Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2008) 1202–1212
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Binding mode analyses and pharmacophore model development for sulfonamide chalcone derivatives, a new class of α-glucosidase inhibitors

Kavitha Bharatham [a,1], Nagakumar Bharatham [a,1], Ki Hun Park [b], Keun Woo Lee [a,*]

[a] *Division of Applied Life Science (BK21 Program), Environmental Biotechnology National Core Research Center, Gyeongsang National University, Jinju 660-701, Republic of Korea*
[b] *Division of Applied Life Science (BK21 Program), Department of Agricultural Chemistry, Institute of Agriculture & Life Sciences, Gyeongsang National University, Jinju 660-701, Republic of Korea*

Received 24 September 2007; received in revised form 7 November 2007; accepted 7 November 2007
Available online 17 November 2007

Abstract

Sulfonamide chalcone derivatives are a new class of non-saccharide compounds that effectively inhibit glucosidases which are the major targets in the treatment of Type 2 diabetes and HIV infection. Our aim is to explore their binding mode of interaction at the active site by comparing with the sugar derivatives and to develop a pharmacophore model which would represent the critical features responsible for α-glucosidase inhibitory activity. The homology modeled structure of *Saccharomyces cerevisiae* α-glucosidase was built and used for molecular docking of non-sugar/sugar derivatives. The validated docking results projected the crucial role of NH group in the binding of sugar/non-sugar derivatives to the active site. Ligplot analyses revealed that Tyr71, and Phe177 form hydrophobic interactions with sugar/non-sugar derivatives by holding the terminal glycosidic ring mimics. Molecular dynamic (MD) simulation studies were performed for protein alone and with chalcone derivative to prove its binding mechanism as shown by docking/Ligplot results. It would also help to substantiate the homology modeled structure stability. With the knowledge of the crucial interactions between ligand and protein from docking and MD simulation studies, features for pharmacophore model development were chosen. The *CATALYST*/HipHop was used to generate a five featured pharmacophore model with a training set of five non-sugar derivatives. As validation, all the crucial features of the model were perfectly mapped onto the 3D structures of the sugar derivatives as well as the newly tested non-sugar derivatives. Thus, it can be useful in virtual screening for finding new non-sugar derivatives as α-glucosidase inhibitors.

Keywords: α-Glucosidase; GOLD molecular docking; MD simulation; Pharmacophore model development; Sulfonamide chalcone derivatives; Ligplot

1. Introduction

α-Glucosidases (EC 3.2.1.20; α-D-glucoside glucohydrolase) are a group of exo-acting enzymes that play essential roles in carbohydrate metabolism, and in glycoprotein processing and quality control. Though members of this class of enzymes share the ability to release a terminal glucose moiety from the non-reducing end of their substrates, they display significant diversity in amino acid sequence and aglycon specificity [1]. Glycosidases are involved in the biosynthesis and processing of the oligosaccharide chains of N-linked glycoproteins in the endoplasmic reticulum (ER) [2]. Inhibition of these glycosidases, especially α-glucosidases, has a profound effect on the glycan structure which consequently affects the maturation, transport, secretion, and function of glycoproteins, and could therefore alter cell–cell or cell–virus recognition processes [3–5]. The mature human immunodeficiency virus type-1 (HIV-1) envelope glycoprotein (Env) originates from the intracellular cleavage of precursor gp160 into outermembrane gp120 and transmembrane gp41 subunits. Both species remain noncovalently linked within oligomeric structures during routing to the cell surface [6]. After virus budding, gp120 exposed on the viral surface mediates HIV binding to CD4+ lymphocytes through interaction with cell surface antigen CD4. Gp120 domains, and in particular the variable V1, V2, and V3 loops,

* Corresponding author.
E-mail address: kwlee@gnu.ac.kr (K.W. Lee).
[1] These authors equally contributed.

1093-3263/$ – see front matter
doi:10.1016/j.jmgm.2007.11.002

interact then with virus coreceptors, including chemokine receptors which induce conformation changes of Env resulting in unmasking of the gp41 fusion peptide and triggering of HIV/cell fusion [7].

For AIDS therapy, there are currently a number of compounds available for multiple targets approved by the FDA and in clinical trials, e.g. protease inhibitors, reverse transcriptase inhibitors (NRTI, NNRTI) and fusion inhibitors, CCR4, CCR5 among others [8–11]. In some cases, the action against the virus occurs before the entrance of HIV into the host cells. Among the recent targets used for drug development in AIDS treatment, glucosidases have been recently explored as novel fusion targets because the synthesis of viral glycoproteins depend on the activity of enzymes, such as glucosidase and transferase, for the elaboration of the polysaccharides [12]. Type 2 diabetes is a metabolic disorder that results from complex interactions of multiple factors and is characterized by two major defects: decreased secretion of insulin by the pancreas and resistance to the action of insulin in various tissues (muscle, liver and adipose), which results in impaired glucose uptake [13]. The oral hypoglycemic agents (OHA) competitively inhibit enzymes in the small intestinal brush border that are responsible for the breakdown of oligosaccharides and disaccharides into monosaccharides suitable for absorption [14]. It works primarily on α-glucosidase, which is found predominantly in the proximal half of the small intestine. The intestinal absorption of carbohydrates is therefore delayed.

Saccharide-like compounds are potential therapeutic agents for the treatment of AIDS, as well as Gaucher's disease, cancer, diabetes, and can disrupt glycoprotein processing via direct-site irreversible glucosidase inhibition. Related compounds, such as hydroxycyclohexanes and their derivatives play an important role in this process because they mimic the disaccharide unit which is cleaved by glucosidases [15]. Most of the glucosidase inhibitors are sugars or derivatives of sugar moieties and there are only few non-saccharide compounds which effectively inhibit glucosidases [16]. Recently our report on sulfonamide chalcone derivatives demonstrated better inhibitory activities in micro molar range against α-glucosidase and other closely related targets [17].

The aim of present work is to investigate and compare the binding properties of sulfonamide chalcone derivatives with other known compounds as well as leads which are proven by experimental studies. We have modeled *Saccharomyces cerevisiae* (baker's yeast) α-glucosidase structure using homology modeling to study the binding mechanism of glucosidase inhibitors. Molecular docking and molecular dynamics (MD) simulation studies were performed to investigate the crucial interactions between the ligand the protein. Based on these results, finally we have developed pharmacophore models using the *CATALYST*/*HipHop* methodology.

2. Methods and materials

2.1. Homology modeling: 3D structure generation using computational methods

The identification of homologues of α-glucosidase (Swissprot code P53341/P38158) was carried out by performing sequence database searches with standard tools, such as PSI-BLAST and blastp. The coordinates of the crystal structure of *B. cereus* oligo-1,6-glucosidase (PDB ID: 1UOK, 2.00 Å resolution) [18] was used as template to build the initial *S. cerevisiae* α-glucosidase structure. The most significant step in homology modeling process is to obtain the correct sequence alignment of the target sequence with the homologues and the reliable sequence–structure mapping has allowed us to identify important structural features. Sequence alignment was performed by using Aign2d program [19]. The 3D model of the *S. cerevisiae* α-glucosidase was built by *MODELLER* module [20,21] in *INSIGHTII* software [22] based on the *B. cereus* oligo-1,6-glucosidase structure. The final protein model was validated by a well-established program, PROCHECK [23] for the evaluation of Ramachandran plot.

