Articles

A Combined STD-NMR/Molecular Modeling Protocol for Predicting the Binding Modes of the Glycosidase Inhibitors Kifunensine and Salacinol to Golgi α-Mannosidase II[†]

Xin Wen,[‡] Yue Yuan,[‡] Douglas A. Kuntz,[§] David R. Rose,[§] and B. Mario Pinto*[,‡]

Department of Chemistry, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6, and Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada M5G 2M9

Received January 7, 2005; Revised Manuscript Received March 14, 2005

ABSTRACT: A combined STD-NMR/molecular modeling protocol to probe the binding modes of the glycosidase inhibitors kifunensine and salacinol to *Drosophila melanogaster* Golgi α-mannosidase II (dGMII) was tested. Saturation-transfer difference (STD) NMR experiments were carried out for the complexes of dGMII with these two inhibitors. The program AutoDock 3.0 was then used to optimize the interactions of the inhibitors with the residues in the active site of dGMII. Theoretical STD effects of the ligand protons in the complexes were calculated for the different binding modes with the recently developed CORCEMA-ST protocol. Comparison of experimental and theoretical effects then permitted selection of the likely binding modes of the ligands. The more rigid kifunensine was used initially to test the protocol. Excellent correlation between experimental and theoretical data was obtained for one of the binding modes that also corresponded to that observed in the crystal structure of the complex. The protocol was then extended to the more flexible salacinol. For the selected binding mode, good correlation of experimental and theoretical data for the five-membered ring was obtained; however, poor correlation for protons on the acyclic chain was obtained, suggesting flexibility in this portion of the molecule. Comparison of the selected binding mode with that from a crystal structure of salacinol with dGMII showed excellent superimposition of the five-membered ring but another orientation of the acyclic chain. The results suggest that reliable structural binding modes of a ligand to protein in aqueous solution can be provided with the combined use of STD-NMR spectroscopy, molecular modeling, and CORCEMA-ST calculations, although highly flexible portions of the ligand may be poorly defined.

Biological processes are often mediated by ligand−macromolecule interactions, and understanding of the latter is of critical importance in drug, vaccine, and diagnostic design. The topic has been the subject of intense experimental and theoretical investigation, and, in the absence of crystallographic information, several protocols have been developed that provide structural information of biological macromolecules in complex with complementary ligands (*1*−*5*).

Availability of the structure of the macromolecule confines the problem to the prediction of the binding modes of a putative ligand, normally treated computationally as a docking problem (*6*−*11*). Programs for automated docking fall into two broad categories: shape matching methods, exemplified by DOCK (*8*), and docking simulation methods, exemplified by AutoDock (*9, 12*−*14*). The program AutoDock has proven effective in the study of the interaction of proteins with flexible molecules and in the prediction of energetically favorable associations and stable complexes.

The computational approaches can benefit from experimental input in the form of bound conformations of the ligand and contact points between protein and ligand (*4, 5*). In this regard, saturation transfer difference NMR[1] (STD-NMR) spectroscopy (*15*−*20*) is one of several NMR-based techniques, which shows particular promise. This method has been used for screening limited libraries to identify the bioactive ligand and also for defining the topography or epitope of the ligand that is bound by the protein, which is of considerable interest in determining the bound conformations of lead compounds. A dramatic advance in epitope mapping was provided recently by use of simulated annealing refinement together with the CORCEMA-ST program (com-

[†] Supported by a Strategic Projects Grant administered by the Natural Sciences and Engineering Research Council of Canada.

* Author to whom correspondence should be addressed. Tel: (604) 291-4152. Fax: (604) 291-5424. E-mail: bpinto@sfu.ca.

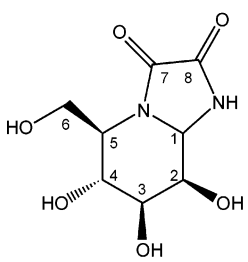
[‡] Simon Fraser University.

[§] Ontario Cancer Institute.

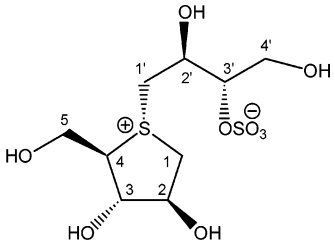
[1] Abbreviations: GMII, Golgi α-mannosidase II; dGMII, *Drosophila melanogaster* Golgi α-mannosidase II; NMR, nuclear magnetic resonance; STD, saturation-transfer difference; TOCSY, total correlation spectroscopy; CORCEMA-ST, complete relaxation and conformational exchange matrix analysis of saturation transfer.

10.1021/bi0500426 CCC: $30.25
Published on Web 04/16/2005

Chart 1



kifunensine **1**



salacinol **2**

plete relaxation and conformational exchange matrix analysis of saturation transfer) (*5*, *21*). The CORCEMA-ST program calculates STD intensities in different models of the complexes. Comparison of these predicted STD effects with the experimental data can then lead to selection of a particular model. The method requires knowledge of the protein structure either from crystallographic or modeling studies. AutoDock could be used to generate models for the complex, which can be evaluated by comparing the experimental STD effects to those obtained by the CORCEMA-ST program to select the optimum model. We present here a test of this protocol with the candidate complexes of the glycosidase inhibitors kifunensine **1** and salacinol **2** to the enzyme, Golgi α-mannosidase II (GMII).

GMII is a class II α-mannosidase acting in the N-glycosylation pathway, that is responsible for the linking (and subsequent processing) of oligosaccharides to specific asparagine residues in nascent proteins (*22*). GMII inhibitors exhibit potent antitumor and antimetastatic activity (*23*). The structures of GMII from *Drosophila melanogaster* (dGMII) and several of its complexes have been solved (*24*−*26*). Kifunensine **1** is a weak inhibitor of dGMII, with a K_i value of 5.2 mM (*25*). Salacinol **2** is one of the active principles in the aqueous extracts of *Salacia reticulata* that are traditionally used in Sri Lanka and India for the treatment of diabetes (*27*). Previous studies also indicated that salacinol **2** can bind weakly to dGMII with a K_i value in the mM range (*26*). The crystal structures of the dGMII−kifunensine and dGMII−salacinol complexes are known (*25*, *26*). These two systems therefore serve as ideal test cases for the validation of the protocol outlined in the preceding section.

MATERIALS AND METHODS

Molecular Modeling. The initial structures of kifunensine **1** and salacinol **2** (Chart 1) for the docking studies were constructed by making all possible ring conformations using molecular mechanics with a random search routine (*28*), as implemented in Sybyl 6.6 (Tripos, Inc.) and conducted as follows: The maximum number of search iterations was set to 5000, with a 3 kcal/mol energy cutoff, and 0.2 rms threshold. The structure energy minimization was performed using the standard Tripos molecular mechanics force field (*29*) and Gasteiger−Marsili charges (*30*, *31*), with a 0.001 kcal/mol energy gradient convergence criterion. The atom charges were retained on **1** and **2** for the docking calculations.

The 1.8 Å resolution crystal structure of dGMII complexed with kifunensine **1** (PDB file code: 1PS3) (*25*) and the 1.87 Å structure of the same protein, complexed with swainsonine (PDB file code: 1HWW) (*24*), were used as the model for the macromolecule in docking studies of **1** and **2**. The models were prepared using the Sybyl 6.6 program package (Tripos, Inc.). The water was removed. Only polar hydrogens were added to the protein, and Kollman united-atom partial charges (*32*) were assigned.

