Toward the Virtual Screening of α-Glucosidase Inhibitors with the Homology-Modeled Protein Structure

Jung-Hum Park, Sungmin Ko, and Hwangseo Park*

*Department of Bioscience and Biotechnology, Sejong University, Seoul 143-747, Korea. *E-mail: hspark@sejong.ac.kr*
Received January 26, 2008

Discovery of α-glucosidase inhibitors has been actively pursued with the aim to develop therapeutics for the treatment of diabetes and the other carbohydrate mediated diseases. As a method for the discovery of new novel inhibitors of α-glucosidase, we have addressed the performance of the computer-aided drug design protocol involving the homology modeling of α-glucosidase and the structure-based virtual screening with the two docking tools: FlexX and the automated and improved AutoDock implementing the effects of ligand solvation in the scoring function. The homology modeling of α-glucosidase from baker's yeast provides a high-quality 3-D structure enabling the structure-based inhibitor design. Of the two docking programs under consideration, AutoDock is found to be more accurate than FlexX in terms of scoring putative ligands to the extent of 5-fold enhancement of hit rate in database screening when 1% of database coverage is used as a cutoff. A detailed binding mode analysis of the known inhibitors shows that they can be stabilized in the active site of α-glucosidase through the simultaneous establishment of the multiple hydrogen bond and hydrophobic interactions. The present study demonstrates the usefulness of the automated AutoDock program with the improved scoring function as a docking tool for virtual screening of new α-glucosidase inhibitors as well as for binding mode analysis to elucidate the activities of known inhibitors.

Key Words : Docking, Scoring function, Virtual screening, Homology modeling, α-Glucosidase inhibitor

Introduction

Glucosidases catalyze the final step in the digestive process of carbohydrates by the hydrolysis of a glycosidic bond in oligosaccharides. They are responsible for the catalytic cleavage of a glycosidic bond with specificity depending on the number monosaccharides, the position of cleavage site, and the configuration of the hydroxyl groups in the substrate.[1] The most extensively studied are α- and β-glucosidases that are known to catalyze the hydrolysis of the glycosidic bonds involving a terminal glucose at the cleavage site through α- and β-linkages at the anomeric center. These two glucosidases differ in how to position their two carboxylic acid sidechains during catalysis:[2] one plays the role of a catalytic nucleophile attacking the anomeric center, and the other acts as an acid catalyst weakening the C-O bond by protonation. Of the two popular glucosidases, α-glucosidase (EC 3.2.1.20) has drawn a special interest of the pharmaceutical research community because it was shown in earlier studies that the inhibition of its catalytic activity resulted in the retardation of glucose absorption and the decrease in postprandial blood glucose level.[3-5] Therefore, effective α-glucosidase inhibitors may serve as chemotherapeutic agents for clinic use in the treatment of diabetes and obesity. Due to the catalytic role in digesting carbohydrate substrates, α-glucosidase has also been well appreciated as a therapeutic target for the other carbohydrate mediated diseases including cancer,[6] viral infections,[7,8] and hepatitis.[9]

Since the discovery of acarbose that is the first member of α-glucosidase inhibitors approved for the treatment of type 2 diabetes,[10] a variety of α-glucosidase inhibitors have been discovered and recently reviewed in an extensive fashion.[11] These include transition state analogues,[12] newly identified synthetic compounds,[13-20] and natural products isolated from a variety of species.[21-23] Most of the α-glucosidase inhibitors reported in the literature stem from either the isolation of new scaffolds by high throughput screening or the generation of the improved derivatives of pre-existing inhibitor scaffolds. So far the rational drug design protocol has not been applied for α-glucosidases because the structural investigations have lagged behind the mechanistic and pharmacological studies. Indeed, structural information of α-glucosidases has thus been limited to those of a few bacterial strains only in ligand-free forms.[24,25] The lack of structural information about the nature of the interactions between α-glucosidases and small molecule inhibitors has thus made it a difficult task to discover good lead compounds based on the structure-based inhibitor design.

In the present study, we address the performance of a computer-aided drug design protocol involving the homology modeling of α-glucosidase and the structure-based virtual screening with docking simulation as a tool for identifying novel classes of potent α-glucosidase inhibitors. Two popular docking programs, FlexX and AutoDock, are used in this work. The characteristic feature that discriminates our virtual screening approach from the others lies in the implementation of an accurate solvation model in calculating the binding free energy between α-glucosidase and its putative ligands, which would have the effect of increasing the hit rate in enzyme assay.[26,27] We select the α-glucosidase from baker's yeast as the target protein because it has been used most extensively in biological assays to

evaluate the newly discovered α-glucosidase inhibitors. To the best of our knowledge, we report the first example for the usefulness of the structure-based virtual screening to identify novel α-glucosidase inhibitors. It will be shown that the docking simulation with the improved binding free energy function can be a valuable tool for enriching the chemical library used in screening assays with molecules that are likely to have a desired biological activity.

