department of Pharmacology, Tulane University, New Orleans, Louisiana 70112

[c]Center for Bioenvironmental Research, Tulane University, New Orleans, Louisiana 70112

[d]College of Pharmacy, Department of Pharmaceutical Sciences, Xavier University, New Orleans, LA 70125

Three different types of Quantitative Structure-Activity Relationship (QSAR) models were constructed for the estrogen receptor (ER) binding activity of 58 myco, phyto and stilbene estrogens reported in the NCTR database (FANG, H., et al., 2001, Chemical Research in Toxicology, 14, 280-294). The purpose of this project was to develop QSAR models that can be used to predict the ER binding of untested compounds in these discrete structure classes. This data set was divided into 50 training set compounds used to build QSAR models and 8 test set compounds to evaluate the predictive capability of each model. A hologram QSAR (HQSAR) model was developed that defines two-dimensional fragments (4-7 atom fragments) responsible for ER activity (Q^2=0.758, R^2=0.915). Three dimensional structures were used to develop Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) QSAR models that define distinct structure features responsible for ER binding activity (Steric, Electrostatic, Hydrophobic). All relevant isomers and enantiomers were modeled to include the stereochemical nature of particular phytoestrogens in these 3D-QSAR models. The optimal CoMFA model displayed a predictive Q^2 of 0.792 (R^2=0.983) while the optimal CoMSIA model produced in a predictive Q^2 of 0.831 (R^2=0.933). This study has produced the most comprehensive QSAR models of ER binding activity for myco, phyto and stilbene estrogens to date. These three models have considerable potential to predict the ER binding activity of myco, phyto and stilbene compounds found through database screening or the analytical separation and identification of plant and fungal extract, as well as being able to develop pharmaceuticals.

MOLECULAR SIMULATION STUDIES OF LIGAND BINDING TO ESTROGEN RECEPTOR. PREDICTION MODELS FOR AFFINITY, SELECTIVITY AND BIOLOGICAL RESPONSE

Lambrinidis George[1], Skaltsounis Alexios Leandros[1], Marakos Panagiotis[1], Pouli Nicole[1], Alexis Michael N.[2], Mikros Emmanuel[1*]

[1]Faculty of Pharmacy, University of Athens, Athens, Greece.

[2]Institute of Biological Research and Biotechnology, National Hellenic Research Foundation, Athens, Greece

Estrogens exert their activity through the interaction with Estrogen Receptor (ER), a member of the nuclear receptor superfamily of transcription factors.[1] Understanding estrogenic activity at the molecular level remains challenging in order to facilitate the discovery of new molecules with optimized pharmacological properties. Crystallographic studies established that the orientation of the C-terminal Helix of ER Ligand Binding Domain (Helix-12) can be associated to the agonist or antagonist activity of a ligand. An equilibrium has been proposed to exist between two stable agonist-antagonist locations of Helix-12 facilitating the interaction with transcriptional co-repressors or co-activators, depending on the relative levels of expression of these co-regulators in a particular cell or tissue.[2] The ability of Selective Estrogen Receptor Modulators (SERMs) to display tissue-specific ER agonist and/or antagonist activity depends on their capability to stabilize distinct receptor conformations.

In the present study we utilized MonteCarlo docking and Molecular Dynamics simulations in order to compare derived models for the prediction of Relative Binding Affinity. Flexible docking calculations revealed the binding mode of a series of known binders of ER as well as a number of novel raloxifene analogues recently synthesized by our group. LIE[3] and MM-PBSA[4] methods were used to create appropriate models for prediction of the Relative Binding Affinity and selectivity between the two subtypes of ER.

Furthermore a model for predicting biological response was created based on conformational flexibility of Helix 12. Four different crystal structures of ERα and ERβ were used, bearing Helix 12 in different conformations (agonist, antagonist, partial agonist, apo conformation). Agonist or antagonist activity can be predicted by comparing the calculated affinity of different ligands such as estradiol, raloxifene, R,R-tetrahydrochrysene genistein and coumestrol towards to all four conformations of both ER subtypes. Since these molecules exhibit different agonist-antagonist activity, establishing a model to predict biological activity upon binding could provide an essential tool in drug design targeting Ers.

References:

1. Katzenellenbogen JA, Katzenellenbogen BS, 1996, 3, 529-536
2. Gangloff M, Ruff M, Eiler S, Duclaud S, Wurtz JM, Moras D, J Biol Chem 2001, 276, 15059–15065
3. Aqvist J, Medina C, Samuelsson JE, Protein Eng 1994, 7, 385–91
3. Wang W, Donini O, Reyes CM, Kolman PA,Ann Rev Biophys Biomol Struct 2001, 30, 211-243

THEORETICAL EVALUATION OF STEROIDS BINDING AFFINITY TO CORTICOSTEROID-BINDING GLOBULIN (CBG) AND SEX HORMONE-BINDING GLOBULIN (SHBG)

Katarina Nikolic, Slavica Filipic, and Danica Agbaba

Institute of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Belgrade, Vojvode Stepe 450, 11000 Belgrade, Serbia

In the presented QSAR study was developed method for theoretical evaluation of steroids binding affinity to corticosteroid-binding globulin (CBG) and sex hormone-binding globulin (SHBG). The binding affinities of steroids to the CBG and SHBG have influence on the half-life, distribution, and efficacy of these agents. Thirty one steroids of different structures were used in theoretical QSAR and docking study of binding affinities to CBG) and SHBG. Plasma CBG and SHBG also bind synthetic glucocorticoids and sex hormones, and therefore influences the half-life, distribution, and efficacy of these group of drugs. Performed theoretical study has developed QSAR models for prediction of the steroids binding affinities for SHBG and CBG.

The compounds examined in the study demonstrated a wide range of experimentally measured binding affinities (K: 10^{-5}–10^{-9} M^{-1}). Constitutional, geometrical, physico-chemical and electronic descriptors were computed for the examined structures by use of the Chem3D Ultra 7.0.0 [1], the Dragon 5.4 [2], the MOPAC2009 [3], and the Chemical Descriptors Library (CDL) [4] program. Partial least squares regression (PLSR), has been applied for selection of the most relevant molecular descriptors and development of the QSAR models.

Optimal QSAR models with six and eight variables, $R^2>0.789$ and cross-validation parameter $Q^2>0.682$, were selected and compared.

Finally, the ligands docking were performed with help of FRED docking program at the CBG and SHBG active sites. Results of the QSAR and docking study were compared in order to define physicochemical, electronical and structural requirements for selective and effective binding to the CBG and SHBG.

References

1. Cambridge Soft Corporation, CS Chem3D Ultra 7.0 Cambridge, MA U.S.A., **2001**, http://www.cambridgesoft.com.
2. Tetko, I.V., Gasteiger, J., Todeschini, R., Mauri, A.; Livingstone, D., Ertl, P.; Palyulin, V.A.; Radchenko, E. V.; Zefirov, N.S.; Makarenko, A.S.; Tanchuk, V.Y.; Prokopenko, V.V. *J. Comput. Aid Mol. Des*. **2005**, 19, 453-463.
3. Stewart J.J.P.; Stewart Computational Chemistry, MOPAC2009 program, Colorado Springs, CO, USA, **2008**. http://openMOPAC.net.
4. Sykora, V.J.; Leahy, D.E.J. *Chem. Inf. Model*. **2008**, 48, 1931–1942. Chemical Descriptors Library (CDL). http://sourceforge.net/projects/cdelib (June 18. 2008).

"LOADING THE DICE": FILE SCREENING TO AN IN VIVO EP2 ANTAGONIST TOOL IN 3 MONTHS

Siew Kuen Yeap, Kevin Dack & Sarah Skerratt

Pfizer Global Research & Development, Sandwich, UK

EP2 is a prostanoid receptor, sharing closest homology with DP1, IP & TP, with no suitable potent and selective tools in the literature. Almost all ligands in the family are acids. The Pfizer 'Global Diversity Representative Subset 1' (GDRSI) is a selection of 100K diverse compounds available on 280 High Throughput Screening plates. Analysis of the acid, base, neutral & zwitterionic populations allowed selection of 22 'acid-rich' GDRS plates. This knowledge-based file screening enabled the rapid identification of lead matter from a minimal screening campaign.

Nearest Neighbour analysis of resulting "hits" then identified additional "actives" with improved potency and selectivity. R1/R2 analysis informed a focused library design, with Hit 3 being identified after just one focused library of 112 compounds, just 3 months after initiation of knowledge-based file screening. Safety profiling of Hit 3 in vivo suggests that a selective EP2 antagonist has potential to be a safe and well tolerated mechanism.

A NEW SERIES OF 2-ARYL-PYRAZOLO-TRIAZOLO-PYRIMIDINES AS POTENT HUMAN A3 ADENOSINE RECEPTOR ANTAGONISTS: QSAR ANALYSIS AND COMFA-GUIDED MODELING OF ANTAGONIST BINDING SITE THROUGH A HUMAN A3 'HYBRID' HOMOLOGY MODEL

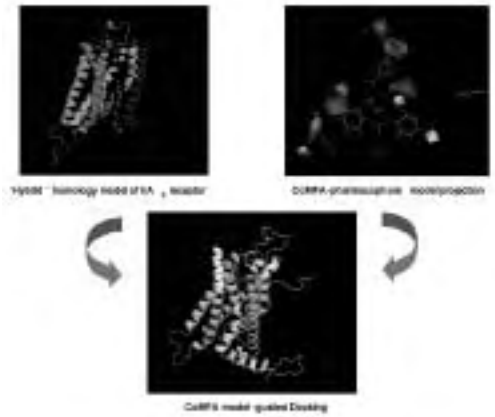
Siew Lee Cheong[a], Vamsi Krishna Kosaraju[b], Chun Wei Yap[c] and Giorgia Pastorin[a]

[a] Department of Pharmacy, National University of Singapore, 3 Science Drive 2, S117543

[b] Department of Chemistry, National University of Singapore, 3 Science Drive 3, S117543

[c] Department of Pharmacy, National University of Singapore, 18 Science Drive 4, S117543

A new series of 2-(*para*-substituted)phenyl-pyrazolo-triazolo-pyrimidine derivatives have been recently synthesized and characterized as potent and highly selective human A_3 adenosine receptor (hA_3AR) antagonists. It was observed that the introduction of an aryl group at position C^2 resulted in good affinity in the low nanomolar range, and it also enhanced the selectivity against other adenosine receptor subtypes. [1] In this study, we proposed a Comparative Molecular Field Analysis (CoMFA)-based QSAR study to explore in-depth the structural features at N^5-, N^8- and in particular the unprecedentedly explored C^2-position of tricyclic nucleus in the new derivatives, which were responsible for the inhibition of hA_3AR. The resultant CoMFA model showed both good correlation and predictability. Moreover, an innovative homology model for the hA_3 receptor was also proposed, built as 'hybrid' from both human A_{2A} adenosine receptor [2] and human β_2-adrenergic receptor [3] templates. This 'hybrid' homology model was used in the following CoMFA model-guided docking study, in which it depicted new insights on the hypothetical binding mode and binding interaction of such derivatives with specific residues. Results obtained from this investigation not only rationalized the experimental data from the bioassay, but also disclosed some residues in the binding site that were speculated to be important for conferring good affinity at the hA_3 receptors.



References:

1) Cheong et al. *J. Med. Chem*. (2010) In Press.

2) Jaakola et al. *Science* 322 (2008) 1211-1217.

3) Rasmussen et al. *Nature* 450 (2007) 383-387.

RETROSPECTIVE MAPPING OF SAR DATA FOR TRANSTHRYRETIN PROTEIN (TTR) IN CHEMICO-BIOLOGICAL SPACE USING LIGAND EFFICIENCY INDICES AS A GUIDE TO DRUG DISCOVERY STRATEGIES

Daniel Blasi[a], Gemma Arsequell[b], Gregori Valencia[b], Joan Nieto[c], Antoni Planas[c], Marta Pinto[d], Nuria B. Centeno[d], Cele Abad-Zapatero[a,e], Jordi Quintana[a]

[a] Drug Discovery Platform (PDD), Parc Cientific de Barcelona, Barcelona, Spain

[b] Unit of Glycoconjugate Chemistry, Institut de Química Avançada de Catalunya, I.Q.A.C.-C.S.I.C., Barcelona, Spain

[c] Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, Barcelona, Spain

[d] Computer-Assisted Drug Design Laboratory, Research Group on Biomedical Informatics (GRIB) IMIM-Universitat Pompeu Fabra, Barcelona, Spain

[e] Center for Pharmaceutical Biotechnology, University of Illinois at Chicago, 900 So. Ashland St, MBRB Building, Room 3020; (M/C870), Chicago, IL 60607-7173, USA

We have previously reported the design and synthesis of ligands that stabilize Transthyretin protein (TTR)[1] in order to obtain therapeutically active compounds for Familial Amyloid Polyneuropathy (FAP). We are hereby reporting a drug design strategy to optimize these ligands to target Familial Amyloid Cardiomyopathy (FAC), through the following steps: a) SAR analyses of the ligands described previously for the TTR tetramer, classified in chemical classes and mapping them in Chemico-Biological Space (CBS) using Ligand Efficiency Indices (LEIs); b) drug design / optimization of TTR ligands through docking in the TTR tetramer three-dimensional structure and through optimization of physicochemical / pharmacokinetic / selectivity properties; c) comparative structural analyses of selected amyloidogenic and non-amyloidogenic TTR mutants and native TTR structures; and d) virtual screening of commercially available ligands and therapeutically active compounds (repurposing) towards wild-type and mutant TTR tetramer structures. First results in step a) and b) of this drug design strategy will be reported.

We have focused in the optimization of TTR amyloid inhibitors exploring more effective ways to guide the proposed drug discovery strategy. The approach is based on using Ligand Efficiency Indices (LEIs) to map the different chemical series in CBS. We use an efficiency index, BEI[2,3], based on the measured binding affinity related to the Molecular Weight (MW) of the compound combined with a surface-binding efficiency index (SEI)[2,3] based on Polar Surface Area (PSA). We will illustrate the use ot these indices, combining three crucial variables (potency, MW and PSA)[4] in a 2D graphical representation of chemical space, to perform a retrospective mapping of SAR data for a current TTR inhibitors database, and we suggest strategies and approaches for future drug design efforts for TTR ligands.

[1]Mairal T, Nieto J, Pinto M, Almeida MR, Gales L, Ballesteros A, Saraiva M, Damas A.M, Planas A, Arsequell G, Valencia G, et al., *PLoS ONE,* ***2009, 4(1)***

[2]L. Hopkins Andrew, R. Groom Colin, A. Alex, *Drug Discov Today,* **2004,** *9, 430*.

[3]Abad-Zapatero, C. and J. T. Metz. Drug Discovery Today, ***2005***. 10 (7), 464-469.

[4]C. Abad-Zapatero, *Expert Opin. Drug Discovery,* ***2007***, 2, 469.

