



Pergamon

# Study on Molecular Mechanism and 3D-QSAR of Influenza Neuraminidase Inhibitors

Xiang Yi, Zongru Guo\* and Feng Ming Chu

*Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China*

Received 4 September 2002; accepted 5 November 2002

**Abstract**—Neuraminidase (NA) is a critical enzyme of the influenza virus and many inhibitors targeting to this enzyme are quite efficient and encouraging as anti-influenza agents. In this paper the binding model of five series of inhibitors to NA was examined using molecular simulation method. The resulted conformation and orientation of the compounds were directly put into CoMSIA study. The most significant amino acid residues at binding sites and the requirement for features of substituents were applied to direct design of new inhibitors. The robust QSAR model and its three-dimensional contour map provided guidelines to building novel compounds with new scaffold and for structural optimization of current molecules.

© 2003 Elsevier Science Ltd. All rights reserved.

## Introduction

Influenza is a respiratory infection associated with significant morbidity in the general population and mortality in elderly and high-risk patients. Current therapeutic measures have only provided limited control. Vaccines are frequently ineffective because influenza viral antigens mutate at rapid rate in forms of antigenic drift and shift.<sup>1</sup> Other options for the therapeutic treatment of influenza are limited to amantadine and rimantadine, which are of restricted usefulness because of the insensitivity to influenza virus B, unwanted side effects, and resistance problems. New influenza drugs with broad-spectrum activity are needed.

The life cycle of the influenza virus provides several potential molecular targets for drug innovation. Among those potential targets, neuraminidase (NA) appears to be particularly attractive. NA is one of two glycoproteins expressed on the surface of virus and is responsible for viral release from infected cells and viral transport through the mucus in the respiratory tract.<sup>2</sup> NA also destroys hemagglutinin receptor on host cells, thus allowing the emergence of progeny virus particles from infected cells. Therefore, compounds that inhibit NA

can protect the host from viral infection and retard its propagation. The catalytic active site of NA is located in a concave cavity on the protein surface, and the residues lining the active site of NA are highly conserved across all influenza A and B virus strains,<sup>3</sup> rendering board-spectrum anti-influenza agents possible.

Structure-based drug design methods have played a critical role in the discovery of NA inhibitors. By exploiting the crystal structure of NA–inhibitors complexes, many compounds were successfully optimized and modified based on the characters of charge and shape in binding site, such as Zanamivir launched in 1999.<sup>4</sup> Rational pharmacophore models of inhibitors like the airplane model<sup>5</sup> and the sketch map of interaction features for strong inhibitor and the binding site<sup>6</sup> were derived for drug design. Many computational methods were applied to research the correlations between binding energy of inhibitors to NA and their inhibitory activity. However, most of these studies were confined to the same scaffold compounds and hence the predictable ability was very restricted.

Hundreds of NA inhibitors and a large body of inhibitory activity data have recently been reported.<sup>7</sup> The purpose of this paper is to examine binding features of inhibitors bearing various scaffolds in the active pocket of NA enzyme, to analyze the interaction strength of different substituents with amino acids residues of binding sites,

\*Corresponding author. Tel.: +86-10-6316-5249; fax: +86-10-8315-5752; e-mail: zrguo@imm.ac.cn

and to propose a presumable model for designing new compounds based on the knowledge of the information of the enzyme and inhibitor structures. Methodologically a docking operation based on genetic algorithm and molecular mechanics simulation was performed to account for the experimental data. Meantime, Comparative molecular similarity analysis (CoMSIA),<sup>8</sup> one of classical QSAR methods, was carried out to reveal the accordance of contour map with the distribution of amino acid residues in binding site of NA.

## Molecular Simulation and QSAR Strategies

### Selection of compounds

The design and synthesis of NA inhibitors have been concentrated on the sialic acid analogues and derivatives, among which Zanamivir and Oseltamivir phosphate have been marketed as an antinfluenza nasal spray (Relenza) and orally administered NA inhibitor, respectively.<sup>9</sup> The emergence of carbocyclic and heterocyclic counterpart greatly encouraged medicinal chemists since these potent NA inhibitors are chemically more accessible. Recently a cyclopentane derivative, Peramivir, has been reported with high NA inhibitory activity and undergone phase III clinical trials.<sup>10,11</sup> Figure 1 illustrates the structures of these molecules.

The common characters of all inhibitors could be described as follows: (1) a cyclic scaffold connected to three to four substituents; (2) a carboxy group at C2 position (see Zanamivir in Fig. 1) being absolutely necessary for strong electrostatic interactions with a triad of arginine residues (Arg118, 292 and 371) of NA; (3) two hydrophobic groups at C5 and C6 position (see Zanamivir in Fig. 1). To obtain the quantitative relationship between diverse structures and inhibitory activities, it is considered that the homogenous repartition of the activities and the structural diversity are both important to a meaningful QSAR model. In this paper, 32 compounds of training set were selected from five different series and five compounds as test set were randomly picked up and used to access the predictive power of the obtained model (Table 1). The inhibitory activity data were collected from refs 4, 12–16 and are comparable. To normalize the experimental data we took the IC<sub>50</sub> values from ref 15 as the standard.

### Molecular modeling

It is well known that the selection of bioactive conformer and ascertain of alignment rule are the most critical factors to the 3D-QSAR. In principle, the actual bioactive conformation and the best alignment can only be derived from the complex structure of ligand and receptor. As for NA, X-ray structures of the enzyme liganded with different molecules have been published<sup>17,18</sup> The coordinates of some complexes can be taken from the Protein DataBank. For those absent in PDB, the position and orientation of ligands in NA binding site can be calculated accurately according to various docking programs. Here we applied Autodock 3.0 program to dock 26 compounds and identified their conformations as well as their orientations. The referenced complex codes for the investigated 37 compounds and complex sources along with pIC<sub>50</sub> values are listed in Table 2.

As indicated in Table 2, five crystal structures of NA-inhibitor complexes retrieved directly from the PDB (Nos. 29, 30, 35–37). Six crystal structures of complexes of benzoic acid derivatives (Nos. 16, 17, 19, 22–24) were kindly provided by Dr. Babu of Biocryst Pharmaceuticals, Inc. In setting up the structures all water molecules and heavy atoms observed in the crystal structures were deleted, hydrogen atoms were added, and charges of Merck Molecular Force Field (MMFF) were loaded. The whole complex was subjected to a minimization using MMFF in Sybyl software, keeping the macromolecule atoms at fixed positions. In the process, 1000 steps were conducted with the Powell method and convergent standard set at 0.05 kcal/mol Å.