Automated docking was performed with AutoDock 3.0 (*9*). In docking of **1**, each initial structure obtained from conformational analysis was superimposed over the equivalent part of **1** from the crystal structure (*25*). In docking for **2**, the five-membered ring of each initial structure derived from conformational analysis was positioned on the coordinates of the five-membered ring of swainsonine in the active site of the 1HWW structure. Atomic affinity grid maps were computed for each atom type in compounds **1** and **2**, as well as an electrostatics grid map using AutoGrid 3.0 (*9*). The grid maps were constructed using 60 × 60 × 60 points, with grid point spacing of 0.375 Å, and centered on the ligand, which was manually positioned within the binding site. A Lamarckian Genetic Algorithm (LGA) was used to search each conformation space for low-energy binding orientations. The size of the initial random population was 50 individuals, the maximal number of energy evaluations was 1.5×10^6, the maximal number of generations was 27000, the elitism was 1, the probability that a gene would undergo a random change was 0.02, and the crossover probability was 0.80. For each initial structure of **1**, 100 LGA docking runs were performed (500 runs for **2**). After docking, the resulting structures for each job were clustered using an rmsd (root-mean-square positional deviation) cutoff of 2.0 Å with respect to the starting position.

NMR Spectroscopy. Kifunensine **1** was purchased from Toronto Research Chemicals (Toronto, Canada), and salacinol **2** was synthesized as described previously (*33*). The dGMII protein sample was prepared according to our previously published method (*18*). To a sample of dGMII (2.8 mg) in phosphate-buffered saline solution (50 mM K_2HPO_4/KH_2PO_4, 25 mM NaCl, 0.02% NaN_3, 99% D_2O, pH 6.0) was added either kifunensine **1** (0.56 mg) or salacinol **2** (0.80 mg). The final ligand concentration was 4 mM at a ratio of 100:1 ligand:protein. The ligand resonances were assigned using TOCSY and ROESY NMR spectroscopy.

The 1D STD-NMR and 1D TOCSY-STD-NMR spectra were recorded on a Bruker AMX-600 NMR spectrometer at 295 K, as described in previous publications (*16*, *18*) with minor modifications. 1D STD-NMR spectra were recorded with 1024 scans and selective saturation of protein resonances at 10 ppm (30 ppm for reference spectra) using a series of 40 Gaussian shaped pulses (50 ms, 1 ms delay between pulses, $\gamma B_1/2\pi = 110$ Hz), for a total saturation time of 2.04 s. The protein resonances are broad and have significant intensity in the region downfield from 10 ppm and even at negative ppm values (*16*). Thus, irradiation of 10 ppm is expected to result in saturation of all protein resonances, from the aromatic to the aliphatic. Irradiation at 10 ppm was also considered prudent to achieve selective saturation of the protein resonances since a ligand resonance was present at 5.5 ppm. For certain experiments, a 10 ms spin-lock pulse ($\gamma B_1/2\pi = 11$ kHz) was applied after excitation to reduce the intensity of broad protein resonances. Subtraction of saturated spectra from reference spectra was performed by phase cycling (*16*−*18*). Measurement of enhancement intensities was performed by direct comparison of STD-NMR

spectra and reference 1D spectra. All 1D NMR spectra were recorded with 64 scans and 32K data points, and processed by zero-filling to 64K points and multiplication by an exponential function, followed by Fourier transformation. Data processing was performed using XWINNMR (Bruker) software.

CORCEMA-ST Calculations. The underlying theory of CORCEMA-STD has been described previously (*5, 21*), as have the details of executing the CORCEMA-ST protocol (*5, 19, 20*). For the ideal case where there is infinite delay between each scan, the magnetizations in STD experiment are given by

$$\mathbf{I}(t) = \mathbf{I}_0 + [1 - \exp\{-\mathbf{D}t\}]\mathbf{D}^{-1}\mathbf{Q} \qquad (1)$$

where $\mathbf{I}(t)$ is a column matrix containing the magnetizations for the ligand and for those protein protons that do not experience a direct rf saturation.

Q is a column matrix containing cross-relaxation terms between the protein protons that experience a direct rf saturation and the rest of the protons. The number of protons in kifunensine **1** or salacinol **2** and the protons of the amino acid residues within the dGMII binding site, the number of protein protons that experience direct rf irradiation and their identities were read into the program on the basis of the PDB coordinates of the kifunensine **1**−dGMII and salacinol **2**−dGMII complexes generated from the AutoDock algorithm. To speed up the computation of the matrix, spectral densities were calculated for only those proton pairs having a distance of 3 Å or less. In the calculations, kifunensine **1** and salacinol **2** and the 15 amino acid residues (His90, Asp92, Trp95, Asp204, Phe206, Arg228, Ser268, Tyr269, Asp340, Asp341, Trp415, His471, Asp472, Arg876, Leu878) were included. These residues were selected because they were within 3 Å of the ligands.

The dynamic matrix **D** is a square matrix and is a sum of the relaxation rate matrix **R** and the exchange matrix **K**. "*t*" is the time period for which the protein proton(s) experience rf irradiation. The CORCEMA-ST program also has a provision for taking into account the effect of finite delays (t_d) between scans in calculating the STD effects, and this finite delay was taken into account in the current analysis. The parameters related with relaxation were set up as suggested in previous publications. The order parameter S^2 was set to 0.25 for the methyl group (*34*) while for methyl-X relaxation S^2 was generally kept in the range of 0.85. To account for the effect of internal motions of the methyl groups, the corresponding spectral densities were calculated using the model-free formalism (*35*). For Tyr and Phe, a simple $\langle 1/r^6 \rangle$ average was used for the dipolar relaxation between the aromatic and other protons. The protein concentration was kept fixed at 40 μM and the ratio of ligand:protein was kept at 100:1 as used in our NMR experiments. The correlation times (τ_c) and equilibrium constants (K_{eq}) were tested to find the best fit between experimental data and predicted STD values. The correlation time (τ_c) was set to 0.1−2 ns for the ligand in the free state and 20−60 ns for ligand-bound states. Since kifunensine **1** and salacinol **2** are weak inhibitors of dGMII with K_i values in the mM range (*25, 26*), K_{eq} values were tested in the range of 10^3−10^6. In our calculations, the predicted STD values did not change significantly with different τ_c and K_{eq} values in the chosen ranges. Thus, in the calculations for all structures, τ_c values were set as 0.1 and 20 ns for the free state and the bound state, respectively. K_{eq} values were set in the range of 10^6.

From the intensity matrix $\mathbf{I}(t)$, the fractional intensity changes $[(I_{0(k)} - I(t)_{(k)})/I_{0(k)}]$ for different ligand protons k were calculated, and compared to the experimental STD values using an NOE *R*-factor defined as (*21*)

$$R\text{-factor} = \sqrt{\frac{\sum(S_{\text{expt},k} - S_{\text{calc},k})^2}{\sum(S_{\text{expt},k})^2}} \qquad (2)$$

In these equations $S_{\text{expt},k}$ and $S_{\text{calc},k}$ refer to experimental and calculated STD values for proton k. Since the protein signals at 10 ppm were irradiated for the STD experiment, we made the reasonable assumption that the aromatic protons in His, Trp, Phe, and Tyr were instantaneously saturated, and that magnetization would take a finite time to spread to other protein and ligand protons (bound and free) through dipolar networks and chemical exchange.

