QSAR study of structurally unrelated substrates of MDR efflux pump VmrA from *V. parahaemolyticus*

Rudolf Kiralj and Márcia Miguel Castro Ferreira. ***marcia@iqm.unicamp.br, rudolf@iqm.unicamp.br, http://lqta.iqm.unicamp.br***

Laboratório de Quimiometria Teórica e Aplicada (LQTA), Instituto de Química, Universidade Estadual de Campinas, Campinas SP, 13084-971, Brazil

THE OBJECTIVES OF THIS WORK

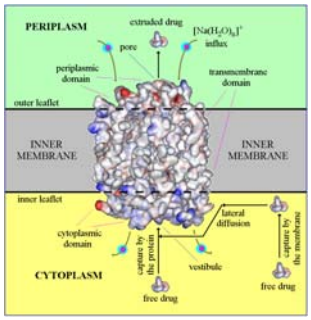
1) ***V. parahaemolyticus***, like other non-cholera ***Vibrio*** species, contaminates most marine animals in coastal waters, and causes frequently food poisoning (associated gastroenteritis), wound and soft tissue infections, septicemia, and other infections. Among several thousand of infected persons worlwide per year, there are over 10% cases with severe diseases that may end in death when immunocompromized persons are infected. To get more insight into the multidrug resistance (MDR) mechanism of this microbe, particularly VmrA efflux pump and its function, is one of the objectives;

2) To perform QSAR (Quantitative Structure-Activity Relationship) & chemometric study of structurally unrelated substrates of the VmrA, as extruded by ***E. coli*** strains: KAM32 and KAM32/pVCJ6 (with VmrA);

3) To rationalize the results of these studies in terms of molecular features that are responsible for elevated MDR of the VmrA to some of the drugs.

THE VmrA PUMP, ITS MULTIDRUG RESISTANCE (MDR) EFFLUX MECHANISM, AND ITS SUBSTRATES

1



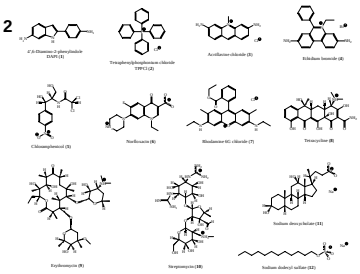
Proposed structure of the Na⁺/multidrug transporter VmrA and VmrA-mediated drug efflux from the cytoplasm to periplasm in a Gram-negative cell of *V. parahaemolyticus* strain AQ334 [M. M. C. Ferreira, R. Kiralj, unpublished results]. The VmrA structure was modeled from its primary structure [J. Chen *et al.*, *J. Bacteriol.*, 184 (2002) 572-576.].

Influx of Na⁺ ions in salty medium (coastal waters) causes complex formation of Na⁺ ions with certain Asp and Glu resides in all three domains of the VmrA (transmembrane, cytoplasmic, and periplasmic). This provokes allosteric changes in the structure of this MDR efflux pump, and opens its periplasmic window, the pore, to extrude drug molecules.

Independently on the influx of Na⁺ ions, drug molecules are captured by the inner leaflet of the inner membrane and by lateral diffusion or movement are brought to the vestibule. Drug molecules may be captured directly by the cytoplasmic domain and brought to the vestibule, a relatively large window through which they enter into the central cavity where they accumulate before being extruded from the pump.

VmrA is a secondary active transporter that uses electrochemical potential of Na⁺ across the membrane as its energy source. It is a defense mechanism of *V. parahaemolyticus* against several structurally unrelated drugs and xenobiotics.

2



The only VmrA substrates with known efflux activities [J. Chen *et al.*, *J. Bacteriol.*, 184 (2002) 572-576.] (Table 1). The active organic parts of these twelve substrates (neutral pH) belong to twelve distinct classes of organic compounds.