IDENTIFICATION OF NOVEL HIV-1 INTEGRASE INHIBITORS USING SHAPE-BASED SCREENING AND QSAR APPROACH

Pawan Gupta[1], Nilanjan Roy[1, 2] and Prabha Garg[1,*]
[1]Centre for Pharmacoinformatics; [2]Department of Biotechnology
National Institute of Pharmaceutical Education and Research (NIPER)
Sector-67, S.A.S. Nagar 160062, Punjab, India

Shape based screening is an emerging tool for the identification of novel molecules as potential inhibitors for specific targets. The rationale of this approach is that similar molecules have high degree of shape conservation and occupy the same region in the binding sites. Thus they may induce the same biological effect and provide support for the application of shape matching in the drug design process. Here, shape based screening has been implemented to find out the computationally novel inhibitors for HIV-1 integrase (IN) followed by docking and QSAR methodologies. In brief, ROCS (Rapid Overlay of Chemical Structure) program is used for shape based screening of ZINC database using best docked conformation of template molecule (highly active molecule of benzodithiazine series). In ROCS program, two keywords ImplicitMillsDean and shape tanimoto has been used for screening of the molecules. The ImplicitMillsDean keyword is related to color force field which encodes six different types of chemical functionality (hydrogen bonding donor and acceptor, hydrophobe, anion, cation and ring) and shape tanimoto is the coefficient which quantitative measure for the shape overlap of two molecules. These keywords analyze chemical functionality and shape of template with database molecules and calculate the score. Output molecules were ranked using combo score which have almost similar chemical features and shape as template molecule. Further, these molecules were filtered out on the basis of ADME and toxicity studies using Discovery Studio 2.5 program. The QSAR model was built using series of benzodithiazine derivatives. The screened molecules were subjected for QSAR prediction. Best predicted molecules were validated using docking studies (FlexX and AutoDock program). Some of the screened molecules were docked in a similar manner as benzodithiazine derivative. The results of docking studies for the screened molecules are in good agreement with published docking results. The new hits which may be potential inhibitor for HIV-1 IN are needed to be tested experimentally. We believe that this methodology can also be used for the identification of novel inhibitors for different targets.

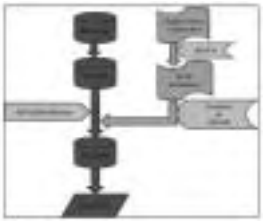


Figure: Flowchart for computational methodologies used for identification of novel HIV-1 IN inhibitors.

References

1. Gupta, P.; Roy, N.; Garg, P. *Eur. J. Med. Chem*. **2009**, 44, 4276.
2. Sotriffer, C. A.; Ni, H.; McCammon, J. A. *J. Med. Chem*. **2000**, 43, 4109.
3. Dayam, R.; Neamati, N. *Bioorg. Med. Chem*. **2004**, 12, 6371.
4. Serrao, E.; Odde, S.; Ramkumar, K.; Neamati, N. *Retrovirology* **2009**, 6, 25.

THEORETICAL vs EXPERIMENTAL MODELS: THE EXPERIENCE FROM HIV-1 INTEGRASE RESEARCH

Nunzio Iraci,[*] Maria Letizia Barreca,[*] Laura De Luca,[‡] Alba Chimirri.[‡]

[*] Dipartimento di Chimica e Tecnologia del Farmaco - Università degli Studi di Perugia, Italy

[‡] Dipartimento Farmaco-Chimico – Università degli Studi di Messina, Italy

HIV-1 integrase (IN) belongs to the superfamily of polynucleotidyl transferases and is a key enzyme in the viral replication cycle of HIV.[1]

Although a wide variety of compounds has been reported as IN inhibitors, raltegravir is the only drug active against this enzyme that has been approved for treating HIV-infected patients. In fact, the drug discovery process has been hampered for years by the lack of structural information about the enzyme and its interaction with DNA substrates, metal co-factors and inhibitors.

Raltegravir belongs to the family of integrase strand transfer inhibitors (InSTIs), which bind IN in complex with DNA rather than binding the free enzyme in solution, and to date represent the major leads in the development of anti-HIV-1 IN drugs.[2]

In this context, a three-dimensional model of a complex between HIV-1 IN, viral DNA, and metal ions was built and later used as a target for induced-fit docking studies of six different InSTIs.[3,4]

The in silico findings provided additional insight into the possible mechanism of action and binding mode of this class of IN inhibitors, and were consistent with the available drug resistance mutation data.

In February 2010, the crystal structures of the full-length IN from the prototype foamy virus in complex with its cognate DNA and two InSTIs (raltegravir and elvitegravir) have been reported by Hare et al.[5]

This is definitely a breakthrough in the field of IN research, since the disclosed complexes can be considered the best surrogate structures for reliably modelling the binding of these inhibitors to the HIV-1 enzyme.

Beside the hope of concretely boosting the development of new antiretroviral agents, this novel structural information has given us the opportunity to understand limits and strength points of the theoretical models and of the molecular modelling techniques applied in our project.

In particular, the experimental work by Hare et al. highlight that “the two InSTIs seem to have very similar modes of binding and action, involving an induced fit mechanism. Their metal chelating oxygen atom orient towards the metal cofactors of the active site, whereas their halobenzyl groups fit within a tight pocket created by displacement of the 3′ adenosine”.

Of note, in agreement with the experimental results, in our paper we stated that “remarkably, two IFD solutions of the most advanced inhibitors raltegravir and elvitegravir revealed an unexplored binding mode with respect to previously developed InSTIs, with the substituted benzyl group deeply occupying a pocket adjacent to, but distinct from, the one of the active site loop. Notably, during this IFD run, the 3′-terminal adenine nucleotide A20 underwent a dramatic conformational movement in order to allow insertion of the substituted benzyl moiety between the two viral DNA strands”.

1) E. Asante-Appiah, A. M. Skalka, Adv. Virus Res. 1999, 52, 351–369.
2) Pommier et al., Nat. Rev. Drug Discovery 2005, 4, 236–248.
3) Barreca et al., ChemMedChem 2009, 4, 1446 – 1456
4) Ferro et al., J. Med. Chem. 2009, 52, 569–573.
5) Hare et al., Nature 2010, 464, 232-236.

QSAR ANALYSIS OF SOME 1-(3, 3-DIPHENYLPROPYL)-PIPERIDINYL AMIDES AND UREAS AS CCR5 INHIBITORS AS ANTI HIV AGENTS USING GENETIC ALGORITHM- LEAST SQUARE SUPPORT VECTOR MACHINE

Razieh Sabet[a,b*], Mohsen Shahlaei [a,c], Afshin Fassihi [a,d] Lotfollah Saghaie [a,d]

[a]Department of Medicinal Chemistry, Faculty of Pharmacy, Isfahan University of Medical Sciences, 81746-73461, Isfahan, Iran

[b]Department of Medicinal Chemistry, Faculty of Pharmacy, Shiraz University of Medical Sciences, 81746-73461, Shiraz, Iran

[c]Department of Medicinal Chemistry, Faculty of Pharmacy, Kermanshah University of Medical Sciences, Kermanshah, Iran

dIsfahan Pharmaceutical Sciences Research Center, 81746-73461, Isfahan, Iran

Quantitative relationships between structures of sixty seven of 1-(3, 3-diphenylpropyl)-piperidinyl Amides and Ureas as CCR5 inhibitors and their activities were investigated by GA-MLR and GA-LS-ASVM. As a preliminary step, linear dependence was established by MLR approach, selecting the relevant descriptors by genetic algorithm feature selection. For LS-SVM model, non-linear GA feature selection was also applied. Comparison of the GA–MLR and GA–LS-SVM models disclose superiority of the GA–LS-SVM over the GA–MLR model. With respect to obtained results, it can be deduced there is a non-linear relationship between the pIC50s and the calculated structural descriptors of the 1-(3, 3-diphenylpropyl)-piperidinyl Amides and Ureas. The accuracy and predictability of the proposed models were illustrated using various validation techniques.

QSAR STUDY OF ANTIMICROBIAL 3-HYDROXYPYRIDINE-4-ONE AND DERIVATIVES

Razieh Sabet [a,c], Bahram Hemmateenejad [b,c], Afshin Fassihi [a,d], Lotfollah Saghaie [a,d], Ramin Miri [c]

[a]Department of Medicinal Chemistry, Faculty of Pharmacy, Isfahan University of Medical Sciences, 81746-73461, Isfahan, Iran

[b] Department of chemistry , Shiraz University, 71454, Shiraz, Iran

[c] Medicinal & Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

[d]Isfahan Pharmaceutical Sciences Research Center, 81746-73461, Isfahan, Iran

Iron is the fourth abundant metal on earth and fundamental factor for sustaining life. It is one trace elements inevitably required for survival and proliferation of all living things including microorganism. While the presence of Fe in the body is essential in the context of biological functions, such as oxygen sensing and transport, electron transfer, catalysis, DNA synthesis, energy production . Pathogenic bacteria require various kinds of nutrient to proliferate and cause infectious disease in human body. Bacteria growth is subjected to organic growth factors such as vitamins and trace elements such iron because they function as cofactors of most enzymes inevitably required during metabolism. Many bacteria synthesize small molecules known as siderophores to scavenge iron.This low molecular weight chelating agent excreted by microorganisms under Fe deficiency and possessing high affinity to Fe. There are many reports of the antimicrobial activity of chelating agents with different chemical structures. Maltol, 3-hydroxy-2-methyl-4-pyrone, is one of several hydroxypyrones long know for high bioavailability and favourable toxicity profile. Few reports of antimicrobial studies of 3-hydroxypyridine-4-one and 3-hydroxypyran-4-one derivatives are available and in those they were not the subject of QSAR studies. A series of 3-hydroxypyridine-4-one derivatives was subjected to quantitative structure-antimicrobial activity relationships (QSAR) analysis. These compounds were synthesized in our lab and were evaluated against nine pathogenic bacteria and fungi. The large number of molecular descriptors was calculated using Hyperchem, Dragon 2.1 and Gaussian 98 packages. Dragon software calculated different functional groups, topological, geometrical and constitutional descriptors for each molecule. Gaussian 98 was employed for calculation of different quantum chemical descriptors. A collection of chemometrics methods, including multiple linear regression (MLR), factor analysis-based multiple linear regression (FA-MLR), principal component regression (PCR) and partial least squares combined with genetic algorithm for variable selection (GA-PLS) were employed to make connections between structural parameters and antimicrobial activity. The results of the investigation is supported be done.

QUANTITATIVE STRUCTURE ACTIVITIES RELATIONSHIPS OF SOME 2-MERCAPTOIMIDAZOLES AS CCR2 INHIBITORS USING GENETIC ALGORITHM- ARTIFICIAL NEURAL NETWORKS

Mohsen Shahlaei [a,b*], Razieh Sabetb[c*], Afshin Fassihib,[d]

[a]Department of Medicinal Chemistry, Faculty of Pharmacy, Kermanshah University of Medical Sciences, Kermanshah, Iran

[b]Department of Medicinal Chemistry, Faculty of Pharmacy, Isfahan University of Medical Sciences, 81746-73461, Isfahan, Iran

[c]Department of Medicinal Chemistry, Faculty of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran

[d]Isfahan Pharmaceutical Sciences Research Center, 81746-73461, Isfahan, Iran

Quantitative relationships between structures of twenty six of 2-Mercaptoimidazoles as CCR2 inhibitors and their activities were investigated by GA-MLR and GA-ANN. Comparison of the GA–MLR and GA–ANN models reveal superiority of the GA–ANN over the GA–MLR model. With respect to obtained results, it can be deduced there is a non-linear relationship between the pIC50s and the calculated structural descriptors of the 2-Mercaptoimidazoles. The obtained models were able to describe about 78% and 93% of the variance in the experimental activity of molecules in training set, respectively. The accuracy and predictability of the proposed models were illustrated using various evaluation techniques. Some of them were: cross validation, validation through an external test set, and Y-randomization

INTERACTION OF AMINOADAMANTANE DERIVATIVES WITH THE INFLUENZA A VIRUS M2 CHANNEL – COMPARATIVE DOCKING CALCULATIONS USING A PORE BLOCKING MODEL AND AN ALLOSTERIC MODEL

Stelios Eleftheratos[1], Antonios Kolocouris,[1*] Nikolaos Zervos[1], Gabriella Ortore[2], Adriano Martinelli[2], Alan Hay[3]

[1]Department of Pharmaceutical Chemistry, University of Athens, Panepistimioupolis-Zografou, 15771 Athens, Greece

[2] Dipartimento di Scienze Farmaceutiche, UniVersità di Pisa, Via Bonanno 6, 56126 Pisa, Italy

[3] MRC National Institute for Medical Research,The Ridgeway, Mill Hill, London NW7 1AA, UK

Interaction of aminoadamantanes with the influenza A virus M2 proton channel was analysed by docking simulations of a series of synthetic aminoadamantane derivatives (Scheme 1), of different binding affinity, into the crystal structure of the transmembrane (M2TM) pore. The pore blocking model tested readily accounted for the relative binding affinities of the compounds [although a series of highly hydrophobic ligands which seem to have little capacity for different specific interactions with their receptor]. The docking calculations predicted poses in which the adamantane ring is surrounded mainly by the alkyl side chains of Val27 or Ala30 and the ligand's amino group is generally hydrogen bonded with hydroxyls of Ser31 (and carbonyls of Val27) or carbonyls of Ala30, the former (Ser31) being the most stable and most frequently observed (Figure 1).

Several observations have important implications for the design of aminoadamantane drugs. The binding affinity of the ligand is a compromise between hydrogen bonding ability, which is elevated by a primary amino group, and van der Waals enthalpic and entropic contributions, which are increased by the ability of the lipophilic moiety to adequately fill a hydrophobic pocket within the M2TM pore. A delicate balance of these hydrophobic contributions is required for optimal interaction.

Scheme 1

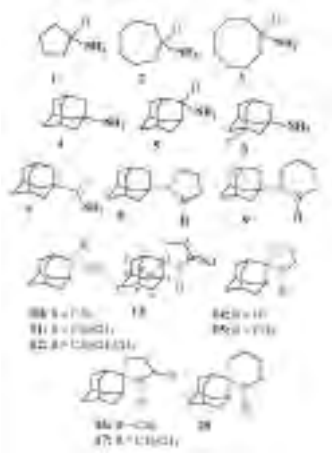
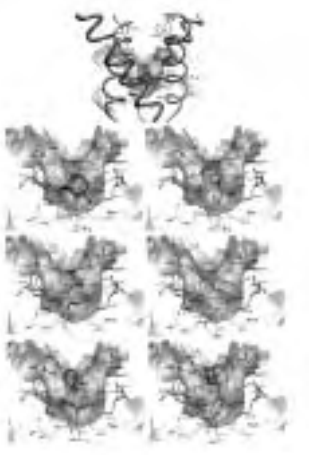


Figure 1



EXPLORING PROTEIN FLEXIBILITY IN 4D-QSAR: APPLICATION TO A SET OF TRYPANOTHIONE REDUCTASE INHIBITORS

Euzébio G. Barbosa[1], Kerly F. M. Pasqualoto[2], Márcia M. C. Ferreira[1]

[1]Universidade Estadual de Campinas, Campinas-SP, Brazil

[2]Universidade de São Paulo, USP, Brazil

In 3D-QSAR analysis, the interactions between ligands and chemical probes are mapped onto a surface or grid surrounding a set of compounds, which are superimposed in 3D space. This surface or grid represents a hypothetical binding site of some biological target. The quality of the QSAR model depends critically on the correct superimposition of the ligands. However, the absence of structural information from the target makes the correct ligands' superimposition almost impossible. The accommodation of ligands in a binding site employing methodologies such as automated docking is one way to handle such situations. But, the docking approaches does not take into account the protein flexibility,[1] which can be considered the major drawback. Nowadays, flexible docking methods and/or molecular dynamics simulations of complexes having docked conformations[2] have been used to overcome these issues.

In this study the advantages of docking method followed by the ligand-receptor molecular dynamics simulations were considered to test such 4D-QSAR approach. The conformational ensemble profile (CEP) of each ligand was aligned based upon both the common structural features from the investigated set of ligands and the relevant amino acid residues in the binding site (see fig. 1a). A preliminary set of 33 phenothiazine derivatives acting as inhibitors of trypanothione reductase (TR) from *Trypanosoma cruzi* was selected from ref. 3. TR enzyme is considered a potential target for the rational design of new selective anti-T. cruzi agents due to the trypanothione redox-defense system, which is a fundamental metabolic difference between the mammalian host and trypanosomal parasite.