Computational Methods

Homology modeling of yeast α-glucosidase. Although the X-ray crystal structures of a few bacterial α-glucosidases have been reported, structural information is still unavailable for the eukaryotic α-glucosidase enzymes commonly used in biological assays, such as that from baker's yeast. In order to obatin the three-dimensional structure of α-glucosidase from baker's yeast, therefore, we carried out the homology modeling using the X-ray structure of oligo-1,6-glucosidase from *Bacillus cereus* as the template.[28] This homology modeling started with the retrieval of the amino acid sequence of the α-glucosidase MAL12 from baker's yeast that comprises 584 amino acid residues from the SWISS-PROT protein sequence data bank (http://www.expasy.org/sprot/; accession number P53341).[29] In order to find a proper structural template for homology modeling, we searched for the Protein Data Bank (PDB) at National Center for Biotechnology and Information (NCBI) using BLAST and PSIBLAST algorithms with the amino acid sequence of the target as input. The results showed that oligo-1,6-glucosidase from *Bacillus cereus* reveals the highest sequence identity (38.5%) with the target. Therefore, its X-ray crytal structure (PDB ID: 1UOK) was selected as the template for homology modeling. Although 4-α-glucanotransferase from *Thermotoga maritima* revealed a sequence identity of about 30% with the target protein, it was not used in homology modeling because the number of aligned amino acids amount to at most 300 as compared to 575 in case of oligo-1,6-glucosidase from *Bacillus cereus*. Sequence alignment between α-glucosidase from baker's yeast and oligo-1,6-glucosidase from *Bacillus cereus* was then obtained with the ClustalW package[30] using the BLOSUM matrices for scoring the alignments. The parameters of GAP OPEN, GAP EXTENTION, and GAP DISTANCE were set equal to 10, 0.05, 8, respectively. Opening and extension gap penalties were changed systematically, and the obtained alignment was inspected for violation of structural integrity in the structurally conserved regions. Based on the best-scored sequence alignment, the three dimensional structure of α-glucosidase from baker's yeast was constructed using the MODELLER 6v2 program.[31] In this model building, we employed an optimization method involving conjugate gradients and molecular dynamics to minimize violations of the spatial restraints. With respect to the structure of gap regions, the coordinates were built from a randomized and distorted structure that is located approximately between the two anchoring regions as implemented in MODELLER 6v2. To increase the accuracy of calculated structure, the loop modeling was also performed with the enumeration algorithm.[32] Then, we calculated the conformational energy of the predicted structure of α-glucosidase with ProSa 2003 program[33] for the purpose of a final evaluation.

Construction of a docking library. The docking library for α-glucosidase comprises its own 20 known inhibitors as well as 980 common compounds selected from the MDL Drug Data Report (MDDR) database. This selection was based on drug-like filters that adopt only the compounds with physicochemical properties of potential drug candidates[34] and without reactive functional group(s). All of the compounds included in the docking library were then subjected to the Corina program to generate their 3-D coordinates, followed by the assignment of Gasteiger-Marsilli atomic charges.[35] The chemical structures of the 20 known inhibitors of α-glucosidase seeded in the docking library are shown in Supporting Information.

Virtual screening of α-glucosidase inhibitors with AutoDock. We used the automated version of the AutoDock program[36] in the structure-based virtual screening of α-glucosidase inhibitors because the outperformance of its scoring function over those of the others had been shown in several target proteins.[37] The atomic coordinates of α-glucosidase obtained from the homology modeling were used as the receptor model in the virtual screening with docking simulations. A special attention was paid to assign the protonation states of the ionizable Asp, Glu, His, and Lys residues. The side chains of Asp and Glu residues were assumed to be neutral if one of their carboxylate oxygens pointed toward a hydrogen-bond accepting group including the backbone aminocarbonyl oxygen at a distance within 3.5 Å, a generally accepted distance limit for a hydrogen bond of moderate strength.[38] Similarly, the side chains of Lys residues were protonated unless the NZ atom was in a close proximity of a hydrogen-bond donating group. The same procedure was also applied to determine the protonation states of ND and NE atoms in the side chains of His residues.

In the actual docking simulation of the compounds in the docking library, we used the empirical AutoDock scoring function improved by the implementation of a new solvation model for a compound. The modified scoring function has the following form:

$$\Delta G_{bind}^{aq} = W_{vdW}\sum_{i=1}\sum_{j>i}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond}\sum_{i=1}\sum_{j>i}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec}\sum_{i=1}\sum_{j>i}\frac{q_i q_j}{\varepsilon(r_{ij})r_{ij}} + W_{tor}N_{tor} + W_{sol}\sum_{i=1}S_i\left(Occ_i^{\max} - \sum_{j>i}V_j e^{-\frac{r_{ij}^2}{2\sigma^2}}\right) \quad (1)$$

where W_{vdW}, W_{hbond}, W_{elec}, W_{tor}, and W_{sol} are the weighting factors of van der Waals, hydrogen bond, electrostatic interactions, torsional term, and desolvation energy of inhibitors, respectively. r_{ij} represents the interatomic distance, and A_{ij}, B_{ij}, C_{ij}, and D_{ij} are related to the depths of the potential energy well and the equilibrium separations between the two

