Journal of Structural Chemistry. Vol. 48, No. 1, pp. 155-160, 2007

MODELING OF DRUG MOLECULE ORIENTATION WITHIN A RECEPTOR CAVITY IN THE BiS ALGORITHM FRAMEWORK

V. A. Potemkin, M. A. Grishina, and E. V. Bartashevich

UDC 541.69:541.572+615.781.6

The BiS algorithm is suggested for modeling the drug molecule orientation within a receptor cavity. It is based on the assumption of complementarity of the field created by biologically active compounds and the field of the responsive receptor. The comparison of predicted orientations of various biologically active compounds on the relevant receptors with the data of X-ray structural studies (Protein Data Bank) reveals that the results obtained with this approach surpasses those reported in the literature. The suggested technique made it possible to elucidate the details of the action mechanism of DNA antimetabolites, dihydrofolate reductase inhibitors. The dependence of the activity on the structural parameters of "ligand–receptor" complexes is determined.

Keywords: biological activity, 3D QSAR approach, structures of receptor–ligand complexes, guest–host.

INTRODUCTION

The investigation of interactions of compounds in biological systems and also the mechanisms of medicament action demands for the structural examination of "ligand–receptor" complexes which are in their essence the guest-host complexes. Nowadays, there is a substantial pool of information concerning X-ray structural studies (XRD), electron and neutron diffraction, NMR spectroscopy of these systems [1]. Apart from experimental exploration, a theoretical study of regularities providing the biological effects arising from the ligand–receptor interactions is required to understand the character of interactions within these systems. This question has both fundamental — the ascertainment of the mechanisms of biological activity of compounds, and applied importance — the prediction of novel medications. Apparently, the determination of the drug molecule orientation within a receptor cavity is non-trivial, and in many ways is similar to molecular packing problems. One of the primary tasks in resolving this question is the determination of a molecule orientation within a receptor cavity which provides the determination of force centers and the nature of interactions, along with the detection of pharmacophore fragments.

Currently, a number of approaches have been suggested for the solution of the orientation problem. They are based mostly on the assumption of geometrical correspondence of the pharmacophore part of the molecule to the receptor site. To "arrange" within the receptor cavity, as in many conventional techniques, the space-filling molecule model is used as a combination of van der Waals spheres. Both rigid spheres [2-6] and spheres of varying rigidity [7-12] have been used in different variations of the solution. These multiple efforts, as a rule, were successful in the prediction of molecular orientation only in a series of structurally similar molecules. When applied to compounds belonging to different classes, they gave

Chelyabinsk State University; maria_grishina@csu.ru. Translated from *Zhurnal Strukturnoi Khimii*, Vol. 48, No.1, pp. 153-158, January-February, 2007. Original article submitted April 20, 2006; revised June 14, 2006.

TABLE 1. Rms Deviations of Experimental and Calculated Atomic Coordinates in the Molecules (σ), Maximum Displacements of Atomic Coordinates (Δr_{max}), and the Angles Between the Primary Rotation Axes for the Experimental and Predicted Orientations (α_1, α_2, α_3)