LQTAQSAR[4] models were built employing the ordered predictor selection[5] algorithm for variable selection and partial least squares method for regression. The y–randomization and leave–N–out cross-validation procedures were carried out in addition to the external validation. PLS models provided the following statistics: $Q^2 = 0.84$, $R^2 = 0.91$ for 8 variables selected and 3 latent variables. The selected descriptors presented in fig. 1b provided information regarding both intrinsic ligand structural information as well as key interaction sites within TR binding pocket. The results are promising and reinforce the importance of induced fit for building grid based QSAR models.

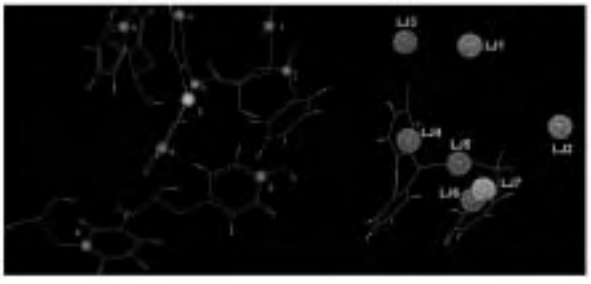


Figure 1. (a) Atoms selected to align the CEP of each ligand of the investigated set. (b) Visualization of the 3D interaction descriptors of the best QSAR model.

References:

1. Teague S.J., Nat Rev Drug Discov, 2003;2(7):527-541. **2**. Hernán A., Andrey A.B., Jill E.G., Med. Res. Rev., 2006;26(5):531-568. **3**. Khan M.O.F., Austin S.E., Chan C., Yin H., Marks D., Vaghjiani S.N., Kendrick H., Yardley V., Croft S.L., Douglas K.T. J., Med. Chem., 2000;43(16):3148-3156. **4**. Martins J.P.A., Barbosa E.G., Pasqualoto K.F.M., Ferreira M.M.C., J Chem Inf Model, 2009;49 (6):1428-1436. **5**. Reinaldo F.T., João Paulo A.M., Márcia M.C.F., J. of Chemometrics 2009; 23(1):32-48.

GROWTH RESPONSE OF TRICHODERMA SPECIES TO ORGANIC SOLVENTS

Gyula Oros[1], Zoltán Naár[2] and Tibor Cserháti[3]

[1] Plant Protection Institute, Hungarian Academy of Sciences, 1525 Budapest 114, Pf. 102, Hungary

[2] Eszterházy Károly College, 3300 Eger, Leányka u. 6, Hungary

[3] Chemical Research Centre, HAS, 1025 Budapest, Pusztaszeri út 59-67, Hungary

The growth response of 25 Trichoderma strains to 26 alkanols and 7 other organic solvents was examined in vitro. The sensitivity of strains considerably varied depending on their taxonomic position, thus the strains of *Longibrachiatum* section proved to be more tolerant than those of *Pachybasium* and *Trichoderma* sections.

Significant relationship was revealed between structure of C_1-C_{24} alkanols and their growth inhibitory effect to *Trichodermas*, both efficacy and selectivity of C_1-C_3 and C_{14}-C_{24} alkanols failed off the C_9-C_{11} alkanols. The non-alkanol solvents tested were non-toxic.

The C_8-C_{11} alkanols used in agricultural practices exhibited noticeable toxicity to all Trichoderma strains, therefore pesticidal preparations containing these alkanols cannot be used simultaneously with Trichoderma based biopreparations.

THE EFFECT OF CARRIER ON 4-AMINOQUINOLINE ANTIMALARIALS

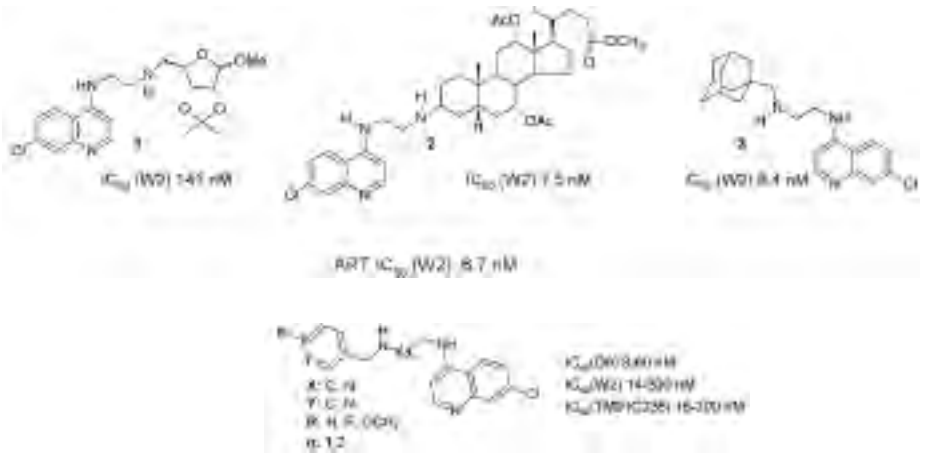
Bogdan A. Šolaja,[a] Igor Opsenica,[a] Nataša Terzić,[b] Dejan Opsenica,[b] Kirsten S. Smith[c]

[a]Faculty of Chemistry, University of Belgrade, Belgrade, Serbia. [b]Institute of Chemistry, Technology and Metallurgy, Belgrade. [c]Division of Experimental Therapeutics, WRAIR, Washington, DC 20307-5100, USA

Malaria is one of the most deadly diseases of the developing world, affecting 300–500 million people annually with an estimated death toll of 1–3 million people. It is caused by an one-celled parasite of the genus *Plasmodium*. The synthesis and testing of various peroxide classes in antimalarial screen, in order to overcome the chloroquine (CQ) resistance, brought to daylight several types of compounds possessing O-O linkage, with 1,2,4,5-tetraoxacyclohexane (tetraoxane) moiety being the promising one.[1] In addition, the introduction of new aminoquinoline antimalarials in order to overcome the drug resistance are also highly desired because of their known ability to inhibit the hemozoin formation.[2]

Here, we show the results of our research aiming at the introduction of sugar, steroid, and adamantane carriers of aminoquinoline pharmacophore.

Target compounds represented by **1, 2** and **3** were prepared by reductive amination.



We will discuss the QSAR depending on a carrier type taking into account the results against CQ-sensitive (D6), CQ-resistant (W2) and multidrug-resistant (TM91C235) P. falciparum strains.

---

[1]. (a) Šolaja B. A., Terzić N., Pocsfalvi G., Gerena L., Tinant B., Opsenica D., Milhous W. K.: J. *Med. Chem.*, **2002**, 45, 3331. (b) Opsenica, I.; Opsenica, D.; Jadranin, M.; Smith, K. S.; Milhous, W. K.; Stratakis, M.; Solaja, B. *J. Serb. Chem. Soc.*, **2007**, 71, 1181. (c) Opsenica, D.; Milić, D.; Tinant, B.; Smith, K. S.; Milhous, W.K.; Šolaja, B.A. *J. Med. Chem.* **2007**, 50, 5118.

[2]. Zearick, K., Ekoue-Kovi, K., Iwaniuk, D. P., Natarajan J. K., Alumasa, J., De Dios, A. C., Roepe, P. D., Wolf, C.: *J. Med. Chem.*, **2008**, 51, 1995–1998.

QSAR STUDIES OF NEW ANTITUBERCULAR AGENTS

Vasyl Kovalishyn[a,b], Filomena Martins[c], Cristina Ventura[c,d], Ruben Leitão[c,e], João Aires de Sousa[a]

[a]REQUIMTE, Departamento de Quimica, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

[b]Institute of Bioorganic Chemistry & Petroleum Chemistry, Murmanskaya, 1, 02660, Kyiv, Ukraine

[c]Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa and Centro de Química e Bioquímica (CQB), Campo Grande, 1749-016 Lisboa, Portugal

[d]Instituto Superior de Educação e Ciências, Alameda das Linhas de Torres, 179, 1750 Lisboa, Portugal

[e]Departamento de Engenharia Química, Instituto Superior de Engenharia de Lisboa (ISEL), Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro, 1, 1950-062 Lisboa, Portugal

Tuberculosis (TB) is one of the most destructive infectious diseases, showing over time increasingly high mortality levels. The motivation for developing new antitubercular drugs comes from the growing appearance of multi-drug resistant strains to the commonly used drugs, and the substantially longer durations of both the therapy and the recovery of disease, particularly in immuno-compromised patients. Some of us have previously reported QSAR models for the hydrazide family, using MLR.[1] The goal of the present study is the exploration of other families of compounds and the use of a variety of machine learning techniques for the development of QSAR models that could further guide the search for new potent antitubercular molecules.

A wide range of machine learning techniques have been explored, such as Random Forest (RF)[2], Multiple Linear Regression Analysis (MLRA), Associative Neural Network (ASNN), k-Nearest Neighbors (kNN)[3]. The models were built with molecular descriptors calculated by Dragon[4] and ADRIANA.Code,[5] and datasets collected from various sources, including the data set from the Institute of Allergy and Infection Diseases Web site.[6,7] The influence of the descriptor set on the model accuracy was evaluated. Descriptors were selected by their importance for Random Forests[2] and by "pruning methods" implemented in ASNN software.[8] The overall best performance was attained by the ASNN and RF methods. The accuracy of all individual models was estimated using cross-validation procedures. The applicability domain of models was estimated by using several measures of "distance to model": number of k-neighbors in training set, proximity, and Euclidean distances a test molecule to the training set using the selected descriptors.

Separate models were developed for benzimidazoles, isoniazide and indole derivatives. R^2 coefficients obtained for independent test prediction sets were in the range 0.75-0.80. The limitations and advantages of the proposed approach are discussed.

References:

1. Ventura K., Martins F. Application of Quantitative Structure Activity Relationships to the Modeling of Antitubercular Compounds. 1. The Hydrazide Family. J. Med. Chem. 2008, 51, 612–624.
2. http://www.r-project.org/
3. http://ochem.eu/
4. http://www.talete.mi.it/products/dragon_new.htm
5. http://www.molecular-networks.com/products/adrianacode
6. National Institutes of Health, National Institute of Allergy and Infectious Diseases (NIAID), http://www.niaid.nih.gov/about/organization/daids/Pages/default.aspx
7. Prathipati P., Ma N.L., and Keller T.H. Global Bayesian Models for the Prioritization of Antitubercular Agents. J. Chem. Inf. Model. 2008, 48, 2362–2370.
8. Kovalishyn V.V., Tetko I.V., Luik A.I., Kholodovych V.V., Villa A.E.P., Livingstone D.J. Neural Network Studies. 3. Variable Selection in the Cascade-Correlation Learning Architecture. J. Chem. Inf. Comp. Sci. 1998, 38, 651-659.

INVESTIGATION OF STRUCTURE ACTIVITY RELATIONSHIPS OF PYRROLE HYDRAZONES AS NEW ANTI-TUBERCULOSIS AGENTS

Iglika Lessigiarska[a], Ilza Pajeva[a], Penka Prodanova[b], Maya Georgieva[b], Atanas Bijev[b]

[a]Centre of Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., [b]l. 105, 1113 Sofia, Bulgaria; bUniversity of Chemical Technology and Metallurgy, dep. Organic synthesis and fuels, Blv. Kliment Ohridski 8, 1756 Sofia, Bulgaria

The search for novel anti-tuberculosis drugs becomes an important task nowadays due to the increasing number of people affected by tuberculosis worldwide and the resistance of the tuberculosis bacilli to some anti-tuberculosis drugs currently used. Preliminary investigations of a research team from the University of Chemical Technology and Metallurgy – Sofia [1] have shown that some pyrrole hydrazones strongly inhibit the tuberculosis bacilli, thus representing a new perspective for development of anti-tuberculosis agents.

In this work a series of pyrrole hydrazones with anti-tuberculosis activity was investigated by classical QSAR analyses and pharmacophore modelling. The compounds were synthesised in the University of Chemical Technology and Metallurgy – Sofia. Their activity varied from 0 to 100 % inhibition of *Mycobacterium tuberculosis* $H_{37}Rv$ at concentrations of 6.25 μg/mL [2]. In addition, for some of the compounds IC_{50} values were available [2].

For the classical QSAR analysis approximately 250 constitutional, topological, physicochemical, and quantum-mechanical structural descriptors were calculated. Before the analysis the logit transformation was applied to the percentage bacillus inhibition corrected to the molar concentrations. This transformation approximates logIC50 correctly in the range of 10-90% inhibition, thus only compounds within this range were included in the analysis (30 compounds). Multiple linear regression was used to derive the QSAR models. The best models included molecular flexibility (Kier molecular flexibility index) and shape (globularity index), and magnitudes of charged molecular surfaces areas and hydrophobic volumes. A classification model of the compounds as active and inactive was further developed using the classification trees approach. For this purpose the compounds were grouped as active with inhibitory effect $\geq$ 90%, and inactive with inhibitory effect $\leq$ 10%: 28 compounds in total (8 active, 20 inactive). The descriptors that best distinguished between the two groups were the presence/absence of chlorine substituent in the structures and the molecular globularity.

Next, a pharmacophore analysis was applied in order to outline structural features responsible for the compound interactions with their putative biological target. A possible pharmacophore consists of three hydrophobic centres, and projected locations at the binding site of two potential H-bond acceptor or one H-bond acceptor and a potential metal ligator. This pharmacophore was present in all active compounds, and distinguished them from the non-active compounds by an accuracy of 0.7 (1.0 corresponding to fully distinguishable active and inactive compounds and 0.0 meaning that the model is completely inaccurate). These results are intended to be used for a rational design of pyrrole hydrazones as novel anti-tuberculosis drugs.

We acknowledge the financial support from the National Science Fund of Bulgaria (grant No. DTK02/58).

[1] Bijev A., Arzneimittelforschung (2006), **56**, 96-103.

[2] The inhibitory effect has been tested by TAACF (Tuberculosis Antimicrobial Acquisition and Coordinating Facility): http://www.taacf.org/

STRUCTURE–ACTIVITY RELATIONSHIPS OF THE SALICYLANILIDE ESTERS GROUP ACTIVE AGAINST MDR-TB

Martin Krátký [a], Jiřina Stolaříková [b], Jarmila Vinšová [a]

[a]Faculty of Pharmacy, Charles University, Department of Inorganic and Organic Chemistry, Heyrovského 1203, 500 05 Hradec Králové, the Czech Republic

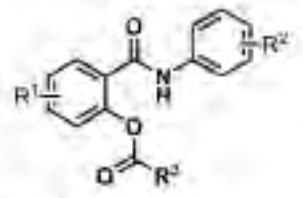
[b]Laboratory of TBC, Regional Health Institute in Ostrava, 702 00 Ostrava, the Czech Republic

According to WHO, tuberculosis is still harsh global public health problem. It is estimated that nearly 9 million of new cases develop each year. Increasing emergence of drug resistant tuberculosis, especially multidrug resistant tuberculosis (MDR-TB) and most recently the extremely drug-resistant tuberculosis (XDR-TB) is alarming. Every year almost 500,000 people are infected with MDR-TB. A multidrug resistant *Mycobacterium tuberculosis* strain is defined as resistant to at least the two basic anti-TB drugs, isoniazid and rifampicin. These reasons have made clear the pressing need for the evolution of new and more powerful drugs, especially with new mechanisms of action.[1]

Salicylanilides (2-hydroxy-N-phenylbenzamides) have been the subject of interest in medicinal chemistry, mainly due to their antimycobacterial and antifungal activity.[2] They exhibit a high activity against *Mycobacterium tuberculosis*, M. avium and M. kansasii, with MIC in the range of 1 – 32 μmol.[3] To improve their bioavailability, we have designed perspective models of salicylanilide amino acid esters and carbamate prodrugs with activity as well against MDR strains.[4,5]

Here we are presenting structure – activity relationships of the three series of halogenated salicylanilide esters (47 compounds). The esterification was performed with N-acetyl-L-phenylalanine, benzoic acid and pyrazine-2-carboxylic acid and the derivatives were assayed for their in vitro activity against one drug sensitive strain and six MDR-TB strains with different resistance profile. All of the compounds are active against *M. tuberculosis* with the MIC values from 0.125 to 8 μmol/L and against MDR-TB (0.125 – 4 μmol/L). On the salicylic ring, independently on the acid used for esterification, 4-chloro substitution is the most convenient for the antimycobacterial activity, slightly worse is 4-Br and the lowest is the contribution of 5-Cl (but these derivatives are still better than non-substituted one).