2.2. Molecular docking simulation

The binding interactions can be ascertained by docking the inhibitors into the active site of the protein. The *GOLD* 3.01 [24] program was used to calculate the docking modes of α-glucosidase inhibitors into the active site of the homology modeled protein structure. It employs genetic algorithm in which the information about the ligand conformation and hydrogen bonding is encoded in chromosome. *GOLD* considers complete ligand flexibility and partial protein flexibility and the energy functions are partly based on conformational and non-bonded interactions. Several types of scoring functions such as GoldScore, ChemScore and User-defined score are available. The following default genetic algorithm parameters were used: 100 population size, 1.1 for selection, 5 number of islands, 100,000 number of genetic operations and 2 for the niche size. The pseudo-atom was created at the center of the catalytic triad, i.e. Asp214, Glu276, Asp349 and the active site defined as 10 Å around it. The GoldScore was opted to rank order the docked conformations.

2.3. Molecular dynamics simulation

MD simulations were performed using GROMACS program (version 3.3.1) with the GROMACS force field [25,26]. The starting geometries for the simulation were prepared from the homology modeled α-glucosidase structure and the *GOLD* docked conformation of the compound **3**. Therefore two MD simulations were performed with and without the compound **3**. The MD simulation protocol that was used is as follows. All hydrogens were added and the protonation state of ionizable groups was chosen appropriate to pH 7.0. Each system was inserted in a water box where the layer of the water molecules was equal to 12 Å. The entire systems were neutralized by adding Na^+ counterions by replacing water molecules. The systems were subjected to a steepest descent energy minimization process until a tolerance of 1000 kJ/mol, step by step. Then the protein backbone was frozen and the solvent molecules with counterions were allowed to move during a 50 ps position restrained MD run under NPT conditions at 300 K. The production run was for 3 ns for all systems. The

sizes of the studied systems are around 103,600 atoms each. The simulation period was chosen as a compromise between the quality of configuration space sampling and the calculation length. The time step for the simulations was 2 fs. The electrostatic interactions were calculated by using the Particle-mesh Ewald (PME) algorithm, with interpolation order of 4 and a grid spacing of 0.12 nm. All simulations were run under periodic boundary conditions with NPT ensemble by using Berendsen's coupling algorithm for maintaining the temperature (300 K) and the pressure constant pressure (1 bar). The SHAKE algorithm with a tolerance of 10^{-5} Å was applied to fix all bonds containing hydrogen atoms. The Van der Waals forces were treated by using a cutoff of 14 Å and the coordinates were stored every 1 ps. All the analyses of the MD simulations were carried out by GROMACS and all the computations were performed in a high performance Linux cluster computer.

2.4. Pharmacophore model generation

Pharmacophore hypotheses generation was carried out by *CATALYST/HipHop* [27] using non-sugar derivatives for further use in the screening for new lead derivatives. As there are only few non-sugar derivatives available, *HipHop* methodology is opted as it can generate common feature pharmacophores from the training set. Five non-sugar derivatives were selected as training set to develop pharmacophore models and all 3D structures of the molecules were built using 2D/3D editor sketcher in the *CATALYST* 4.11 software package [28]. Conformational models for the molecules was developed by poling algorithm implemented in *CATALYST*, which seeks to provide a broad coverage of conformational space using the best conformer generation method with a maximum conformational energy of 20 kcal/mol above the lowest energy conformation found [29,30]. The number of conformers generated for each compound was limited to a maximum of 250. *HipHop* has taken into account the features of five compounds and found the 3D arrangement of functional groups common to all that interact with the α-glucosidase enzyme, thus carrying out a ''qualitative model''.

3. Results and discussion

3.1. Protein 3D structure generation using homology modeling method

As most of the biological testing of α-glucosidase inhibitors were carried out on *S. cerevisiae*, it would be appropriate to investigate the binding mode of sulfonamide chalcone derivatives with *S. cerevisiae* structure. Unfortunately the 3D structure of the protein is not available so far and hence a homology modeled structure was developed. The structure of *B. cereus* oligo-1,6-glucosidase was taken as template because blastp showed it as the best hit with 38.5% sequence identity and 58.4% similarity. Although 4-α-glucanotransferase from *T. maritima* revealed a sequence identity of about 30%, it was excluded in homology modeling because of few reasons: (i) the number of aligned amino acids amounts to atmost 300 as compared to 575 in case of oligo-1,6-glucosidase from *B. cereus* [31]; (ii) both 1,4 and oligo-1,6-glucosidases identify the terminal glucose moiety and break the glycosidic bond unlike *T. maritima* 4-α-glucanotransferase. The catalytic triad comprises of the same Asp-Glu-Asp residues responsible for catalytic

```
P_1UOK  --------MEKQWWKESVVYQIYPRSFMDSNGDGIGDLRGIISKLDYLKELGIDVIWLSP
YEAST   MTISDHPETEPKWWKEATIYQIYPASFKDSNNDGWGDLKGITSKLQYIKDLGVDAIWVCP
                 *  ****    ***** ** *** ** *** ** *** * * ** * **  *
P_1UOK  VYESPNDDNGYDISDYCKIMNEFGTMEDWDELLHEMHERNMKLMMDLVVNHTSDEHNWFI
YEAST   FYDSPQQDMGYDISNYEKVWPTYGTNEDCFELIDKTHKLGMKFITDLVINHCSTEHEWFK
         * **  * ***** * *     ** **  **    *   **   *** ** * ** **
P_1UOK  ESRKSKDNKYRDYYIWRPGK----EGK--EPNNWGAAFSGSAWQYDEMTDEYYLHLFSKK
YEAST   ESRSSKTNPKRDWFFWRPPKGYDAEGKPIPPNNWKSFFGGSAWTFDETTNEFYLRLFASR
        *** ** *  **   *** *    ***   ****   * ****  ** * * ** **
P_1UOK  QPDLNWDNEKVRQDVYE-MMKFWLEKGIDGFRMDVINFISKEEGLPTVETEEEGYVSGHK
YEAST   QVDLNWENEDCRRAIFESAVGFWLDHGVDGFRIDTAGLYSKRPGLPDSPIFDKTSKLQHP
        * **** **  *    *    ***  * **** *     **  ***           *
P_1UOK  HF--MNGPNIHKYLHEMNEEVLSHY----DIMTVGEMPGVTTEEAKLYTGEERKELQMVF
YEAST   NWGSHNGPRIHEYHQELHRFMKNRVKDGREIMTVGEV--AHGSDNALYTSAARYEVSEVF
           *** **    *              ******         ***   * *   **
P_1UOK  QFEHMDLDSGEGGKWDVKPCSLLTLKENLTKWQKALEHT-GWNSLYWNNHLQPRVVSRFG
YEAST   SFTHVEVGTSPFFRYNIVPFTLKQWKEAIASNFLFINGTDSWATTYIENHDQARSITRFA
         * *                *   **              *     **** *   **
P_1UOK  NDG-MYRIESAKMLATVLHMMKGTPYIYQGEEIGMTNVRFESIDEYRDIETLNMY---KE
YEAST   DDSPKYRKISGKLLTLLECSLTGTLYVYQGQEIGQINFKEWPIEKYEDVDVKNNYEIIKK
         *   **  * * *        ** * *** ***  *     *  * *    * *   *
P_1UOK  KVMERGEDIEKVMQSIYIKGRDNARTPMQWD-DQNHAGFTTG--EPWITVNPNYKE-INV
YEAST   SFGKNSKEMKDFFKGIALLSRDHSRTPMPWTKDKPNAGFTGPDVKPWFLLNESFEQGINV
                    *    **  **** *  *   ****      **   *      ***
P_1UOK  KQAIQNKDSIFYYYKKLIELRKN-NEIVVYGSYDLILE-NNPSIFAYVRTYGVEKLLVIA
YEAST   EQESRDDDSVLNFWKRALQARKKYKELMIYGYDFQFIDLDSDQIFSFTKEYEDKTLFAAL
              **     *    **   *   **           **     *    *
P_1UOK  NFTAEECIFELPEDISYSEVELLIHNYDVENGPIENITLRPYEAMVFKLK
YEAST   NFSGEEIEFSLPR--EGASLSFILGNYD--DTDVSSRVLKPWEGRIYLVK
        **  **  * **             ***          * * *      *
```