3

Table 1. Experimental data for drugs 1 - 12

No.	CSD source[a]	Formula[b]	pMIC(pVC6)[c]	pMIC(KAM)[d]	pMICΔ[e]
1	ASUMEG	$C_{24}H_{29}N_2$	3.242	5.961	2.719
2	DURDAV01	$[C_{24}H_{32}P]^+$	3.468	4.672	1.204
3[f]	-	$[C_{14}H_{14}N_2]^+$	3.909	5.114	1.205
4	ETHDIB	$[C_{22}H_{28}N_2]^{2+}$	4.392	4.994	0.602
5	CLMPCL02	$[C_{17}H_{17}Cl_2N_2O_3S]^+$	5.810	5.810	0
6	XAYGEJ	$C_{18}H_{12}FN_3O_3$	7.027	7.027	0
7	QIMMEE	$[C_{28}H_{33}N_2O_2]^+$	4.777	4.777	0
8	TETCYH10	$[C_{22}H_{25}N_2O_8]^+$	5.949	5.949	0
9	NAVTEJ	$[C_{27}H_{45}O_{12}N]^+$	5.264	5.264	0
10	STOSEH10	$[C_{28}H_{41}O_{12}N_2]^+$	5.464	5.464	0
11	GOLWIN	$[C_{34}H_{57}O_4]^-$	5.318	5.318	0
12	SATLUU	$[C_{32}H_{37}O_4S]^-$	6.461	6.461	0

[a]CSD codes for the structures retrieved from the CSD database. Complete or partial structures were used in molecular modeling. [b]Formula for the drug's organic component in neutral, cationic/protonated (+), anionic (–) or zwitterionic (±) state as applied in molecular modeling. [c]Efflux activity pMIC of *E. coli* strain KAM32/pVCJ6. [d]Efflux activity pMIC of *E. coli* strain KAM32. [e]Difference between the two efflux activities. [f]The simplest drug that was not found in the CSD, and was modeled by Titan program.

The modeling and biological activity data for drugs 1-12. Appropriate structures from the Cambridge Structural Database (CSD) were retrieved and modified, then optimized at PM3 semi-empirical level. The activities are pMIC = -log(MIC/mol·l⁻¹), where MIC is Minimal Inhibitory Concentration of the drugs as extruded by two strains of *E. coli*: KAM32 strain without VmrA, and KAM32/pVCJ6 with VmrA (VmrA from *V. parahaemolyticus* AQ3334). The pMICΔ is a measure of the MDR effect of VmrA.

MOLECULAR DESCRIPTORS

IMPORTANT MOLECULAR DESCRIPTORS

4

Table 2. Molecular descriptors above the cut-off (0.500) in correlation with the activities