RESULTS AND DISCUSSION

*Binding Mode of Kifunensine **1**.* In order to test the combined STD-NMR/molecular modeling protocol, the binding mode of kifunensine **1** to dGMII was studied initially, and the results were then compared to the crystal structure of the complex (*25*).

A random sampling of the conformational space of **1** using the random search routine (*28*) identified six principal conformations of the six-membered ring moiety of **1**, namely, the ^{1,4}B, ^{1}C$_4$, ^{1}H$_N$, ^{1}S$_3$, ^{2}H$_1$, and ^{3}H$_2$ conformations, the fused five-membered ring being planar. Of these, the ^{1}C$_4$ conformer was observed in the single-crystal structure determined by X-ray crystallography (*36*). Each of these conformations was subsequently docked into the active site of dGMII identified from the high-resolution X-ray structure of the dGMII−kifunensine **1** complex (*25*). For each starting conformer, one hundred docked structures, i.e., 100 runs, were obtained by using the Lamarckian Genetic Algorithm searches and were clustered according to the results differing in positional root-mean-square deviation (rmsd). The lowest docked energy and the average energy of the best cluster (that included the lowest docked energy conformation) for each conformer of **1**, together with the number of structures in each cluster, are listed in Table 1. The docked energy is the sum of the intermolecular energy and the internal energy of the ligand. Intermolecular energy refers to the potential energy of the interaction between the protein and ligand (van der Waals, electrostatic, hydrogen bonding, and desolvation free energy components); internal energy of the ligand represents the contribution to the total energy due to changes in ligand conformation.

Analysis of the docking results indicated that six conformers could be complexed with dGMII in three different binding modes (Figure 1): mode I (^{1,4}B, ^{1}H$_N$, ^{1}S$_3$, and ^{3}H$_2$), mode II (^{1}C$_4$), and mode III (^{2}H$_1$), based on the distance between the zinc ion and the hydroxyl groups of **1** (Table 2). The interactions between kifunensine **1** and dGMII identified from the docking studies and the X-ray structure (*25*),

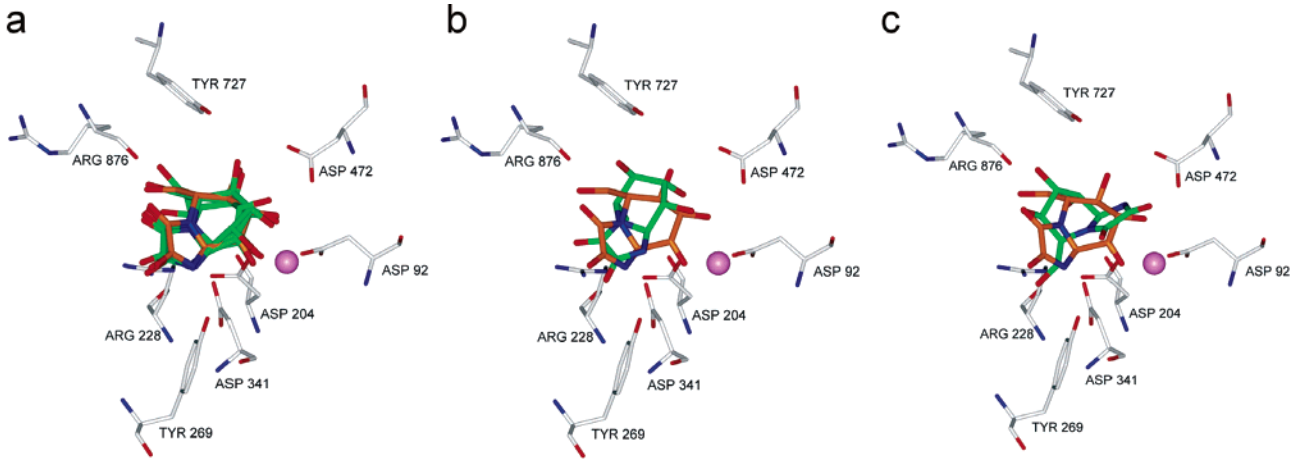


FIGURE 1: Three binding modes of kifunensine **1** in the active site of dGMII from docking calculations: (a) mode I, (b) mode II, and (c) mode III. The zinc ion is shown as a pink ball, and the relevant side chains are shown as sticks. Atoms in **1** are shown in different colors (red, oxygen; blue, nitrogen; and green, carbon). Each binding mode of kifunensine **1** to dGMII (green carbon atoms) is superimposed on the ligand in the corresponding crystal structure (orange carbon atoms) in the active site of dGMII.

Table 1: Calculated Energies of Kifunensine **1** Docked in the Active Site of dGMII[a]

mode	conformation[b]	no. of clusters	no. in the best cluster	docked energy (kcal mol^{-1})[c]	intermolecular energy (kcal mol^{-1})[d]	internal energy of ligand (kcal mol^{-1})[d]
I	^{1,4}B	2	99	−9.55 (−9.46)	−9.05	−0.50
I	$^{1}H_N$	12	38	−9.69 (−9.43)	−9.41	−0.28
I	$^{1}S_3$	12	46	−9.24 (−8.91)	−8.68	−0.56
I	$^{3}H_2$	7	48	−9.84(−9.62)	−9.35	−0.49
II	$^{1}C_4$	17	35	−8.42 (−7.95)	−8.22	−0.20
III	$^{2}H_1$	14	13	−7.09 (−6.34)	−6.82	−0.27

[a] Docked structures are grouped into clusters, the lowest energy cluster being shown. [b] Conformation of six-membered ring for **1**. [c] Of lowest docked energy conformation in the cluster; the energy of the cluster average is given in parentheses. [d] Of lowest energy conformation in the cluster.

Table 2: Interactions of Polar Groups of Kifunensine **1** with Residues in the Active Site of dGMII

groups of kifunensine **1**	residues in dGMII active site	distance (Å) mode I[a]	mode II[b]	mode III	cryst struct[c]
OH-2	Asp92OD1	2.65			2.75
	Asp204OD1	3.06			2.92
	Zn				2.41
OH-3	Asp472OD1		2.84		
	Asp472OD2	2.83			2.37
	Arg876O			2.47	
	Zn				2.37
OH-4	Asp472OD1	2.61			2.61
	Tyr727OH	3.03	2.60		2.72
	Arg876O		2.69	2.47	
OH-6	Arg228NE		2.50		
	Tyr269OH			2.62	
	Arg876O	3.16			2.76
O-7	Asp341OD2	2.99		2.77	
O-8	Asp341OD2		2.70		
	Asp472OD2			2.46	
N	Asp204OD1			2.95	
N9	Tyr269OH	2.79			3.00
	Asp341OD2	2.79	2.85		2.74

[a] Average distances for the lowest energy docked conformers for ^{1,4}B, $^{1}H_N$, $^{1}S_3$, and $^{3}H_2$ in mode I. The distances between the zinc ion and 2- or 3-hydroxyl groups on the six-membered ring of **1** are as follows. Conformer ^{1,4}B: OH-2···Zn = 2.25 Å, OH-3···Zn = 2.21 Å. Conformer $^{1}H_N$: OH-3···Zn = 1.98 Å. Conformer $^{1}S_3$: OH-2···Zn = 1.83 Å. Conformer $^{3}H_2$: OH-3···Zn = 2.02 Å. [b] OH-2···Zn = 3.31 Å. [c] From the crystal structure of the complex of dGMII with kifunensine **1** (*25*).

including hydrogen bonding interactions with putative catalytic residues, Asp204, Asp341, Asp472, are summarized in Table 2.