For those structures which lack crystallographic data in Table 1, the complexes were generated and manually adjusted to be as similar as possible to the crystal conformation shown in ref 19. After energy minimization with Tripos force field, MOPAC was used to generate atom partial charge for the inhibitor molecules. For the ligands with flexibility the rotatable bonds were defined with AutoTors program, and the nonpolar hydrogen atoms of the inhibitors were removed and merged their charges to those of their bonded carbon atoms. The protein structures also contained only polar hydrogen atoms and were assigned Kollman charges with SYBYL software. These molecules were coded with '1nnc' and

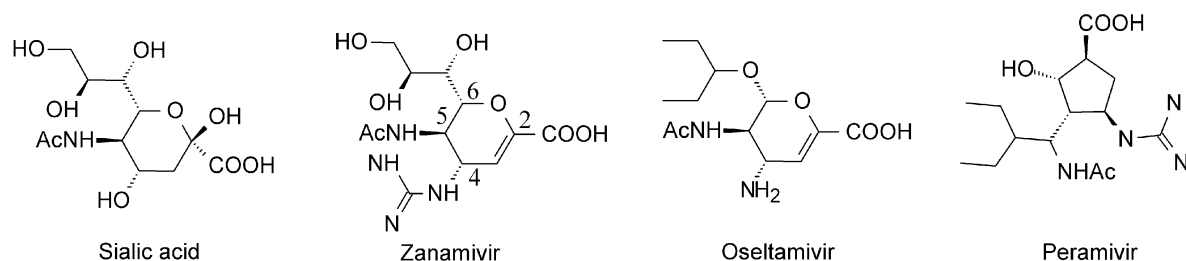
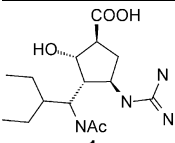
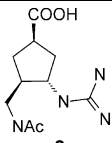
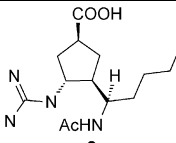
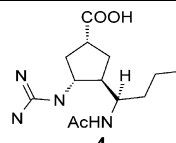
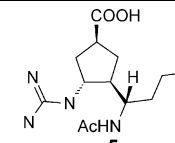
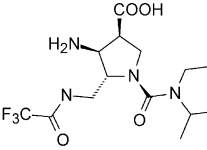
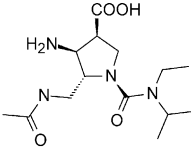
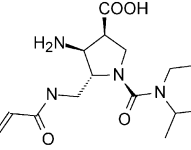
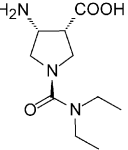
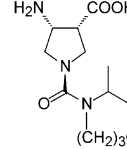
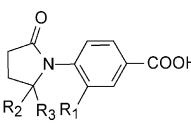
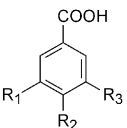
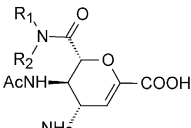
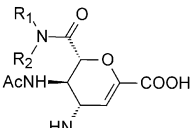
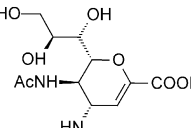
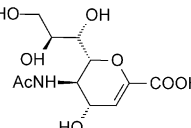


Figure 1. Structures of sialic acid and inhibitors.

**Table 1.** Structures of inhibitors

						
						
	$R_1$ <b>11</b> –NHCH(CH <sub>2</sub> CH <sub>3</sub> ) <sub>2</sub> <b>12</b> –N=C (NH <sub>2</sub> ) <sub>2</sub> <b>13</b> –H <b>14<sup>a</sup></b> –N=C (NH <sub>2</sub> ) <sub>2</sub> <b>15</b> –NHCH(CH <sub>2</sub> CH <sub>3</sub> ) <sub>2</sub>	$R_2$ –CH <sub>2</sub> OH –CH <sub>2</sub> OH –CH <sub>2</sub> OH –H –H	$R_3$ –CH <sub>2</sub> OH –CH <sub>2</sub> OH –CH <sub>2</sub> OH –H –H			
	$R_1$ <b>16</b> –N=C (NH <sub>2</sub> ) <sub>2</sub> <b>17</b> –CH=NOH <b>18</b> –CH <sub>2</sub> CH <sub>2</sub> NO <sub>2</sub> <b>19<sup>a</sup></b> –N=C (NH <sub>2</sub> ) <sub>2</sub> <b>20<sup>a</sup></b> –CH=NOH	$R_2$ –NHAc –NHAc –NHAc –NHSO <sub>2</sub> CH <sub>3</sub> –NHSO <sub>2</sub> CH <sub>3</sub>	$R_3$ –H –H –H –H –H	$R_1$ <b>21</b> –N=C(NH <sub>2</sub> ) <sub>2</sub> <b>22</b> –N=C(NH <sub>2</sub> ) <sub>2</sub> <b>23</b> –N=C (NH <sub>2</sub> ) <sub>2</sub> <b>24</b> –N=C (NH <sub>2</sub> ) <sub>2</sub> <b>25<sup>a</sup></b> –NH=C (NH <sub>2</sub> ) <sub>2</sub>	$R_2$ –NHCOCH(CH <sub>2</sub> ) <sub>2</sub> –CONHCH <sub>3</sub> –SO <sub>2</sub> NH <sub>2</sub> –CH <sub>2</sub> SOCH <sub>3</sub> –CH <sub>2</sub> SO <sub>2</sub> CH <sub>3</sub>	$R_3$ –H –H –H –H –H
	$R_1$ <b>26</b> –CH <sub>3</sub> <b>27</b> –CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub> <b>28</b> –CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	$R_2$ –CH <sub>3</sub> –CH <sub>3</sub> –CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>		$R_1$ <b>29</b> –CH <sub>2</sub> CH <sub>3</sub> <b>30</b> –CH <sub>2</sub> CH <sub>2</sub> Ph <b>31</b> –CH(CH <sub>3</sub> )CHCHCH(CH <sub>3</sub> )–	$R_2$ –CH <sub>2</sub> CH <sub>3</sub> –CH <sub>2</sub> CH <sub>3</sub>	
	$R_1$ <b>32</b> –CH <sub>2</sub> CH <sub>2</sub> NH <sub>2</sub> <b>33</b> –CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub> <b>34<sup>a</sup></b> –CH <sub>3</sub> <b>35</b> –CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	$R_2$ –H –H –CH <sub>3</sub> –CH <sub>3</sub>				

<sup>a</sup>These compounds are in the test set.

produced their binding modes, except that the complexes of benzoic acid derivatives with code 'bc1'. The Lamarckian genetic algorithm in AUTODOCK 3.0 was applied to search the conformational and orientational space of the inhibitors while keeping the protein structures rigid.<sup>20</sup> The interaction energy of ligand and protein was calculated using atom affinity potentials with AutoGrid program using the default parameters. The orientation with the lowest docked energy in the top-ranked cluster was chosen and these conformations were relaxed into a neighbouring local energy minimum, molecular mechanics was applied with keeping protein structure fixed as the above mentioned.