Fig. 1. Sequence alignment result between the template protein (*B. cereus* oligo-1,6-glucosidase [PDBID:1UOK]) and the target protein (*S. cerevisiae* α-glucosidase represented as YEAST). The asterisks (*) indicate the conserved residues between the two proteins. The highlighted residues in red form the catalytic triad.

activity in both template as well as *S. cerevisiae* α-glucosidase. Two His residues near to active site which may be required for substrate binding are also conserved in the two proteins.

Fig. 1 shows the sequence alignment between the template protein (*B. cereus* oligo-1,6-glucosidase; PDBID:1UOK) and the target protein, *S. cerevisiae* α-glucosidase (represented as YEAST) by Align2d method of *INSIGHTII* package. In the recent publication, modeling was performed similarly with the same template but they have used ClustalW alignment unlike the present study [31]. The total sequence alignment is the same with minor differences in the position of gaps. The final homology modeling calculation achieved by *MODELLER* has generated a reliable 3D structure of *S. cerevisiae* α-glucosidase (Fig. 2). The PROCHECK, a protein structure validation program, predicted that 87.6% of residues of the 3D structure lied in most favored regions unlike the template protein which has 86.3% of residues. The root mean square deviation (RMSD) between the template and the target structures is 0.34 Å. The active site was defined by considering 10 Å around the pseudo-atom which was defined at the center of catalytic triad. The comparison between the template and homology model at the active site region revealed that 71% residues were identical. The Asp199, Glu255 and Asp329 residues form the catalytic triad in the template protein [18] while the Asp214, Glu276 and Asp349 form in *S. cerevisiae* α-glucosidase, respectively. Two more residues, His103 and His328 of oligo-1,6-glucosidase which may be involved in substrate binding are also conserved in α-glucosidase (His111 and His348, respectively) [17]. Some of the residues are not conserved at the binding pocket. As an example the Val200 in oligo-1,6-glucosidase which exists next to the catalytic Asp residue is replaced with Thr215 in *S. cerevisiae* α-glucosidase. This residue is important for substrate specificity of α-glucosidases which is proven by experimental studies [32].

3.2. Molecular docking and validation

Docking studies were performed to gain insight into the most probable binding conformation of sulfonamide chalcone derivatives and to compare with the other known sugar and non-sugar derivatives which are in clinical use/trails (Fig. 3). Compounds **1** and **2** (acarbose and voglibose, respectively) are known as sugar derivatives and are extensively studied by experiments. Compounds **3** and **4** are sulfonamide chalcone derivatives discovered recently as potent, non-sugar α-glucosidase inhibitors with IC50 values 0.4 and 0.98 μM, respectively [17]. Compounds **5**–**7** are iminocyclitol derivatives, selective and potent glycosidase inhibitors with experimental activity in micromolar range (0.15, 0.28 and 1.2 μM, respectively) [16]. Compounds **1**, **2**, **3** and **5** were used for the docking experiments and then for the comparison (Fig. 4a–d). The docking scores evaluated from *GOLD* for the four compounds are 67.03, 57.78, 62.31 and 58.85 respectively. Acarbose showed a higher dock score, very likely due to the high number of hydroxyl groups as well as hydrophobic interactions. As acarbose resembles the substrate it may bind selectively with the active site residues. The binding mode analyses of the acarbose with the active site residues provided important information of catalytic site. The NH group of the acarbose formed a H-bond with one of the catalytic triad residue, Asp349 (Fig. 4a). The hydroxyl groups in the terminal ring formed H-bonds with Arg212 and His348 residues. This is compatible with the previous experimental and theoretical results [33,34].

For validating the docking simulation protocol, the docked conformation of acarbose is compared with the crystal structure of *Thermotoga maritima* 4-α-glucanotransferase (PDBID:1LWJ) [33] by aligning their structures (Fig. 10 in supplementary data). In the crystal structure the pseudoacarbose

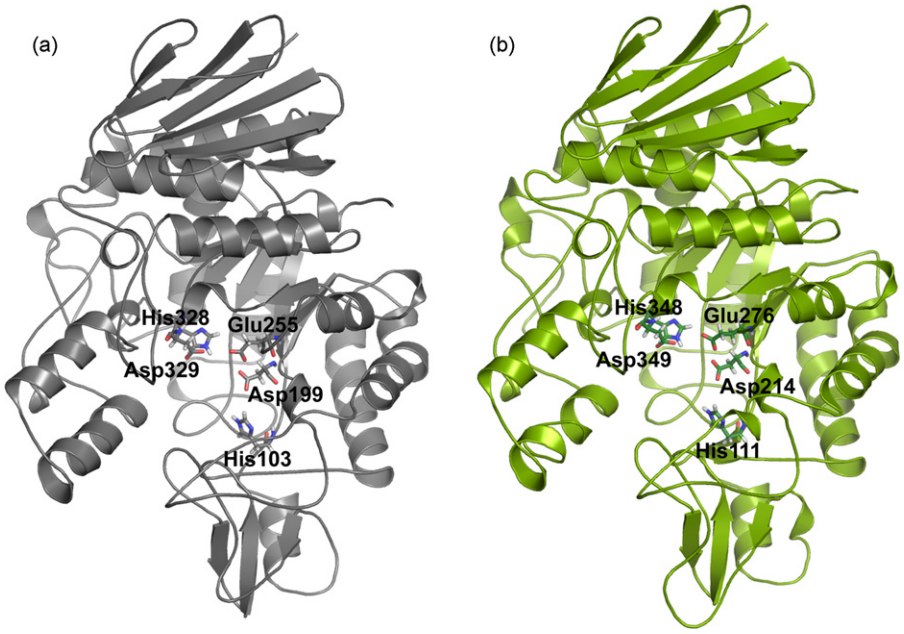

Fig. 2. Ribbon diagrams of the template structure of *B. cereus* oligo-1,6-glucosidase [PDBID:1UOK] (a), the homology modeled structure *S. cerevisiae* α-glucosidase (b) with the conserved catalytic residues displayed as sticks.