atoms. In this study, AMBER force field parameters were assigned for calculating the van der Waals interactions and the internal energy of a ligand as implemented in the AutoDock program. The hydrogen bond term has an additional weighting factor, $E(t)$, representing the angle-dependent directionality. With respect to the distant-dependent dielectric constant, $\varepsilon(r_{ij})$, a sigmoidal function proposed by Mehler *et al*.[39] was used in computing the interatomic electrostatic interactions between a receptor protein and a ligand molecule. In the entropic term, N_{tor} is the number of sp^3 bonds in the ligand. In the desolvation term, S_i and V_i are the solvation parameter and the fragmental volume of atom i,[40] respectively, while Occ_i^{max} stands for the maximum atomic occupancy. In the calculation of molecular solvation free energy term in equation (1), we used the atomic parameters recently developed by Kang *et al*.[41] because those of the atoms other than carbon were unavailable in the current version of AutoDock. This modification of the solvation free energy term is expected to increase the accuracy in virtual screening, because the underestimation of ligand solvation often leads to the overestimation of the binding affinity of a ligand with many polar atoms.[27]

The docking simulation of a compound in the docking library started with the calculation of the three-dimensional grids of interaction energy for all of the possible atom types present in chemical database. These uniquely defined potential grids for the receptor protein were then used in common for docking simulations of all compounds in the docking library. As the center of the common grids in the active site, we used the center of mass coordinates of the docked structure of the probe molecule, acarbose, whose binding mode had been known in the active site of 4-α-glucanotransferase that is closely similar in structure to the template (oligo-1,6-glucosidase) used in the homology modeling.[42] The calculated grid maps were of dimension 61 × 61 × 61 points with the spacing of 0.375 Å, yielding a receptor model that includes atoms within 22.9 Å of the grid center. For each compound in the library, 10 docking runs were performed with the initial population of 50 individuals. Maximum number of generations and energy evaluation were set to 27,000 and 2.5×10^5, respectively.

Virtual screening of α-glucosidase inhibitors with FlexX. All default parameters, as implemented in Sybyl 6.9, were used for all target proteins and compounds in docking simulations. The active site and the interaction surface of the receptor were defined by using the reference ligand, acarbose, whose binding mode had been calculated with docking simulations and cutoff distance of 6.5 Å. The conformational flexibility of a ligand was modeled by a discrete set of preferred torsional angles for acyclic single bonds. Base fragments were then selected automatically with the maximum number of 4. A base fragment was placed into the active site based on the two algorithms. The first one superimposes triplets of interaction centers of the base fragment with triples of compatible interaction sites. Second, the matching algorithm was used when the base fragment had fewer than three interaction centers. The empirical scoring function given in equation (2) was used for ranking the binding modes of each ligand in the prepared compound databases:[43]

$$\Delta G_{bind} = \Delta G_0 + W_{hbond} \sum_{hbonds} f(\Delta R,\Delta\alpha) + W_{ionic} \sum_{ionic} f(\Delta R,\Delta\alpha) + W_{aro/aro} \sum_{aro/aro} f(\Delta R,\Delta\alpha) + W_{lipo} \sum_{lipo} f^*(\Delta R) + W_{tor} N_{tor}. \quad (2)$$

```
MAL12     9  TEPKWWKEATIYQIYPASFKDSNNDGWGDLKGITSKLQYIKDLGVDAIWVCPFYDSPQQD  68
O16GB     1  MEKQWWKESVVYQIYPRSFMDSNGDGIGDLRGIISKLDYLKELGIDVIWLSPVYESPNDD  60

MAL12    69  MGYDISNYEKVWPTYGTNEDCFELIDKTHKLGMKFITDLVINHCSTEHEWFKESRSSKTN 128
O16GB    61  NGYDISDYCKIMNEFGTMEDWDELLHEMHERNMKLMMDLVVNHTSDEHNWFIESRKSKDN 120

MAL12   129  PKRDWFFWRPPKGYDAEGKPIPPNNWKSFFGGSAWTFDETTNEFYLRLFASRQVDLNWEN 188
O16GB   121  KYRDYYIWRPGK----EGK--EPNNWGAAFSGSAWQYDEMTDEYYLHLFSKKQPDLNWDN 174

MAL12   189  EDCRRAIFESAVGFWLDHGVDGFRIDTAGLYSKRPGLPDSPIFDKTSKLQHPNWGSHNGP 248
O16GB   175  EKVRQDVYEMMK-FWLEKGIDGFRMDVINFISKEEGLPTVETEEEGYVSGHKHF--MNGP 231

MAL12   249  RIHEYHQELHRFMKNRVKDGREIMTVGEVAHGS--DNALYTSAARYEVSEVFSFTHVEVG 306
O16GB   232  NIHKYLHEMN----EEVLSHYDIMTVGEMPGVTTEEAKLYTGEERKELQMVFQFEHMDLD 287

MAL12   307  TSPFFRYNIVPFTLKQWKEAIASNFLFINGTDSWATTYIENHDQARSITRFADDSPKYRK 366
O16GB   288  SGEGGKWDVKPCSLLTLKENLTKWQKALEHTG-WNSLYWNNHDQPRVVSRFGNDG-MYRI 345

MAL12   367  ISGKLLTLLECSLTGTLYVYQGQEIGQINFKEWPIEKYEDVDVKNNYEIIKKSFGKNSKE 426
O16GB   346  ESAKMLATVLHMMKGTPYIYQGEEIGMTNVRFESIDEYRDIETLN---MYKEKVMERGED 402

MAL12   427  MKDFFKGIALLSRDHSRTPMPWTKDKPNAGFTGPDVKPWFLLNESFEQGINVEQESRDDD 486
O16GB   403  IEKVMQSIYIKGRDNARTPMQWD-DQNHAGFT--TGEPWITVNPNYKE-INVKQAIQNKD 458

MAL12   487  SVLNFWKRALQARKKYKELMIYGYDFQFIDLDSDQIFSFTKEYEDKTLFAALNFSGEEIE 546
O16GB   459  SIFYYYKKLIELRKN-NEIVVYG-SYDLILENNPSIFAYVRTYGVEKLLVIANFTAEECI 516

MAL12   547  FSLPREG--ASLSFILGNYDDTD--VSSRVLKPWEGRIYLVK 584
O16GB   517  FELPEDISYSEVELLIHNYDVENGPIENITLRPYEAMVFKLK 558
```