A key feature of binding mode I is that the distance between the zinc ion and the 2- or 3-hydroxyl groups of **1** is less than 2.3 Å (Table 2). For mode II (conformer $^{1}C_4$), the closest distance between any polar group of **1** to the zinc ion was 3.31 Å. Mode III (conformer $^{2}H_1$) displayed a very different binding mode to I and II. In this case, the orientation of **1** is such that the ligand is turned over in the active site, and all polar groups of **1** are distant from the zinc ion. Previous studies of complexes of dGMII inhibitors with different binding affinities have suggested the importance of coordination with the zinc ion to give a high-affinity inhibitor (*24*, *25*). We infer, therefore, that the binding modes II (conformer $^{1}C_4$) and III (conformer $^{2}H_1$) could be excluded because the distances between the zinc ion and polar groups of **1** in these two modes are too long to form coordinate bonds to zinc.

The mode of binding predicted by the modeling study (mode I) was confirmed by STD-NMR experiments performed on the complex of kifunensine **1** with dGMII. In these studies, 20−30% STD enhancements were observed for seven un-exchangeable protons in kifunensine **1**, as shown in Figure 2. Theoretical STD effects for these seven protons were calculated with CORCEMA-ST (*5*, *21*) for the three binding modes obtained from molecular modeling. As shown in Figure 3a, all the conformations in mode I gave 20−40% STD enhancements for each proton, consistent with the fact that the distances between these protons to the closest residue in dGMII are less than 3 Å. Consequently, a low *R*-factor (0.24) was generated for mode I, which suggested an excellent match between the predicted value and experimental data. However, the $^{1}C_4$ (mode II) and $^{2}H_1$ (mode III) conformations exhibited different patterns (2 and 3, respectively, shown in Figure 3a). The predicted STD values for protons H2 and H3 are much larger than the observed values, the differences in STD intensities between experimental and

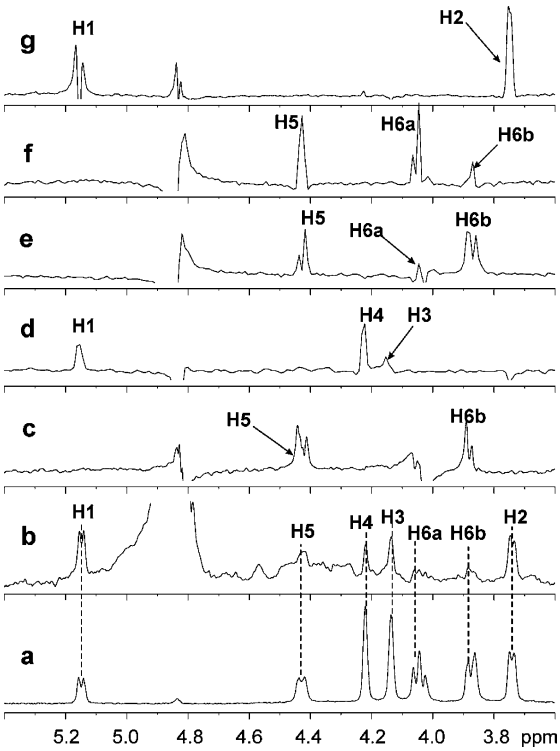


FIGURE 2: (a) Expansions of the reference 1D ^{1}H NMR spectra of of kifunesine **1** in the presence of dGMII, at 600 MHz and 295 K. (b) The corresponding STD-NMR spectrum recorded with a $T_{1\rho}$ filter eliminating the protein signals. (c−g) The corresponding 1D TOCSY-STD-NMR spectra. Arrows indicate the irradiated resonance.

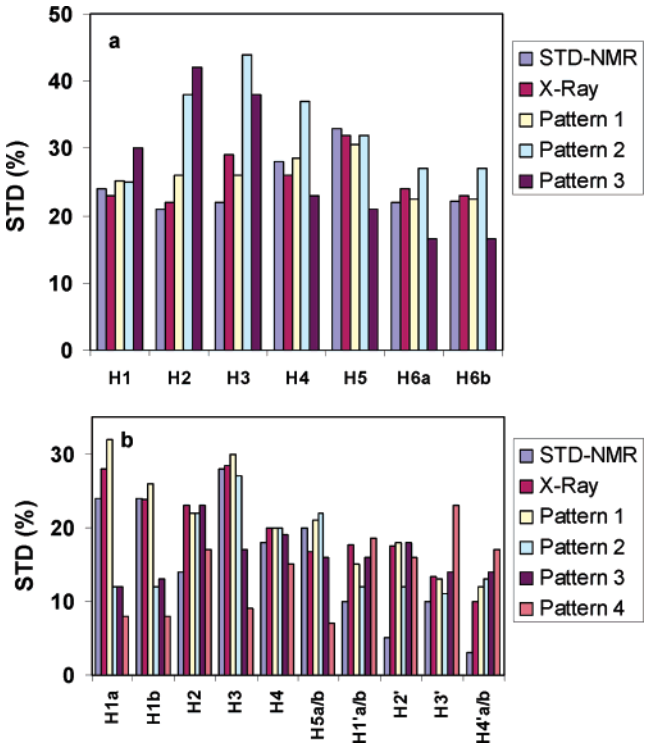


FIGURE 3: Comparison of experimental and predicted STD values from the CORCEMA-ST protocol for (a) kifunensine **1** and (b) salacinol **2** in the presence of dGMII. Experimental STD values (colored in blue) were calculated as $[((I_{0(k)} - I(t)_{(k)})/I_{0(k)}) \times 100]$, with $I_{0(k)}$ being the intensity of the signal of proton k without saturation transfer at time $t = 0$, and $I(t)_k$ being the intensity of proton k after saturation transfer during the saturation time t. Theoretical STD values were predicted by the CORCEMA-ST protocol based on the crystal structures (colored in red) of kifunensine **1**−dGMII and salacinol **2**−dGMII complexes and the modeling results generated from the AutoDock (colored in yellow, green, purple, and orange).

predicted values being greater than 20% for these two protons, which is larger than the experimental error (<10%) in our STD experiments. Thus, as predicted earlier, modes II and III should be excluded.

The four conformers in mode I exhibited very similar binding modes in the active site of dGMII, including the possible coordination with the zinc ion by the 2- or 3-hydroxyl groups on the six-membered ring; they also give similar docked energies (Table 1). In the free ligand state, the ^{1,4}B, ^1H_N, $^3H_2$, and $^1S_3$ conformations give similar calculated energies. It is interesting to note that, of the four conformers, the boat conformation ^{1,4}B was observed in the crystal structure of the complex with dGMII (PDB entry 1PS3) (*25*). Superimposition of conformers ^1H_N, $^3H_2$, and $^1S_3$ on ^{1,4}B (just heavy atoms, not including the 6-OH) gave root-mean-square deviation of 0.448 (^1H_N), 0.256 ($^3H_2$), and 0.411 ($^1S_3$), suggesting that the four conformers of **1** are very similar. We presume that dGMII selects the ^{1,4}B conformation from an equilibrium mixture of the four conformers in solution.