For the substitution of the aniline ring, the optimal moieties are 4-CF_3 and 3,4-diCl, followed by 3-Br and 3-CF_3, which are still enhancing markedly the antimycobacterial activity against drug sensitive and MDR strains; the moderate contribution has shown 4-Br. The monosubstitution by chlorine or fluorine provides minimal benefit when compared with others.



R^1 = 4-Cl, 4-Br, 5-Cl

R^2 = 3-Cl, 4-Cl, 3,4-diCl, 3-Br, 4-Br, 3-F, 4-F, 3-CF_3, 4-CF_3

R^3 = phenyl, (*S*)-1-acetamido-2-phenylethan-1-yl, pyrazin-2-yl

When we evaluated the MDR-TB efficacy of esters with respect to the used organic acid, we found pyrazine-2-carboxylic acid superior than benzoic acid and *N*-acetyl-L-phenylalanine, which lowest MICs were comparable. This result may be caused by the synergic activity of salicylanilides and pyrazin-2-carboxylic acid, an active form of pyrazinamide. In general, the most active compound in our investigation exhibited the MIC value against MDR-TB strains in the range of 0.125 – 0.25 μmol/L.

This work was financially supported by the Research project MSM 0021620822, GAUK 27610/2010 and IGA NS 10367-3.

References:

[1] Vinšová, J.; Krátký, M. Tuberculosis – The Development of New MDR-TB Drugs (chapter in book Drug-Resistant Tuberculosis). Nova Science Publishers, 2009, ISBN 978-1-60876-055-8.
[2] Vinšová, J.; Imramovský, A. *Cesk. Slov. Farm.* 2004, 53, 294-299.
[3] Waisser, K.; Bureš, O. et al. *Arch. Pharm*. 2003, 336, 53-71.
[4] Imramovský, A.; Vinšová, J. et al. *Bioorg. Med. Chem*. 2009, 17, 3572-3579.
[5] Férriz, JM.; Vávrová, K. et al. *Bioorg. Med. Chem*. 2010, 18, 1054-1061.

IN SILICO SCREENING METHODOLOGIES FOR THE DESIGN OF NOVEL TryS INHIBITORS WITH POTENTIAL ANTILEISHMANIAL ACTIVITY

Panagiotis Zoumpoulakis[1], Constantinos Potamitis[1] , Giorgio Maccari[2], Theodora Calogeropoulou[1], Maurizio Botta[2]

[1]Institute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation. 48 Vas. Constantinou Ave., 11635, Athens, Greece

[2] Department of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Siena. via Aldo Moro , 53100, Siena, Italy

Protozoan parasites of the trypanosomatidae family are the causative agents of various forms of Leishmaniasis (Leishmania species). Trypanothinone Synthetase (TryS) is the sole enzyme responsible for the biosynthesis of trypanothione in the human pathogenic parasites *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. Therefore, TryS inhibition is considered to be a particularly attractive strategy to fight leishmanial infections. A model of Leishmania major-TryS has been produced, based on the crystal structures of Leishmania major – Trypanothione Synthetase Amidase (TSA)[1] and E. coli – glutathionylspermidine synthetase (GSPS). Based on the docked complex of the TryS with glutathione, a pharmacophore model was generated (Figure 1). Pharmacophore-based *in silico* screening of commercially available compound libraries in combination with molecular docking studies has led to the identification of putative inhibitor candidates.

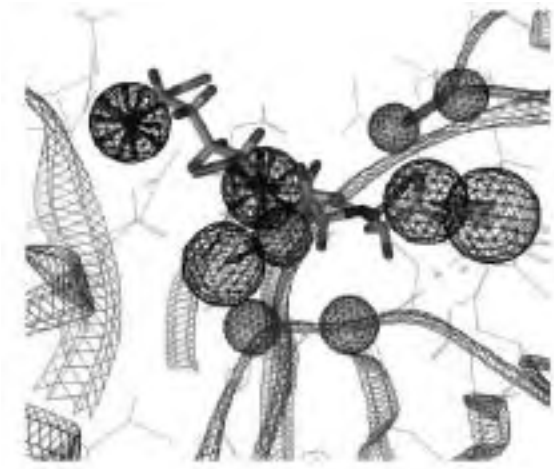


Figure 1: Generated pharmacophore model for the complex of glutathione-TryS, characterized by negative ionisable areas, H-bond donor, metal binding feature and exclusion volume spheres.

Reference

1. Fyfe, P.K.; Oza S.L.; Fairlamb A.H.; Hunter W. N. *Leishmania* Trypanothione Synthetase-Amidase Structute Reveals a Basis for Regulation of Conflicting Synthetic and Hydrolytic Activities. *J. Biol. Chem.* **2008**, 283, 17672-17680.

POSTER PRESENTATION II-89

DOCKING, CHEMICAL SIMILARITY AND HQSAR STUDIES FOR THE IDENTIFICATION OF *Trypanosoma cruzi* CRUZAIN INHIBITORS

Helton J. Wiggers[a,b], Josmar R. Rocha[a], Juliana Cheleski[a], Carlos A. Montanari[a,b*]

[a] Grupo de Estudos em Química Medicinal de Produtos Naturais - NEQUIME-PN, Instituto de Química de São Carlos - Universidade de São Paulo, 13560-970 – São Carlos – SP, Brazil

[b] Departamento de Química Orgânica da Universidade Federal de São Carlos – UFSCar, 13565-905 – São Carlos – SP, Brazil

Chagas disease is caused by the trypanosomatid parasite *Trypanosoma cruzi*. The drugs available are ineffective and have serious side effects. A validated target that has been studied for the discovery of new agents for chemotherapy treatment of Chagas Disease is the cysteine protease cruzain (EC 3.4.22.51). Cruzain is essential for the development and survival of the parasite within the host cells.

The objective of this work was to evaluate the performance of structure and ligand based virtual screening methods (SBVS and LBVS) to identify cruzain inhibitors. Database construction was carried out by selecting inhibitors from literature with IC_{50} or K_i < 100 nM, which resulted in a collection of 27 compounds. 1D property of these compounds were calculated and used for selecting 326 compounds from ZINC database. Cruzain 3D structure PDB 1ME4 was used for docking the dataset by programs with different search algorithms and scoring functions (FRED2.02, FlexX2.2.1 and Autodock Vina). 3D chemical similarity search using ROCS (Rapid Overlay of Chemical Structure) was also employed; the ligand from 1ME4 structure was used as query for searching the database. A HQSAR model was built using the PLS method with leave-one-out cross-validation. A series of 41 inhibitors containing *pKi* values ranging from 6.5 to 9 was divided in training (29) and test sets (12). The fit parameters of the best model were: $r^2 = 0.95$, $q^2 = 0.79$, $r^2_{pred} = 0.97$. In addition, an external validation was carried out using an independent test set of 10 compounds, which resulted in r^2_{pred} of 0.96. Methods performance was evaluated by the *Receiver Operating Characteristic* (ROC) curve[1].

The areas under the ROC curve (AUC) were: HQSAR 0.99, ROCS 0.84 with ColorTanimoto and Scaled color, FlexX 0.76 with PMF_SCORE, FRED 0.73 with Chemgauss3 and Autodock Vina 0.69. The HQSAR method has shown the best performance when compared to Docking and ROCS programs. LBVS has shown a better performance compared to SBVS docking approaches. However, by applying a rank-by-rank consensus docking, FlexX AUC increased to 0.89. Consensus of ROCS and FlexX has shown higher performance than applying each method alone, which indicates that by combining LBVS and SBVS methods the hit rate of selected compounds to be biochemically screened might be substantially increased.

Reference:

[1]Freitas, R.F. et al, Bioorg. Med. Chem. 16, p.838 (2008)

VIRTUAL SCREENING, CALORIMETRIC ASSAYS AND X-RAY CRYSTALLOGRAPHY FOR *Trypanosoma cruzi* DHODH INHIBITOR DISCOVERY

Juliana Cheleski[a], Josmar R. Rocha[a], Matheus P. Pinheiro[b], Helton José Wiggers[a], Maria C. Nonato[b], Carlos A. Montanari[a*]

[a] Grupo de Estudos em Química Medicinal de Produtos Naturais - NEQUIME-PN, Instituto de Química de São Carlos - Universidade de São Paulo, 13560-970 – São Carlos – SP, Brazil

[b]Laboratório de Cristalografia de Proteínas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, 14040-903 Ribeirão Preto, SP, Brazil

Chagas disease, caused by the protozoan parasite *Trypanosoma cruzi*, is one of the major health problems in several countries of Latin America.[1] *T. cruzi* dihydroorotate dehydrogenase enzyme (TcDHODH) is a promising target for the design of novel trypanocidal agents. This enzyme catalyzes the oxidation of L-dihydroorotate to orotate with concomitant reduction of fumarate to succinate in the de novo pyrimidine biosynthetic pathway. In this work, we report the discovery of novel inhibitors of TcDHODH identified by a combination of virtual screening and Isothermal Titration Calorimetry (ITC) methods.[2]

The structure-based drug design led us to identify novel TcDHODH and LmDHODH inhibitors acting at the catalytic site at low micromolar levels. For the first time, we have provided fundamental structural information of a new DHODH region of potential interaction with ligands that extends the catalytic site. This druggable region could perhaps be used in the design of Class 1A DHODH inhibitors. Overall, the orthogonal analysis (ITC and X-ray crystallography) enabled an unequivocal identification of the mode of action of these compounds as competitive inhibition, which is of remarkable importance in advancing the medicinal chemistry of such enzyme inhibitors.

References:

[1] Petherick, A. Nature, 465, S10-S11 (2010).

[2] Cheleski J. et al, Analytical Biochemistry, 399, 13-22 (2010).

IDENTIFICATION STUDIES OF POTENTIAL TARGETS FOR AZOMETHINIC AND OXADIAZOLINIC NITROCOMPOUNDS WITH DUAL ANTIFUNGAL AND ANTI-T. *cruzi* ACTIVITIES

Ieda Yuriko Sonehara, Alex Alfredo de Oliveira, Salomão Dória Jorge, Fanny Palace-Berl, Marina Ishii and Leoberto Costa Tavares

Faculty of Pharmaceutical Science – São Paulo University- Brazil

The studied compounds belong to four closely related series: **AzoO series:** 4-substituted 5-nitro-2-furfurylidene benzhydrazides; **AzoS Series:** 4-substituted 5-nitro-2-thiophylidene benzhydrazides; **OxaO Series:** 4-substituted 3-acetyl-oxadiazolyne 2,5-furfurylidene benzhydrazides; and **OxaS Series: 4**-substituted 3-acetyl-oxadiazolyne 2,5-thiophilydene benzhydrazides. These compounds were synthesized by the group, characterized spectrometrically, and found to be chemically pure.

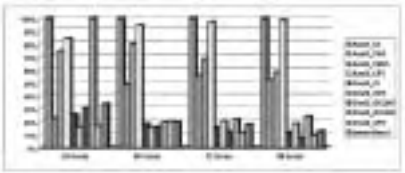
The compounds showed antifungal activity against Candida albicans, ATCC 537Y strain, with highest activity presented by the methyl-substituted oxadiazolinic compound in S configuration, OxaS-CH_3 (MIC 3,28 µg/mL). It was also demonstrated that the compounds inhibit growth in epimastigote forms of T. cruzi in concentrations of 10%, 20%, and 30%, in levels comparable to benznidazole.

All compounds were characterized by their SHED (*Shannon Entropy Descriptors*) profiles. This set of descriptors aims to describe chemical compounds according to the distribution of their hydrophobic, aromatic, and hydrogen bond donor and acceptor characteristics. SHED profiles were used as parameters for virtual profiling and in studies having as objective the identification of possible targets for the compounds under study.

Analysis of virtual profiling results pointed among the possible targets the P450 enzymes CYP19A (aromatase) and CYP3A4, which also interact with azole antifungal compounds whose target is CYP51 (14α-demethylase). The inhibition of CYP51 is responsible for the antifungal and *anti-T. cruzi* activity of azole compounds, and was thus investigated as a possible target responsible for the dual action of the studied compounds.

Characterization studies of CYP51 catalytic cavity considering not only SHED descriptors but also steric characteristics, allied to docking studies, confirmed the possibility of interaction of compounds from the oxadiazolynic series with CYP51, in an orientation similar to that of azole antifungals. The azomethynic series do not present an adequate conformation for the interaction as proposed; however, the interaction of compounds in a different way from that of oxadiazolynic compounds cannot be excluded for there have been reports showing an alternative binding mode for compounds unrelated to the azole family. The results are also coherent with the higher antifungal activity detected in compounds from the oxadiazolynic series.

Docking result of oxadiazolynic compounds at the CYP51 catalytic site



Growth inhibition results against epimastigote forms of T.cruzi. Compound concentration was 10µM.

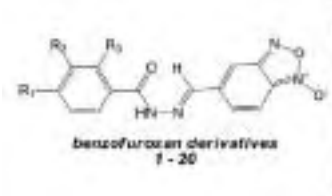
ACKNOWLEDGMENTS: The authors thank CNPq, CAPES, and FAPESP for the financial support. The authors also thank Dr. Jordi Mestres from the Research Unit on Biomedical Informatics (GRIB/IMIM), Barcelona, Spain, for kindly providing access to the tools necessary for the determination of SHED profiles, virtual profiling, docking, and cavity characterization studies.

STRUCTURE-ACTIVITY RELATIONSHIPS AND *IN VITRO* STUDIES OF A SET OF BENZOFUROXAN DERIVATIVES AS POTENTIAL ANTI-*Trypanosoma Cruzi* AGENTS

Salomão Dória Jorge, Marina Ishii, Fanny Palace-Berl, Alex Alfredo de Oliveira, Ieda Yuriko Sonehara, Kerly Fernanda Mesquita Pasqualoto and Leoberto Costa Tavares

Faculty of Pharmaceutical Sciences, University of São Paulo, Brazil

Chagas's disease – also called as American trypanosomiasis – is caused by the flagellate *Trypanosoma cruzi* and affects an estimated 8 to 14 million people in Latin America being 14,000 deaths associated to this disease every year [1]. The therapeutic strategies are restricted to benznidazole and nifurtimox, which are drugs with severe side effects. The present study involved a search for new leads with better pharmacological profile, and the *in vitro anti-T.cruzi* activity of a series of twenty *N'*-(benzofuroxan-5-yl)methylene benzohydrazides was investigated. The in vitro assay was carried out on epimastigote forms of *T. cruz*i Y strain and the results were expressed as the concentration of compound which causes a fifty percent of parasite growth inhibition (IC_{50}) [2].