No.	Symbol	Definition	R[KAM]*	R[pVC6]*
8	Iy	2nd principal moment of inertia	-0.262	**-0.621**
9	Iz	3rd principal moment of inertia	-0.216	**-0.654**
10	FF	log(Dip+1), Dip is molecular dipole moment (Titan)	0.431	**0.721**
14	Nh	No. hydrophobic carbon atoms	**-0.601**	-0.200
15	Na	No. aromatic carbon atoms	-0.469	**-0.645**
17	Np	No. polar (not hydrophobic) atoms (non-H)	0.203	**0.517**
26	Nr	No. ring atoms (non-H)	**-0.570**	-0.410
27	wh	Nh/Nt, number fraction of hydrophobic atoms; Nt is total No. non-H atoms	**-0.566**	**-0.726**
28	wa	Na/Nt, number fraction of aromatic atoms	-0.387	**-0.754**
29	wb	Nb/Nt, number fraction of hydrogen bonding non-H atoms; Nb is No. HB donors/acceptors	**0.562**	**0.579**
30	wp	Np/Nt, number fraction of polar atoms	**0.566**	**0.726**
31	wl	Nl/Nt, number fraction of planes atoms; Nl is No. non-H atoms in all planar fragments	-0.177	**-0.553**
40	wON	NON/Nt, number fraction of O/N atoms; NON is No. oxygen and nitrogen atoms	0.454	**0.558**
45	Bat	B/Nt, No. bonds per atom; B is No. bonds (non-H)	-0.415	**-0.603**
46	wr	Nr/Nt, number fraction of ring atoms	-0.425	**-0.713**
50	DM	molecular dipole moment (E4 method, MOPAC)	**0.694**	**0.788**
54	E4	average polarizability (E4 method, MOPAC)	**-0.611**	-0.227
55	BT	β hyperpolarizability along the dipole moment (E4 method, MOPAC)	**-0.663**	-0.456
59	GM	absolute average γ hyperpolarizability (E4 method, MOPAC)	-0.108	**-0.508**
63	Wi3	W(3d)', normalized Wiener index; W is H-depleted Wiener index	**0.510**	0.221
64	HOMO-1	energy of HOMO-1 orbital	**0.584**	0.375
65	HOMO	energy of HOMO orbital	**0.588**	0.376
68	Q-	the most negative ESP atomic charge (non-H)	**-0.672**	**-0.675**
69	Q+	the most positive ESP atomic charge (non-H)	0.426	**0.541**
74	QdF	(Q+) – (Q-), the largest ESP charge difference	**-0.539**	**-0.636**
76	%hspk	hspk/S, surface area fraction of hydrophobic C/H atoms; hspk and Sspk are CPK surface areas of hydrophobic C/H atoms and molecule, respectively	**-0.508**	**-0.508**
78	Mrefi	Mref/N, molecular refractivity per atom; Mref is molecular refractivity (ClogP method, CHEMDRAW 5) at No. all atoms	-0.259	0.606
79	Eshl	eHOMO-1/N, HOMO-1 orbital energy per atom	0.583	0.538
80	Esh	eHOMO/N, HOMO orbital energy per atom	0.582	0.529
83	Esh-11	[(HOMO-1)+(HOMO)]/N, frontier orbital energy sum (HOMO-1)+(HOMO) per atom	0.582	0.534
84	Esh-11	[(HOMO-1)+(LUMO)]/N, frontier orbital energy sum (HOMO-1)+(LUMO) per atom	0.328	0.501
87	HNn	HN/N, HM per atom	**0.819**	0.788
88	HFn	HF/N, hyperpolarizability β per atom	**-0.743**	**-0.881**
89	Bpa	B/Nt; No. bonds per atom; B is No. bonds	-0.165	**0.633**
90	L	No. non-H atoms along the longest bond chain	0.215	**0.549**
93	Xmin	minimum X coordinate	**-0.598**	**-0.528**
97	Xmax	maximum X coordinate	**0.638**	**0.538**
98	Zmax	maximum Z coordinate	**-0.577**	**-0.645**
100	DY	Ymax-Ymin, molecular box width	**-0.543**	**-0.674**
101	DZ	Zmax-Zmin, molecular box height	**-0.538**	**-0.615**
104	sighb	Nb/Sm, HB donors/acceptors surface density; Sm is molecular surface area	0.496	**0.533**
105	sighyd	Nh/Sm, hydrophobic carbons surface density	**0.660**	**0.752**
106	sica	Na/Sm, aromatic carbons surface density	**-0.347**	**0.686**
107	sigp	Np/Sm, polar atoms surface density	**0.494**	**0.669**
108	sigr	Nr/Sm, ring atoms surface density	**-0.426**	**-0.671**
109	sigon	ON/Sm, ON atoms surface density	**0.375**	**0.536**
110	orh	Nh/N, non-H atoms per bond	**0.469**	**0.629**
111	Np2	eNp - 12nP, square function of Np	**-0.527**	**-0.707**
112	ON2	eON - 10nP, square function of ON	**-0.463**	**-0.593**
113	wh2	ewh - 0.6nP, square function of wh	**-0.682**	**-0.784**
114	wa2	ewa - 0.3nP, square function of wa	**-0.668**	**0.627**
115	wp2	ewp - 0.4nP, square function of wp	**-0.682**	**-0.784**
116	Mrefi2	eMrefi - 0.19nP, square function of Mrefi	**-0.485**	**-0.745**
117	sigh2	esig - 0.022nP, sig = Nw/Sm, valence electron surface density; Nwe is No. valence electrons	**-0.669**	**-0.674**
118	RD	ratio of actual and standard bond lengths sum	0.232	**0.662**
119	RD2	eRD - 0.96nP, square function of RD	**0.583**	**0.787**
120	RRD	1/RD, inverse of RD	-0.243	**-0.673**

*Correlation coefficient with activity pMIC(KAM). [a]Correlation coefficient with activity pMIC(pVC6). [b]The bond lengths sums exclude H atoms. Actual bond lengths are from drugs 1 - 12 and standard bond lengths are analogue single bond lengths are from the simplest molecules: C-C (CH3-CH3), C-N (CH3-NH2), C-P (CH3-PH2), S-O (HS-OH), C-F (CH3F), C-Cl (CH3Cl), N-O (H2N-OH), C-Cl (CH3-Cl).