Binding Mode of Salacinol 2. The inherent flexibility of five-membered ring systems due to the facile interconversion of several twist and envelope conformations via pseudorotation (*37*) makes it necessary to consider the different conformers of the ring moiety of **2** for the docking studies.

Twenty unique structures of **2**, representing 10 envelope and 10 twist ring conformations, were obtained based on the random search method (*28*). The conformation of the five-membered ring of **2**, observed in the X-ray crystal structure of salacinol **2** (*38*), was the $^3T_2$ conformation, one of the 20 chosen ring conformations. Each of the conformers of **2** was docked into the dGMII active site using AutoDock 3.0 (*9*). For each starting conformer, five hundred docked structures, i.e. 500 runs, were obtained by using the Lamarckian Genetic Algorithm searches and clustered according to the results differing in positional root-mean-square deviation (rmsd). The lowest docked energy for each conformer of **2** is listed in Table 3. The average values of docked energy for the best cluster that include the lowest energy docked conformation for each conformer of **2**, together with the number of structures in the cluster, are also shown in Table 3.

Analysis of the docking results of salacinol **2** in the active site indicated that 20 unique structures of **2** could dock into the active site of dGMII, but their orientations are quite distinct from one other. According to the orientation of the five-membered ring of **2** in the active site, 20 conformers could be complexed with dGMII in three different binding modes, which differed mainly in the interactions of the zinc ion with either the 2- or 5-hydroxyl groups of the five-membered ring.

The binding mode I included conformers ^3E, E_2, E_4, E_1, ^5E, $^5T_4$, $^1T_2$, and $^5T_1$. In this case, the five-membered ring moieties of **2** appeared in almost the same position in the active site of dGMII in different conformations (see Figure 4a), but the acyclic, sulfated chain appeared in different orientations. For all eight conformers, the 2-hydroxyl groups

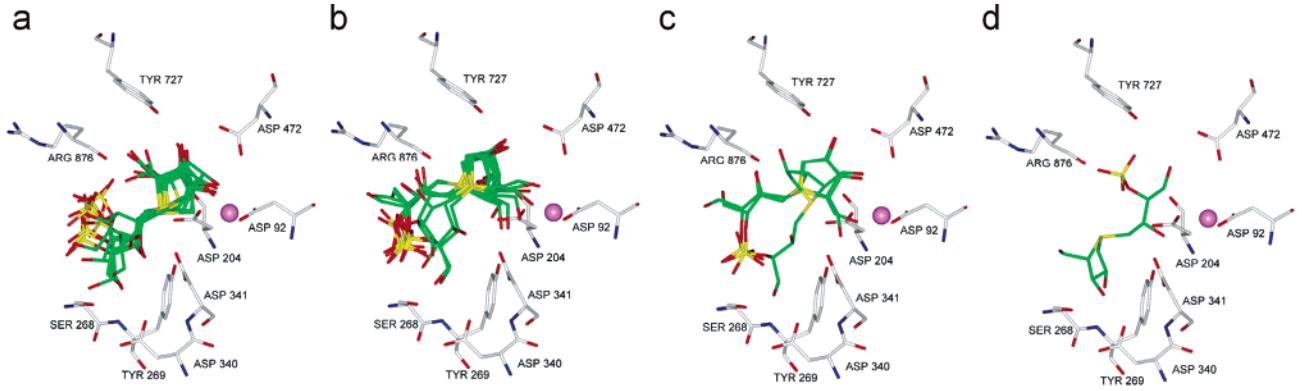


FIGURE 4: Three binding modes of salacinol **2** in the active site of dGMII from docking calculations: (a) mode I, (b, c) mode II, and (d) mode III. The zinc ion is shown as a pink ball, and the relevant side chains are shown as sticks. Atoms in **2** are shown in different colors (red, oxygen; blue, nitrogen; yellow, sulfur; and green, carbon).

Table 3: Calculated Energies of Salacinol **2** Docked in the Active Site of dGMII[a]

mode	conformation[b]	no. of clusters	no. in the best cluster	docked energy (kcal mol^{-1})[c]	intermolecular energy (kcal mol^{-1})[d]	internal energy of ligand (kcal mol^{-1})[d]
I	E_4	31	296	−11.71 (−10.29)	−11.00	−0.71
I	^5E	30	349	−11.48 (−10.07)	−9.98	−1.50
I	$^5T_4$	42	302	−11.44 (−10.10)	−10.66	−0.78
I	^3E	3	495	−11.03 (−10.18)	−9.76	−1.27
I	$^5T_1$	41	186	−10.60 (−8.91)	−9.11	−1.49
I	$^1T_2$	36	360	−10.27 (−9.21)	−8.89	−1.37
I	E_2	48	269	−10.23 (−8.90)	−8.83	−1.39
I	E_1	45	51	−10.02 (−8.12)	−9.36	−0.65
II	$^3T_4$	32	114	−11.38 (−9.89)	−10.83	−0.55
II	$^4T_3$	46	90	−10.98 (−9.60)	−10.42	−0.56
II	$^2T_3$	42	63	−10.82 (−9.32)	−10.28	−0.54
II	$^3T_2$	49	80	−10.67 (−9.41)	−10.08	−0.59
II	$^4T_5$	64	101	−10.52 (−9.18)	−9.98	−0.54
II	^4E	37	133	−10.37 (−8.79)	−9.95	−0.42
II	E_3	24	180	−10.34 (−8.78)	−9.87	−0.47
II	$^1T_5$	56	76	−10.27 (−8.86)	−10.13	−0.13
II	^2E	37	108	−10.26 (−8.41)	−9.82	−0.44
II	$^2T_1$	49	111	−10.21 (−8.58)	−9.36	−0.85
II	E_5	52	47	−9.81 (−8.01)	−9.64	−0.17
III	^1E	59	45	−9.50 (−7.74)	−9.06	−0.45

[a] Docked structures are grouped into clusters, the lowest energy cluster being shown. [b] Conformation of five-membered ring for **2**. [c] Of lowest docked energy conformation in the cluster; the energy of the cluster average is given in parentheses. [d] Of lowest docked energy conformation in the cluster.