Figure 1. Sequence alignment between α-glucosidase (MAL12) and oligo-1,6-glucosidase (O16GB). The identity and the similarity between the corresponding residues are indicated in red and green, respectively. The active site residues are indicated in a blue rectangular box.

Here, $f(\Delta R,\Delta\alpha)$ is a scaling function penalizing deviations from the ideal distances and angles and $f^{*}(\Delta R)$ penalizes the forbiddingly close contacts for lipophilic interactions of nonaromatic groups.

Results and Discussion

Homology Modeling of α-glucosidase. Figure 1 displays the sequence alignment between α-glucosidase MAL12 from baker's yeast and oligo-1,6-glucosidase from *Bacillus cereus* (O16GB). According to this alignment, the sequence identity and the similarity amount to 38.5% and 58.4%, respectively. Judging from such a high sequence homology, a high-quality 3D structure of α-glucosidase can be expected in the homology modeling. It is indeed well known that a homology-modeled structure of a target protein can be accurate enough to be used in docking studies once the sequence identity between target and template approaches 40%.[44] Based on the sequence alignment shown in Figure 1, ten structural models of α-glucosidase were calculated and the one with the lowest value of MODELLER objective function was selected as the final model to be used in the virtual screening.

Figure 2 shows the structure of α-glucosidase obtained from the homology modeling in comparison with the X-ray crystal structure of oligo-1,6-glucosidase that was used as the template. The target and the template possess a very similar folding structure and are superimposable over the main chain atoms. The two enzymes also share the catalytic residues that are situated in their respective active sites in a similar fashion. This is not surprising because both enzymes catalyze the hydrolysis of terminal glycosidic bond of carbohydrates.[44] The van der Waals volumes of α-glucosidase and oligo-1,6-glucosidase are calculated to be 40,928 and 41,073 Å^3, respectively. Such a close similarity in the van der Waals volume and the possession of 18 additional amino acids in the sequence alignment indicate that α-glucosidase should be more compact in amino acid packing than oligo-1,6-glucosidase. Such a structural difference may be related with the differentiations of the active site geometry and substrate specificity as well as the catalytic efficiency. Indeed, glucosidases have generally exhibited a high specificity in enzyme catalysis by cleaving only one type of glycosidic linkage in a given anomeric configuration.[45]

The final structural model of α-glucosidase obtained from homology modeling was tested with the ProSa 2003 program by examining whether the interaction of each residue with the remainder of the protein is maintained favorable. This program calculates the knowledge-based mean fields to judge the quality of protein folds, and has been widely used to measure the stability of a protein conformation. More specifically, the energy profile of a protein is calculated using the potential of mean forces derived from a large set of known protein structures. The main criterion is that the interaction energy of each residue with the remainder of the protein should have a negative value. Figure 3 shows the ProSa 2003 energy profile of the homology-modeled α-glucosidase in comparison to that of the X-ray structure of oligo-1,6-glucosidase. We note that the ProSa energy of α-glucosidase remains negative for all amino acid residues except for a few around the residue number of 210, indicating the acceptability of the homology modeled structure. This result supports the possibility that the homology modeling with a high sequence identity and a high-quality template structure can produce a 3-D structure of a target protein comparable in accuracy to that determined from X-ray crystallography.[44]

As a further evaluation of the homology-modeled structure of α-glucosidase, the final model obtained with MODELLER was subject to stereochemical analysis with the PROCHECK program. The results show that the backbone F and Y dihedral angles of 69.3%, 25.1%, and 5.6% of the residues are located within most favorable, additionally allowed, and generously allowed regions of the Ramachandran plot, respectively, with no residue in disallowed region. This good stereochemical quality is not surprising for the high sequence identity (38.5%) and similarity (58.4%) between the template and the target as illustrated in Figure 1.