SELECTED MOLECULAR DESCRIPTORS

5

Table 3. Selected molecular descriptors for drugs 1 - 12

No.	FF	Nh	wr	E47/a.u.	HOMO-7/eV	sighyd-7/Å	Np2	wh2	Mpolh2-7/cm³·mol⁻¹
1	0.4236	14	0.8824	163.649	-7.857	0.0540	81	0.0500	0.2048904
2	0.0004	24	0.9600	215.097	-13.003	0.0622	121	0.1296	0.385641
3	0.4807	14	0.8235	203.707	-11.666	0.0527	81	0.0500	0.163216
4	0.5931	21	0.8333	241.188	-11.251	0.0580	81	0.0756	0.1660001
5	0.8224	10	0.3000	129.924	-10.264	0.0303	4	0.0100	0.1705659
6	1.6850	14	0.6957	173.499	-6.345	0.0396	9	0.0001	0.0155625
7	0.6896	27	0.6923	356.642	-11.168	0.0569	36	0.0176	0.0376536
8	1.1533	19	0.5625	219.420	-11.366	0.0155	1	0	0.002601
9	2.3617	35	0.5098	316.371	-11.867	0.0166	16	0.0071	0.695181
10	1.7979	20	0.1250	249.815	-13.862	0.0329	64	0.0100	0.6768178
11	1.3923	23	0.6071	176.495	-4.333	0.0526	49	0.0490	0.0378961
12	1.6981	12	0	110.602	-0.124	0.0424	49	0.0132	0.0663504

The descriptors were calculated by Titan, MOPAC and Chem3D using optimized geometry, and from molecular formula.

PLS (PARTIAL LEAST SQUARES) AND PCR (PRINCIPAL COMPONENT REGRESSION) MODELS

6

Table 4. PLS and PCR regression vectors[a] and statistics[b]

Parameter	PLS (pVC6)	PCR (pVC6)	PLS (KAM)	PCR (KAM)
PCs (PCs)	1 (70.0)	1 (70.0)	1 (52.8)	1 (55.6)
SEV	0.721	0.683	0.452	0.413
SEP	0.562	0.565	0.368	0.389
Q	0.893	0.818	0.762	0.763
R	0.889	0.888	0.866	0.849
Δ	0.410	0.410	0.288	0.270
FF	0.206	0.209		
Nh			-0.228	-0.260
wr	-0.201	-0.212		
E4			-0.231	-0.261
HOMO			0.223	0.190
sighyd	-0.215	-0.217	-0.250	-0.258
Np2	-0.225	-0.215		
wh2			-0.258	-0.250
Mrefi2	-0.213	-0.210		

[a]Regression vectors and statistical parameters are given for PLS and PCR modelling of efflux activity of the *E. coli* strains KAM32/pVCJ6 (pVC6) and KAM32 (KAM). [b]Parameters: PCs (PCs) – the number of principal components and % of the total variance that they contain; SEV – standard error of validation; SEP – standard error of prediction; Q – correlation coefficient of validation; R – correlation coefficient of prediction; Δ – average absolute deviation of predicted from experimental activity values.

The PLS & PCR regression statistics for modeling of the both activities shows that there is no significant difference between the PLS and PCR models. The regression coefficients are in favour for pMIC(pVCJ6), whilst the errors are smaller for pMIC(KAM). The regression vectors agree with the correlation analysis (Table 2) and chemical interpretation of the MDR phenomenon.

7

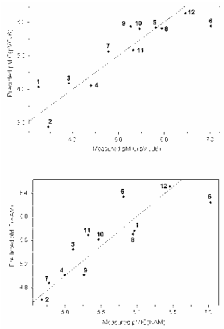
Table 5. Predicted[a] efflux activities with absolute and relative deviations[b] and difference[c]

No.	pMIC(pVC6)	Δ(pVC6)	%Δ(pVC6)	pMIC(KAM)	Δ(KAM)	%Δ(KAM)	pMICΔ
1	4.076/4.076	0.834/0.834	25.7/25.7	5.768/5.796	0.198/0.165	3.4/2.8	1.687/1.720
2	2.884/2.895	0.584/0.573	16.8/16.5	4.254/4.716	0.078/0.044	1.7/0.9	1.710/1.821
3	4.187/4.186	0.278/0.277	7.1/7.1	5.116/5.497	0.332/0.383	6.5/7.5	1.259/1.311
4	4.113/4.112	0.279/0.280	6.4/6.4	5.016/5.063	0.022/0.069	0.4/1.4	0.903/0.951
5	5.844/5.844	0.034/0.034	0.6/0.6	6.358/6.356	0.532/0.546	9.1/9.4	0.454/0.512
6	5.899/5.883	1.128/1.141	16.1/16.3	6.239/6.196	1.128/0.831	16.1/11.8	0.340/0.313
7	5.122/5.112	0.345/0.335	7.2/7.0	4.876/4.885	0.099/0.098	2.1/2.1	0.246/0.227
8	5.826/5.809	0.123/0.110	2.1/2.1	5.703/5.671	0.216/0.278	4.1/5.1	0.123/0.138
9	5.887/5.884	0.623/0.620	11.8/11.8	5.015/4.918	0.249/0.346	4.7/6.6	0.872/0.968
10	5.810/5.824	0.346/0.360	6.3/6.6	5.614/5.594	0.150/0.130	2.7/2.4	0.196/0.230
11	5.170/5.169	0.148/0.149	2.8/2.8	5.693/5.688	0.375/0.350	7.1/6.6	0.523/0.499
12	6.266/6.289	0.196/0.172	3.0/2.7	6.512/6.500	0.051/0.039	0.8/0.6	0.247/0.211