Table 4: Interactions of Polar Groups of Salacinol **2** with Residues in the Active Site of dGMII[a]

groups of salacinol **2**	residues in dGMII active site	distance (Å)			
		mode I[a,c]	mode II[b,c]	mode III	complex structure[d]
OH-2	Asp92OD1				2.85
	Asp92OD2	2.85			3.07
	Asp204OD1				2.76
	Asp340OD1			2.78	
	Asp472OD1		2.77		
	Asp472OD2	2.83			2.58
	Tyr727OH		2.76		
	Zn	2.42			2.15
OH-3	Asp204OD2		2.78		
	Ser268OG			2.74	
	Asp472OD1	2.76			2.60
	Tyr727OH	2.88			2.73
OH-5	Asp204OD1		2.94		
	Arg876O				2.76
	Zn		2.91		
S^+	Asp204OD2	2.84			3.30
	Arg876O		3.43		
OH-2′	Asp204OD1			2.83	
	Tyr269OH			2.57	
	Asp341OD2			2.99	
OH-4′	Asp472OD1			2.85	

[a] Average distances for the lowest energy conformers for ^3E, E_2, E_4, E_1, ^5E, $^5T_4$, $^1T_2$, and $^5T_1$ in mode I. [b] Average distances for the lowest energy conformers for E_3, ^2E, ^4E, E_5, $^2T_1$, $^4T_3$, $^2T_3$, $^3T_4$, $^3T_2$, and $^4T_5$ in mode II. [c] The interactions of polar groups on the acyclic chain of **2** with residues in the active site of dGMII have not been included. [d] From the crystal structure of the complex of dGMII with salacinol *(26)*.

on the five-membered-ring portion of **2** coordinated to the zinc ion with a T_5, square-based pyramidal geometry, with an average OH···Zn distance of 2.42 Å. Furthermore, the sulfonium center was significantly closer to the residue Asp204, the putative catalytic nucleophile (*39*), with an average distance of 2.84 Å. In addition, some common interactions were also observed for all the conformers, such as a hydrogen bond between the 2-OH group and residue Asp92 or Asp472, and the 3-OH groups and residue Asp472 or Tyr727, as summarized in Table 4.

In contrast, a very different binding mode of **2** was shown in mode II, which included eleven conformers (E_3, ^2E, ^4E, E_5, $^2T_1$, $^4T_3$, $^1T_5$, $^2T_3$, $^3T_4$, $^3T_2$, and $^4T_5$). In this case, the five-membered rings of almost all conformers overlapped with each other (see Figure 4b), with the exception of $^1T_5$, $^3T_4$, and $^3T_2$. (Although the conformers $^1T_5$, $^3T_4$, and $^3T_2$ indeed belong to mode II and keep the key features presented in mode II, their five-membered rings do not overlap one another in mode II. Thus, their binding modes in the active site of dGMII are shown in Figure 4c in order to display these features more clearly.) The orientation of the five-membered rings of **2** was flipped in contrast to mode I. Consequently, the 5-OH groups of the ring, but not the 2-OH groups, provided closest contacts with the zinc ion in the active site (average distance OH···Zn = 2.91 Å). In fact, the 2-OH groups in this mode were quite distant from the zinc ion. In addition, the sulfonium ion center could not interact electrostatically with residue Asp204 because of the long distance (>4 Å). In this case, the average distance between the sulfonium center and the closest polar atom (the oxygen of Arg876) in the active site was 3.43 Å. Furthermore, there were some common ligand−enzyme interactions, such as an interaction between the 3-OH groups and residue Asp204, the 2-OH groups and Asp472 or Tyr727 for almost

all conformers expect for $^1T_5$. These interactions are also listed in Table 4.

The mode III with only one conformer ^1E showed a completely different binding mode, compared to the other two modes. There was no interaction between the hydroxyl groups in the five-membered ring with any residue in the dGMII active site. In contrast, the acyclic sulfated chain appeared in the active site instead of the five-membered ring, with interactions between the hydroxyl groups of the acyclic chain and residues Asp204, Asp341, Asp472, and Tyr269. In comparison with other binding modes, the conformer ^1E exhibited the highest docked energy (-9.50 kcal mol^{-1}), indicating that this is an unlikely binding mode to dGMII.

Analysis of the three binding modes suggested that the likely binding mode for salacinol **2** to dGMII in would be mode I. The justification for this choice follows. Compound **2** is a glycomimetic bearing a permanent positive charge on the sulfonium ion that interacts electrostatically with active site carboxylate residues such as Asp204 and Asp341, and makes a dominant contribution to the interaction energy (*26*). This interaction mimics that of the positively charged protonated nitrogen atom in swainsonine (*24*) and/or the transient positive charge of a putative oxacarbenium ion transition state in the glycosidase-mediated hydrolysis reaction (*39, 40*). There was no interaction between the sulfonium center of **2** and the negatively charged carboxylate groups of dGMII in binding modes II and III. In mode I, however, the sulfonium center of **2** was significantly closer to the residue Asp204, the putative catalytic nucleophile, with an average distance of 2.84 Å. In addition, the average docked energy for every conformer in mode I (-10.86 kcal mol^{-1}) was lower than that in mode II and mode III (-10.51 and -9.50 kcal mol^{-1}, respectively). On the basis of the analyses above, we conclude that binding mode I is more favorable than mode II or III.

The experimental STD-NMR results for the dGMII−salacinol **2** complex were examined next, together with the theoretical STD values of 20 different conformations calculated by the CORCEMA-ST protocol. As shown by the 1D STD-TOCSY spectra, significant STD effects for the ring protons of **2** were observed (Figure 5). These results correlated well with several of the modes shown by molecular modeling, in which the distances between these protons and the closest residue in dGMII were less than 3 Å. For the protons in the acyclic chain, STD enhancements in the range of 15−20% were predicted for some conformations. However, only very weak STD signals (<5%) were detected in the STD-NMR experiments, suggesting either that this chain is quite flexible in solution within the active site and does not make significant contacts with the protein for extended times or that the chain protrudes from the binding site. However, the latter possibility may be discounted on the basis of the modeling results, and we conclude that the actual distances between protons on the ligand and the protein are much larger than calculated, leading to weak STD effects. That this acyclic chain has mobility within the active site is also corroborated by the X-ray crystallographic data in which structure refinement for the dGMII−salacinol **2** complex proved difficult and the electron density for the acyclic chain was not well defined (*26*).

A more detailed comparison between the experimental and predicted STD values was then used to discriminate between

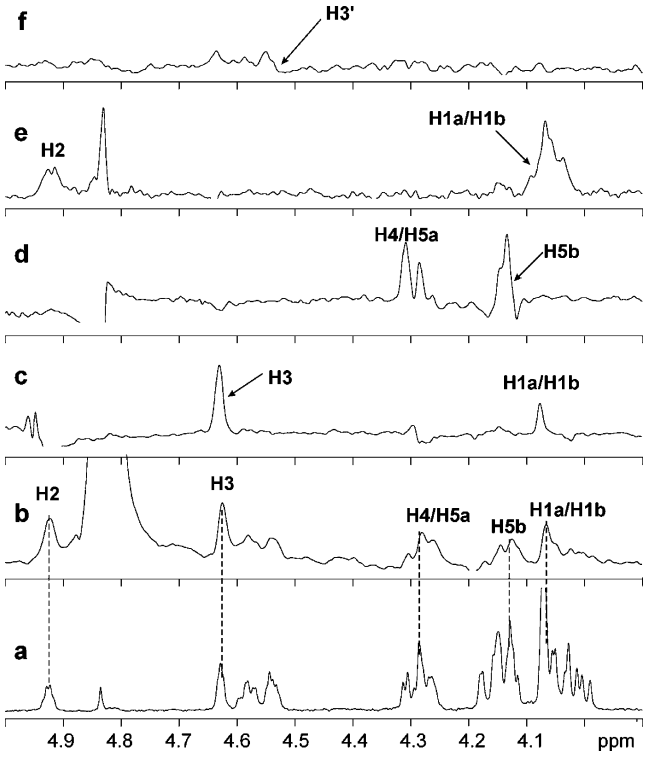


FIGURE 5: (a) Expansions of the reference 1D ^{1}H NMR spectra of of salacinol **2** in the presence of dGMII, at 600 MHz and 295 K. (b) The corresponding STD-NMR spectrum recorded with a $T_{1\rho}$ filter eliminating the protein signals. The STD signals for the side-chain protons are very weak. (c−f) The corresponding 1D TOCSY-STD-NMR spectra. Arrows indicate the irradiated resonance.

different binding modes. The CORCEMA-ST calculations for conformation ^1E, shown as pattern 4 in Figure 3b, predicted very weak STD effects for all ring protons and strong STD effects for the protons in the acyclic chain; therefore, this conformation and mode III can be discarded. The three conformers $^1T_5$, $^3T_4$, and $^3T_2$ in mode II were also judged unlikely to be representative of the bound ligand conformation. Thus, as indicated in pattern 3 in Figure 3b, weak STD effects for H3 and H1a/H1b but strong STD effects for H2 were predicted by the CORCEMA-ST calculation for these three conformers. However, in the 1D STD-NMR experiments, STD effects for H3 and H1 were stronger than that observed for H2. Thus, these three conformers were discounted as candidates within mode II.