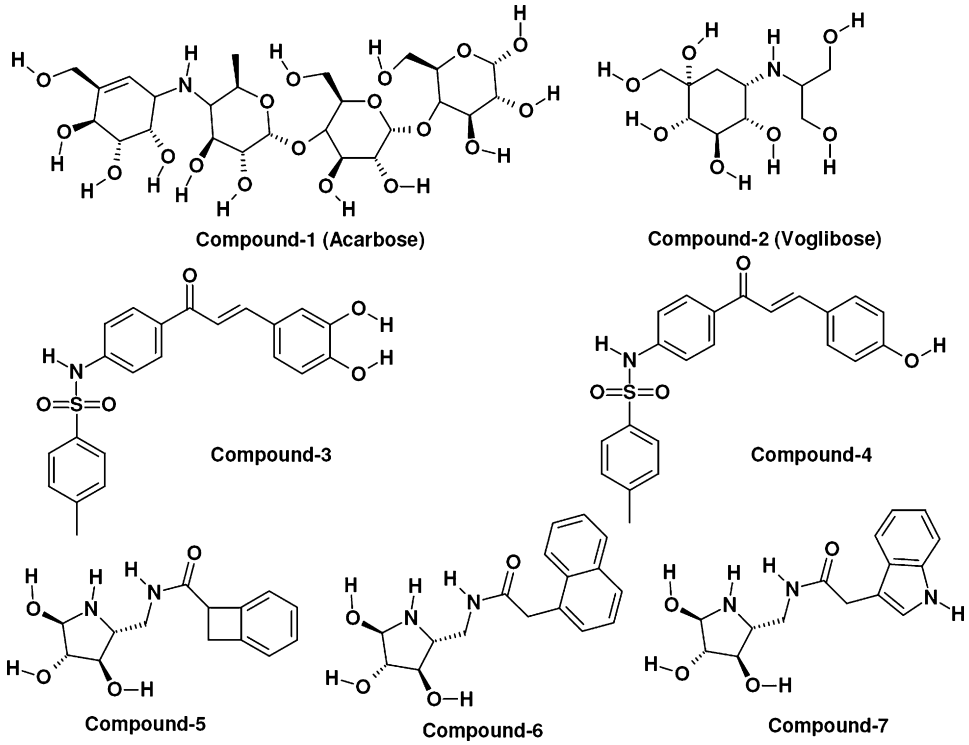

Fig. 3. Molecular structures of sugar derivatives (compounds **1** and **2**), sulfonamide chalcone derivatives (compounds **3** and **4**) and iminocyclitol derivatives (compounds **5**–**7**).

interacts similarly with the corresponding active site residues including a H-bond between its amino group and Asp278. This validation has proved that the docking protocol was reasonable in identifying the binding conformation accurately. Voglibose has also showed similar type of interactions (Fig. 4b). The NH group formed a H-bond with Asp349 and the hydroxyl groups formed with His348 and Arg212. The Arg439 in the active site also formed another H-bond with one of the

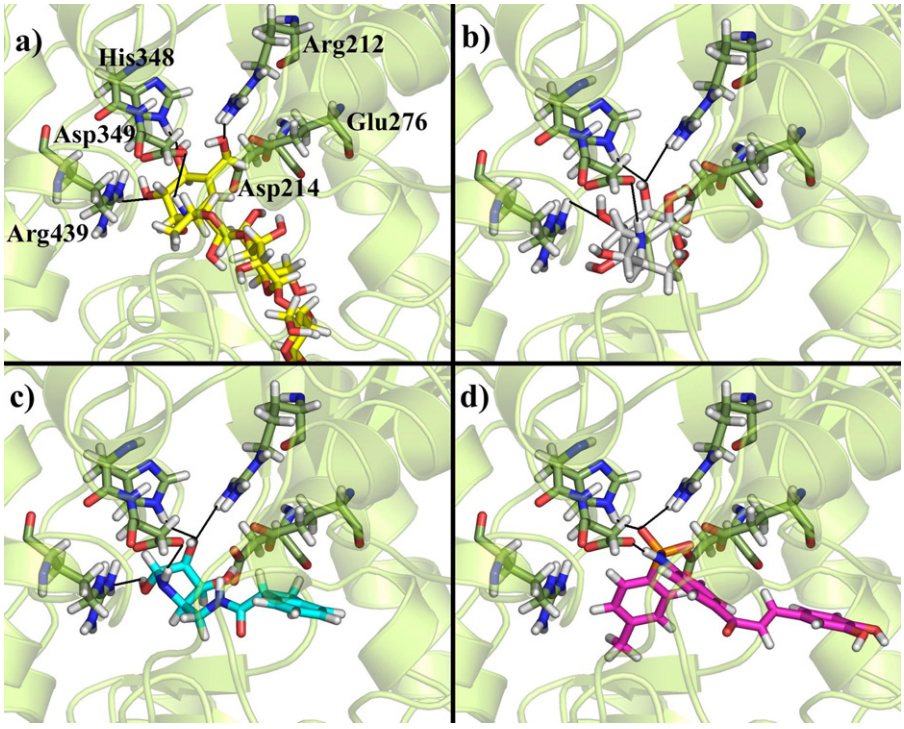

Fig. 4. The molecular docking results. The binding conformations of acarbose (a), voglibose (b), compound **5** (c), and compound **3** (d) interacting with the active site residues of the modeled *S. cerevisiae* α-glucosidase structure.

hydroxyl group of glycosidic ring. Fig. 4c represents binding pattern of iminocyclitol derivative (compound **5**), in which NH present in the iminocyclitol ring formed H-bonding with Asp349 and the hydroxyl groups form with His348 and Arg212. The Arg439 similarly formed another H-bond with one of the hydroxyl group of iminocyclitol ring. In a similar manner sulfonamide chalcones were also able to bind with the active site amino acids. Analyses of interactions between compound **3** and the active site revealed that NH group of sulfonamide interacts through H-bonding with Asp349. The two oxygens in SO_2 group formed H-bonds with His348 and Arg212 (Fig. 4d). Based on the binding conformation observations, the Arg212 and Arg439 which are forming H-bonding with hydroxyl groups of terminal rings may have important role either in catalytic mechanism or in substrate binding.

