Structural chemometrics applied to bioactive conformation of some lignans

Rudolf Kiralj and Márcia Miguel Castro Ferreira

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

E-mails: rudolf@iqm.unicamp.br, marcia@iqm.unicamp.br, URL: http://lqta.iqm.unicamp.br

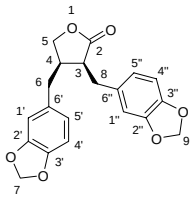
THE OBJECTIVES OF THIS WORK

1) To see which of the four conformer types of (-)-hinoquine with very close energy, previously obtained by quantum-chemical methods, are preferrential for intermolecular interactions → the use of the Cambridge Structural Database (CSD) for this class of compounds;

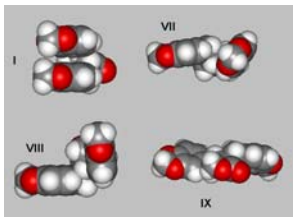
2) To explain the chemical background of the four conformer types;

3) To determine which of the four conformer types is the closest to the bioactive conformation of (-)-hinoquine, what is important for design of drugs active against Chaga's disease → use of protein-drug complexes from the Protein Data Bank (PDB), where drug is crystal violet or related compound → use of structural chemometrics for the final determination of the bioactive conformation → docking of (-)-hinoquine instead of crystal violate to protein QacR.

CONFORMER TYPES OF (-)-HINOQUINE AS OBTAINED FROM QUANTUM-CHEMICAL CONFORMATIONAL SEARCH AND THE CSD SEARCH



Molecular structure of (-)-3,4-di-(3,4-methylenedioxybenzyl)-γ-butyrolactone or (-)-hinoquinin (HQ) with atomic numbering.



Four conformer types illustrated by space filling model of selected conformers: stacked (I), T-shaped (VII), intermediate (VIII) and extended (IX). This classification is based on the mutual position of the methylenedioxybenzyl rings, obtained from semi-empirical (PM3) conformational search around bonds C4-C6 and C3-C8.

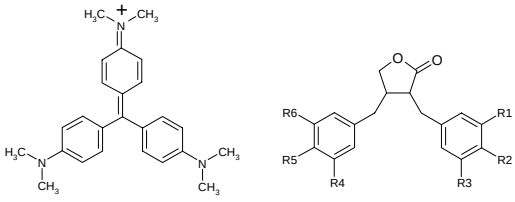
Conformer	Type	Θ_1	Θ_2	E / kcal mol^{-1}
I	stacked	48.88°	45.68°	-169.668
II	T-shaped	52.66°	183.96°	-169.650
III	T-shaped	51.33°	325.66°	-170.206
IV	T-shaped	185.86°	51.64°	-170.027
V	extended	188.06°	187.41°	-168.773
VI	intermediate	187.85°	327.85°	-169.170
VII	T-shaped	310.00°	50.00°	-170.795
VIII	intermediate	309.10°	185.96°	-169.547
IX	extended	308.82°	327.51°	-169.606

Nine conformers (I-IX) close in energy have been obtained from the PM3 systematic conformational search. The two torsion angles were defined as Θ_1(C6' - C6 - C4 - H4) and Θ_2(C6'' - C8 - C3 - H3), where H3 and H4 were hydrogen atoms at C3 and C4, respectively. One stacked (I), four T-shaped (II-IV, VII), two extended (V, IX) and two intermediate (VI, VIII) conformers were obtained. It is not obvious that T-shaped type is the most probable conformation for intermolecular interactions with participation of (-)-HQ. Systematic search in the CSD (the next table) also does not prove that T-shaped is the most probable conformation type but only predominant (8 from 19 hits).