For the remaining eight conformers in each of modes I and II, patterns 1 and 2, respectively, were calculated using CORCEMA-ST. These patterns showed strong STD values for H3. The STD effects predicted for H2 and H1a/H1b showed significant differences between patterns 3 and 4 (above), and also correlated well with the different orientations of the five-membered rings shown in modes I and II. For the eight conformations in mode II, the predicted STD values for H2 were larger than those for H1a/H1b. However, for those conformations in mode I stronger STD effects for H1a/H1b compared to that for H2 were predicted; the latter results gave a closer match with the experimental data. Therefore, on the basis of the combined analysis of experimental STD-NMR effects and the predicted effects from CORCEMA-ST calculations, we propose that the orientation of the five-membered ring of salacinol **2** in the dGMII active site would be as shown in mode I.

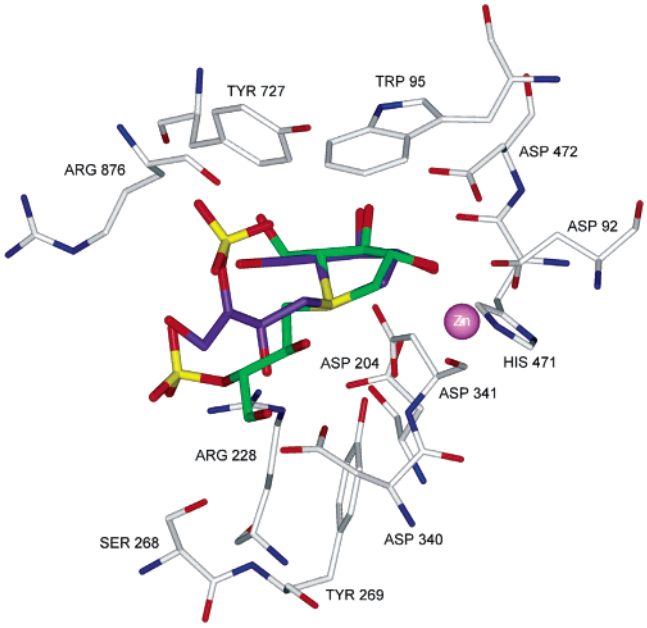


FIGURE 6: Binding modes of conformer ^{4}E of salacinol **2** in the active site of dGMII, as predicted by AutoDock 3.0. The calculated structure (green carbon atoms) is superimposed on the crystal structure of **2** (amethyst carbon atoms) in the active site of dGMII. Atoms in **2** and contact residues of dGMII are shown in different colors (red, oxygen; blue, nitrogen; and yellow, sulfur).

Of the different conformers represented by mode I, conformer E_4 in binding mode I exhibited the most favorable docked energy (-11.71 kcal mol^{-1}). This conformation was the same as that of salacinol **2** in the complex with dGMII in the crystal structure (see Figure 6) (*26*). Thus, we predict that salacinol **2** binds to dGMII in mode I and, further, that the five-membered ring of **2** adopts an E_4 envelope conformation that resembles the half-chair conformation of a putative oxacarbenium ion transition state; the transient positive charge of that state is mimicked by the permanent positive charge provided by the sulfonium ion of **2**.

In addition to understanding the mode of inhibition of dGMII by novel compounds, these results can be extended to yield insights into the conformational itinerary of the ligand during the course of the enzyme-mediated hydrolysis reaction. For dGMII, Numao et al. (*41*) have suggested an itinerary for the deglycosylation step that passes through a $B_{2,5}$ conformation based on a $^1S_5$ covalent intermediate. Interestingly, the ^{1,4}B conformation predicted by the present study and observed in the dGMII−kifunensine **1** complex (*25*) (and mimicked by the results with salacinol **2**), in the context of a natural mannose substrate, would fit well into a sampling of the same conformational itinerary. Application of the approach described here to a true substrate could substantiate this proposal further. Finally, these results emphasize the important role of the Zn atom in inducing the correct substrate conformation for catalysis.

CONCLUSIONS

A protocol that combines the use of molecular modeling by AutoDock, experimental STD-NMR effects, and calculated STD effects with CORCEMA-ST has yielded reliable models for the binding modes of kifunensine **1** and salacinol **2** in the active site of *D. melanogaster* Golgi α-mannosidase II (dGMII). The structures predicted for the complexes of **1** and **2** with dGMII compare favorably with those obtained by X-ray crystallography (*25*, *26*), and lend credence to the validity of the protocol. In addition, discrepancies in experimental and calculated STD effects for the acyclic chain in **2** suggest that this chain exhibits flexibility within the active site.

ACKNOWLEDGMENT

We are grateful to Dr. N. R. Krishna for providing the CORCEMA-ST program.

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published 04/16/05 with Figures 5 and 6 incorrectly labeled. The corrected version was published 04/19/05.

REFERENCES

1. Kelly, M. D., and Mancera, R. L. (2004) Expanded interaction fingerprint method for analyzing ligand binding modes in docking and structure-based drug design, *J. Chem. Inf. Comput. Sci. 44*, 1942−1951.
2. Ota, N., and Agard, D. A. (2001) Binding mode prediction for a flexible ligand in a flexible pocket using multi-conformation simulated annealing pseudo crystallographic refinement, *J. Mol. Biol. 314*, 607−617.
3. Klebe, G. (2000) Recent developments in structure-based drug design, *J. Mol. Med. 78*, 269−281.
4. Johnson, M. A., and Pinto, B. M. (2004) NMR spectroscopic and molecular modeling studies of protein-carbohydrate and protein-peptide interactions, *Carbohydr. Res. 339*, 907−928.
5. Jayalakshmi, V., and Krishna N. R. (2004) CORCEMA refinement of the bound ligand conformation within the protein binding pocket in reversibly forming weak complexes using STD-NMR intensities, *J. Magn. Reson. 168*, 36−45.
6. Taylor, R. D., Jewsbury, P. J., and Essex, J. W. (2002) A review of protein-small molecule docking methods, *J. Comput.-Aided Mol. Des. 16*, 151−166.
7. Halperin, I., Ma, B., Wolfson, H., and Nussinov, R. (2002) Principles of docking: An overview of search algorithms and a guide to scoring functions, *Proteins: Struct., Funct., Genet. 47*, 409−443.
8. Ewing, T. J. A., Makino, S., Skillman, A. G., and Kuntz, I. D. (2001) DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases, *J. Comput.-Aided Mol. Des. 15*, 411−428.
9. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem. 19*, 1639−1662.
10. Rarey, M., Kramer, B., Lengauer, T., and Klebe, G. (1996) A fast flexible docking method using an incremental construction algorithm, *J. Mol. Biol. 261*, 470−489.
11. Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997) Development and validation of a genetic algorithm for flexible docking, *J. Mol. Biol. 267*, 727−748.
12. Goodsell, D. S., and Olson, A. J. (1990) Automated docking of substrates to proteins by simulated annealing, *Proteins 8*, 195−202.
13. Goodsell, D. S., Lauble, H., Stout, C. D., and Olson, A. J. (1993) Automated docking in crystallography: analysis of the substrates of aconitase, *Proteins 17*, 1−10.
14. Goodsell, D. S., Morris G. M., and Olson, A. J. (1996) Automated docking of flexible ligands: applications of AutoDock, *J. Mol. Recognit. 9*, 1−5.
15. Meyer, B., and Peters, T. (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors, *Angew. Chem., Int. Ed. 42*, 864−890.