Activity	Compound	σ, Å	Δr_{max}, Å	α_1, deg	α_2, deg	α_3, deg	Activity	Compound	σ, Å	Δr_{max}, Å	α_1, deg	α_2, deg	α_3, deg
HRV14 rhinovirus inhibitors (**1ruc**)*	**1hri**	0.14	0.15	5.7	6.9	5.9	CDK2 inhibitors (**1bkx**)	**1agw**	0.80	1.24	8.8	8.5	3.6
	1hrv	0.28	0.41	4.2	4.1	2.1		**1aq1**	0.66	1.23	1.9	4.6	4.3
	1r09	0.29	0.48	2.6	4.8	4.3		**1atp**	0.70	1.36	2.6	2.7	0.9
	1ruc	0.00	0.00	0.0	0.0	0.0		**1bkx**	0.00	0.00	0.0	0.0	0.0
	1rue	0.02	0.04	3.4	5.5	4.7		**1ckp**	0.26	0.56	7.0	8.9	4.9
	1rug	0.00	0.00	0.0	0.0	0.0		**2csn**	0.74	1.09	7.2	7.2	1.2
	1vrh	0.35	0.64	2.2	2.0	1.4		**1fgi**	0.81	1.64	3.3	4.3	2.7
	2r04	0.18	0.22	0.3	1.3	1.2		**2hck**	0.81	1.63	1.8	1.7	1.0
	2r06	0.00	0.00	0.0	0.0	0.0		**1ian**	0.24	0.43	2.5	2.7	1.6
	2r07	0.08	0.15	0.2	1.4	1.4		**1stc**	0.72	1.46	7.9	8.6	6.3
	2rs5	0.14	0.15	0.1	0.8	0.8		**1ydr**	0.97	1.81	1.6	5.7	5.9
	1r08	0.31	0.42	0.5	1.7	1.7		**1yds**	0.68	1.11	1.1	4.3	4.5
	2hwb	0.18	0.26	4.7	5.0	3.0	DNA antimetabolites (**1ims**) Dihydrofolate reductase inhibitors (**1rg7**)	**1ydt**	1.14	2.17	1.1	5.7	5.5
	2hwc	0.14	0.24	4.9	7.0	6.0		**1d35**	0.24	0.35	6.4	6.5	1.4
	1rud	0.29	0.35	6.4	6.5	8.7		**1d37**	0.61	1.02	8.8	8.9	1.4
	1ruh	0.30	0.38	6.5	5.8	9.0		**1ims**	0.00	0.00	0.0	0.0	0.0
	1rui	0.29	0.35	6.4	6.5	8.7		**1rg7**	0.00	0.00	0.0	0.0	0.0
	2rm2	0.32	0.42	7.0	7.0	0.5		**Hfp1**	1.05	2.02	8.0	9.0	5.8
	2rr1	0.28	0.32	6.4	6.6	8.6							
	2rs1	0.29	0.36	6.4	6.5	8.7							
	2rs3	0.36	0.40	6.2	7.7	4.2	p38 MAP inhibitors (**1bl7**)	**1a9u**	0.77	1.37	5.2	7.8	7.6
Thermolysin inhibitors (**5tmn**)	**1thl**	0.87	1.93	2.2	1.8	1.8		**1bl6**	0.51	0.92	1.2	8.3	8.4
	1tmn	0.45	0.92	4.0	4.3	3.9		1**bmk**	0.50	1.00	2.6	6.8	7.2
	1tlp	0.80	1.71	6.0	4.4	6.6		**1di9**	0.80	1.28	5.2	6.9	8.3
	5tln	1.26	2.25	7.7	2.5	7.6		**1ian**	1.36	2.74	5.9	1.8	6.2
	4tmn	0.18	0.31	6.5	2.7	6.3		**1kv1**	0.00	0.00	0.0	0.0	0.0
	3tmn	0.94	1.66	5.0	5.0	3.2		**1kv2**	1.07	2.03	6.9	7.4	7.4
	5tmn	0.00	0.00	0.0	0.0	0.0	Elastase inhibitors (**1eau**)	**1ela**	0.73	1.33	2.8	3.3	1.9
	6tmn	0.01	0.02	0.1	0.1	0.1		**1elb**	0.80	1.30	6.1	7.1	7.5
Elastase inhibitors (**1eau**)	**1b0e**	0.72	1.20	4.2	5.0	5.4		**1elc**	0.86	1.74	7.0	8.4	7.0
	1bma	0.84	1.59	8.6	7.8	4.8		**1eld**	0.97	1.86	5.2	4.4	2.9
	1eas	0.95	1.73	7.0	2.4	7.4		**1ele**	0.98	1.92	5.0	6.4	4.6
	1eat	0.96	1.80	6.6	2.2	6.9		**1inc**	0.83	1.28	3.0	4.9	4.8
	1eau	0.00	0.00	0.0	0.0	0.0		**4est**	0.20	0.35	1.2	2.4	2.4
								7est	0.98	1.49	5.0	3.3	3.7

*Compound names are given according to the Protein Data Bank, standard names are given in parentheses.