	R_1	R_2	R_3	IC_{50} (μM)
1	H	H	H	16.92
2	CH_3	H	H	11.06
3	NH_2	H	H	116.96
4	OH	H	H	
5	F	H	H	15.29
6	CN	H	H	26.58
7	CH_2CH_3	H	H	9.91
8	OCH_3	H	H	13.2
9	Cl	H	H	9.96
10	$COCH_3$	H	H	16.01
11	$N(CH_3)_2$	H	H	
12	OCH_2CH_3	H	H	9.62
13	NO_2	H	H	17.84
14	CF_3	H	H	9.62
15	Br	H	H	7.38
16	SO_2NH_2	H	H	60.38
17	I	H	H	6.11
18	Cl	Cl	H	9.54
19	Cl	H	Cl	9.07
20	NO_2	H	CF_3	15.93
benznidazole				20.84

The two most active compounds were the 4-I (**17**) and 4-Br (**15**) substituted derivatives having IC_{50} values of 6.11 μ*M* (approximately 2.50 μg/mL) and 7.38 μ*M* (2.80 μg/mL), respectively. Otherwise, the lowest activity was observed for the 4-NH_2 (**2**) and 4-SO_2NH_2 (**16**) substituted derivatives, which presented IC50 values of 116.96 μ*M* (34.74 μg/mL) and 60.38 μ*M* (21.80 μg/mL), respectively. The 4-OH (**4**) and 4-$N(CH_3)_2$ (**11**) substituted derivatives did not show any anti-*T.cruzi* activity at the concentrations tested.

The influence of the substituent's physicochemical properties on *in vitro* activity of the compounds was also investigated. In general, the presence of hydrophobic and electron-withdrawing substituent at **4** position in the benzene ring resulted in more active compounds than the reference drug benznidazole (IC_{50} = 20.84 μ*M*; 5.42 μg/mL). The three 3,4-disubstituted derivatives (**18,19,20**) also showed better anti-*T. cruzi* activity than benznidazole.

Hydrophobic functional groups, such as 4-CH_3 (**2**) and 4-OCH_3 (**8**), attached to the benzene moiety provided a favorable influence on anti-T.cruzi activity. However, hydrophilic groups, such as 4-NH_2 (**2**), 4-CN (**6**) and 4-SO_2NH_2 (**16**), had an unfavorable influence on activity. The increase of the alkyl chain employing the 4-CH_2CH_3 (**2**) and 4-OCH_2CH_3 (**8**) groups improved the *in vitro* activity. The electronic effect also played an important role in activity, as can be observed for the compounds containing an electron-withdrawing substituent (4-Cl (**9**), 4-Br (**15**), and 4-I (**17**)).

The findings highlight the benzofuroxan derivatives as potential leads for designing new future anti-*T.cruzi* drug candidates.

Acknowledgments: The authors thank Brazilian Governmental Agencies CNPq, FAPESP and CAPES for the financial support.

References

[1] Ribeiro, I.; Sevcsik, A.; Alves, F.; Diap, G.; Don, R.; Harhay,M. O.;Chang, S.; Pecoul, B. *PLoS Negl. Trop. Dis*. 3(7), 2009, e484.
[2] Bollini, M.; Casal, J. J.; Alvarez, D. E.; Boiani, L.; Gonzalez, M.; Cerecetto, H.; Bruno, A. M. *Bioorg. Med. Chem*. 17 (4), 2009, 1437-1444.

COMPUTATIONAL DISCOVERY OF NOVEL TRYPANOSOMICIDAL DRUG-LIKE CHEMICALS BY USING BOND-BASED NON-STOCHASTIC AND STOCHASTIC QUADRATIC MAPS AND LINEAR DISCRIMINANT ANALYSIS

Juan Alberto Castillo-Garit,[a,b,c*] Maria C. Vega,[d] Miriam Rolon,[d] Yovani Marrero-Ponce,[b,c,f] Vladimir V. Kouznetsov,[g] Diego Fernando Amado Torres,[g] Alicia Gómez-Barrio,[d] Alfredo Alvarez Bello,[b] Alina Montero,[b] Francisco Torrens,[c] and Facundo Pérez-Giménez.[f]

[a]Applied Chemistry Research Center, Faculty of Chemistry-Pharmacy and Department of Drug Design, Chemical Bioactive Center, Central University of Las Villas, Santa Clara, 54830, Villa Clara, Cuba.
[b]Unit of Computer-Aided Molecular "Biosilico" Discovery and Bioinformatic Research (CAMD-BIR Unit), Faculty of Chemistry-Pharmacy. Central University of Las Villas, Santa Clara, 54830, Villa Clara, Cuba.
[c]Institut Universitari de Ciència Molecular, Universitat de València, Edifici d'Instituts de Paterna, P.O. Box 22085, E-46071, València, Spain.
[d]Departamento de Parasitología, Facultad de Farmacia, UCM, Pza. Ramón y Cajal s/n, 28040 Madrid.
[f]Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Departamento de Química Física, Facultad de Farmacia, Universitat de València, Spain.
[g]Laboratorio de Química Orgánica y Biomolecular, Escuela de Química, Universidad Industrial de Santander, Bucaramanga, Colombia.

Herein we present results of a Quantitative Structure-Activity Relationship (QSAR) studies to classify and design, in a rational way, new antitrypanosomal compounds by using non-stochastic and stochastic bond-based quadratic indices.[1] A data set of 440 organic chemicals, 143 with antitrypanosomal activity and 297 having other clinical uses, is used to develop QSAR models based on Linear Discriminant Analysis (LDA). Non-stochastic model correctly classifies more than 93% and 95% of chemicals in both training and external prediction groups, respectively. On the other hand, the stochastic model shows an accuracy of about the 87% for both series. As an experiment of virtual lead generation, the present approach is finally satisfactorily applied to the virtual evaluation of 8 already synthesized *in house* compounds. The in vitro antitrypanosomal activity of this series against epimastigote forms of *Trypanosoma cruzi* is assayed.[2] The model is able to predict correctly the behaviour for the majority of these compounds. Three compounds (FER16, FER32 and FER33) showed more than 70% of epimastigote inhibition at a concentration of 100μg/mL (86.74%, 78.12% and 88.85%, respectively) and two of these chemicals, FER16 (78.22% of AE) and FER33 (81.31% of AE), also showed good activity at a concentration of 10μg/mL.[3] At the same concentration, compound FER16 showed lower value of cytotoxicity (15.44%), and compound FER33 showed very low value of 1.37%. Taking into account all these results, we can say that these three compounds can be optimized in forthcoming works, but we consider that compound FER33 is the best candidate. Even though none of them resulted more active than Nifurtimox, the current results constitute a step forward in the search for efficient ways to discover new lead antitrypanosomals.[3]

1. Marrero-Ponce, Y., *Molecules* **2003**, 8, 687-726.
2. Vega, M.C., Rolon, M., Martinez-Fernandez, A.R., Escario, J.A., Gomez-Barrio, A., *Parasitol. Res*. **2005**, 95, 296-298.
3. Castillo-Garit J.A., Vega M.C., Rolon M., Marrero-Ponce Y., Kouznetsov V., Torres D.F., et al. *Eur J Pharm Sci*; **2010**, 39, 30-6.

COMBINING MOLECULAR DOCKING AND QSAR STUDIES FOR MODELLING THE ANTIGYRASE ACTIVITY OF CYCLOTHILIDINES DERIVATIVES: AN ALTERNATIVE APPROACH FOR GETTING INSIGHT INTO THE STRUCTURE-ACTIVITY RELATIONSHIPS

Liane Saiz-Urra[1,2], Eveline Lescrinier[2], Miguel Ángel Cabrera Pérez[1], Aliuska Helguera[1,3], Matheus Froeyen[2]

[1] Chemical Bioactive Center, Central University of Las Villas, Santa Clara, Villa Clara, C.P. 54830, Cuba.
[2] Laboratory for Medicinal Chemistry, Rega Institute for Medical Research, Katholieke Universiteit Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium.
[3] CIQ, Department of Chemistry, University of Porto, 4169-007 Porto, Portugal

The increasing worldwide death toll in the world caused by bacterial infections diseases together the emergence of multi-drug resistant strains encourages the search for new antimicrobial agents. DNA gyrase is a well-established antibacterial target consisting of two subunits, GyrA and GyrB, in a heterodimer A_2B_2. GyrA is involved in DNA breakage and reunion while GyrB catalyzes the hydrolysis of ATP. Cyclothialidine (Ro 09-1437) has been considered as a promising inhibitor whose modifications might lead to more potent compounds against the enzyme.

Herein, is reported for the first time, QSAR studies regarding to ATPase inhibitors of DNA Gyrase which were developed by using constitutional, 1D, 2D and 3D descriptors from DRAGON software, and a set of 61 cyclothialidine derivatives to model the inhibition of the DNA GyrB subunit, expressed by the maximum noneffective concentration (MNEC) in μg/mL. Multiple Regression Analysis (MRA) was the technique applied to build the models and the variables were selected using a genetic algorithm as implemented in the MobyDigs software. OpenEye package was used to generate the 3D structure of the compounds. Based on the core of the cyclothialidine GR1222222, different conformations were created by using OMEGA, then FRED was used to dock these conformers in the cavity of the GyrB subunit to choose the best conformations taking special attention to the 12 members ring.

Three QSAR models were developed attending to the dimension of the descriptors. The models were robust and predictive and good in statistical significance, over 70% of the experimental variance was explained. Interpretability of the models was possible by extracting the SAR(s) encoded by these predictive models. Generally, the most important structural features from the compounds that are involved in non covalent interactions with the receptor were elucidated in agreement with the x-ray structure described in the literature. Analysis of the interactions of most active compounds in the binding pocket was done. The fact that 3D descriptors are conformation dependent can be a problem when performing ligand based approaches, but taking the input 3D structures generated by the abovementioned technique, allowed to increase the reliability of the information obtained for the QSAR model.

Maxwell, A., DNA gyrase as a drug target. Trends Microbiol 1997, 5, (3), 102-9.

Angehrn, P. et al., New antibacterial agents derived from the DNA gyrase inhibitor cyclothialidine. J Med Chem 2004, 47 (6), 1487-513.

Acknowledgments: We are grateful to Flemish Interuniversity Council (Belgium) for financial support to the project "Strengthening postgraduate education and research in Pharmaceutical Sciences" as well as to the AECID project: "Montaje de un laboratorio de química computacional, con fines académicos y científicos, para el diseño racional de nuevos candidatos a fármacos en enfermedades de alto impacto social (D/024153/09))". Finally we want to thank to Prof. Wigley for the contribution to this work with the x-ray structure of the protein ligand to the GR1222222.

SYNTHESIS AND ANTI-INFECTIVE PROPERTIES OF SOME SUBSTITUTED PYRAZINEDINITRILES

Martin Doležal[1], Ondřej Janďourek[1], Jiří Kačírek[1], Marcela Vejsová[1], Michaela Svobodová[2]

[1] Faculty of Pharmacy in Hradec Kralove, Heyrovskeho 1203, 500 05 Hradec Kralove, Czech Republic
[2] Department of Microbiology, Regional Hospital, Kyjevska 44, Pardubice, 532 03, Czech Republic

The widespread of Mycobacteria strains resistant to current treatment methods is an important epidemiological problem recognized by WHO[1]. Pyrazinamide, the first line antituberculotic substance, was used as a model compound for other chemical modifications in our research project. Some pyrazinamide analogues, 3-arylaminopyrazine-2,5-dicarbonitriles, have already demonstrated good antimycobacterial and/or antifungal activity in a previous study[2]. To prepare position isomers of these pyrazinedicarbonitriles, 5-chloro-6-methylpyrazine-2,3-dicarbonitrile[3] (I) was treated with various non-aromatic/aromatic amines and 5-methyl-6-(alkyl/arylamino)pyrazine-2,3-dicarbonitriles (II) were yielded. Prepared compounds were chemically characterized and screened for their antimycobacterial (*Mycobacterium tuberculosis RvH37, M. tuberculosis wild stem, M. kansasii, M. avium 80/72, M. avium 152/73*), antibacterial (*Staphylococcus aureus* CCM 4516/08, *Escherichia coli* CCM 4517, *Pseudomonas aeruginosa* CCM 1961, *Staphylococcus aureus* H 5996/08 – methicilin resistant, *Staphylococcus epidermidis* H 6966/08, *Enterococcus sp.* J 14365/08, *Klebsiella pneumoniae* D11750/08, *Klebsiella pneumoniae* J 14368/08 – ESBL positive) and antifungal activity (*in vitro* testing against *Candida albicans, C. tropicalis, C. crusei, C. glabrata, Trichosporon asahii, Trichophyton mentagrophytes, Aspergillus fumigatus, and Absidia corymbifera*).

5-Methyl-6-(octylamino)pyrazine-2,3-dicarbonitrile (II, R^1 = H, R^2 = octyl) possessed better antimycobacterial effect (MIC = 8 μg.ml^{-1}) against *M. tuberculosis RvH37* than pyrazinamide (MIC = 32 μg.ml^{-1}), an interesting activity against *Staphylococcus aureus* CCM 4516/08 (MIC = 15.62 μmol.l^{-1}, comparable with the effect of bacitracin), and specific activity against *Trichophyton mentagrophytes* (MIC = 1.96 μmol.l^{-1}, standard amphotericine B: MIC = 0.98 μmol.l-1), and did not inhibit the growth of other tested fungi species (MIC > 250 μmol.l^{-1}).

This study was supported by the Ministry of Health of the Czech Republic (IGA NS 10367-3).

References:

1. Anti-tuberculosis drug resistance in the world, Fourth Global Report, WHO/HTM/TB/2008.394, Accessed at http://www.who.int/tb/publications/2008/drs_report4_26feb08.pdf, November 5th, 2008
2. Palek, L.; Dvořák, J.; Svobodová, M.; Buchta, V.; Jampílek, J.; Doležal, M. *Arch. Pharm*. **2008**, 341, 61.
3. Nakamura, A.; Ataka, T.; Segawa, H.; Takeuchi, Y.; Takematsu, T. Agric. *Biol. Chem*., **1983**, 47, 1561.
4. Jones, R. N.; Barry, A. L. *J. Clin. Microbiol*., **1987**, 25, 1920-1925.

EVALUATION OF ANTISEPTIC ACTIVITY OF DISINFECTANT AQUEOUS SOLUTIONS

Biljana Gjorgjeska, Uzunoska Zora
Faculty of medical science, University "Goce Delcev"-Stip

Knowing antiseptic activity of chemical disinfectant substances has great practical value. It is evidential that there is the need for defining standard technique for quantitative determination of bactericidal activity of chemical disinfectant substances, as well as the need for defining parameter for comparing various chemical disinfectants. Solution of phenol (5%) was considered as referent standard for evaluation of efficacy of disinfectant aqueous solutions. Phenol coefficient shows how many times bactericidal activity of examined disinfectant is greater or lower than bactericidal activity of standard phenol solution (5%). However, phenol coefficient gives only limited information. Suitability of phenol coefficient for evaluation of nonphenolic disinfectants is still opened question. On the other side the methods for evaluation of antiseptic activity of disinfectant aqueous solutions are microbiological methods.