16. Mayer, M., and Meyer, B. (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy, *Angew. Chem., Int. Ed. 38*, 1784−1788.
17. Mayer, M., and Meyer, B. (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor, *J. Am. Chem. Soc. 123*, 6108−6117.
18. Johnson, M. A., and Pinto, B. M. (2002) Saturation transfer difference 1D-TOCSY experiments to map the topography of oligosaccharides recognized by a monoclonal antibody directed against the cell-wall polysaccharide of group A *Streptococcus*, *J. Am. Chem. Soc. 124*, 15368−15374.
19. Bhunia, A., Jayalakshmi, V., Benie, A. J., Schuster, O., Kelm, S., Krishna, N. R., and Peters, T. (2004) Saturation transfer difference NMR and computational modeling of a sialoadhesin-sialyl lactose complex, *Carbohydr Res. 339*, 259−267.
20. Jayalakshmi, V., Biet, T., Peters, T., and Krishna, N. R. (2004) Refinement of the conformation of UDP-galactose bound to galactosyltransferase using the STD NMR intensity-restrained CORCEMA optimization, *J. Am. Chem. Soc. 126*, 8610−8611.
21. Jayalakshmi, V., and Krishna, N. R. (2002) Complete relaxation and conformational exchange matrix (CORCEMA) analysis of intermolecular saturation transfer effects in reversibly forming ligand-receptor complexes, *J. Magn. Reson. 155*, 106−118.
22. Rabouille, C., Kuntz, D. A., Lockyer, A., Watson, R., Signorelli, T., Rose, D. R., van den Heuvel, M., and Roberts, D. B. (1999) The *Drosophila* GMII gene encodes a Golgi α-mannosidase II, *J. Cell Sci. 112*, 3319−3330.
23. Goss, P. E., Reid, C. L., Bailey, D., and Dennis, J. W. (1997) Phase IB clinical trial of the oligosaccharide processing inhibitor swainsonine in patients with advanced malignancies, *Clin. Cancer Res. 3*, 1077−1086.
24. van den Elsen, J. M., Kuntz, D. A., and Rose, D. R. (2001**)** Structure of Golgi α-mannosidase II: A target for inhibition of growth and metastasis of cancer cells, *EMBO J. 20*, 3008−3017.
25. Shah, N., Kuntz, D. A., and Rose, D. R. (2003) Comparison of kifunensine and 1-deoxymannojirimycin binding to class I and II alpha-mannosidases demonstrates different saccharide distortions in inverting and retaining catalytic mechanisms, *Biochemistry 42*, 13812−13816.
26. Kuntz, D. A., Ghavami, A., Johnston, B. D., Pinto B. M. and Rose, D. R. (2005) Crystallographic analysis of the interactions of *Drosophila melanogaster* Golgi α-mannosidase II with the naturally occurring glycomimetic salacinol and its analogues, *Tetrahedron: Asymmetry 16*, 25−32.
27. Yoshikawa, M., Murakami, T., Shimada, H., Matsuda, H., Yamahara, J., Tanabe, G., and Muraoka, O. (1997) Unique antidiabetic principle from ayurvedic traditional medicine in Sri Lanka and India, *Tetrahedron Lett. 38*, 8367−8370.
28. Saunders, M., Houk, K. N., Yun-Dong, W., Still, W. C., Lipton, M., Chang, G., and Guida, W. C. (1990) Conformations of cycloheptadecane. A comparison of methods for conformational searching, *J. Am. Chem. Soc. 112*, 1419−1427.
29. Clark, M., Cramer, R. D. 3rd, and Opdenbosch, N. V. (1989) Validation of the general purpose Tripose 5.2 Force Field, *J. Comput. Chem. 10*, 982−1012.
30. Gasteiger, J., and Marsili, M. (1981) Prediction of proton magnetic resonance shifts: the dependence on hydrogen charges obtained by iterative partial equalization of orbital electronegativity, *Org. Magn. Reson. 15*, 353−360.
31. Gasteiger, J., and Marsili, M. (1980) Iterative partial equalization of orbital electronegativity − a rapid access to atomic charges, *Tetrahedron 36*, 3219−3228.
32. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids and organic molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
33. Ghavami, A., Johnston, B. D., and Pinto, B. M. (2001) A new class of glycosidase inhibitor: synthesis of salacinol and its stereoisomers, *J. Org. Chem. 66*, 2312−2317.
34. Dellwo, M. J., and Wand, A. J. (1993) The influence of methyl rotor dynamics on hydrogen relaxation networks: derivation of spectral densities in model-free form, *J. Am. Chem. Soc. 115*, 1886−1893.
35. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546−4559.
36. Kayakiri, H., Takase, S., Shibata, T., Okamoto, M., Terano, H., and Hashimoto, M. (1989) Structure of kifunensine, a new immunomodulator isotated from an actinomycete, *J. Org. Chem. 54*, 4015−4016.
37. Kilpatrick, J. E., Pitzer, K. S., and Spitzer, R. (1947) The thermodynamics and molecular structure of cyclopentane, *J. Am. Chem. Soc. 69*, 2483−2488.
38. Yoshikawa, M., Murikama, T., Matsuda, H., Tanabe, G., and Muraoka, O. (2002) Absolute stereostructure of potent α-glucosidase inhibitor, salacinol, with unique thiosugar sulfonium sulfate inner salt structure from salacia reticulate, *Bioorg. Med. Chem. 10*, 1547−1554.
39. Numao, S., He, S., Evjen, G., Howard, S., Tollersrud, O. K., and Withers, S. G. (2000) Identification of Asp197 as the catalytic nucleophile in the family 38 α-mannosidase from bovine kidney lysosomes, *FEBS Lett. 484*, 175−178.
40. Lillelund, V. H., Jensen, H. H., Liang, X., and Bols, M. (2002) Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin, *Chem. Rev. 102*, 515−553.
41. Numao, S., Kuntz, D. A., Withers, S. G., and Rose, D. R. (2003) Insights into the mechanism of *Drosophila melanogaster* Golgi α-Mannosidase II through the structural analysis of covalent reaction intermediates. *J. Biol. Chem. 278*, 48074−48083.

BI0500426