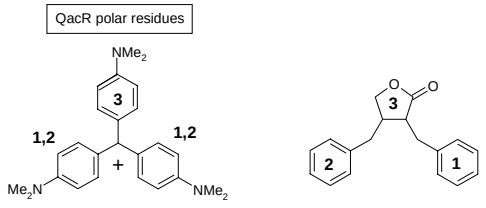
Cryst. structure[a]	Space group[b]	ring config.[c]	abs. config.[d]	conformer type[e]
CEKVET	$P2_1$	*trans*	*R,R*	T-shaped
CEPTEW	$P2_1/n$	*trans*	*R,R*	T-shaped
		trans	*S,S*	T-shaped
HXBZBL10	$P2_12_12_1$	*trans*	*S,S*	intermediate
JIFKAK	*Cc*	*cis* (epimer 1)[f]	*R,S*	extended
		cis (epimer 2)[f]	*S,R*	extended
JIJJUH	$P2_1$	*trans*	*R,R*	T-shaped
QOQFEH	$P2_12_12_1$	*trans*	*R,R*	T-shaped
UBULOS	$P2_12_12_1$	*trans*	*R,R*	extended
VAWTUI	$P2_1/c$	*trans*	*R,R*	intermediate
		trans	*S,S*	intermediate
VEJBUH	$P2_1/c$	*cis*	*R,S*	T-shaped
		cis	*S,R*	T-shaped
DOHROH	$P2_12_12_1$	*trans*	*S,S*	T-shaped
FAFYAM	*Fdd2*[g]	*trans*	*R,R*	stacked
KIFLAM	*Fdd2*[g]	*trans*	*R,R*	stacked
XIHYUI	$P2_1/n$	*trans*	*R,R*	stacked
		trans	*S,S*	stacked

[a]Crystal structures' CSD REFCODE. The last four crystal structures are of tetrahydrofuran derivatives, while the majority of structures is for γ-butyrolactones. [b]Crystal symmetry (space group). Space groups with mirror or/and inversion symmetry operations allow packing of mirror-related isomers if the molecules are not positioned on symmetry elements. [c]The γ-butyrolactone or tetrahydrofuran ring conformation at C3 and C4: phenyl-containing substituents are either in *cis* or *trans* position. [d]The absolute configuration at C3 and C4. [e]The conformer type according to the convention adopted in the previous table. [f]Isomers with additional chiral centers. [g]Molecule situated on two-fold rotation axis.

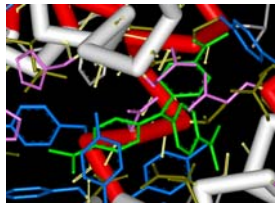
CRYSTAL VIOLET AND ITS STRUCTURAL SIMILARITY WITH (-)-HINOQUINE



Molecular structure of crystal violet cation (CV, left) compared with general structure of γ-butyrolactone lignans (right). CV is represented only by one resonance structure. The structural similarity between CV and the lignans is obvious: there are three rings connected through a small fragment which is rigid in CV, and two analogous fragments which are flexible in lignans. Furthermore, polar groups in CV are positioned at the ends of the rings, and similar arrangement can be found in (-)-HQ. Lignans may adopt conformation type which is similar to that of CV.

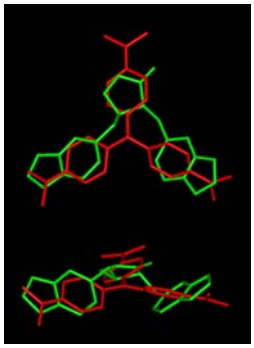


Numbering of rings from CV complexed with QacR and of HQ conformers. The two CV rings are numbered arbitrarily as 1 or 2, due to similar character of the QacR residues that interact with these rings. High degree of rigidity of CV can lead to similar, bioactive conformation of lignans, since CV is commonly used as a standard drug in bioassays for drugs against Chaga's disease.



CV (green) at the active site of QacR. The ligand interacts with predominantly polar residues (basic: blue, acidic: pink) rather than with hydrophobic (gray; other residues: brown-green) from three α-helices. (PDB: 1JTX. M. A. Schumacher *et al*., *Science*, 294, 2001, 2158). Similar arrangement can be noticed for malachite green as ligand of QacR (PDB: 1JUP. The same publication as for 1JTX).