The aim of this study is to develop a new empirical coefficient which is capable to express the various physic-chemical properties of disinfectant solutions on bactericidal activity. The basic duty of this parameter (Disinfection Activity Coefficient of Solution - DACS) is to express capability for comparison and prediction of disinfectant activity. The DACS index, which is the sum of four terms (fluidity, surface tension, redox potential and osmolality), results in good correlation with the activity at different disinfectant aqueous solutions. The DACS index can be calculated using additive and statistical models. Statistical model is adequate for evaluation of different disinfectant solutions because of better expressing the bactericidal activity then additive model. For analyze of various dilutions of one disinfectant there is no significant difference between this two models. The usefulness of DACS is demonstrated for analyze of bactericidal activities on different disinfectant solutions containing boric acid, chlorhexidne, chlorhexidine with cetrimide, chloroxylenol, chlorophen, eosin, hydrogen peroxide, phenyl mercury borate, povidon-iodine, thiomersal, tosilchloramide and phenol. Results for bactericidal activities obtained from microbiological tests on Staphylococcus aureus was compared with activities predicted with DACS. As the conclusion, it is considered good correlation between experimental and calculated values for bactericidal activity.

ARTIFICIAL NEURAL NETWORK PREDICTION OF PSYCHOMETRIC ACTIVITIES OF PHENYLALKYLAMINES BY USING DFT CALCULATED MOLECULAR DESCRIPTOR

Mina Haghdadi [1*] and Mohammad H. Fatemi [2]

[1]Department of Chemistry, Islamic Azad University, Babol Branch, Iran

[2]Department of Chemistry, Mazandaran University, Babolsar, Iran

In the present work, a quantitative structure–activity relationship (QSAR) method was used to predict the psychometric activity values (as mescaline unit, logMU) of 48 phenylalkylamine derivatives from their density functional theory (DFT) calculated molecular descriptors and artificial neural network (ANN). In the first step the molecular descriptors were obtained by DFT calculation at 6-311G* level of theory. Then the stepwise multiple linear regression method was employed to screen descriptor spaces. These parameters are: hardness, chemical potential, lowest unoccupied molecular orbital energy, dipole moment and vibrational energy. In the next step, the artificial neural network and multiple linear regressions (MLR) models were developed to construct the nonlinear and linear QSAR models, respectively. The standard errors in the prediction of the log MU by MLR model are; 0.398, 0.443 and 0.427 for training, internal and external test sets, respectively, while these values for the ANN model; 0.132, 0.197 and 0.202, respectively. The predictivity of these models were further examined by cross-validation test which produce the statistics of $Q^2 = 0.623$ and SPRESS= 0.472 for the MLR model and $Q^2 = 0.911$ and SPRESS= 0.147 for the ANN model. The obtained results show the applicability of QSAR approaches by using ANN techniques in prediction of log MU of phenylalkylamine derivatives from their DFT calculated molecular descriptors.

References:

1. M. Thakur, A. Thakur, P. V. Khadikar, Bioorg. Medicin. Chem. 12 (2004) 825.
2.H. Fatemi, S. Garaghani, Bioorg. Medicin. Chem. 15(2007) 7746

MODELLING OF A SHORT-CHAIN DEHYDROGENASE/REDUCTASE AND IN SILICO SCREENING FOR POTENTIAL LIGANDS

Juliane Fischer[a], Nicole Reinhardt[b], Birgit Dräger[b], Wolfgang Brandt[a], Ludger Wessjohann[a]

[a]Leibniz-Institute of Plant Biochemistry, Weinberg 3, D-06120 Halle (Saale), Germany
[b]Institute of Pharmacy, Faculty of Science I, Martin Luther University Halle-Wittenberg, Hoher Weg 8, D-06120 Halle (Saale), Germany

The enzyme family of short-chain dehydrogenases/reductases (SDR) contains proteins with a sequence length of 250 to 350 amino acids and unique motifs. Tropinone reductase-like SDRs from the model species *Arabidopsis thaliana* were reported from gene annotation. Tropinone reductases participate in the biosynthesis of tropane alkaloids in plants like deadly nightshade (*Atropa belladonna*) or thornapple (*Datura stramonium*), however, Arabidopsis thaliana is a member of Brassicaceae which does not contain tropane alkaloids. Therefore, the native function of the tropinone reductase-like SDRs is unknown. In biochemical analyses some of these enzymes, obtained by heterologous proteins synthesis, did not reduce tropinone to tropine or pseudotropine. In order to gather more information on the function of these enzymes in the Arabidopsis plant, protein models were created by homology modelling. The models formed the foundation for the *in silico* screening of possible ligands. Pharmacophore models were used for the search in data bases. The hits from the search were analysed with the docking program PLANTS. Furthermore plant metabolites which appeared as possible substrates, e.g. monoterpene and sesquiterpene ketones or alcohols, were docked. Docking hits were analysed in enzyme incubation assays. The experiments show some new hints for putative functions of the tropinone reductase-like SDRs in *Arabidopsis thaliana*.

INTERVET/SCHERING-PLOUGH ANIMAL HEALTH'S BIOINFORMATICS PORTAL - A CUSTOMIZED GLOBAL PP WEBPORTAL

Frank Oellien, Andreas Rohwer, Hon Q. Tran, Andreas Krasky, Richard J. Marhöfer, Paul M. Selzer

Intervet Innovation GmbH, D-55270 Schwabenheim, Germany

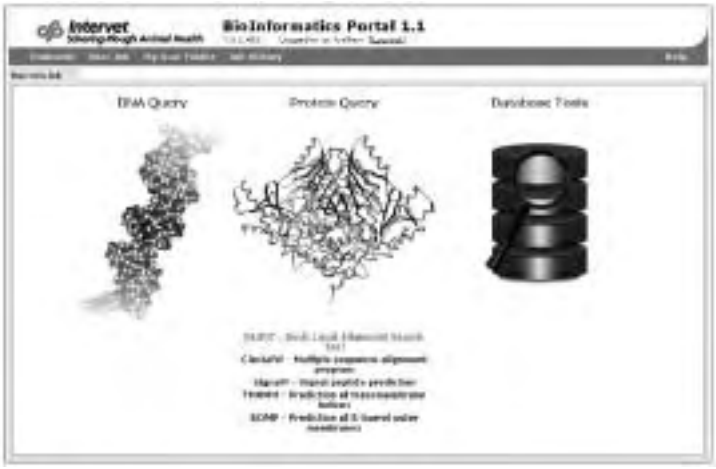
We present SP Intervet's BioInformatics Portal which is based on Pipeline Pilot's webport solution. The portal has replaced our former proprietary bioinformatics suite and is now globally used for all common bioinformatics tasks at Intervet / Schering-Plough Animal Health.

Many functionalities are based upon existing webport and Pipeline Pilot technologies. However, to fulfil other criteria like corporate identity, comparability (usability and functionality) with our former bioinformatics application, and appearance like a real desktop application, we had to extend and customize the out-of-the-box webport client and also large parts of the existing bioinformatics components. The functionalities are accessible via a completely new webport view (see figure). Furthermore, we have added a Windows-like explorer view to the webport that allows the user full access to his data.

Almost all NCBI BLAST functionalities have been re-implemented in the portal and can be used with the NCBI look and feel. The most important public sequence databases are available within the application and will be updated daily. This allows our users to use the intranet solution instead of the NCBI tools which prevents the usage of proprietary sequence information in public search services.

Furthermore, users can upload their own proprietary sequence databases which can automatically be accessed by all implemented functions right after the upload has been completed. In addition, we have added a relational database to the webport which is storing proprietary data sets like in-house genomes and EST data.

We will close the presentation with a short outlook to our current developments: i) building a PP solution which will replace our self-developed application for comparative genome analysis, ii) PP solutions that allow an automation of our Isentris-based compound registration.



POSTER PRESENTATION II-100

TRYPANOTHIONE REDUCTASE: A TARGET PROTEIN FOR A COMBINED IN SILICO AND *IN VITRO* SCREENING APPROACH

Mathias Beig[1,2,3], Frank Bender[1], Frank Oellien[1], Andreas Rohwer[1], Michael Gaßel[1], Richard J. Marhöfer[1], Gottfried Unden[2], R. Luise Krauth-Siegel[3], Paul M. Selzer[1]

[1] Intervet Innovation GmbH, Zur Propstei, Schwabenheim, Germany

[2] Institut für Mikrobiologie und Weinforschung, Johannes Gutenberg Universität Mainz, Germany

[3] Biochemie-Zentrum der Universität Heidelberg, Heidelberg, Germany

Trypanosomes are the causative agents of severe tropical diseases such as African sleeping sickness, Chagas' disease and Nagana cattle disease. New therapeutic compounds are urgently needed, because most available pharmaceuticals are outdated, too expensive, or afflicted with serious side effects. The parasites lack both glutathione reductase and thioredoxin reductase but have a phylogenetically related trypanothione reductase (TR) which catalyzes the NADPH dependent reduction of trypanothione disulfide [TS_2] to the dithiol trypanothione [bis(glutathionyl)spermidine, $T(SH)_2$]. By inhibiting this essential flavoenzyme, the parasite becomes vulnerable to oxidative stress induced by the host defense system and drug treatment. Here we present a combination of an *in silico* and *in vitro* screening approach to identify novel TR inhibitors. For the in vitro experiments, the NADPH-linked photometric assay was adapted to 384 well microtiter plates. A TS_2 concentration of 150 µM was chosen with the aim to mimic highly oxidative stress conditions and to discriminate against weak competitive inhibitors. As variation of a typical target-based approach, we started with the in vitro screening of a highly diverse compound library because the huge TS_2-binding site of the enzyme with a dimension of 2.2 x 2.0 x 2.8 nm^3 is not well suited for virtual screening experiments. Screening of about 3000 compounds against *T. cruzi* TR resulted in promising hits with 19 compounds showing an even better inhibition than chlorhexidine, a known competitive inhibitor with a K_i of 2 µM. The 64 best hits were verified by determination of IC50 values and a quality control through LC/MS and NMR measurements. Afterwards, the 22 best hits were used for similarity searches with several fingerprints on a compound library containing more than 200.000 compounds. In vitro testing of virtual hits resulted in compounds showing IC50s up to nM range.

INHIBITION OF *EIMERIA TENELLA* CDK-RELATED KINASE 2: BIOINFORMATICS ANALYSIS AND BIOCHEMICAL CHARACTERIZATION

María L. Suárez Fernández[1,2], Kristin Engels[1,a], Frank Bender[1], Michael Gaßel[1], Richard J. Marhöfer[1], Gottfried Unden[2], Jeremy C. Mottram[3] and Paul M. Selzer[1,3,4]

[1]Intervet Innovation GmbH, Zur Propstei, 55270 Schwabenheim, Germany;
[2]Institut für Mikrobiologie und Weinforschung, Johannes-Gutenberg-Universität, Mainz, Germany;
[3]Wellcome Trust Centre for Molecular Parasitology and Division of Infection & Immunity, Faculty of Biomedical and Life Sciences, University of Glasgow, G12 8TA, UK;
[4]Eberhard-Karls Universität Tübingen, Interfakultäres Institut für Biochemie, Hoppe-Seyler-Str. 4, 72076 Tübingen, Germany; (a) current address: Bayer Schering Pharma AG, Medicinal Chemistry, Aprather Weg 18a, 42096 Wuppertal, Germany

Apicomplexan parasites enclose several human-pathogenic as well as animal-pathogenic protozoans, like *Eimeria tenella, Toxoplasma gondii and Plasmodium falciparum*. The animal-pathogenic representative, *E. tenella* provokes coccidiosis a poultry disease, which causes tremendous losses to the world poultry industry. Considerable increase of drug resistance makes it necessary to develop and pursue new therapeutic strategies. Cyclin-dependent kinases (CDKs) are key molecules in the regulation of the cell cycle and are therefore prominent target proteins in parasitic diseases. To date several proteins from apicomplexan parasites, homologous to mammalian CDKs, have been characterized using classical molecular biology techniques. Our extensive bioinformatics analysis revealed additional candidate proteins in *E. tenella*. Besides EtCRK2, the only *E. tenella* CRK identified so far (Kinnaird et al., 2004. Int. J. Parasitol. 34, 683-692), three new CDK-like proteins were identified. Cyclins (CDK binding regulatory subunits) are required for the activation of CDKs. Regarding these activating cyclins of E. tenella, bioinformatics analysis revealed a situation similar to that found in *P. falciparum*. For biochemical characterization of these cyclins, one of them EtCYC3a, was cloned *in vitro*, expressed i*n Escherichia coli* and purified in complex with EtCRK2. Using the non-radioactive TR-Fret assay, we demonstrated the ability of EtCYC3 to activate EtCRK2. Furthermore we found that XlRINGO (Xenopus laevis rapid inducer of G2/M progression in oocytes) which is a non-cyclin activator (Ferby *et al*., 1999. Genes Dev. 13, 2177-2189), is also able to activate EtCRK2. IC50 determination of known CDK inhibitors (Dai *et al*., 2003. *Curr. Opin. Pharmacol*. 3, 362-370) resulted in similar values for the EtCRK2/XlRINGO and EtCRK2/EtCYC3a complex.

• Literature: Kinnaird, K.; Bumstead, J.; Mann, D.; Ryan, R.; Shirley, M.; Shiels, B.; Tomley, F. (2004) EtCRK2, a cyclin-dependent kinase gene expressed during the sexual and asexual phases of the Eimeria tenella life cycle. *Int. J. Parasitol*. 34, 683-692
• Ferby, I.; Blazquez, M.; Palmer, A.; Eritja, R.; Nebreda, A. (1999) A novel p34cdc2-binding and activating protein that is necessary and sufficient to trigger G2/M progression in *Xenopus oocytes. Genes Dev*. 13, 2177-2189
• Dai, Y.; Grant, S. (2003) Cyclin-dependent kinase inhibitors. *Curr. Opin. Pharmacol*. 3, 362-370

MHC CLASS II BINDING PREDICTION BY MOLECULAR DOCKING

Atanasova, M., Dimitrov, I., Flower, D. R., Doytchinova, I.
Faculty of Pharmacy, Medical University of Sofia, 2 Dunav str, 1000 Sofia, Bulgaria

Major histocompatibility complex (MHC) proteins are glycoproteins that bind self and nonself antigens (peptide fragments or epitopes) within the cell and present them at the cell surface for recognition by T cells. T cells cause activation of immune response. All T-cell epitopes are MHC binders but not all MCH binders are T-cell epitopes. The MHC class II proteins (found only on B lymphocytes, macrophages, and other cells that present antigens to T cells) present peptides which have been digested from external sources. MHC class II proteins are extremely polymorphic. The polymorphic residues are clustered in the peptide-binding region and are responsible for the binding peptide requirements. The peptide binding site of MHC proteins consist of 5 binding pockets. In the present work we model the interactions peptide - MHC class II proteins from locus DRB1 by molecular docking (GOLD suite v. 4.1.2, CCDC). A combinatorial peptide library was generated by mutation of residues at peptide positions corresponding to the binding pockets (so called anchor positions). We used several different protocols and the binding affinities were assessed using scoring functions. The normalized scoring functions for each amino acid at each anchor position were used to construct quantitative matrices (QM) for MHC class II binding prediction. Models were validated by external test sets. Good correlations between predicted and experimental IC50 values were achieved.

MHC CLASS II BINDING PREDICTION BY PROTEOCHEMOMETRICS

Dimitrov, I., Garnev, P., Flower, D. R., Doytchinova, I
Faculty of Pharmacy, Medical University of Sofia, 2 Dunav str, 1000 Sofia, Bulgaria

T-cell epitope identification is a critical immunoinformatic problem for vaccine design. To be an epitope, a peptide must bind an MHC protein. Here we present EpiTOP, the first server predicting MHC class II binding based on proteochemometrics, a QSAR approach for ligands binding to several related proteins. EpiTOP uses a quantitative matrix to predict binding to 12 HLA-DRB1 alleles. It identifies 90% of the known epitopes in the top 20% of the predicted binders, reducing laboratory labour, materials, and time by 80%. EpiTOP is easy to use, gives comprehensive quantitative predictions, and will be updated with new quantitative matrices over time. EpiTOP is freely accessible at http://www.pharmfac.net/EpiTOP.