molecular orientations strongly deviating from the experimental location of the molecule in the receptor cavity. The origin of this discrepancy is rather obvious: the geometrical shape of a molecule is an important but not the single feature determining the molecule arrangement in the receptor cavity. Beyond doubt, the molecule orientation is governed by the whole spectrum of van der Waals, Coulomb, and specific interactions with the receptor. Taken together, the interactions of these types determine a molecular field which must provide the maximal complementarity of biologically active compounds to

receptors. It should be taken into account that every molecule possessing the biological activity given contains fragments determining the binding to the receptor active centers. This part of the molecule can be termed the pharmacophore. The binding efficiency of the pharmacophore part of the molecule with the receptor will determine the biological activity value. The residual atoms lacking interactions with the receptor may serve as the "ballast" part of the molecule. Different molecules may bind to different active sites of the receptor. Therefore, the entire assembly of active molecules must adequately represent the field of the receptor. Hence, to solve the problem of molecular orientation within a receptor cavity, it is necessary to determine the cumulative field for an ensemble of molecules. This problem can be solved with 3D QSAR algorithm BiS [13-17]. The applicability of this algorithm to the analysis of biological activity was earlier tested on a substantial amount of ensembles of compounds. Among them were antituberculous, antibacterial, antiinflammatory, antitumoral drugs, P450 cytochrome substrates, etc. [13-17, 19, 20]. Here the compounds of the diverse composition and structure were considered involving both typical organogens and heavier elements, including transition metals [21]. The BiS algorithm was employed to reveal the pharmacophore fragments of the molecules and determine the relation of biological activity to the characteristics of the molecular fields. It is shown that correlation coefficients of the derived relationships are 0.90-0.99. In this contribution the BiS algorithm was used to find the orientation of molecules in the cavity of a real receptor. The quality of the prediction was checked in comparison with the X-ray diffraction data on the complexes of drugs with the receptor (Table 1).

BiS ALGORITHM

Within the frame of the BiS algorithm, the molecular field is determined with taking into account the Coulomb potential and the potential of van der Waals interactions affecting a given point *m* of the molecular surface. The Coulomb potential is calculated from the canonical equation

$$\varphi_m^q = \sum_{i=1}^{N} \frac{q_i}{R_{im}} k, \tag{1}$$

where N is the number of atoms in the structure in question; q_i is the charge of the i atom; R_{im} is the distance from the point m to the atom i; k is the scaling coefficient.

To calculate the van der Waals potential at the point m, similarly to Coulomb interactions, the following equation was suggested [17]:

$$\varphi_m^{\mathrm{VDW}} = -2\sum_{i=1}^{N} V_{im} \frac{2^3 r_i^3}{R_{im}^6}, \tag{2}$$

where V_{im} is the potential energy minimum of the Lennard-Jones potential; r_i is the van der Waals radius of the i atom. The quantities V_{im}, r_i, and q_i can be calculated with the MERA model [18]. To simplify the calculations, equation (2) includes only the attraction term of the van der Waals interactions.

The complementary receptor field can be represented as a combination of pseudo-atoms (test spheres having some charge and radius) contacting the molecular surface. To determine the characteristics of the complementary field, the Coulomb and van der Waals potentials are calculated for the first molecule of the ensemble. The derived potentials allow us to find the characteristics of the pseudo-atom located at a point m (charge q_m and radius r_m) yielding the maximum complementarity to the molecule in the given point of the field. A set of pseudo-atoms represents the model of the receptor. Characteristics of each test sphere m can be calculated as:

$$q_m = -\frac{\varphi_m^q}{\sum_{i=1}^{N} \frac{k}{R_{im}}}; \quad r_m = \sqrt[3]{\frac{\varphi_m^{\mathrm{VDW}}}{-2^3 \sum_{i=1}^{N} \frac{2V_{im}}{R_{im}^6}}}.$$

Then the orientation of the second molecule is optimized in the complementary field by combined simplex and quasi-Newton methods to reach the minimal overall probability (P) of the contact of its atoms with all pseudo-atoms:

$$P = 1 - \prod_{m=1}^{M}(1 - p_m), \quad \text{where } p_m = \exp\left(-\frac{E_m}{RT}\right), \quad M \text{ is the number of test spheres;}$$

$$E_m = \sum_{i=1}^{N}\left(\frac{kq_iq_m}{R_{im}} - 2V_{im}\frac{(r_m + r_i)^6}{R_{im}^6} + V_{im}\frac{(r_m + r_i)^{12}}{R_{im}^{12}}\right). \quad (3)$$

The orientation thus found is used to improve the complementary receptor field by addition of the potential field of the second molecule to the previous one: $\varphi_m^q = \varphi_m^q + \varphi_m^{q\prime}$; $\varphi_m^{\mathrm{VDW}} = \varphi_m^{\mathrm{VDW}} + \varphi_m^{\mathrm{VDW}\prime}$.