### CoMSIA analyses

During Comparative Molecular Similarity Index Analysis (CoMSIA) generally only bioactive conformers of compounds are fitted according to proper alignment rules, the character of each compound structure can be

described properly with potential function. All the minimized complex structures were superimposed and their ligands were extracted keeping the atom coordinates unchangeable. Since the conformation and orientation of all inhibitors originated from experiment or rational computational simulation, it is reasonable to apply the information of structures directly to CoMSIA analyses.<sup>21</sup>

The CoMSIA studies were carried out with the QSAR model of SYBYL running on a Silicon Graphics Octane 2-CPU workstation. To derive the CoMSIA descriptor fields, a 3D cubic lattice with grid spacing of 1 Å and extending 4 Å units beyond the NA inhibitors in all directions was created to encompass all the aligned molecules. A distance-dependent Gaussian-type function form expressed in eq 1 was used to calculate the CoMSIA similarity indices ( $A_F$ ) for a molecule  $j$  with atoms  $i$  at a grid point  $q$ . The default value of 0.3 was used as the attention factor ( $\alpha$ ).

**Table 2.** The complex of NA and inhibitor and activity

No.	Code of complex	Complex source	pIC <sub>50</sub> <sup>exp a</sup>	pIC <sub>50</sub> <sup>pre b</sup>	No.	Code of complex	Complex source	pIC <sub>50</sub> <sup>exp a</sup>	pIC <sub>50</sub> <sup>pre b</sup>
1	lnnc	Autodock	10.08	10.16	20	bc1	Autodock	2.78	4.37
2	lnnc	Autodock	3.97	3.97	21	bc1	Autodock	3.09	3.16
3	lnnc	Autodock	6.44	6.41	22	bc4	Dr. Babu	5.46	5.41
4	lnnc	Autodock	8.89	8.83	23	bc7	Dr. Babu	5.19	5.25
5	lnnc	Autodock	4.83	4.86	24	bc8	Dr. Babu	3.97	3.92
6	lnnc	Autodock	6.13	6.06	25	bc1	Autodock	3.63	3.39
7	lnnc	Autodock	4.79	5	26	lnnc	Autodock	5.89	5.94
8	lnnc	Autodock	3.76	3.66	27	lnnc	Autodock	7.06	7
9	lnnc	Autodock	4.36	4.33	28	lnnc	Autodock	8.30	8.27
10	lnnc	Autodock	5.31	5.37	29	2qwj	PDB	8.93	9.35
11	lnnc	Autodock	7.38	7.41	30	1bji	PDB	8.7	8.76
12	lnnc	Autodock	5.34	5.28	31	lnnc	Autodock	8.7	8.66
13	lnnc	Autodock	3.14	3.09	32	lnnc	Autodock	4.93	4.91
14	lnnc	Autodock	3.63	2.39	33	lnnc	Autodock	6.60	6.65
15	lnnc	Autodock	3.68	3.61	34	lnnc	Autodock	7.97	8.4
16	bc1	Dr. Babu	5.77	5.82	35	2qwi	PDB	8.80	8.82
17	bc2	Dr. Babu	2.33	2.36	36	lnnc	PDB	8.70	8.7
18	bc1	Autodock	2.68	2.71	37	lnnb	PDB	5.04	4.99
19	bc3	Dr. Babu	4.12	5.02					

<sup>a</sup>Experimental data were taken from the references noted.

<sup>b</sup>The predicted activities pIC<sub>50</sub> were calculated using CoMSIA model D.

$$A_{F,K}^q(j) = -\sum_{i=1}^n \omega^{\text{probe},k} \omega^{i,k} e^{-\alpha^i i q^2} \quad (1)$$

Five physicochemical properties *k* (steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor) were evaluated using the probe atom. In CoMSIA, the steric indices are related to the third power of the atomic radii, the electrostatic descriptors are derived from atomic partial charges, the hydrophobic fields are produced from atom-based parameters developed by Viswanadhan et al., and the hydrogen bond donor and acceptor indices are obtained by a rule-based method derived from experimental values.<sup>22</sup>

To choose the appropriate components and check the statistical significance of the models, leave-one-out cross-validations were performed by the enhanced version of PLS, the SAMPLE method.<sup>23</sup> Then the conventional correlation coefficient *r*<sup>2</sup> and standard errors were computed with non-cross-validated calculation.

To attenuate the effect of some lattice points on CoMSIA, region focusing was applied to subsequent analyses. Based on the above PLS model and CoMSIA columns the StDev\*Coefficient values were selected as the weights to produce the new region. The final 3D-QSAR model was obtained with the same regression method as before.