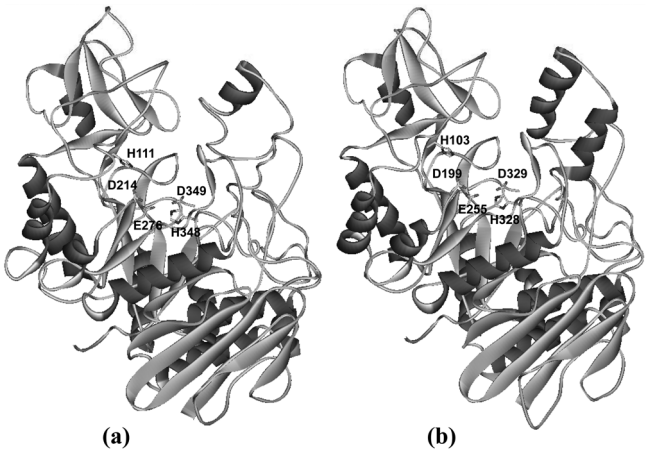

Figure 2. Comparative view of (a) homology-modeled structure of α-glucosidase and (b) X-ray crystal structure of oligo-1,6-glucosidase.

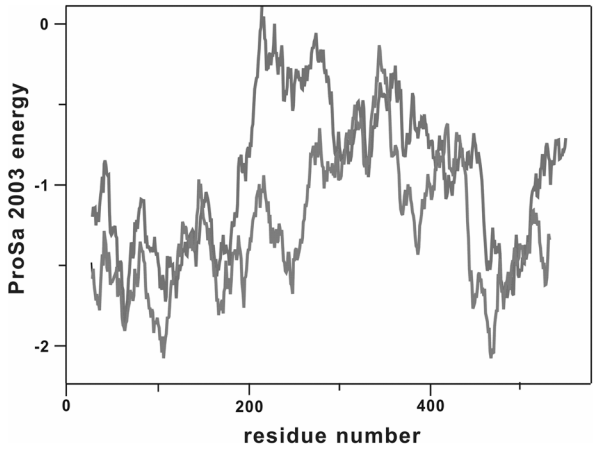

Figure 3. Comparison of the ProSa energy profiles for the homology-modeled structure of α-glucosidase (red) and the X-ray structure of oligo-1,6-glucosidase (green).

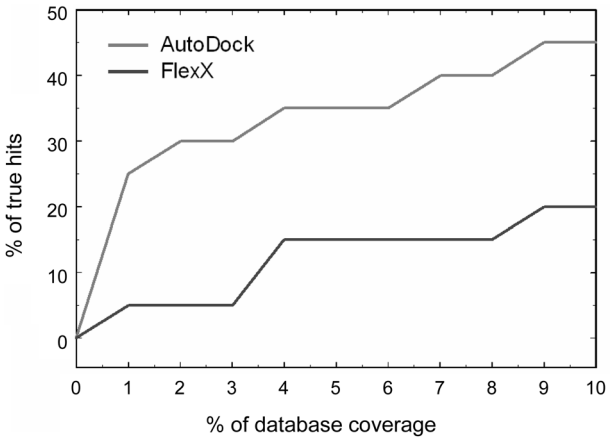

Figure 4. The cumulative percentage of known α-glucosidase inhibitors recovered by virtual screening as a function of the top-scoring fraction of database selected for generating a hit list.

Virtual screening. We have tested the performances of the automated AutoDock and FlexX in the structure-based virtual screening of α-glucosidase inhibitors. This comparative evaluation was done with the homology-modeled structure of α-glucosidase as the target protein and the docking library that contains 980 randomly chosen drug-like molecules and 20 known inhibitors. Compared in Figure 4 are the percentages of true hits retrieved by the AutoDock and FlexX in increasing fractions of the starting database. The horizontal and vertical axes represent the top percentage of one thousand of all tested compounds and the percentage of the known inhibitors in a given top percentage, respectively. We note that the AutoDock performs better than FlexX in providing the highest enrichment at every fraction cutoff. It picks 5 actives seeded in top 1% of the database as compared to 1 for FlexX. The performance of the AutoDock becomes clearer when one compares the ability to pick out the most actives out of a cumulative total of 20 used in this study. When 10% of the database is considered, for example, the AutoDock retrieved a total of 9 actives out of the total 20 known inhibitors, contrary to only 4 actives by FlexX. Thus, the outperformance of the automated AutoDock reveals a consistency for all cutoffs, indicating that it can be a promising docking tool for virtual screening of α-glucosidase inhibitors.

The difference in the accuracies of AutoDock and FlexX in database screening can be understood by comparing their respective scoring functions. It is common to the two docking programs that their scoring functions include the angle-dependent directionality of a hydrogen bond and entropic penalty for the formation of a protein-ligand complex. On the other hand, there are two characteristic features that discriminate the scoring function of AutoDock from that of FlexX: the use of a sigmoidal distance-dependent dielectric function in the electrostatic term and desolvation cost for complexation of a ligand in the binding site. The former has an effect of modeling solvent screening in the electrostatic interactions between charged atoms.[39] This is important because the top-scored ligands obtained with a small value of dielectric constant tend to possess many atoms with high partial charges as a consequence of the overestimation of electrostatic interactions. The effect of ligand solvation is also important, particularly in comparing many putative ligands that differ in polarity and size. The hit compounds may have a severe charge separation on their molecular structures or be larger than expected unless the energy of the solvated state is considered in docking simulations.[27] Thus, a significant outperformance of AutoDock over FlexX should be attributed to the inclusion of solvation term in the scoring function as well as a more proper description of electrostatic interactions between protein and ligand atoms.