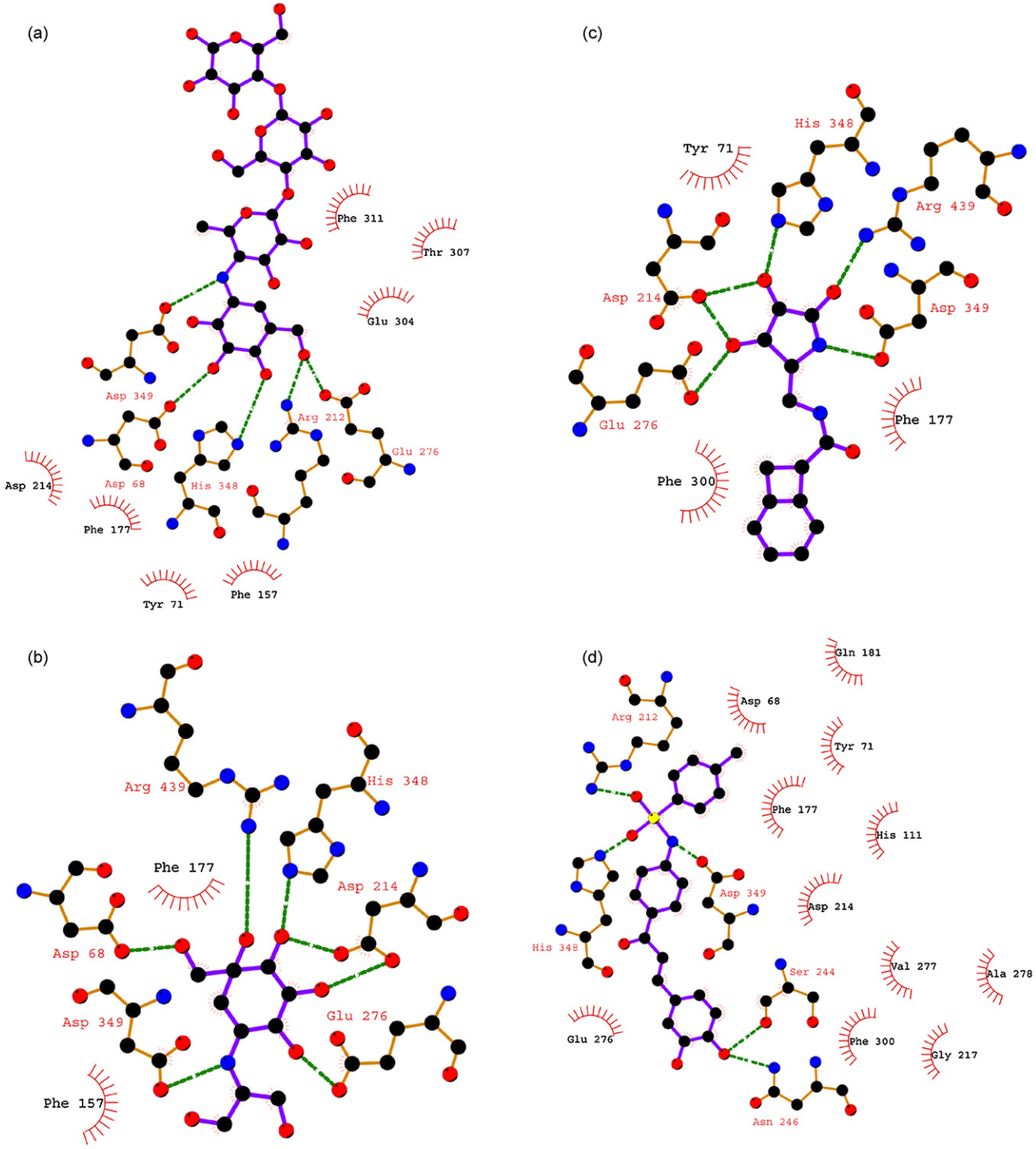

Fig. 5. Ligplot analyses results. 2D representation of ligand–protein interactions were analyzed between acarbose (a), voglibose (b), compound **5** (c), compound **3** (d) and *S. cerevisiae* α-glucosidase. Hydrogen bond forming residues were shown in lines with hydrogen bonds shown as dotted lines and residues interacting by hydrophobic interactions were represented as lines in red.

The Ligplot analyses were introduced to understand the indepth interaction pattern between the docked ligands and the active site residues. Ligplot is an essential tool to understand hydrophobic interactions as well as hydrogen bonding pattern [35]. Fig. 5 shows the Ligplot analyses for the four compounds which were discussed previously in docking simulations. The same H-bond interactions as seen in docking results were obtained for the four compounds. The two residues Tyr71 and Phe177 which are proved to hold the terminal ring to be cleaved in the catalytic reaction by forming a hydrophobic patch are also observed to be present for the four compounds [33,34]. The Phe300 is also observed to be forming hydrophobic interactions. In the case of acarbose, it can be seen that a hydrophobic patch comprising of Tyr71, Phe177 along with Phe157 surround and hold the terminal ring of acarbose. Thus, Ligplot analyses were especially useful in knowing the hydrophobic interaction pattern.

3.3. MD simulation studies

MD simulations were performed (i) to determine the stability of the homology modeled structure of *S. cerevisiae* α-glucosidase and (ii) to prove the binding mechanism as shown in the docking results. The trajectories were stable during the whole production part of the 3 ns MD simulation run. The trajectory stability was monitored and confirmed by the analysis of backbone RMSD for both the systems, protein with and without compound **3**. The average RMSD when measured from 500 ps for protein and protein with compound **3** are 0.23 and 0.21 nm, respectively (Fig. 6a). The stability of the simulations can be further viewed by analyzing the total energy as a function of time. The total energy (kJ/mol) of the protein with compound **3** is relatively lower (i.e. more stable) when compared with that of protein alone throughout the simulation (Fig. 6b). The total number of hydrogen bonds between the compound **3** and the protein was measured for all snapshots of the simulation (Fig. 6c). The average number from 500 ps to the end is around 3.5 which indicate that the compound **3** is stably forming H-bonds in the active site. The major contributors of these are the NH and SO_2 group as previously shown in the docking studies. The NH forms H-bond with the carboxyl group of Asp349 while the SO_2 forms with Arg212 and His348/His111.

Distance between the NH of the compound **3** and the two oxygen atoms of the carboxyl group of the key catalytic residue, Asp349 were measured throughout the simulation time (Fig. 6d). Initially the OD2 was near but within 50–100 ps OD1 interacted with NH and the distance was maintained around 0.28 nm for the entire simulation. The NH of compound **3** plays an important role as an H-bond donor similar to that present in acarbose and voglibose which are shown in docking results. Docking studies showed that the His348 interacts with SO_2 group while the MD simulation study revealed that either His348 or His111 can interact with SO_2. Both the residues were proved to be crucial for substrate binding. This can be explained by the presence of hydroxyl groups which can reach both the histidines in acarbose or pseudoacarbose [33,34]. In the case of compound **3** SO_2 contacts any one of the histidines. This can be viewed by measuring the distance between SO_2 and His348/His111 (Fig. 7a and b). The compound **3** is also stabilized by hydrophobic interactions on the two farthest ends. On one end the 3,4-dihydroxy phenyl ring is making stacking interaction with Phe300 as observed by their orientation with an average distance of 0.3 nm (Fig. 7a and c). This can also be viewed in the 2D Ligplot analyses for acarbose and compound **3**. Thus, Phe300 may play a key role in orientating the ligand at the active site by hydrophobic interactions. Similarly on the other end the *p*-toulene group is surrounded by four hydrophobic residues, Tyr71, Phe157, Phe158 and Phe177 (Fig. 7d). Among them, Tyr71 and Phe177 were experimentally determined to be crucial for the terminal ring recognition of substrates by α-glucosidase [33]. Minimum distance between the four residues and compound **3** was measured and were around 0.35 nm for the total simulation time (Fig. 7e). Their relative orientation

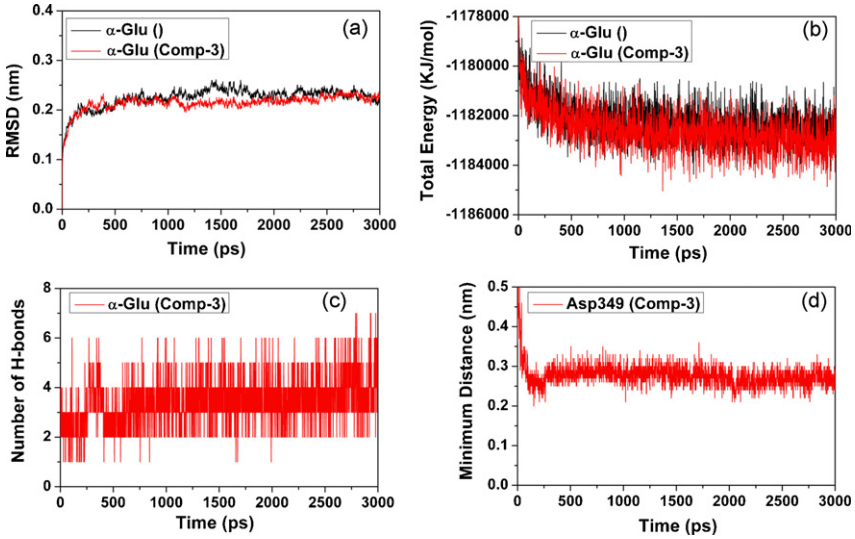

Fig. 6. MD simulation results. RMSD (a), total energy (b), total number of H-bonds between compound **3** and protein (c) and the distance between Asp349 and compound **3** (d) for 3 ns MD with reference to initial structure.