## Results

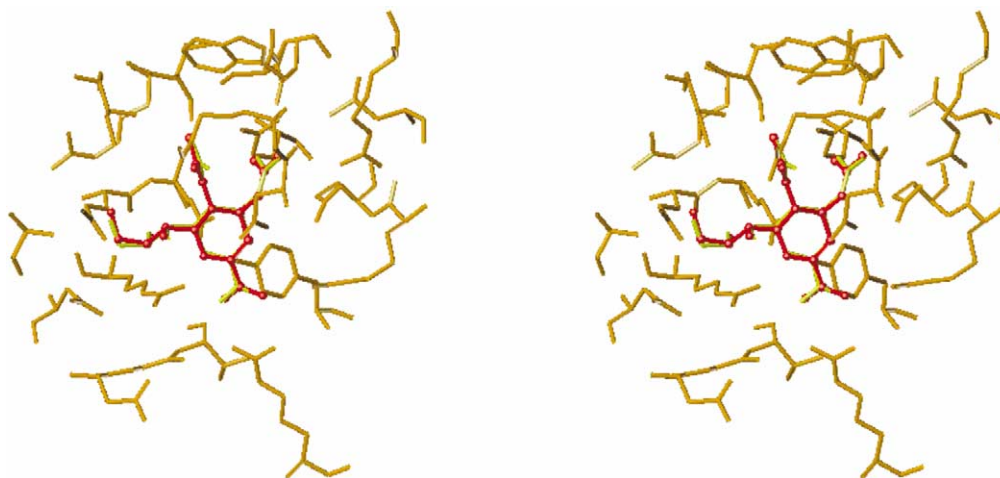
### The binding features of inhibitors with various scaffolds and different substituents to NA enzyme possess an evident similarity

A scrutiny of the crystal structures revealed that the NA complexed with an inhibitor (including Zanamivir and benzoic acid derivative BCX-140) showed a nearly identical three-dimensional structure. Several other

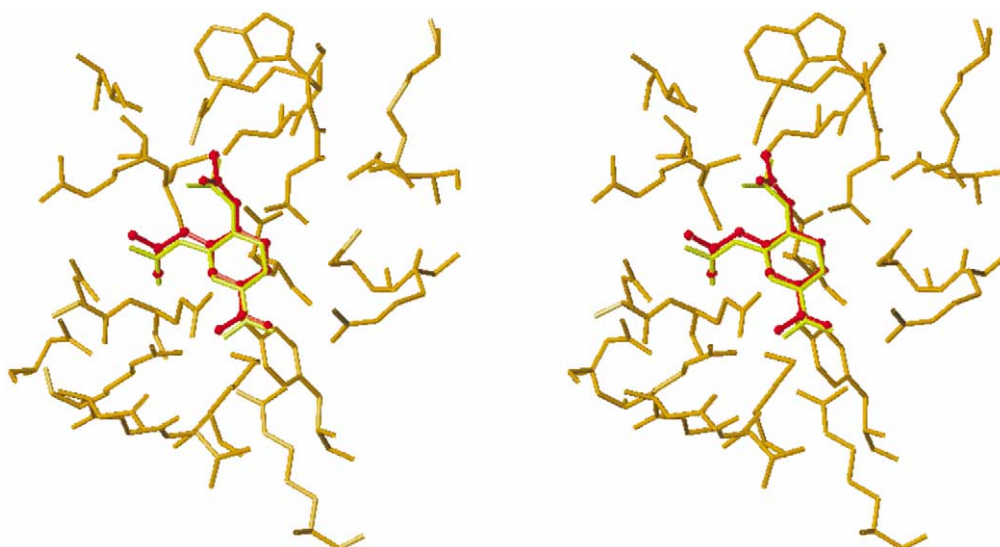
potential inhibitors were reported to orient almost in the same way in the active pocket as Zanamivir by application of soaking crystal method.<sup>24,25</sup> Thus in the present study the crystal structure of NA–Zanamivir complex (PDB code lnnc) was assigned to the investigation of molecular simulation along with benzoic acid derivative referred to NA–BCX140 complex. Another reason that these two complex structures were taken as the templates for docking procedure lied in the correct reproduction of crystal structures when the two ligands were chosen as the positive control to test the program. AutoDock considers only the polar hydrogen atoms during docking process, and uses simple potential functions for the calculation of binding energy. Therefore, the obtained enzyme-inhibitor complexes were further refined using a more sophisticated optimization procedure. All hydrogen atoms were added to the enzyme structure and the resulting complexes were minimized using MMFF94 force field, which has been parameterized for a wide variety of chemical systems and equally applicable to proteins and small molecules.<sup>26,27</sup> During the optimization process only the ligand structure was allowed to relax.

The complexes for Zanamivir and BCX-140 showed in Figures 2 and 3 are the results of calculation with AutoDock/MMFF procedure. The superimposition of the structures with the corresponding crystal structures (lnnc for zanamivir and bc1 for BCX-140) revealed that the root-mean square deviation (RMSD) values of heavy atoms between the observed and calculated position were 0.364 for Zanamivir and 0.311 for BCX-140, indicating that AutoDock was successful in reproducing the experimentally found binding position and the MMFF force field was able to correctly rank the obtained complexes.

The next step is to generate the inhibitor-enzyme complexes for 26 ligands by automated docking procedure. For all ligands the complexes which possessed the lowest interaction energies after MMFF94 optimization



**Figure 2.** The stereo-viewed comparison of position of Zanamivir calculated by the Autodock/MMFF program (yellow) and observed in the crystal structure (red). RMSD = 0.364.



**Figure 3.** The stereo-viewed comparison of position of BCX-140 (compound 16) calculated by the Autodock/MMFF program (yellow) and the one observed in the crystal structure (red). RMSD = 0.311.

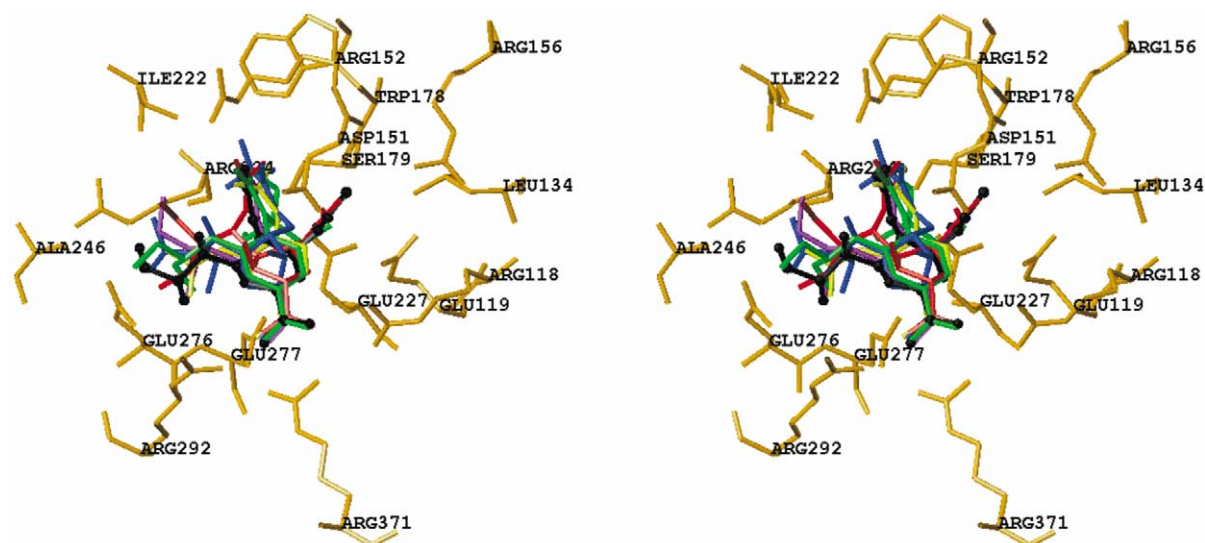
were selected. To further validate the reasonable orientation and favorable interaction between inhibitor and enzyme, the binding pocket was analyzed using the DOCKING program in SYBYL software. The results of energy calculation indicated that there existed strong interaction between inhibitor and NA, especially the significant electrostatic contribution in each case. The binding modes of five series inhibitors generated from molecular simulation method were described in the picture of Figure 4.