Field potentials $\varphi_m^{q\prime}$ and $\varphi_m^{\mathrm{VDW}\prime}$ of the second molecule are calculated similarly to the first molecule by equations (1), (2). The same procedure is carried out for the third, fourth, and all the following molecules of the ensemble. When it is completed, the positions of the first, second, and the following molecules are re-analyzed in the improved complementary field. The iterations are terminated when the difference in the atomic coordinates at the current and previous steps is below a predetermined threshold.

After optimization of the orientation is completed, a linear relationship is derived between the biological activity value and the parameters of interactions between the pseudo-atoms of the model receptor and the molecule. The parameters can include interaction energies calculated according to (3) and forces F_m

$$F_m = \sum_{i=1}^{N}\frac{dE_m}{dR_{im}}.$$

Those pseudo-atoms, which have characteristics linearly related to the activity and, hence, determine it, can model the receptor active sites. The drug fragments situated near the active sites determine the pharmacophore part of the molecule.

Therefore, the application of the BiS algorithm finally determines the orientation of molecules in the cavity of the model receptor, finds the dependence of the biological activity on the parameters of molecular fields (characteristics of interactions between molecules and the model receptor), and reveals the pharmacophore fragments of the molecules.

ORIENTATION OF MOLECULES IN THE CAVITY OF A REAL RECEPTOR

The XRD data on a “receptor–ligand” complex at least for one molecule of the ensemble and the mutual orientation of the molecules found from modeling (BiS algorithm) provide the spatial arrangement of all the molecules of the ensemble in the real receptor (Fig. 1). To test the molecular orientation quality within the framework of the approach suggested, the XRD data on the receptor — drug complexes were retrieved from the Protein Data Bank [1]. Among them are DNA antimetabolites, p38 MAP kinase, HRV14 rhinovirus, thermolysin, elastase, cyclin-dependent kinase (CDK2), and dihydrofolate reductase inhibitors. The comparison of the experimental and theoretical orientation (Table 1) revealed that the orientation accuracy is much better than in the known analogs. Thus, in the work [11] the positional atomic deviations ($\Delta r_{\max}$) in the ensemble of HRV14 rhinovirus inhibitors amounts to 4.77 Å (compound **1r09)**, for thermolysin inhibitors to 11.14 Å (compound **5tln**), for elastase to 10.01 Å (compound **1inc**), and for CDK2 inhibitors to 18.70 Å (compound **1ian**). For the orientations of compounds from the same ensembles the BiS algorithm yields, as a rule, the deviations $\Delta r_{\max}$ less than 2 Å (Table 1).

A quite accurate determination of the structures of the complexes allows us to find quantitative relationships between the parameters of the interactions in the complex “real receptor–ligand” with the biological activity. Thus, it was demonstrated by the example of DNA antimetabolites that the activity depends on the interaction energy of the “DNA–drug molecule” complexes with the water environment ($R = 0.92$). Moreover, it is shown that the activity of highly effective DNA antimetabolites correlates with the deformation energy of the bond angles of the molecules situated in the receptor cavity ($R = 0.90$). This permits the supposition that the molecule embedding into the receptor can be followed by a covalent binding with DNA, as confirmed by the XRD data for a number of highly active DNA antimetabolites.

