Bioorganic & Medicinal Chemistry Letters 20 (2010) 4420–4423

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Docking and SAR studies of salacinol derivatives as α-glucosidase inhibitors

Shinya Nakamura [a,*], Kazunori Takahira [a], Genzoh Tanabe [a], Toshio Morikawa [b], Mika Sakano [a], Kiyofumi Ninomiya [b], Masayuki Yoshikawa [c], Osamu Muraoka [a], Isao Nakanishi [a]

[a] Faculty of Pharmacy, Kinki University, 3-4-1 Kowakae, Higashi-osaka, Osaka 577-8502, Japan
[b] Pharmaceutical Research Institute, Kinki University, 3-4-1 Kowakae, Higashi-osaka, Osaka 577-8502, Japan
[c] Kyoto Pharmaceutical University, 1 Shichono-cho, Misasagi, Yamashina-ku, Kyoto 607-8412, Japan

ARTICLE INFO

Article history:
Received 26 April 2010
Revised 3 June 2010
Accepted 9 June 2010
Available online 12 June 2010

Keywords:
Docking study
Structure–activity relationship
α-Glucosidase
Antidiabetic drug

ABSTRACT

Salacinol is a potent α-glucosidase inhibitor isolated from *Salacia reticulata*, and a good lead compound for an antidiabetic drug. It is essential to clarify the binding state of salacinol to α-glucosidase for efficient optimization study using structure-based drug design. Redocking simulations of two inhibitors, acarbose and casuarine whose complex structures are known, were performed to assure the appropriate docking pose prediction. The simulation reproduced both experimental binding states with accuracy. Then, using the same simulation protocol, the binding mode of salacinol and its derivatives has been predicted. Salacinol bound to the protein with a similar binding mode as casuarine, and the predicted structures could explain most of the structure–activity relationships of salacinol derivatives.

The roots and stems of certain types of plants belonging to the genus Hippocrateaceae (e.g., *Salacia reticulata*, *Salacia oblonga*, and *Salacia prinoides*), found in the forests of India and Sri Lanka, have been known to consist an antidiabetic medicine as per the Indian medical lore called Ayurveda. As the active ingredients in sugar absorption control based on α-glucosidase inhibition, some compounds such as salacinol[1] (**1**) and kotalanol[2] (**2**) have been extracted and isolated from these plants (Fig. 1).

These compounds have unique structures, which are composed of cyclic thiosugar sulfonium and a sulfate group in the side chain. Interestingly, salacinol formed an intramolecular salt bridge between the sulfonium center and sulfate anion in the isolated state.[1b] On the contrary, voglibose (**3**) and acarbose (**4**), the current antidiabetic medicines in the market, are amino sugars that possibly form a tertiary ammonium ion moieties in the binding state.

Recently complex structures of acarbose and casuarine (**5**) (Fig. 1) with α-glucosidase have been reported[3,4] and are available in the Protein Data Bank archive (2QMJ and 3CTT, respectively). They both possess an ammonium ion moiety of acarbose and casuarine and form a salt bridge with different amino acid residues, while hydroxy groups interact with the same amino acid residues, as shown in Figure 2. Therefore, it is interesting to know with which amino acid residue the sulfonium ion part of salacinol or kotalanol form the salt bridge, and if hydroxy groups makes hydrogen bonds with the same residues as acarbose and casuarine. Furthermore, it is very important to reveal if the sulfonic cation part forming the intramolecular salt bridge in the binding state from the viewpoint of drug design. In this study, we predicted the binding mode of salacinol with the N-terminal catalytic domain of maltase–glucoamylase (NtMGAM) by docking simulation and examined the ability of the predicted binding mode to explain the structure–activity relationship data of salacinol derivatives and a newly designed compound.

Molecular modeling and visualization were applied using MOE2008[5] and PYMOL0.99,[6] and docking simulations were performed using ASEDOCK[7] implemented in MOE. The atomic coordinates of NtMGAM (3CTT) were employed for the target protein structure. All

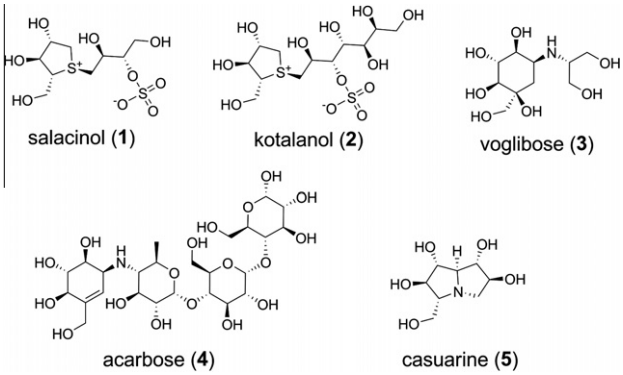


Figure 1. α-Glucosidase inhibitors.

* Corresponding author. Tel.: +81 6 6721 2332; fax: +81 6 6721 2505.
E-mail address: nakas@phar.kindai.ac.jp (S. Nakamura).