Compared with the binding interaction of Zanamivir (a ball-stick structure), the NA inhibitors with various scaffold structures described by a stick structure in Figure 4 exhibited the similar binding way with substituent groups extruding into a special pocket. The central ring in dihydropyran (**29**) and benzoic acid derivatives (**11** and **16**) are located as the same way as Zanamivir. In

cyclopentane series (**1**), however, a slight torsion occurred, while in pyrrolidine series (**6**) inhibitors suffered from obvious shift in space. Taking compounds **6** and **10** as examples, the ring plane unexpectedly rotated by about  $39^\circ$  and  $34^\circ$ , respectively, relative to that of Zanamivir (Fig. 5). In spite of the different position of scaffold the expected interaction of the carboxylate of all inhibitors with the positively charged site composed of Arg118, Arg292 and Arg371 of neuramidase was readily observed. Even for the acyclic series,<sup>28</sup> carboxylate group was also reported to interact with the three residues. So it could be suggested that to ensure potential inhibitory activity the carboxylate is the essential substituent.

The fact that the presence of a basic substituent (e.g., an amino or guanidine group) at C4-position efficiently enhanced the overall binding<sup>4</sup> was proved in the present





**Figure 4.** Stereoview of comparison between the binding modes of Zanamivir (in black) and Compound 1 (in red), compound 6 (in blue), compound 11 (in green), compound 16 (in yellow) and compound 29 (in violet) which were taken as an example of various scaffold structures. The principal interactions between the inhibitor and the active site residues were shown. The carboxylate was located in a positively charged pocket formed by Arg118, Arg292 and Arg371, while the binding site at C4-position was composed of residues Asp151, Glu119 and Glu227.

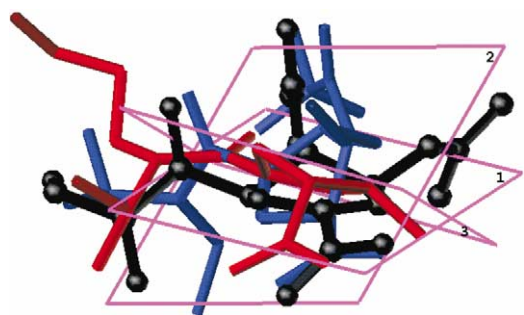
molecular simulation. Analyzing the structure–activity relationships of all the series, we gained an insight into that cyclopentane and dihydropyran derivatives exhibited comparatively potential inhibition than the other series owing to the favorable orientation of guanidine group in these molecules. In cyclopentane and dihydropyran series this group extruded into the space through the electronic interaction with residue Asp151, Glu119 and Glu227, while in pyrrolidine series the amino group turned away from the position for optimal interaction with the three acidic residues due to the twist of the center ring. However, for benzoic acid derivatives, the guanidine group entered into another pocket where glyceride group of Zanamivir located.

The hydrophobic moiety of pharmacophoric models in reported literatures was documented by the binding to the pocket consisted of Ile222 and Trp178. In the crystallographic structure of Zanamivir and NA this region was filled with terminal methyl of acetylamino group, for which our molecular simulation did not disclose much room to modify methyl group. However, at the

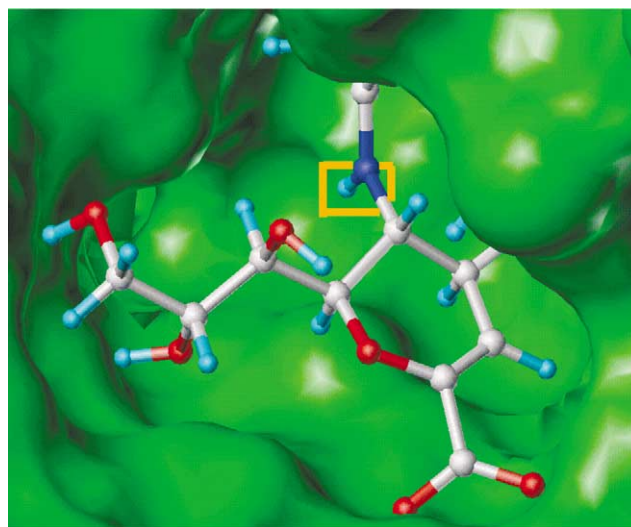
bottom of nitrogen atom there existed a region close to Glu227 and Glu277, which could accommodate hydrophilic group and thus enhance the binding affinity of inhibitors, as illustrated in Figure 6.

#### The CoMSIA method characterized by five force fields exhibits a statistical significance and a high predictability for activity

Based on the orientation and conformation of aforementioned inhibitors identified by docking computation, a 3D-QSAR were analyzed, and the produced contour map was tested for the ability to predict the activity of test set, that may validate the rationality and reliability of our molecular simulation. The superimposition of the



**Figure 5.** Comparing the binding conformation of Zanamivir (black), compounds 6 (blue) and 10 (red). The plane of central ring in the three inhibitors was marked in order as 1, 2 and 3, respectively. The rotation between ring planes was obviously observed.



**Figure 6.** The binding mode of Zanamivir with active pocket of NA. Another binding site was found around the region of an orange rectangle which was able to be accommodated with hydrophilic group.