POSTER PRESENTATION II-104

QSAR STUDY ON POLYPHENOLS AND THEIR DERIVATES ISOLATED FROM PLANT MATERIALS AND FOOD AS INHIBITORS OF α-GLUCOSIDASE

Sonja Nikolić[1], Vesna Rastija[2], Drago Bešlo[2], Dragan Amić[2]
[1] The Rugjer Boskovic Institute, P.O. Box 180, Zagreb, 10 002, Croatia
[2] Faculty of Agriculture, Department of Chemistry, Trg Sv. Trojstva 3, Osijek, 31 000, Croatia

Diabetes mellitus is a chronic disease caused by inherited and/or acquired deficiency in insulin secretion and/or by decreased responsiveness of the organs to secreted insulin. Such a deficiency results in increased blood glucose levels that in turn can damage many of the body's systems, including blood vessels and nerves. Small intestinal α-glucosidase and pancreatic α-amylase are key enzymes of dietary carbohydrate digestion in humans. Inhibitors of these enzymes may be effective in retarding carbohydrate digestion and glucose absorption to suppress postprandial hyperglycaemia. Some polyphenols are well known inhibitors of intestinal α-glucosidase. The aim of this study was to derive a Quantitative Structure-Activity Relationship (QSAR) analysis for α-glucosidase inhibition potential of 48 polyphenols and their derivates.
"Two-dimensional" (2D), "three-dimensional" (3D) molecular descriptors, quantum chemical descriptors, and physicochemical parameters, derived from optimised three-dimensional structure, have been calculated applying the PCLIENT (an on-line version of Dragon program) and HyperChem 8.0. Multiple regression analysis was performed using STATISTICA 7.0. Selection of predictor variables for multiple regression was performed by best-subset method. Dataset was divided into training set and test by manual selection and by cluster analysis, and then both method have been compared. The generated QSPR models were validated by leave-one-out cross-validation procedure using the CROMRsel (Rugjer Bošković Institute, Zagreb).
Presented QSAR could be a useful for the design of new molecules more efficient in inhibition of enzymes responsible for the glucose production from carbohydrates. Also, this study will be helpful in illumination of antihyperglycemic effects of certain medicinal plants and food.

ELECTROSPRAY MASS SPECTROMETRY AS A POTENTIAL TOOL FOR DETERMINING THE LIPOPHILIC PROPERTIES OF A SERIES OF STRUCTURALLY RELATED COMPOUNDS

E. Gikas[1], F. Bazoti[2], V. Myrianthopoulos[1], E. Mikros[1], A. Tsantili-Kakoulidou[1]

[1]National and Kapodistrian University of Athens, School of Pharmacy, Division of Pharmaceutical Chemistry, Athens, 157 71, Greece

[2]GAIA Research Center, Bioanalytical Department, The Goulandris Natural History Museum, 13 Levidou Street, Kifissia 145 62, Greece

Assessing the lipophilicity is of paramount importance for the rational development of new bioactive compounds. However, most of the established methods for accurate experimental determination of this physicochemical parameter are combatible only with a low- to medium-throughput project framework. Thus, there is a lack of methodology concerning the determination of lipophilicity of a large library of compounds in a highthroughput fashion. It would be highly desirable to introduce a fast and rigorous method for lipophilicity studies on the basis of a sensitive and widely available analytical instrumentation. Electrospray mass spectrometry (ESI/MS) could offer an attractive alternative to the existing methodologies. The idea behind the use of ESI/MS is the differential disorption of the studied compounds when coexisting within mixtures with an internal standard of known lipophipicity. The response, expressed as the logarithm signal intensity ratio of the two substances analyzed was found to be well correlated with logP values with an R^2 value of 0.82

AUTHOR INDEX

A

Abad-Zapatero, C., 157, 268, 312
Adcock, S.A., 91, 141
Afantitis, A., 138, 212
Agbaba, D., 309
Agrafiotis, D., 15, 23, 50
Agüero-Chapín, G., 281
Ajmani, S., 22
Akamatsu, M., 186, 187
Aki, E., 18, 20, 84
Alberts, I., 156
Aldini, G., 70
Aleksiev, B., 176
Alexiou, P., 251
Alexis, M.N., 308
Alloatti, D., 276
Almeida, P., 136
Almerico, A.M., 112, 275
Alov, P., 236
Altiparmak, G., 300
Alvarado, Y., 152
Amaral do, A.T., 230
Amić, D., 344
Andersson, D., 74, 241, 242
Andersson, I.E., 241
Aqvist, J., 247
Argyriou, A., 271
Aristei, Y., 55
Arrault, A., 129
Arsequell, G., 249, 312
Artursson, E., 242
Atanasova, M., 342
Audouze, K., 140, 211
Aviram, M., 259

B

Bak, A., 126
Ballante, F., 116
Baltyn, K.K., 211
Banai, A., 162
Barbosa, E.G., 262, 319
Baringhaus, K., 58
Barlow, S., 205
Baroni, M., 101
Barreca, M.L., 314
Barton, N., 63
Baruchello, R., 276
Baskin, I., 93, 97, 100
Bassan, A., 205
Bausen, M., 181
Bautista, M.T., 281, 282
Bazoti, F., 345
Beccari, A.R., 145
Beig, M., 340
Bello, A.A., 333
Bender, A., 83
Bender, F., 87, 340, 341
Bendix, F., 55
Benigni, R., 18, 75
Bentrop, D., 94, 271, 272
Bermejo Sanz, M., 171, 172
Bernshausen, T., 181
Bertrand, M., 129
Bešlo, D., 344
Bettinelli, I., 264
Betul Aycan, M., 302
Beyer, A., 287
Beyer, C., 87
Bhhatarai, B., 38, 174
Bienfait, B., 182
Bijev, A., 324
Blasi, D., 157, 249, 312
Böcker, A., 144
Bocquet, N., 271
Bojić, M., 237
Bolelli, K., 84
Boobis, A., 205
Bortolato, A., 85
Borzilleri, K., 59
Boss, C., 106
Botta, M., 327
Boyer, S., 19, 48, 136
Boz, M., 300
Bozec, S., 268
Brandt, W., 338
Bravi, G., 119
Bruczko, M., 274
Brunak, S., 37, 140, 211
Brunetti, T., 276
Buchwald, F., 15, 62, 133, 134
Bujons, J., 247
Burden, F., 16, 65
Buttar, D., 135

C

Cabrera-Pérez, M.A., 171, 172, 281, 282, 334
Cabri, W., 276
Calogeropoulou, T., 327
Canales, A., 285
Carlander, D., 213
Carlsson, L., 48, 136
Carminati, P., 276
Carniato, D., 268
Carotti, A., 14, 82, 243
Carrieri, A., 107
Carvalho, M.A., 262
Casas Cardoso, G., 152
Castillo-Garit, J.A., 333
Castillo-González, D., 281, 282
Castorina, M., 276
Catto, M., 243
Cellamare, S., 243
Centelles Sangüesa, I., 171
Centeno, N.B., 249, 312
Champness, E., 110
Chang, J., 59
Charon, C., 268
Chasapis, C.T., 94, 257, 271, 272
Chatzianastasiou, A., 299
Chekmarev, D., 131
Cheleski, J., 328, 329
Chen, B., 74
Cheng, F., 202
Cheong, S.L., 311
Chiba, P., 57
Chimirri, A., 314
Chin Yee, L., 198
Choi, S., 177
Christmann-Franck, S., 268
Chrunyk, B., 59
Chrysanthakopoulos, M., 167
Chun Wei, Y., 198
Chupakhin, V.I., 245
Ciacci, A., 276
Ciana, C.L., 106

AUTHOR INDEX

Ciepielski, J., 304
Clark, R.D., 175
Coderch, C., 285
Consonni, V., 305
Cooper, T., 63
Cordomí, A., 266
Corringer, P.J., 271
Court-Marques, D., 213
Cramer, R., 20, 81, 114
Cronin, M.T.D., 200, 201
Cruciani, G., 101
Cserháti, T., 321
Cubillán Acosta, N., 150, 152
Cumming, J., 135
Custodi, C., 56

D

Dack, K., 310
Dalkas, G.A., 257, 272
Dapkunas, J., 178, 179
Davis, A., 135
De Cuyper, M., 168
De Moor, B., 121
De Toma, C., 264
De Winter, H., 121
Debeljak, Ž., 237
Defour, L., 168
Degliesposti, G., 143
Delfaud, F., 91, 141
Delgado, A., 247
Demopoulos, V.J., 166, 250, 251
Denny, W.A., 290
Detsi, A., 231
Devillers, J., 19, 76, 224
Díaz, L., 247
Didziapetris, R., 179, 180
Dietrich, A., 209
Dimitropoulos, N., 257
Dimitrov, I., 342, 343
Dimopoulos, N.J., 251
Dixon, S., 64
Dmitriev, A., 60
Doare, L., 268
Doi, I., 105
Doležal, M., 335
Doppelt-Azeroual, O., 91, 141
Dorostkar, F., 229
Doytchinova, I., 342, 343
Dräger, B., 338
Drakulić, B.J., 244, 278
Dreger, M., 258
Dreher, J., 292
Duan, J., 64
Dzhambazov, B., 241
Dvaz, F., 285

E

Ebert, R.U., 158, 159, 160
Ebrahimi, P., 162, 163
Ebrahim-Magham, B., 162
Echeverría Díaz, Y., 152
Ecker, G., 14, 19, 31, 57
Edsgard, D., 140
Edvinsson, L., 242
Ehrlich, K., 267
Ekström, F., 242
Eleftheratos, S., 318
Emel, E., 277
Engels, K., 87, 341
Episkopou, V., 94
Ertan-Bolelli, T., 84
Evers, A., 209

F

Fabian, B., 115
Fabian, E., 181
Fallah, E., 163
Fang, W.S., 285
Fassihi, A., 315, 316, 317
Fatemi, M.H., 173, 229, 337
Fedorov, D.G., 280
Feng, C., 183
Fernandes, P.A., 194, 196
Ferreira, M.M.C., 262, 319
Filimonov, D., 15, 142, 192, 218
Filipiak, K., 304
Filipic, S., 309
Filipsson, M., 210
Finking, R., 181
Fioravanzo, E., 205
Fischer, J., 338
Fischer, J.J., 258
Flanagan, J.U., 290
Flanagan, M., 59
Flocco, M., 101
Flohé, L., 139
Flower, D. R., 342, 343
Fousteris, M., 299
Fraczkiewicz, R., 175
Franck, S., 209
Freyss, J., 106
Froeyen, M., 334
Frydenvang, K., 273
Fuguet, E., 215, 216, 217

G

Gabano, E., 283
Gago, F., 285
Gao, H., 17, 69
García-Sosa, A.T., 111
Gardikas, A.V., 272
Garg, P., 313
Gariboldi, M.B., 283
Garnev, P., 343
Garriguez-Pelufo, T., 171, 172
Gašić, M.J., 303
Gaßel, M., 87, 340, 341
Gasteiger, J., 181, 182
Gaviglio, L., 283
Geçen, N., 228, 302
Georgieva, M., 324
Geronikaki, A., 113
Gessner, C., 260
Ghorbanzad'e, M., 173
Ghosh, I., 184, 203
Giaginis, C., 166, 167, 255, 256
Giangreco, I., 82, 243
Giannini, G., 276
Giegerich, C., 58
Gikas, E., 295, 345
Gilbert, G.E., 260
Girschick, T., 62, 133, 134
Gjorgjeska, B., 336
Gkazonis, P.V., 272
Glinski, M., 258

AUTHOR INDEX

Globisch, C., 223
Gloriozova, T., 60, 142
Gnoth, M.J., 169
Gokce, M., 301
Göller, A.H., 170
Gómez-Barrio, A., 333
González-Alvarez, I., 171, 172
Gouda, H., 105
Graebner, O., 258
Gramatica, P., 18, 38, 174, 214, 283, 305
Graziani, D., 264
Green, D., 63, 119
Grigor'ev, V., 77
Grishina, M., 113
Grisolia, G., 276
Grombacher, T., 103
Guba, W., 54, 102
Gudrun, L., 86
Gupta, P., 313
Gupta, V.B., 263
Gutierrez de Teran, H., 247
Guzel, Y., 302

H

Hadjipavlou-Litina, D., 16, 235, 238, 239, 286
Haghdadi, M., 337
Hamodrakas, S.J., 248
Han, S., 59
Hannongbua, S., 287, 288
Harada, T., 186
Härkönen, H.H., 284
Harris, M., 59
Hashida, M., 183
Hashimoto, T., 110
Hay, A., 318
Hee Lee, J., 177
Helguera, A., 334
Hemmateenejad, B., 316
Hessler, G., 15, 58
Hetényi, C., 111
Hidalgo, M., 215, 216, 217
Hillertz, P., 268
Hillisch, A., 170
Hirai, N., 186
Hirono, S., 105
Hobbs, H., 63
Holmdahl, R., 241
Honary, S., 162
Hong, X., 119
Horvath, D., 96, 97, 98, 99
Hubesch, B., 124
Hussain, J., 119

I

Ibis, G., 55
Iconomidou, V.A., 248
Igglessi-Markopoulou, O., 138
Ijzerman, A.P., 83
Ikushiro, S., 186
Ilouga, P., 144
Iraci, N., 314
Ishii, M., 330, 331
Istace, F., 213

J

Jabeen, I., 15, 57
Jacobs, M., 213
Jaeger, T., 139
Jandourek, O., 335
Japertas, P., 178, 179, 180
Jaworska, J., 124
Jeliazkov, V., 176
Jeliazkova, N., 124, 125, 174, 176
Jens, E., 131
Jensen, T.S., 140, 211
Jereva, D., 147
Jianxin, D., 16
Johal, M., 190
Johansen, T.N., 273
Jónsdóttir, S.O., 131
Juncker, A.S., 211
Juranić, I.O., 244, 278
Juska, L., 180

K

Kačírek, J., 335
Kairane, C., 267
Kalinić, M., 164
Kanaoka, S., 187
Kandias, N.G., 94
Kapedanovska Nestororvska, A., 273
Karl, J., 209
Kato, H., 123
Katsori, A.M., 286
Keiser, M.J., 73
Kenny, P., 16, 33
Khatib, S., 259
Kheawsod, C., 288
Kholodovych, V., 131
Kihlberg, J., 241
Kim, B., 177
Kimura, Y., 187
Kirchhoff, C.J., 144
Kitaura, K., 280
Klein, R., 86, 146
Kleinöder, T., 182
Klepsch, F., 57
Klett, J., 285
Knape, K., 193
Koch, O., 139
Kochev, N., 125, 174
Koester, H., 258
Kohtaro, Y., 219
Kokel, D., 73
Kokotos, G., 232, 233, 235
Kolocouris, A., 318
Kontogiorgis, C., 238
Kop, T., 303
Korff von, M., 106
Kosaraju, V.K., 311
Koukoulitsa, C., 167
Kouskoumvekaki, I., 21, 89, 197
Kouznetsov, V.V., 333
Kovalishyn, V., 323
Kovarich, S., 214
Kramer, S., 62, 133, 134, 207
Krasky, A., 339
Krátký, M., 325
Krauth-Siegel, L., 340
Kristl, A., 18, 71
Kroll, F., 258
Kubbutat, M.H.G., 292