Molecular modeling studies of the known inhibitors. Shown in Figure 5 are the chemical structures of the known α-glucosidase inhibitors in the top 1% of all of the tested compounds obtained with AutoDock (**1**-**5**) and FlexX (**6**). It is noted that none of the six inhibitors is retrieved by both of the two virtual screening programs. This is not surprising due to the difference in the scoring functions of the two programs as shown in equations (1) and (2). Virtual screening with AutoDock predicts that the two compounds (**1** and **2** in Figure 5) are the strongest binders in the active site of α-

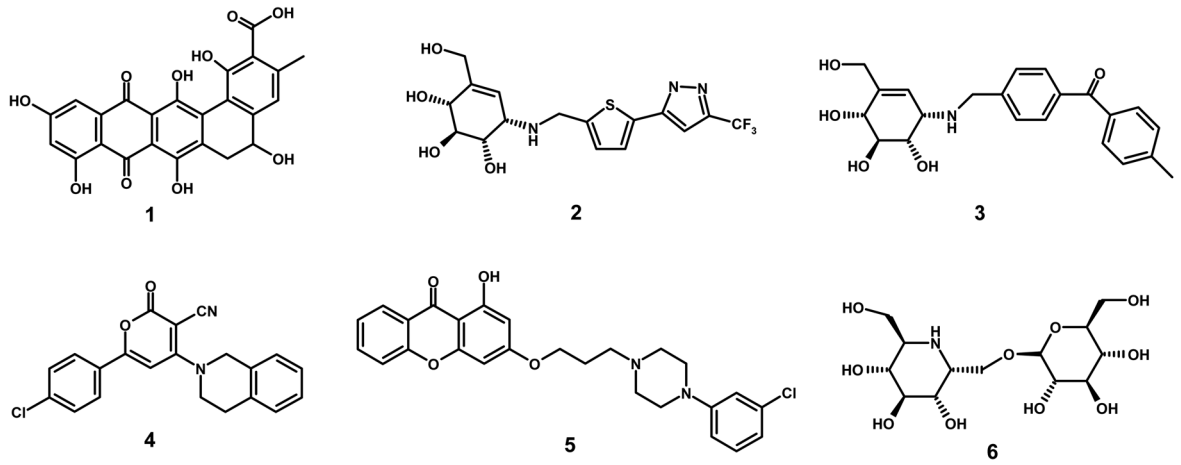

Figure 5. Chemical structures of the top-scored α-glucosidase inhibitors retrieved in virtual screening with AutoDock (**1**-**5**) and FlexX (**6**).

glucosidase among the 20 known inhibitors under consideration, the systematic names of which are 1,5,7,9,11,14-hexahydroxy-3-methyl-8,13-dioxo-5,6,8,13-tetrahydro-benzo-[*a*]naphthacene-2-carboxylic acid and 4-hydroxymethyl-6-{[5-(5-trifluoromethyl-2*H*-pyrazol-3-yl)-thiophen-2-ylmethyl]-amino}-cyclohex-4-ene-1,2,3-tripol, respectively. It is noted that both inhibitors possess many hydroxyl groups, indicating the involvement of multiple hydrogen bonds in their interactions with the active site of α-glucosidase. It is also a common structural feature of the two inhibitors that the polar groups are attached to a hydrophobic backbone. These hydrophobic moieties seem to be stabilized at the active site through the interactions with the nonpolar groups of α-glucosidase.

To gain more structural insight into the inhibitory mechanism for α-glucosidase, the binding modes of **1** and **2** were examined using the AutoDock program with the procedure described in the previous section. The calculated binding modes of the two inhibitors in the active site of α-glucosidase are compared in Figure 6. It is seen that the 5,7-dihydroxy-[1,4]naphthoquinone moiety of **1** resides in close proximity to the catalytic residues including Asp214, Glu276, His348, and Asp349, indicating that it can serve as a surrogate for the terminal glucose with anomeric center in the substrate. Four hydrogen bonds are established between the phenolic and carbonyl oxygens of **1** and the side chains of Asp214, Thr215, Ser244, and Arg312. This is consistent with recent computational studies on the inhibition of β-glucosidase in which the formation of multiple hydrogen bonds in the active site was shown to be a significant binding force.[46] We also note that the tetrahydro-benzo[*a*]naphthacene backbone of **1** forms hydrophobic contacts with the side chains of Tyr71, Phe157, His279, Phe300, and Phe311, indicating that van der Waals interactions would also play a significant role in stabilizing the enzyme-inhibitor complex. Therefore, **1** is most likely to be capable of inhibiting the catalytic action of α-glucosidase by a tight binding in the active site through the multiple hydrogen bond and hydrophobic interactions in a cooperative fashion.

Four hydrogen bonds are also observed in the calculated binding mode of **2** between the hydroxyl groups of the inhibitor and the side chains of Asp68, His111, Asp214, and Arg349. Hence, the formation of multiple hydrogen bonds seems to play a role of anchoring the inhibitors to the enzymatic active site. The thiophenylpyrazole moiety of **2** is stabilized by hydrophobic contacts with Tyr71, Phe157, His279, Phe300, Thr307, Phe311, and Arg312 in a stronger way than the hydrophobic interactions of **1** in the active site of α-glucosidase. On the other hand, the backbone scaffold of **1** has no torsional degree of freedom, indicating a substantial decrease in entropic penalty for the formation of enzyme-inhibitor complex as compared to the other ligands with rotatable bonds. The entropic contribution has indeed been shown to be the most significant ingredient in the binding free energy function.[35] This affinity-enhancing factor reflected in **1** seems to compensate for its relatively weak van der Waals interactions in the active site, which can be an explanation for the similarity in the calculated binding free energies of **1** and **2**.