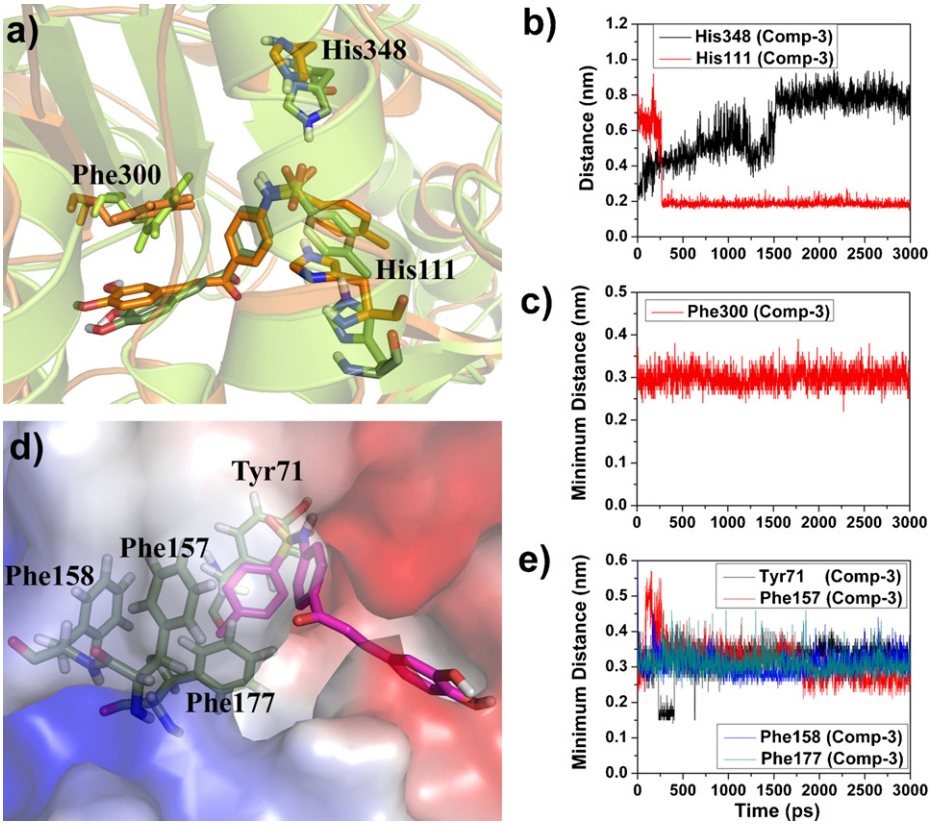

Fig. 7. Initial (green) and final (orange) snapshots of 3 ns MD simulation were superimposed to show the exchange of histidines in interacting with compound **3** and the position of Phe300 (a). The distances between compound **3** and the histidines (b) and Phe300 (c) were measured. Electrostatic potential surface map of *S. cerevisiae* α-glucosidase displaying the role of the four hydrophobic residues in holding the terminal ring of compound **3** (d). Minimum distance between four hydrophobic residues and compound **3** (e).

around the *p*-toulene group also makes it plausible to propose that the Phe157 and Phe158 may have some important role in surrounding the terminal ring of the ligand/substrate for the glucosidase activity.

3.4. Pharmacophore model generation and validation

From the results of molecular docking, Ligplot and MD simulation analyses a clear idea about the critical groups responsible for binding between ligand and the protein active site were obtained. This knowledge can be applied in generating more reasonable 3D-pharmacophore models from the existing non-sugar derivatives for the purpose of virtual screening. The compounds **3**–**7** were taken as a training set to generate pharmacophore models by the *HipHop* methodology. The compound **3**, reported earlier was considered as reference compound, which was allowed to map all features, while the other four molecules were allowed to map partially on the hypothesis. Finally the 10 five-feature hypotheses were generated. The ranking scores of hypotheses range from 67.55 to 68.45 as shown in Table 1. The first hypothesis (Hypo1) was selected for further analysis because of its best score. Direct hit mask and partial hit mask revealed that all the five features namely three hydrogen bond acceptors (HBA), one hydrogen bond donor (HBD) and one ring aromatic (RA) were mapped onto five training set compounds. Fig. 8 shows the mapping of Hypo1 onto the compounds **3**–**6**. All the features mapped well on compounds with reasonable fit values (3.5–5). The compound **3** taken as reference mapped well with all features with a fit value 5. The HBD feature mapped on NH group of the compound **3** represents the key interaction with

Table 1
Results of pharmacophore hypotheses generated by *catalyst/HipHop*

Hypotheses[a]	Features[b]	Rank[c]	Direct hit mask[d]	Partial hit mask[e]
Hypo1	RDAAA	68.4501	11111	00000
Hypo2	RDAAA	68.4501	11111	00000
Hypo3	RDAAA	68.4501	11111	00000
Hypo4	RDAAA	67.9125	11111	00000
Hypo5	RDAAA	67.9125	11111	00000
Hypo6	RDAAA	67.9125	11111	00000
Hypo7	RDAAA	67.9125	11111	00000
Hypo8	RDAAA	67.9125	11111	00000
Hypo9	RDAAA	67.5495	11111	00000
Hypo10	RDAAA	67.5495	11111	00000

[a] Numbers for the hypothesis are the numeration as obtained by the hypothesis generation.
[b] R: ring aromatic; D: H-bond donor; A: H-bond acceptor.
[c] The ranking score for each hypothesis. Best hypotheses have highest ranks.
[d] Mapping of each feature onto training set molecule. 1 means yes and 0 means no. For our 10 hypotheses, all 5 compounds mapped all hypothesis features.
[e] A training set molecule mapped all but one feature in the hypothesis. 1 means yes and 0 means no.