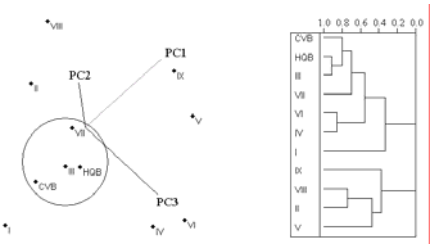
FINDING THE PROBABLE BIOACTIVE CONFORMATION OF (-)-HQ



Overlap of the two ligands, CV (red) and (-)-HQ (green) in their bioactive conformation (CVB and HQB, respectively). The CVB structure was obtained by geometry optimization of the QacR-CV complex at molecular mechanics level (force field MMFF94, chains A and B of QacR were used). The HQB structure was modeled from CV at the active site, following then the same procedure as for CVB. Both ligands were extracted from the complexes and refined at PM3 semi-empirical level with essential torsion angles constrained. Hydrogen atoms are removed because of clarity.

Conf.	D1/Å	D2/Å	D3/Å	SD/Å	A1/°	A2/°	A3/°	G1/°	G2/°	G3/°	3W/Å
I	4.084	4.866	4.823	13.733	7.93	29.24	33.15	64.53	65.62	49.85	337.650
II	6.345	4.007	4.830	15.182	65.53	60.41	46.27	49.55	39.15	91.30	363.824
III	5.592	4.930	4.841	15.366	66.14	40.44	43.75	54.38	55.83	69.79	363.261
IV	6.223	4.861	4.088	15.112	69.73	47.11	72.48	40.66	51.32	88.02	362.681
V	7.858	4.052	4.036	15.966	57.58	71.84	71.65	14.26	14.24	151.50	384.500
VI	6.764	4.947	4.116	15.827	60.81	45.35	76.73	37.24	46.65	96.11	377.435
VII	6.275	4.856	4.819	15.950	44.32	37.93	15.38	49.30	49.82	80.88	375.126
VIII	7.135	3.998	4.827	15.960	84.22	68.66	16.63	40.16	32.29	167.55	382.460
IX	8.406	4.931	4.826	18.163	55.51	37.75	17.91	30.14	30.87	118.98	416.468
HQB	5.821	4.948	4.953	15.722	67.52	35.42	52.25	54.03	53.95	72.01	368.835
CVB	4.944	4.969	4.953	14.866	53.29	44.61	47.08	59.95	60.27	59.78	353.606

Eleven structural parameters for PM3 geometry-optimized conformers of (-)-HQ (I-IX, HQB) and CV (CVB): the distances between the ring centroids D1(1,2), D2(1,3), D3(2,3) and their sum SD; the angles between the ring planes A1(1,2), A2(1,3), A3(2,3); the angles of the triangle formed by the ring centroids G1(2-1-3), G2(1-2-3) and G3(1-3-2), and 3D Wiener index 3W. The nature of these parameters (descriptors) is structural and topological. It is obvious that the distances D1 - D3 vary in a narrow range (4.9 - 5.3 Å), the angles G1 - G3 retain the regular triangle arrangement of the ring cetroids (57 - 62º), while considerable variation is noticed for the interplanar angles A1 - A3 (48 - 86º). The conformers among I-IX that are the most similar to HBQ and CVB should exhibit similarity in a structural chemometric analysis applied to this data set (see the next figure).



Structural chemometrics in finding the bioactive conformation type of (-)-HQ and analogous lignans. The 3D scores plot from Principal Component Analysis (left) and the dendogram of samples from the Hierarchical Cluster Analysis with complete linkage (right) applied to the data set from the previous table. Three PCs account for 90.73% of the total variance, meaning that the data set has 3D character. The set of the conformers consists of two clusters: a four-membered with less compact conformers (II, V, VIII, IX), and the greater cluster where the conformers are more condensed and the triangular arrangement of the three rings (benzene rings and the γ-butyrolactone ring) is more regular. It is visible that the T-shaped conformation type of (-)-HQ is the closest one to HQB and CVB. The two conformers III and VII are symmetry-related through the approximate mirror plane that passes through O1 and between C3 and C4 (ignoring the carbonyl group). However, modeling of HQ at the active site of QacR showed that the new bioactive conformer of (-)-HQ should be closer to VII than to III.