AUTHOR INDEX

Kubinyi, H., 14, 28
Kühne, R., 158, 159, 160
Kumkong, A., 288
Kunick, C., 292
Kuntiya, N., 288
Kurilo, M.N., 245
Kuzmanovski, I., 164

L

Lacey, B., 59
Laggner, C., 18, 73
Lagunin, A., 60, 142, 192, 218
Lahtela-Kakkonen, M., 284
Lambrinidis, G., 308
Lanevskij, K., 180
Lange, G., 21, 146
Langenaeker, W., 121
Langer, T., 19, 49
Lannutti, F., 254
Lapa, A.J., 161
Lapenna, S., 220
Lauria, A., 112, 275
Laxdal, E.M., 251
Le Gall, A., 118
Leitão, R., 323
Lemcke, T., 291, 292
Leonardi, A., 264
Leonetti, F., 243
Lepifre, F., 268
Leproult, E., 298
Lescrinier, E., 334
Lessigiarska, I., 324
Li, J., 305
Li, W., 202
Linusson, A., 18, 74, 241, 242
Liplavskiy, Y.V., 224
Lippert, T., 54
Liu, Q., 115
Liu, T., 174
Lobell, M., 17, 67, 169, 170
Löhr, J.H., 291
Lorenzi, S., 145
Loutsidou, A., 94
Loving, K., 156
Lowrie, J., 64
Luca de, L., 314
Luca, A., 96
Luini, M., 283, 305
Luscombe, C., 15, 118, 119

M

Maccari, G., 327
Macchiarulo, A., 14, 56
Macheras, P., 16, 17, 36
Madden, J.C., 200, 201
Mahlapuu, R., 267
Majekova, M., 250
Makhaeva, G.F., 93
Malaki, N., 166
Maldonado, A.G., 109
Malvezzi, A., 161, 230
Manallack, D., 16, 66
Mandić, L.M., 244
Mannhold, R., 16
Marakos, P., 308
Maran, U., 111
Marconi, C., 70, 264
Marcou, G., 96, 97, 98, 99
Marhöfer, R.J., 21, 87, 139, 339, 340, 341
Marrero-Ponce, Y., 148, 150, 152, 333
Marrone, A., 254
Marshall, A.J., 290
Martin, E., 18, 22, 42
Martin, L., 91, 141
Martin, Y., 16, 46
Martinelli, A., 318
Martins, F., 323
Martin-Santamaria, S., 304
Marusczyk, J., 182
Masłyk, M., 304
Mason, J., 101
Matsoukas, M.T., 266
Matsui, M., 186
Mauri, A., 305
Maurici, D., 213
Mavridis, L., 107
Mavromoustakos, T., 21, 167, 232, 233
Maxwell, A., 118
McLachlan, J.A., 307
Medić-Šarić, M., 237
Meijer, L., 295
Melagraki, G., 138, 212
Melnikov, A.A., 93
Meslamani, J.E., 117
Mestres, J., 20, 22, 40
Michaelis, S., 258
Michas, G., 235
Michopoulou, V., 232
Mikata, K., 186
Mikros, E., 14, 293, 295, 308, 345
Milić, D., 303
Miller, D., 175
Miller, R., 59
Mills, J., 101
Minor, D., 73
Miri, R., 316
Miscioscia, T.F., 82
Miteva, M.A., 147
Miyagawa, H., 186
Modina, E.A., 77
Mohimont, L., 213
Montanari, C.A., 328, 329
Montero, A., 333
Monti, E., 283
Montzolis, A., 299
Moorthy, N.S.H., 194, 196
Morales-Helguera, A., 281
Moras, D., 298
Moreira Moraes, J., 161
Moriaud, F., 21, 91, 141
Morikawa, T., 253
Morreale, A., 285
Motta, G., 264
Mottram, J.C., 87, 341
Mouchlis, V., 232, 233
Moutevelis-Minakakis, P., 235
Mukherjee, P., 42
Mulliner, D., 199
Munoz-Muriedas, J., 119
Muraoka, O., 253
Musil, D., 268
Musmuca, I., 116
Myers, K., 190

AUTHOR INDEX

Myrianthopoulos, V., 293, 295, 345

N

Naár, Z., 321
Nakagawa, Y., 187
Nakamura, S., 253
Nakanishi, I., 253
Neophytou, N., 235
Nicholls, A., 15, 32
Nicolotti, O., 20, 82, 243
Nielsen, S.K., 140, 197
Nieto, J., 249, 312
Niinivehmas, S., 127
Nikolaropoulos, S., 299
Nikolic, K., 309
Nikolić, S., 344
Ninomiya, K., 253
Nishioka, K., 186
Noj, M., 122
Nomura, K., 128
Nonato, M.C., 329
Norinder, U., 135
Nuti, R., 56
Nuytten, N., 168

Ö

Öberg, T., 174, 210

O

Oellien, F., 339, 340
Oguievetskaia, K., 91, 141
Olabisi, O., 131
Oliferenko, A.A., 155
Oliveira de, A.A., 330, 331
Oprea, T., 21, 23, 51, 56, 140
Oprisiu, I., 98
Opsenica, D., 322
Opsenica, I., 322
Oros, G., 321
Ortega, B., 148, 150, 152
Ortore, G., 318
Osella, D., 283
Osolodkin, D.I., 155

P

Pajeva, I., 221, 223, 236, 324
Palace-Berl, F., 330, 331
Palyulin, V.A., 18, 22, 93, 100, 154, 155, 245, 246, 309
Pandard, P., 76
Pantano, L., 275
Pantazatos, D., 260
Papa, E., 18, 214, 283, 305
Papakyriakou, A., 257, 299
Papapetropoulos, A., 299
Pardo, L., 266
Parenti, M.D., 143
Pareva, I., 147
Parra-Hernandez, R., 251
Pascual-Teresa de, B., 274, 304
Pascual-Teresa de, S., 304
Pasqualoto, K.F.M., 319, 331
Pastorin, G., 311
Patel, H., 203
Pavan, M., 205
Pedretti, A., 35, 70, 264, 278
Pellicciari, R., 56
Pencheva, T., 147
Pentikäinen, O.T., 127
Pereira, F., 270
Pérez-Giménez, F., 152, 333
Pérez-Machado, G., 281, 282
Pérez-Nueno, V.I., 107
Perruccio, F., 21, 85, 101
Peterlin-Masic, L., 48
Peterson, R.T., 73
Pham The, H., 171, 172
Pheida, D., 248
Pickering, D., 273
Pickett, S., 63
Pinheiro, M.P., 329
Pinto, M., 249, 312
Pisani, L., 243
Pisano, C., 276
Planas, A., 249, 312
Polanski, J., 126
Pongprom, N., 288
Pontiki, E., 239
Pontoppidan, L., 131
Popović, A., 164
Poroikov, V., 19, 22, 60, 142, 192, 218
Portioli, C., 143
Poso, A., 284
Postila, P.A., 127
Potamitis, C., 231, 327
Potemkin, V., 113
Pouli, N., 308
Pourmorad, F., 162
Prakash, O., 184, 203
Prasad, R., 184
Prodanova, P., 324
Prousis, K.C., 231
Puckowska, A., 274
Pukalov, O., 174
Pungpo, P., 287, 288
Punkvang, A., 287, 288
Puri, N., 184

Q

Queiroz, R.F., 230
Quintana, J., 249, 312

R

Radchenko, E.V., 93, 154
Raevskaya, O.E., 224
Raevsky, O.A., 19, 77, 224
Ràfols, C., 215, 216, 217
Ragno, R., 116
Ramos, A., 274, 304
Ramos, M.J., 194, 196
Rarey, M., 54, 79, 86, 102, 103, 137, 146, 292
Rastelli, G., 143
Rastija, V., 344
Ravera, M., 283
Rawle, R., 190
Ray, S., 203
Re, N., 254

AUTHOR INDEX

Rea, C., 119
Reichel, A., 67
Reinhardt, N., 338
Repanas, A., 286
Rezende de, L., 230
Ribeiro Soares, A., 289
Ritchie, D.W., 107
Riva, C., 264
Rocha, J.R., 328, 329
Roche, D., 268
Roche, O., 54
Rodgers, S., 135
Rodrigues, I.C., 188, 296
Rognan, D., 117
Rohwer, A., 339, 340
Rolon, M., 333
Romeiro, N.C., 289
Roque, F., 140, 211
Rosén, J., 132
Rosés, M., 215, 216, 217
Roth, B.L., 73
Roy, N., 313
Rullo, C., 276

S

Sabet, R., 315, 316, 317
Sacher, O., 182
Sadiel, E., 148, 150, 152
Saghaie, L., 315, 316
Sahasrabudhe, P., 59
Şahin, K., 228, 302
Sahlin, U., 210
Saiz-Urra, L., 334
Sakaki, T., 186
Sakano, M., 253
Salomão, D.J., 330, 331
Sander, T., 106
Santos M.H., 188
Santos, P.G., 188
Sanz, F., 19, 20, 39
Saparpakorn, P., 287, 288
Saraiva, L.A., 188, 296
Sârbu, C., 130
Sarimveis, H., 18, 138, 212
Sarıpınar, E., 228
Sarver, R., 59
Sastry, M., 64
Sauer-Hecks, B., 144
Sazonovas, A., 178, 179
Schächtele, C., 292
Schärfer, C., 102, 137
Scherrer, R., 17, 68
Schleifer, K.J., 20, 41
Schmidlin, F., 268
Schmidt, F., 58, 209
Schneider, N., 86, 146
Schrey, A.K., 258
Schulz-Gasch, T., 14, 54, 102
Schüürmann, G., 158, 159, 160, 199
Schwab, C.H., 182
Schwöbel, J.A.H., 200, 201
Sciabola, S., 101
Sebestyen, I., 213
Seebeck, B., 20, 79
Seeland, M., 62, 133
Segall, M., 110
Seidel, T., 55
Selassie, C., 16, 190
Selda, A., 277
Selzer, P.M., 87, 139, 339, 340, 341
Senger, S., 119
Serafimoska, Z., 273
Serra, P., 274
Setola, V., 73
Shahlaei, M., 315, 317
Shanmugasundaram, V., 15, 59
Sharma, M., 263
Shen, J., 202
Sherman, W., 64, 156
Shirotani, N., 186
Shoichet, B.K., 73
Shulga, D.A., 155
Silva, J.M.S., 188
Silveira da, N.J.F., 188, 296
Simeonov, V., 221
Simoni, D., 276
Siqueira, L.J.A., 161
Skaltsounis, A.L., 308
Skerratt, S., 310
Slavica, E., 164
Slavnić, V., 278
Smith, K., 322
Smith, K.M., 115
Sobolev, B., 60
Soenen, S.J.H., 168
Šolaja, B., 303, 322
Sonehara, I.Y., 330, 331
Sonsoles, M.S., 274
Soomets, U., 267
Sopasakis, P., 138
Souccar, C., 161
Sousa de, J.A., 323
Sousa, S.F., 194
Soutter, H., 59
Spjuth, O., 48
Spyropoulos, E., 299
Spyroulias, G.A., 22, 94, 257, 271, 272, 299
Stålring, J., 136
Stanton, R., 101
Stefanachi, A., 243
Stefek, M., 250
Steinkellner, H., 213
Stöckl, S., 159
Stolaříková, J., 325
Stróligo Pecly, J., 289
Suárez Fernández, M.L., 87, 341
Subhash, A., 92
Sullivan, D., 42
Suturkova, L., 273
Svobodová, M., 335
Swanson, S., 118
Szymańska, E., 273

T

Taboureau, O., 140, 197, 211
Tada, Y., 280
Takahashi, Y., 128
Takahira, K., 253
Taminau, J., 121
Tanabe, G., 253
Tang, Y., 202
Tarkhov, A., 181, 182
Tasca, M., 305
Tavares do Amaral, A., 161
Tavares, L.C., 330, 331

AUTHOR INDEX

Tavori, H., 259
Teetz, W., 174
Terfloth, L., 181, 182
Terzić, N., 322
Testa, B., 14, 17, 35
Tetko, I.V., 174
Theocharis, S., 255
Thijs, G., 120, 121
Tiramani, M., 213
Todeschini, R., 305
Togawa, M., 186
Tolia, A., 73
Torrens, F., 152, 333
Torres, D.F.A., 333
Totzke, F., 292
Tran, H.Q., 339
Triantos, N., 250
Trombetta, S., 145
Tropsha, A., 19, 21, 47
Tsakovska, I., 221, 236
Tsakovski, S., 221
Tsantili-Kakoulidou, A., 166, 167, 250, 255, 256, 345
Tsareva, D.A., 155
Tselios, T., 266
Tsopelas, F., 166
Tsvetelina, B., 241
Tuğba, T., 277
Tunemalm, A.K., 242
Tutone, M., 112, 275

U

Uebayasi, M., 280
Ueda, K., 187
Uhrig, U., 115
Unden, G., 87, 340, 341

V

Valencia, G., 249, 312
Valeur, E., 268
Vallianatou, T., 255, 256
Varnek, A., 96, 97, 98, 99
Vaya, J., 259
Vayer, P., 129
Vega, M.C., 333
Vejsová, M., 335
Veloso, M.P., 188, 296
Venkatraman, V., 107
Vennmann, E., 146
Ventura, C., 323
Verma, R.P., 190
Vesci, L., 276
Vinšová, J., 325
Virtanen, S.I., 127
Vistoli, G., 17, 35, 70, 264
Vitorovic-Todorović, M.D., 244
Vlijmen, H.W.T., 83
Volkamer, A., 103
Voronkov, A.E., 246
Vorotyntsev, A., 91, 141
Vrakas, D., 166, 167

W

Wagener, M., 79
Waldmann, H., 14, 30
Wallén, E., 284
Wang, L., 58
Wass, J., 157
Waterbeemd van de, H., 16, 17, 34
Wegner, J.K., 83
Weinhold, N., 140, 211
Weinzinger, A., 193
Welsh, W., 131
Wendt, B., 21, 90
Wessjohann, L., 338
Westen van, G.J.P., 20, 83
Wetwitayaklung, P., 57
Wicker, J., 207
Wiese, M., 223
Wiese, T., 307
Wiggers, H.J., 328, 329
Williams, K.Y.B., 307
Winkler, D., 65
Winter, H., 120
Withka, J., 59
Wolber, G., 14, 55
Wolschann, P., 193, 287, 288
Woltosz, W.S., 175
Wood, D., 135
Woods Jr., V.L., 260
Worth, A., 77, 205, 220, 224
Wurtz, J.M., 298

Y

Yahya, G., 300, 301
Yalçin, I., 20, 84
Yamaotsu, N., 105
Yamashita, F., 183
Yang, C., 205
Yang, Z., 119
Yanmaz, E., 228
Yap, C.W., 311
Yeap, S.K., 310
Yen Ching, L., 198
Yildiz, I., 84
Yilmaz, H., 300, 301, 302
Yilmaz, S., 84
Yoshida, S., 183
Yoshikawa, M., 253
Ysenbaert, T., 168
Yu, H., 160

Z

Žakelj, S., 71
Zakharov, A., 60, 192, 218
Zamora, I., 247
Zapico, J.M., 274
Zefirov, N.S., 93, 100, 154, 155, 245, 246, 309
Zervos, N., 318
Zervou, M., 231
Zheng, M., 235
Zhiryakova, D., 176
Zhokhova, N., 100
Zhou, R., 22, 43
Zilmer, M., 267
Zloh, M., 164, 278
Zora, U., 336
Zoumpoulakis, P., 231, 327