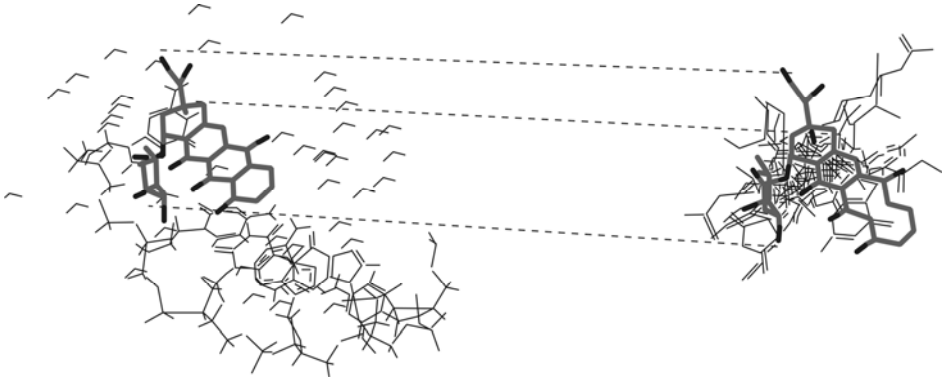

Fig. 1. Determination of molecular orientation in a cavity of the real receptor as exemplified by the DNA antimetabolite. At the right is the result of the relative orientation of molecules obtained by the BiS algorithm. XRD data on the receptor complex are available for the molecule drawn with heavy lines. At the left is the "receptor–ligand" complex according to XRD data.

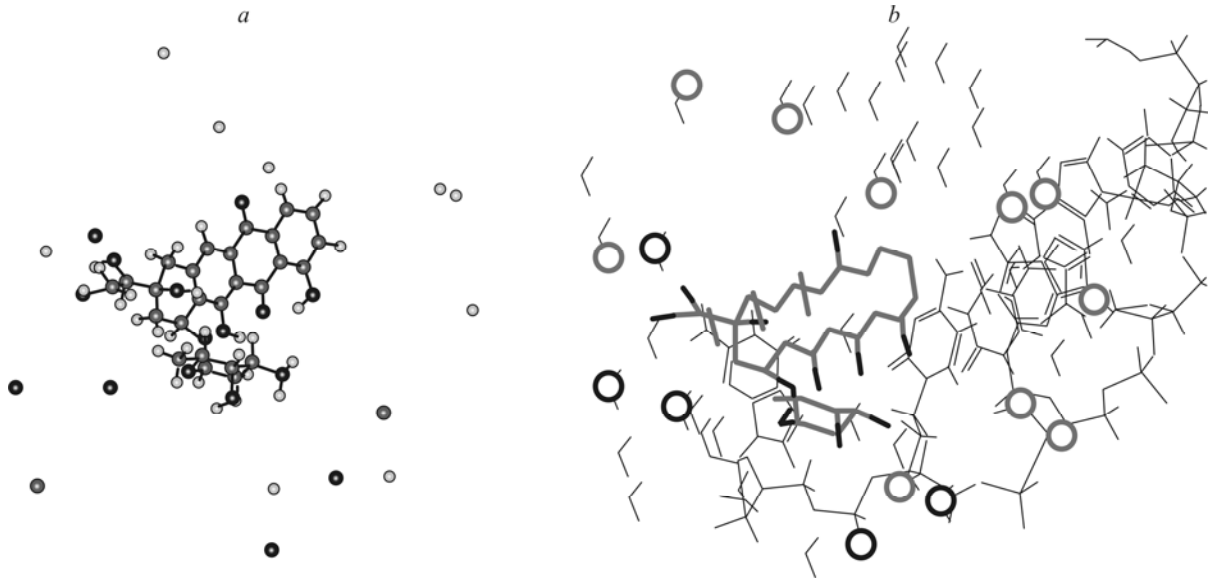

Fig. 2. Arrangement of a highly effective DNA antimetabolite (**1d37)** in the cavity of the model receptor (*a*) and the DNA molecule (*b*). ● — oxygen and nitrogen atoms, ● — carbon atoms, ○ — hydrogen atoms.

The calculations of energetic characteristics of the "real receptor–ligand" complexes for dihydrofolate reductase inhibitors revealed that for this group of drugs the water environment performs another function as compared to the DNA antimetabolites. In this case, the hydrophilicity of the drug molecules reduces the activity of the compounds. The correlation coefficient was 0.91.

The number of atoms in a real receptor is, as a rule, rather large, so the isolation of active binding sites is a difficult problem. To solve this task, it is suggested to compare the real and model receptors as exemplified by DNA antimetabolites in Fig. 2. It is found that the active sites of the model receptor well reproduce the atoms of the real receptor. It is shown for DNA antimetabolites that the receptor active centers are carbon and oxygen atoms of guanosine and cytidine nucleotides. Moreover, oxygen and hydrogen atoms of DNA-bound water molecules intensively interact with the drug molecule, which explains the dependence of the biological activity of DNA antimetabolites on the energy of the "DNA–drug" complex interaction with the water environment.