Conclusions

As a method for the discovery of new novel inhibitors of α-glucosidase, we have addressed the performance of the computer-aided drug design protocol involving the homology modeling of the target protein and the structure-based virtual screening with the two docking tools: FlexX and the automated and improved AutoDock implementing the effects of ligand solvation in the binding free energy function. The homology modeling of α-glucosidase provides a high-quality 3-D structure to the extent of enabling the structure-based inhibitor design. Of the two docking programs under consideration, AutoDock is found to be more accurate than FlexX in terms of scoring putative ligands with 5-fold enhancement of hit rate in database screening when 1% of database coverage is used as a cutoff. The outperformance of the improved AutoDock program in virtual screening of α-glucosidase inhibitors can be attributed to the accuracy in the scoring function in which the effects of ligand solvation in protein-ligand interaction are taken into account. It is also shown from a detailed binding mode analysis of the known inhibitors that their binding in the active site of α-glucosidase can be facilitated by the establishment of multiple hydrogen bonds with the side

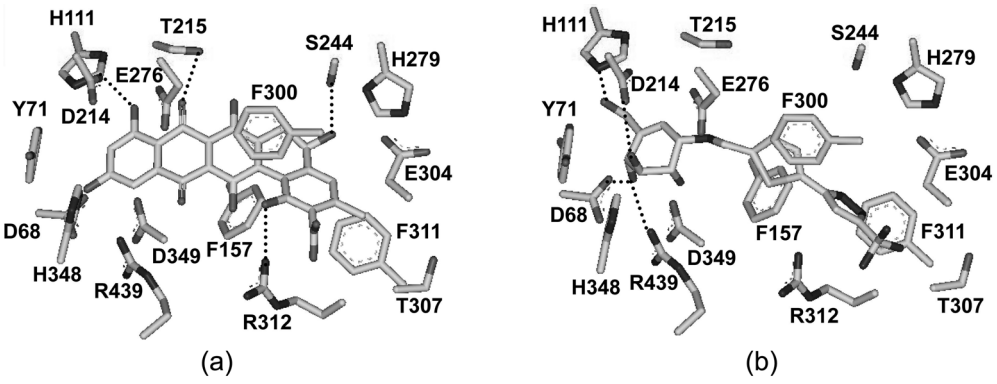

Figure 6. Calculated binding modes of (a) **1** and (b) **2** in the active site α-glucosidase. Carbon atoms of the protein and the ligand are indicated in green and cyan, respectively. Each dotted line indicates a hydrogen bond.

chains of polar residues. Simultaneously, the hydrophobic interactions with the residues near the active site can also play a significant role in stabilizing the inhibitors in the active site of α-glucosidase. The present study demonstrates the usefulness of the automated AutoDock program with the improved scoring function as a docking tool for virtual screening of new α-glucosidase inhibitors as well as for binding mode analysis to elucidate the activities of known inhibitors.

Acknowledgments. This work was supported by the faculty research fund of Sejong University in 2006.

Supporting Information. The chemical structures of the 20 known inhibitors of α-glucosidase seeded in the docking library are available on request from the correspondence author.