**Table 3.** The result of 3D-QSAR

Parameter	Model			
	A: steric + electrostatic	B: steric + electrostatic + hydrophobic	C: steric + electrostatic + H-bond	D: steric + electrostatic + hydrophobic + H-bond
$q^2$	0.701	0.562	0.704	0.651
$r^2$	0.996	0.999	1.000	0.999
SE	0.154	0.078	0.014	0.078
F	676.4	2126.7	49,362.2	2360.8
OPC <sup>a</sup>	9	11	14	10
Fraction				
Steric	0.332	0.267	0.199	0.160
Electrostatic	0.668	0.428	0.351	0.192
Hydrophobic		0.305		0.195
H-bond			0.450	0.453

<sup>a</sup>OPC, optimal principal component.

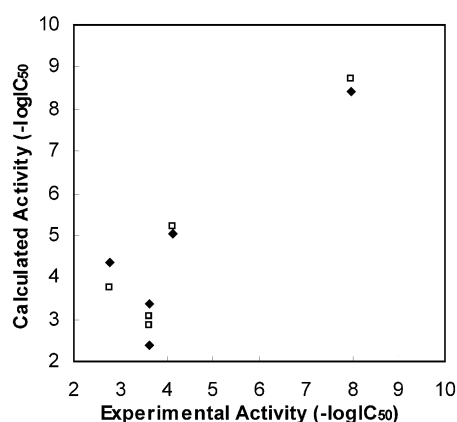
inhibitors derived from the docking and optimization was taken as starting point of a CoMSIA method using the strategy described in the Method section. To obtain the optimal QSAR model all of five fields were applied to calculate molecular similarity index. Model A obtained from steric and electrostatic fields displayed good correlation with the LOO (Leave One Out) cross-validated  $q^2$  of 0.701. However, the quality of the model was impaired because of the higher value of Standard Error (SE=0.154). From Table 3, it can be found that the biological activity could be well-expressed by four force fields including steric, electrostatic, hydrogen acceptor and donor factors, since Model C yielded the best correlation with the LOO cross-validated  $q^2$  of 0.704 and SE 0.014 using 14 principal components. Adding the hydrophobic factor to the fields of Model C, Model D was obtained with the  $q^2$  0.651 and SE 0.078. Although Model D was statistically less significant than Model C, its ability to predict the activity of the test set was stronger than that of Model C, as illustrated in Figure 7. Thus function comprised of all of five fields

was thought as the optimal 3D-QSAR model and was used to produce the contour maps.

### The distribution of CoMSIA contour map manifestly corresponds to the disposition of the amino acid residues in NA binding sites

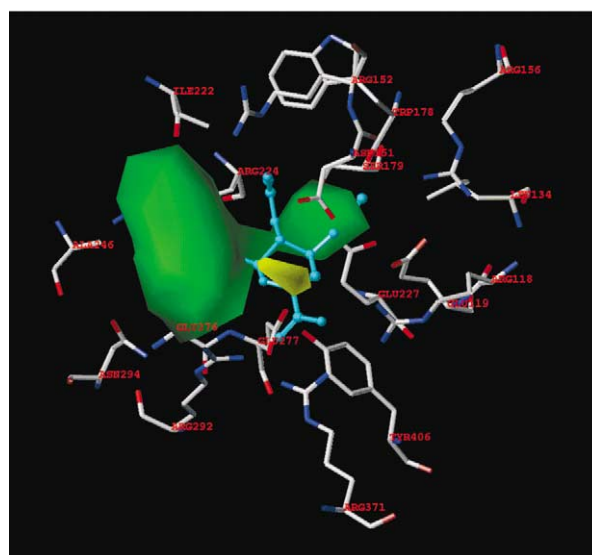
It is well documented that the contour maps reflect those regions in space where the ligand-probe interaction energy is correlated with a variance in the biological activity. Although the contour maps are unable to reveal the low-resolution pictures of the binding site, it provides a possibility to characterize the interaction feature of inhibitor with respect to the receptor environment.

The coefficient contour maps generated from Model D comprised of steric, electrostatic, hydrophobic and hydrogen bond fields were superimposed with the active site of NA. Figures 8–11 showed the grid dots of the PLS coefficients for steric field, electrostatic field,

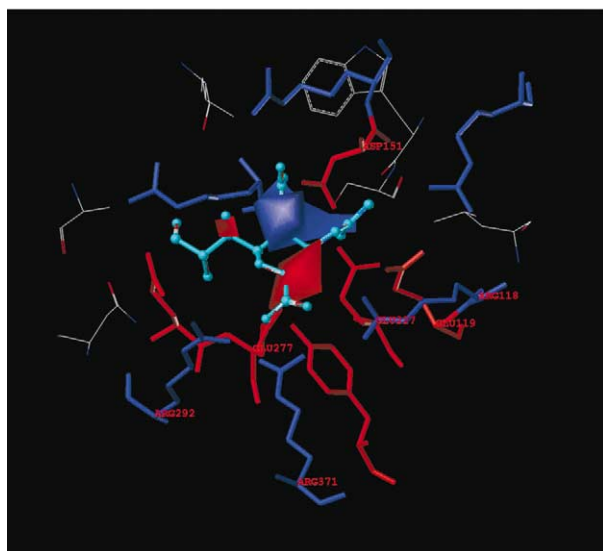


- ◆ **Model C:** Steric, electrostatic and H-bond fields  
 $r^2 = 0.77$  SEE = 1.25
- **Model D:** Steric, electrostatic, hydrophobic and H-bond fields  
 $r^2 = 0.87$  SEE = 0.99

**Figure 7.** The correlation of experimental  $pIC_{50}$  and calculated  $pIC_{50}$  of compounds in test set. The coordinates of dark diamonds and light squares were generated respectively by Model C and D.

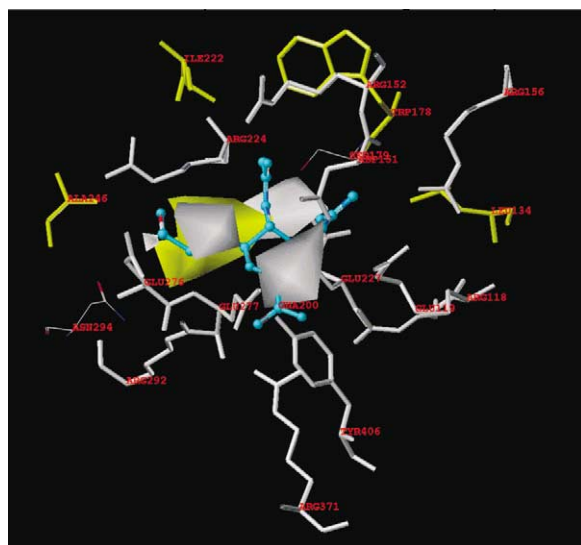


**Figure 8.** Coefficient maps of steric field and the topology of amino acid residues located close to the binding pocket. The green area map contoured at 0.017 delegated bulky group favored the binding, while yellow contour at -0.01 highlighted regions where a bulky substituent affected inhibitory activity in a negative way.