[a]Left and right values from PLS and PCR models, respectively. [b]Efflux activities: pMIC, their absolute (Δ) and relative (%Δ) deviations from experimental activities for *E. coli* strains KAM32/pVCJ6 and KAM32. [c]Absolute difference between the two predicted efflux activities.

Predicted activities, absolute and relative errors for drugs 1-12, and the predicted MDR parameter pMICΔ. There are 4 drugs with relative error >10% for pMIC(pVCJ6), and only 1 drug with such error for pMIC(KAM). The pMICΔ parameter shows correctly the presence and absence of elevated multidrug resistance of VmrA for 9 from 12 drugs (75%). All regression and other chemometric analyses in this work were performed by using programs Pirouette and Matlab on autoscalled data, and leave-one-out crossvalidation.

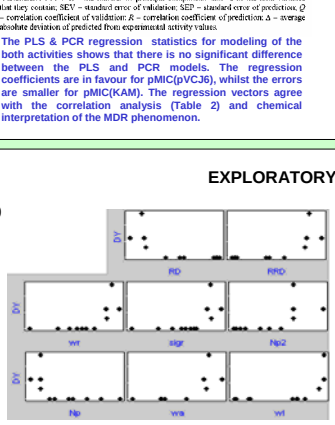
8



The PLS plots for prediction of the two activities: top - pMIC(pVCJ6); bottom – pMIC(KAM).

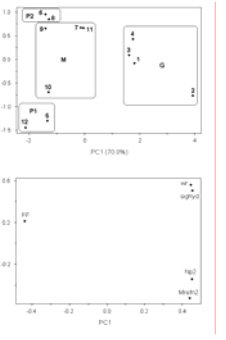
EXPLORATORY ANALYSIS AND ELEVATED MDR EFFECT OF THE VmrA PUMP

9



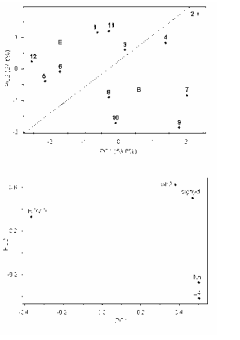
The pMICΔ-descriptor correlation plots that show which molecular features are responsible for distinction of drugs to which VmrA has elevated multidrug resistance (1-4) or there is no significant resistance (5-12). These molecular properties account for the content of unsaturated bonds, ring and planar fragments, aromatic and polar groups.

10



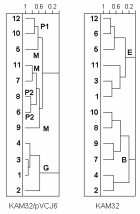
The PCA (Principal Component Analysis) scores (top) and loadings (bottom) plots using the 5 descriptors from regression modeling of pMIC(pVCL6). The activity classes can be observed: good (G), moderately good (M) and poor (P1, P2) substrates of the VmrA pump. Better substrates have higher content of aromatic and hydrophobic groups, and lower content of polar groups, small dipole moments and molar refractivity.

11



The PCA scores (top) and loadings (bottom) plots using the 5 descriptors from regression modeling of pMIC(KAM). Two classes of substrates can be observed: E – more elongated, linear or chain-like species; B – more branched, cyclic or spherical species, as substrates of other pumps in *E. coli* KAM32. B species are in average better substrates than E species. Better substrates may be characterized by higher content of hydrophobic groups, more negative HOMO and smaller polarizability.

12



The HCA dendograms with samples and complete linkage and using the same molecular descriptors as in previous PCA analyses. Two classes of drugs can be clearly observed in each dendogram, whilst sub-clusters may differ from those from the PCA plots. Two-membered sub-clusters reflect the structural similarity of their members.

CONCLUSION:
The KAM32 strain prefers rather hydrophobic species. The KAM32/pVCJ6 strain, and consequently the VmrA efflux pump, prefers rather rigid and condensed heteroaromatic species.

ACKNOWLEDGEMENT: FAPESP