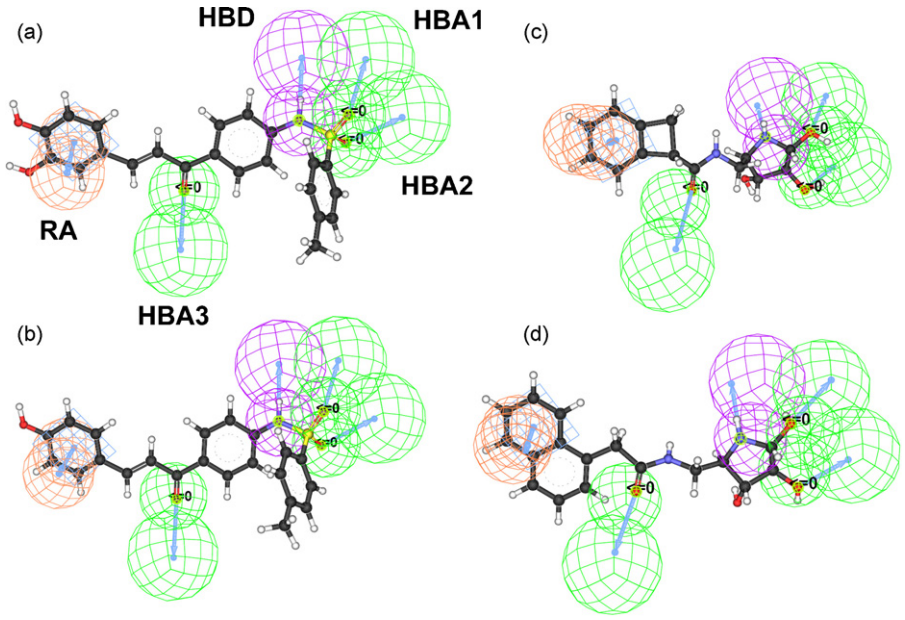

Fig. 8. Mapping of top scoring Hypo1 on training set compound **3** (a), compound **4** (b), compound **5** (c), and compound **6** (d).

Asp349 residue of α-glucosidase while two HBA features mapped on SO_2 group represents H-bond interactions with Arg212 and His111/His348. The RA feature maps on the dihyroxy phenyl group, which has shown hydrophobic interactions with Phe300. The five featured pharmacophore model is in agreement with the docking and MD simulation results.

For validation purpose, Hypo1 was mapped onto two entirely different set of compounds (i) sugar derivatives,

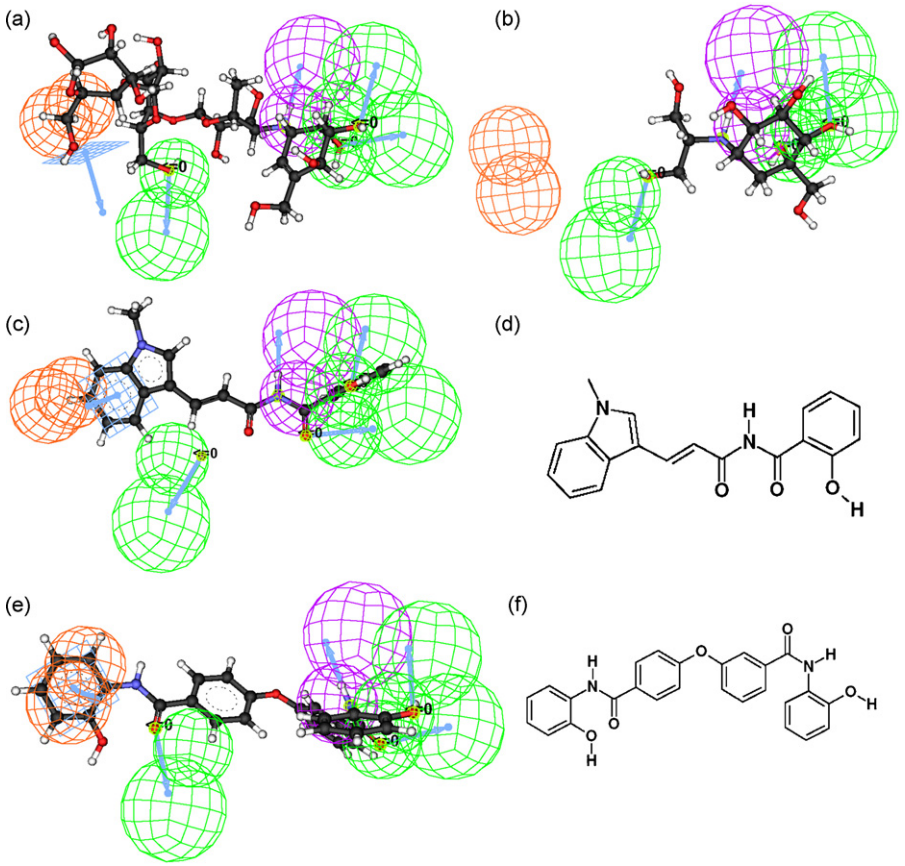

Fig. 9. Validation of Hypo1 by mapping of acarbose (a), voglibose (b) compound **t1** (c), compound **t2** (e) by Hypo1 and the chemical structures of **t1** (d) and **t2** (f).

acarbose and voglibose, and (ii) non-sugar derivatives. All the features except for RA feature were mapped perfectly on both the sugar derivatives (Fig. 9a and b). As there were no aromatic rings in these two compounds, the respective features were not mapped. The HBD feature mapped on the NH group of acarbose and the two HBAs mapped onto two hydroxyl groups of the terminal ring. Second validation of Hypo1 was performed by mapping them onto two non-sugar derivatives (compounds **t1** and **t2**) which were retrieved by virtual docking methodology in a recent publication. They were tested for inhibitory activity against the α-glucosidase from baker's yeast by in vitro enzyme assay and revealed more than 50% inhibition at the concentration of 50 μM [31]. The molecular structures of the two compounds and their mapping by Hypo1 are shown in Fig. 9d, f, c and e, respectively. The Hypo1 mapped four features for compound **t1** and five features for compound **t2** thereby validating our pharmacophore model. Based on these two validations, our pharmacophore model is being used for further studies.

4. Conclusions

Our aim was (i) to find out the binding mechanism of the sugar/non-sugar derivatives to represent the critical features responsible for α-glucosidase inhibitory activity and (ii) to develop a pharmacophore model based on the knowledge from the receptor based analyses. We firstly built the homology modeled structure of *S. cerevisiae* α-glucosidase and used it for the molecular docking and MD simulation studies to determine the interaction mode of the sulfonamide chalcone and sugar derivatives with the protein. The results projected the crucial role of NH group in the binding of sugar/non-sugar derivatives with the active site. The residues Phe157 and Phe158 other than experimentally proved Tyr71, Phe177 also make the hydrophobic patch for holding the terminal glycosidic ring to be cleaved and released as free glucose. The Phe300 may also play a key role in orientating the ligand at the active site by hydrophobic interactions. These results revealed that sulfonamide chalcone derivatives which form a potent non-sugar class of α-glucosidase inhibitors, bind to active site residues in similar manner with the known compounds like the acarbose and the voglibose. With the knowledge of the key interactions between the ligand and protein resulted from the molecular docking and the MD simulation studies, the five featured pharmacophore model was developed from non-sugar derivatives only. The model was then validated by mapping onto the sugar derivatives and newly tested non-sugar derivatives. Interestingly it has covered all the critical features required for the ligand binding. Therefore, our pharmacophore model could be very useful for the virtual screening in the development of new non-sugar α-glucosidase inhibitors.

Acknowledgments

Kavitha Bharatham and Nagakumar Bharatham were recipients of fellowships from the BK21 Programs and this work was supported by grants from the MOST/KOSEF for the Environmental Biotechnology National Core Research Center (grant # R15-2003-012-02001-0).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2007.11.002.