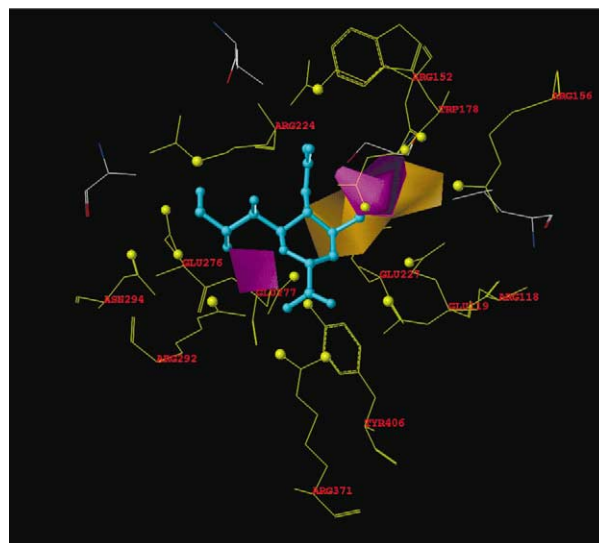


**Figure 9.** Coefficient maps of electrostatic field and the topology of amino acid residues located close to the binding pocket. The favorable electrostatic areas with positive charges were indicated by blue isopleths contoured at 0.04. Whereas the favorable electrostatic areas with negative charges were shown by red isopleths contoured at  $-0.07$ . The surrounded residues with positive and negative properties displayed in stick colored by blue and red respectively.

hydrophobic and hydrogen bond field respectively using Zanamivir as the reference compound. It can be seen that there exists an accordance of the distribution of force fields with the positions of particular amino acid residues in the active pocket. The two bulky favorable areas colored in green in Figure 8 indicated that molecules with bulky groups able to extrude into such space would enhance the biological activity. One of the areas corresponded an obvious pocket surrounded by the amino acid residues Ile222, Arg224, Ala246, Glu276



**Figure 10.** Coefficient maps of hydrophobic field and the topology of amino acid residues located close to the binding pocket. The favorable hydrophobic areas were described by yellow isopleths contoured at 0.04, whereas the disfavored hydrophobic areas were shown by gray isopleths contoured at  $-0.03$ . The surrounded residues with polar and nonpolar properties displayed in stick and colored by gray and yellow respectively.



**Figure 11.** Coefficient maps of hydrogen bond accept and donor field and the topology of amino acid residues located close to the binding pocket. Magenta isopleths contours at 0.04 represented an H-bond acceptor in the receptor favorable for activity. While H-bond donor in the receptor in the range of space described in orange color (isopleths contours at 0.1) contributed to the increase of activity. The surrounded atoms which possibly formed H-bond with inhibitors were highlighted in ball and colored yellow.

and Glu277 of the enzyme. Another space located in the pocket close to Ser179, Trp178, Glu227, into which one of substituents connected with central ring were demonstrated to enter. Compound **11**, **13** and **15**, with different groups thrust into the two binding pockets, displaying distinct inhibitory activity against NA. The most active molecule, compound **11**, favorably occupied both bulky areas simultaneously.

The graphical interpretations of the field contribution of electrostatic properties were shown in Figure 9. It comprised of two areas with different requirement for charge. The blue shadow indicated that substituents with positive charge would be favorable to the enhancement of activity, which is consistent with the space close to negative residues Asp151, Glu119, Glu227 and Glu277. Another field area colored by red corresponded to the space surrounded by Arg118, Arg292 and Arg371, which implied negative group occupying the pocket favored to the binding to the enzyme. The red area overlapping with location of carboxyl group in space and strong electrostatic interaction between carboxyl group and the amino acid residues were taken for granted. However the question is how the negative charge field generates since the carboxyl group existed in all studied inhibitors. The graphic data from molecular simulation gave such an explanation that in order to accommodate to the entire active pocket the inhibitor could sacrifice the binding of one of its substituents, such as carboxyl group. Thus the orientation and position of carboxyl group exhibited to be variable and constituted a large area in the formation of salt bridge, which was observed in all reported crystallographic complex of inhibitor with NA.

Figure 10 showed the hydrophobic std coeff CoMSIA for NA with inhibitor Zanamivir and enzyme topology.



The yellow contour indicated hydrophobically favorable region, while gray contour exhibited where hydrophilic substituents increased activity. Analyzing the contour map and the amino acid feature of binding pocket, we can reach the following knowledge: 1. The gray contours at the distal acetamido and carboxylic as well as C6 substituent groups corresponded to the hydrophilic regions formed by the side chains of Glu227, Ser179, Arg156, Glu277, Glu276 and Arg292, which predominately caused a hydrophilic interaction, while an inhibitor intruding and binding to the active pocket. 2. The gray CoMSIA contour close to C5-position was featured by a relatively large space, where a bulky hydrophilic moiety could occupy and presumably favored hydrogen bond interactions to Ser179 hydroxy group and Glu227 side-chain functionality. This inference may provide a new approach to modification of C5-position. In the structures of most inhibitors in literatures C5-position possessed acetamido group that was not able to be replaced by higher alkyl groups owing to the restricted space for hydrophobic contact. However, the molecular simulation and SAR revealed that a bulky room existed beneath the nitrogen atom of the amido group that can be exploited for introducing a polar substituent. 3. Substituents at C6-position should have both hydrophobic and hydrophilic property instead of only bulky hydrophobic groups as described in literatures.<sup>5</sup> Our results from molecular simulation and CoMSIA indicated that a large yellow region highlighted a hydrophobic requirement at the bottom of the binding pocket composed of Arg224, Ala246 and other hydrophobic side chains of amino acid residues, to which a non-polar segment of inhibitor preferably binds. A gray region at the upper side close to the yellow area signified a hydrophilic area, which corresponded to the contribution of Glu276, Glu277 and Arg292 constituting the entrance of the binding pocket. Therefore a polar moiety is suggested at C6-position, aiding to the interaction with the polar amino acid residues. In addition, modulation and balance of polar and non-polar properties at C6 substituent would enhance not only the inhibitory potency, but pharmacokinetic characters.