CONCLUSIONS

To summarize, the BiS algorithm is suggested to determine the orientation of molecules in the cavity of a real receptor. The application of this approach is demonstrated by ensembles of DNA and DNA/RNA antimetabolites, p38 MAP kinase, HRV14 rhinovirus, thermolysin, elastase, cyclin-dependent kinase (CDK2), and dihydrofolate reductase inhibitors. It is shown that the BiS algorithm can be successfully applied to predict the biological activity of compounds. The quantitative models derived describe the experimental activity values with large correlation coefficients. The comparison of the structures of the "model receptor–ligand" and "real receptor–ligand" complexes provides the isolation of active centers of the receptors and the determination of the nature and specific features of the interaction between the molecules and the real receptor.

This study was supported by the Federal Education Agency (Grant 2.2.2.3.9693), the President of the Russian Federation grant MK-2800.2005.3, the Chelyabinsk regional government grant No. urchel 04-04-96094.

REFERENCES

1. H. M. Berman, J. Westbrook, Z. Feng, et al., *Nucl. Acids Research*, **28**, No. 1, 235-242 (2000).
2. T. R. Stouch and P. C. Jurs, *J. Med. Chem*., **29**, No. 11, 2125-2136 (1986).
3. B. B.Masek, A. Merchant, and J. B. Matthew, *ibid*., **36**, No. 9, 1230-1238 (1993).
4. R. S. Bohacek and C. McMartin, *J. Am. Chem. Soc*., **116**, No. 13, 5560-5571 (1994).
5. H. E. Alper and T. R. Stouch, *J. Phys. Chem*. **99**, No. 15, 5724-5731 (1995).
6. A. S. Ripka, K. A. Satyshur, R. S. Bohacek, et al., *Org. Lett*., **3**, No. 15, 2309-2312 (2001).
7. E. E. Hodgkin and W. G. Richards, *Chem. Commun*., **17,** 1342-1344 (1986).
8. C. Lemmen, T. Lengauer, and G. Klebe, *J. Med. Chem*., **41**, No. 23, 4502-4520 (1998).
9. D. Hoffmann, B. Kramer, T. Washio, et al., *ibid*., **42**, No. 21, 4422-4433 (1999).
10. J. A. Grant, M. A. Gallardo, and B. T. Pickup, *J. Comput. Chem*., **17**, No. 14, 1653-1666 (1996).
11. D. A. Cosgrove, D. M. Bayada, and A. P. Johnson, *J. Comput.-Aided Mol. Design*, **14**, No. 6, 573-591 (2000).
12. R. D. Hull, S. B. Singh, R. B. Nachbar, et al., *J. Med. Chem*., **44**, No. 8, 1177-1184 (2001).
13. M. A. Grishina, V. A. Potemkin, K. M. Mikushina, et al, *Biomed. Khim.*, **50**, Supp. 1., 68-76 (2004).
14. V. A. Potemkin, M. A. Grishina, A. V. Belik, et al, *Khim. Farm. Zh*., **36**, No. 1, 22-25 (2002).
15. V. A. Potemkin, R. M. Arslambekov, E. V. Bartashevich, et al., *Zh. Strukt. Khim.*, **43**, N 6, 1126-1130 (2002).
16. K. Mikuchina, V. Potemkin, M. Grishina, et al., *Arch. Pharm. Pharm. Med. Chem*., **335**, No. 1, C74 (2002).
17. M. A. Grishina, V. A. Potemkin, G. L. Rusinov, et al., From Genes to Drugs via Crystallography, 33rd Crystallographic Course at the E. Majorana Centre, Erice, Italy, 23 May-2 June 2002. Book of Abstracts. Poster Absracts., 48.
18. V. A. Potemkin, E. V. Bartashevich, A. V. Belik, *Zh. Fiz. Khim.*, **72**, No. 4, 650-656 (1998).
19. V. A. Potemkin, M. A. Grishina, O. V. Fedorova, et al., *Khim. Farm. Zh*., **37**, No. 9, 24-28 (2003).
20. V. A. Potemkin, M. A. Grishina, E. V. Bartashevich, et al., *Biofizika*, **50**, No. 3, 418-422 (2005).
21. V. A. Potemkin, V. A. Maksakov, and V. S. Korenev, *Zh. Strukt. Khim.*, **46**, No. 1, 100-105 (2005).