References

1. Kimura, A.; Lee, J.-H.; Lee, I.-S.; Lee, H.-S.; Park, K.-H.; Chiba, S.; Kim, D. *Carbohydr. Res.* **2004**, *339*, 1035.
2. Heightman, T. D.; Andrea, T.; Vasella, A. T. *Angew. Chem. Int. Ed.* **1999**, *38*, 750.
3. Robinson, K. M.; Begovic, M. E.; Rhinehart, B. L.; Heineke, E. W.; Ducep, J. B.; Kastner, P. R.; Marshall, F. N.; Danzin, C. *Diabetes* **1991**, *40*, 825.
4. Braun, C.; Brayer, G. D.; Withers, S. G. *J. Biol. Chem.* **1995**, *270*, 26778.
5. Dwek, R. A.; Butters, T. D.; Platt, F. M.; Nicole Zitzmann, N. *Nature Rev. Drug. Discov.* **2002**, *1*, 65.
6. Humphries, M. J.; Matsumoto, K.; White, S. L.; Olden, K. *Cancer Res.* **1986**, *46*, 5215.
7. Mehta, A.; Zitzmann, N.; Rudd, P. M.; Block, T. M.; Dwek, R. A. *FEBS Lett.* **1998**, *430*, 17.
8. Karpas, A.; Fleet, G. W. J.; Dwek, R. A.; Petursson, S.; Namgoong, S. K.; Ramsden, N. G.; Jacob, G. S.; Rademacher, T. W. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 9229.
9. Zitzmann, N.; Mehta, A. S.; Carrouee, S.; Butters, T. D.; Platt, F. M.; McCauley, J.; Blumberg, B. S.; Dwek, R. A.; Block, T. M. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11878.
10. Yee, H. S.; Fong, N. T. *Pharmacotherapy* **1996**, *16*, 792.
11. de Melo, E. B.; Gomes, A. S.; Carvalho, I. *Tetrahedron* **2006**, *62*, 10277.
12. Lillelund, V. H.; Jensen, H. H.; Liang, X.; Bols, M. *Chem. Rev.* **2002**, *102*, 515.
13. Xu, H.-W.; Dai, G.-F.; Liu, G.-Z.; Wang, J.-F.; Liu, H.-M. *Bioorg. Med. Chem.* **2007**, *15*, 4247.
14. Tanabe, G.; Yoshikai, K.; Hatanaka, T.; Yamamoto, M.; Shao, Y.; Minematsu, T.; Muraoka, O.; Wang, T.; Matsuda, H.; Yoshikawa, M. *Bioorg. Med. Chem.* **2007**, *15*, 3926.
15. Liu, Y.; Ma, L.; Chen, W.-H.; Wang, B.; Xu, Z.-L. *Bioorg. Med. Chem.* **2007**, *15*, 2810.
16. Pandey, J.; Dwivedi, N.; Singh, N.; Srivastava, A. K.; Tamarkar, A.; Tripathi, R. P. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1321.
17. Hakamata, W.; Nakanishi, I.; Masuda, Y.; Shimizu, T.; Higuchi, H.; Nakamura, Y.; Saito, S.; Urano, S.; Oku, T.; Ozawa, T.; Ikota, N.; Miyata, N.; Okuda, H.; Fukuhara, K. *J. Am. Chem. Soc.* **2006**, *128*, 6524.
18. Dai, G.-F.; Xu, H.-W.; Wang, J.-F.; Liu, F.-W.; Liu, H.-M. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2710.
19. Liu, H.; Sim, L.; Rose, D. R.; Pinto, B. M. *J. Org. Chem.* **2006**, *71*, 3007.
20. Seo, W. D.; Kim, J. H.; Kang, J. E.; Ryu, H. W.; Curtis-Long, M. J.; Lee, H. S.; Yang, M. S.; Park, K. H. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5514.
21. Luo, J.-G.; Wang, X.-B.; Ma, L.; Kong, L.-Y. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4460.
22. Saludes, J. P.; Lievens, S. C.; Molinski, T. F. *J. Nat. Prod.* **2007**, *70*, 436.
23. Du, Z.-Y.; Liu, R.-R.; Shao, W.-Y.; Mao, X. P.; Ma, L.; Gu, L.-Q.; Huang, Z.-S.; Chan, A. S. C. *Eur. J. Med. Chem.* **2006**, *41*, 213.
24. Lodge, J. A.; Maier, T.; Liebl, W.; Hoffmann, V.; Sträter, N. *J. Biol. Chem.* **2003**, *278*, 19151.
25. Rajan, S. S.; Yang, X.; Collart, F.; Yip, V. L. Y.; Withers, S. G.; Varrot, A.; Thompson, J.; Davies, G. J.; Anderson, W. F. *Structure* **2004**, *12*, 1619.
26. Zou, X.; Sun, Y.; Kuntz, I. D. *J. Am. Chem. Soc.* **1999**, *121*, 8033.
27. Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. *Proteins* **1999**, *34*, 4.
28. Watanabe, K.; Hata, Y.; Kizaki, H.; Katsube, Y.; Suzuki, S. *J. Mol. Biol.* **1997**, *269*, 142.
29. Bairoch, A.; Apweiler, R. *Nucl. Acids Res.* **1999**, *27*, 49.
30. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucl. Acids Res.* **1994**, *22*, 4673.
31. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
32. Fiser, A.; Do, R. K. G.; Sali, A. *Protein Sci.* **2000**, *9*, 1753.
33. Sippl, M. J. *Proteins* **1993**, *17*, 355.
34. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug. Delivery. Rev.* **1997**, *23*, 3.
35. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
36. 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.
37. Park, H.; Lee, J.; Lee, S. *Proteins* **2006**, *65*, 549.
38. Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997.
39. Mehler, E. L.; Solmajer, T. *Protein Eng.* **1991**, *4*, 903.
40. Stouten, P. F. W.; Frömmel, C.; Nakamura, H.; Sander, C. *Mol. Simul.* **1993**, *10*, 97.
41. Kang, H.; Choi, H.; Park, H. *J. Chem. Inf. Model.* **2007**, *47*, 509.
42. Roujeinikova, A.; Raasch, C.; Sedelnikova, S.; Liebl, W.; Rice, D. W. *J. Mol. Biol.* **2002**, *321*, 149.
43. Böhm, H. J. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243.
44. Baker, D.; Sali, A. *Science* **2001**, *294*, 93.
45. Nishio, T.; Hakamata, W.; Kimura, A.; Chiba, S.; Takatsuki, A.; Kawachi, R.; Oku, T. *Carbohydr. Res.* **2002**, *337*, 629.
46. Zhou, J.-M.; Zhou, J.-H.; Meng, Y.; Chen, M.-B. *J. Chem. Theory Comput.* **2006**, *2*, 157.