From the fraction of fields, it seemed that the hydrogen bond field conveyed significant contribution compared with the other fields. Scrutinizing the residue environment of binding pocket it was visualized that the formation of hydrogen bond between inhibitor and enzyme played a dominating role in the binding. Among the residues within 3 Å of the ligand about three of fourth possess the possibility of forming hydrogen bond with inhibitor. The distribution of these residues was described in Figure 11.

### Conclusion

The application of Autodock and MMFF procedure provided an efficient approach to correctly orienting and positioning each of 37 inhibitors in the binding pocket of NA. Combining the amino acids residue distribution in binding pocket and contour maps from PLS analysis, some explanation could be given for the

variation in the activities of current structures, hence new inhibitors with potentially high activity could be designed. To maintain the inhibitory activity a positive group at C4-position should be in the range of electrostatic interaction with residues Asp151, Glu119 and Glu227, just like the guanidine group in the cyclopetane and dihydropyran derivatives. Moreover, based upon this molecular simulation, imido group at C5-position, which was kept unchangeable in hundreds of investigated inhibitors could be modified by a bulky hydrophilic segment, by which a binding pocket of NA was filled.

Medicinal chemists acquire great inspiration from the binding information of potential inhibitors and NA investigated by molecular simulation. The main point in design of novel NA inhibitors is to search for brand-new scaffolds of structure without or with few chiral centers. In addition, NA exists not only in the influenza virus, but also in bacteria. Sialic acid of bacteria is split by NA that causes a series of biological response.<sup>29</sup> Therefore study on inhibitors of NA isozymes offers more chances for the treatment of various infection diseases.

### Acknowledgements

We are grateful to Dr. Ting Wang and Prof. Rebecca Wade (European Molecular Biology Laboratory) for interesting discussion. We also thank Dr. Y. S. Babu of BioCryst Pharmaceuticals, Inc. for providing the crystal structures of six neuraminidase inhibitor complexes.

### References and Notes

1. Zambon, M. C. *J. Antimicrob. Chemother.* **1999**, *44*, 3.
2. McKimm-Breschkin, J. L. *Antiviral Res.* **2000**, *47*, 1.
3. Colman, P. M. *Protein Sci.* **1994**, *3*, 1687.
4. Von, I. M.; Wu, W. Y.; Kok, G. B. *Nature* **1993**, *363*, 418.
5. Wang, G. T.; Chen, Y. W.; Wang, S.; Gentles, R.; Sowin, T.; Kati, W.; Muchmore, S.; Giranda, V.; Stewart, K.; Sham, H.; Kempf, D.; Graeme, L. W. *J. Med. Chem.* **2001**, *44*, 1192.
6. Wang, T.; Wade, R. C. *J. Med. Chem.* **2001**, *44*, 961.
7. Prous Science, S.A. The Prous Science Ensemble<sup>®</sup> database, June 2002.
8. Bohm, M.; Sturzebecher, J.; Klebe, G. *J. Med. Chem.* **1999**, *42*, 458.
9. Abdel-Magid, A. F.; Maryanoff, C. A.; Mehrman, S. *J. Curr. Opin. Drug. Discov. Devel.* **2001**, *4*, 776.
10. Sidwell, R. W.; Smee, D. F. *Expert Opin. Investig. Drugs* **2002**, *11*, 859.
11. Bantia, S.; Parker, C. D.; Ananth, S. L. *Antimicrob. Agents Chemother.* **2001**, *45*, 1162.
12. Chand, P.; Kotian, P. L.; Dehghani, A.; El-Kattan, Y.; Lin, T. H.; Hutchison, T. L.; Babu, Y. S.; Bantia, S.; Elliott, A. J.; Montgomery, J. A. *J. Med. Chem.* **2001**, *44*, 4379.
13. Atigadda, V. R.; Brouillette, W. J.; Duarte, F. *J. Med. Chem.* **1999**, *42*, 2332.
14. Chand, P.; Babu, Y. S.; Bantia, S. *J. Med. Chem.* **1997**, *40*, 4030.
15. Smith, P. W.; Sollis, S. L.; Howes, P. D. *J. Med. Chem.* **1998**, *41*, 787.
16. Babu, Y. S.; Chand, P.; Bantia, S. *J. Med. Chem.* **2000**, *43*, 3482.

17. Varghese, J. N.; Smith, P. W.; Sollis, S. L.; Blick, T. J.; Sahasrabudhe, J. A.; Mckimm-Breschkin, J. L.; Colman, P. M. *Structure (London)* **1998**, *6*, 735.
18. Varghese, J. N.; Mckimm-Breschkin, J. L.; Caldwell, J. B.; Kortt, A. A.; Colman, P. M. *Proteins: Struct., Funct.* **1992**, *14*, 327.
19. Varghese, J. N.; Epa, V. C.; Colman, P. M. *Protein Sci.* **1995**, *4*, 1081.
20. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639.
21. Sippl, W. *J. Com-Aided Mol. Des.* **2000**, *14*, 559.
22. Hou, T. J.; Li, Z. M.; Liu, J.; Xu, X. J. *J. Chem. Inf. Compt. Sci.* **2000**, *40*, 1002.
23. Bush, B. L.; Nachbar, R. B. *J. Computer-Aided Mol. Des.* **1993**, *7*, 587.
24. Jedrzejewski, M. J.; Singh, S.; Brouillette, W. J. *Biochem.* **1995**, *34*, 3144.
25. Taylor, N. R.; Cleasby, A.; Singh, O. J. *Med. Chem.* **1998**, *41*, 798.
26. Halgren, T. J. *Comp. Chem.* **1999**, *20*, 730.
27. Halgren, T. J. *Am. Chem. Soc.* **1990**, *112*, 4710.
28. Schurdak, M. E.; Voorbach, M. J.; Gao, L.; Cheng, X.; Comess, K. M.; Rottinghaus, S. M.; Warrior, U.; Truong, H. N.; Burns, D. J.; Beutel, B. A. *J. Biomolecular Screening* **2001**, *6*, 313.
29. Taylor, G. *Curr. Opin. Struc. Biol.* **1996**, *6*, 830.