References

[1] H.A. Ernst, L. Lo Leggio, M. Willemoes, G. Leonard, P. Blum, S.J. Larsen, Structure of the *Sulfolobus solfataricus* alpha-glucosidase: implications for domain conservation and substrate recognition in GH31, Mol. Biol. 358 (2006) 1106–1124.
[2] L. Ellgaard, A. Helenius, Quality control in the endoplasmic reticulum, Nat. Rev. Mol. Cell Biol. 4 (2003) 181–191.
[3] N. Asano, Glycosidase inhibitors: update and perspectives on practical use, Glycobiology 13 (2003) 93R–104R.
[4] R.A. Dwek, T.D. Butters, F.M. Platt, N. Zitzmann, Targeting glycosylation as a therapeutic approach, Nat. Rev. Drug Discov. 1 (2002) 65–75.
[5] G.S. Jacob, Glycosylation inhibitors in biology and medicine, Curr. Opin. Struct. Biol. 5 (1995) 605–611.
[6] D. Einfeld, Maturation and assembly of retroviral glycoproteins, Curr. Top. Microbiol. Immunol. 214 (1996) 133–176.
[7] R. Wyatt, J. Sodroski, The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens, Science 280 (1998) 1884–1888.
[8] M.A. Noor, R.A. Parker, E. O'Mara, D.M. Grasela, A. Currie, S.L. Hodder, F.T. Fiedorek, D.W. Haas, The effects of HIV protease inhibitors atazanavir and lopinavir/ritonavir on insulin sensitivity in HIV-seronegative healthy adults, AIDS 18 (2004) 2137–2144.
[9] T.M. Dando, L.J. Scott, Abacavir plus lamivudine: a review of their combined use in the management of HIV infection, Drugs 65 (2005) 285–302.
[10] A.S. Veiga, N.C. Santos, L.M. Loura, A. Fedorov, M.A. Castanho, HIV fusion inhibitor peptide T-1249 is able to insert or adsorb to lipidic bilayers, Putative correlation with improved efficiency, J. Am. Chem. Soc. 126 (2004) 14758–14763.
[11] K. Princen, S. Hatse, K. Vermeire, S. Aquaro, E. De Clercq, L.O. Gerlach, M. Rosenkilde, T.W. Schwartz, R. Skerlj, G. Bridger, D. Schols, Inhibition of human immunodeficiency virus replication by a dual CCR5/CXCR4 antagonist, J. Virol. 78 (2004) 12996–13006.
[12] A. Mehta, N. Zitzmann, P.M. Rudd, T.M. Block, R.A. Dwek, Alpha-glucosidase inhibitors as potential broad based anti-viral agents, FEBS Lett. 430 (1998) 17–22.
[13] A.Y. Cheng, I.G. Fantus, Oral antihyperglycemic therapy for type 2 diabetes mellitus, CMAJ 172 (2005) 213–226.
[14] H.E. Lebovitz, Alpha-glucosidase inhibitors, Endocrinol. Metab. Clin. N. Am. 26 (1997) 539–551.
[15] N. Asano, M. Nishida, A. Kato, H. Kizu, K. Matsui, Y. Shimada, T. Itoh, M. Baba, A.A. Watson, R.J. Nash, P.M. Lilley, D.J. Watkin, G.W. Fleet, Homonojirimycin isomers and N-alkylated homonojirimycins: structural and conformational basis of inhibition of glycosidases, J. Med. Chem. 41 (1998) 2565–2571.
[16] P.H. Liang, W.C. Cheng, Y.L. Lee, H.P. Yu, Y.T. Wu, Y.L. Lin, C.H. Wong, Novel five-membered iminocyclitol derivatives as selective and potent glycosidase inhibitors: new structures for antivirals and osteoarthritis, ChemBioChem 7 (2006) 165–173.
[17] W.D. Seo, J.H. Kim, J.E. Kang, H.W. Ryu, M.J. Curtis-Long, H.S. Lee, M.S. Yang, K.H. Park, Sulfonamide chalcone as a new class of alpha-glucosidase inhibitors, Bioorg. Med. Chem. Lett. 15 (2005) 5514–5516.
[18] K. Watanabe, Y. Hata, H. Kizaki, Y. Katsube, Y. Suzuki, The refined crystal structure of Bacillus cereus oligo-1,6-glucosidase at 2.0 A resolution: structural characterization of proline-substitution sites for protein thermostabilization, J. Mol. Biol. 269 (1997) 142–153.

[19] N. Bharatham, K. Bharatham, K.W. Lee, Pharmacophore identification and virtual screening for methionyl-tRNA synthetase inhibitors, J. Mol. Graph. Model. 25 (2007) 813–823.
[20] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[21] A. Fiser, R.K. Do, A. Sali, Modeling of loops in protein structures, Protein Sci. 9 (2000) 1753–1773.
[22] InsightII, Version 2005.3L, Accelrys Inc., San Diego, 2005, www.accelrys.com.
[23] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[24] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.
[25] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comp. Phys. Commun. 91 (1995) 43–56.
[26] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A.L.T.M. Sijbers, K.A. Feenstra, R. van Drunen, H.J.C. Berendsen, Gromacs User Manual version 3.3, 2005, www.gromacs.org.
[27] R. Dayam, F. Aiello, J. Deng, Y. Wu, A. Garofalo, X. Chen, N. Neamati, Discovery of small molecule integrin alphavbeta3 antagonists as novel anticancer agents, J. Med. Chem. 49 (2006) 4526–4534.
[28] CATALYST 4.11 User Guide, Accelrys Inc., San Diego, CA, USA, 2005.
[29] A. Smellie, S.D. Kahn, S.L. Teig, Analysis of conformational coverage. 1. Validation and estimation of coverage, J. Chem. Inf. Comput. Sci. 35 (1995) 285–294.
[30] A. Smellie, S.D. Kahn, S.L. Teig, Analysis of conformational space. 2. Applications of conformational models, J. Chem. Inf. Comput. Sci. 35 (1995) 295–304.
[31] H. Park, K.Y. Hwang, K.H. Oh, Y.H. Kim, J.Y. Lee, K. Kim, Discovery of novel α-glucosidase inhibitors based on the virtual screening with the homology-modeled protein structure, Bioorg. Med. Chem. 16 (2008) 284–292.
[32] K. Yamamoto, A. Nakayama, Y. Yamamoto, S. Tabata, Val216 decides the substrate specificity of alpha-glucosidase in *Saccharomyces cerevisiae*, Eur. J. Biochem. 271 (2004) 3414–3420.
[33] A. Roujeinikova, C. Raasch, S. Sedelnikova, W. Liebl, D.W. Rice, Crystal structure of Thermotoga maritima 4-alpha-glucanotransferase and its acarbose complex: implications for substrate specificity and catalysis, J. Mol. Biol. 321 (2002) 149–162.
[34] C.H. Tomich, P. da Silva, I. Carvalho, C.A. Taft, Homology modeling and molecular interaction field studies of alpha-glucosidases as a guide to structure-based design of novel proposed anti-HIV inhibitors, J. Comput. Aided Mol. Des. 19 (2005) 83–92.
[35] A.C. Wallace, R.A. Laskowski, J.M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions, Protein Eng. 8 (1995) 127–134.