0960-894X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2010.06.059

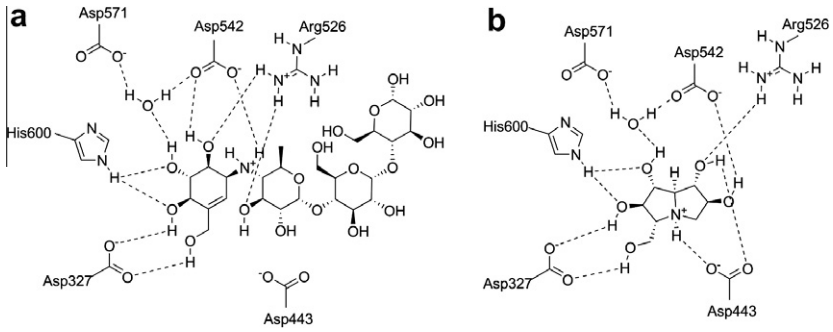
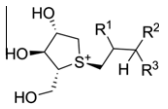


Figure 2. Schematic interaction diagrams between NtMGAM and known inhibitors; (a) acarbose, and (b) casuarine.

Table 1
The chemical structures and potency of salacinol derivatives



	R^1	R^2	R^3	pIC_{50}[a]
Salacinol (**1**)	(*S*)-OH	(*S*)-OSO_3^-	$-CH_2OH$	5.28
Neosalacinol (**6**)	(*S*)-OH	(*S*)-OH	$-CH_2OH$	5.10
Salaprinol (**7**)	(*S*)-OH	-OSO_3^-	–H	ca. 3
8	–H	(*S*)-OSO_3^-	$-CH_2OH$	<3
9	–H	-OSO_3^-	–H	<3
10	(*R*)-OH	(*R*)-OSO_3^-	$-CH_2OH$	<3
11	(*R*)-OH	(*R*)-OH	$-CH_2OH$	<3
12	(*S*)-OH	–H	$-CH_2OH$	3.94

[a] Experimental value as the log scale of the concentration (*M*) for 50% inhibition to maltase.[9]

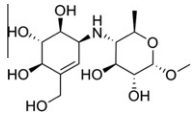


Figure 3. Model compound of acarbose for docking simulation.

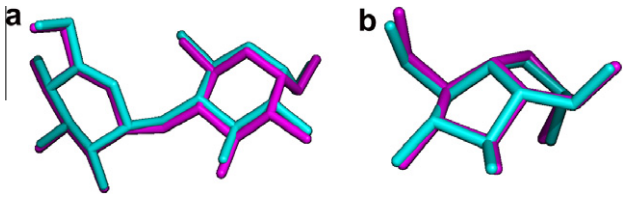


Figure 4. Redocking results (cyan: X-ray structure, magenta: docked structure). Hydrogen atoms are not shown for clarity; (a) acarbose model, and (b) casuarine.

water molecules in the experimental structure were removed. Hydrogen atoms were added using MOE. The protonation states of the amino acid residues of NtMGAM and the direction of the hydrogen atoms associated with the hydrogen bonds were assigned using the Protonate3D algorithm.[8] Ligand molecules such as salacinol (**1**) and its derivatives (**6**–**12** in Table 1),[9,10] were modeled using MOE, and the structures were optimized using the MMFF94x forcefield.[11] We used a truncated model for acarbose in which two sugar units were removed from acarbose as shown in Figure 3, because it is known that the electron densities of two sugar units occupying the outside of the binding pocket were ambiguous from the X-ray crystal structure analysis of the acarbose–NtMGAM complex.[3]

The initial docked poses (50 candidate poses for each compound) were optimized by the MMFF94x forcefield. To consider induced fit, both ligand and protein structures were flexibly treated, tethering on the initial structure during optimization. After optimization, binding energies were evaluated using the U_{dock} score consisting of the protein–ligand interaction energy with conformational strain energy. Then, the top ranked pose was chosen as the predicted binding mode.

The results of the use of acarbose and casuarine for docking simulation are shown in Figure 4. The RMSD values excluding hydrogen atoms between experimental and simulated structures were 0.7 Å and 0.2 Å for the acarbose model and casuarine, respectively. These values were small enough and supported the hypothesis that experimental binding modes could be reproduced with accuracy using this simulation protocol.

The salacinol binding state was explored using this docking protocol. The predicted docking mode for salacinol is shown in Figure 5. The binding mode of the thiosugar ring moiety is almost the same as that for casuarine. All hydroxy groups formed hydrogen bonds with NtMGAM. Three groups on the ring interact with Asp327 and His600 as well as casuarine and additionally interact with Asp542. These amino acid residues are located in the interior of the binding pocket, and such hydrogen bonds appear to be very important for specific molecular recognition and affinity.

In the predicted binding mode, the C2 hydroxy group of salacinol interacts with Asp542 directly by a hydrogen bond. On the other hand, casuarine interacts with Asp542 through an intermediating water molecule as shown in Figures 2b and 6a. This water molecule was omitted in docking simulation to explore the docking pose as a wide binding space and not to exclude the possibility of direct hydrogen bonding. In Figure 6b, this water molecule was overlaid with the predicted salacinol binding mode. We found that there is enough space for one water molecule to exist. Salacinol may also interact with Asp542 similar to casuarine via a water molecule.

Two hydroxy groups at C2′ and C4′ on the methylene chain interacted with Asp203 and Arg526 by hydrogen bonding. These interactions are considered to strengthen the binding affinity. An interesting finding on the contrary is that the sulfate group did not interact significantly with the protein (Fig. 5). It existed at the entrance of binding cavity, and did not form direct polar interaction with protein. This negative charged group may not locate stably at the binding pocket where many acidic residues exist. On the other hand, the sulfonium cation interacts with Asp443 (Fig. 5a and b) by the strong electrostatic interaction similar to

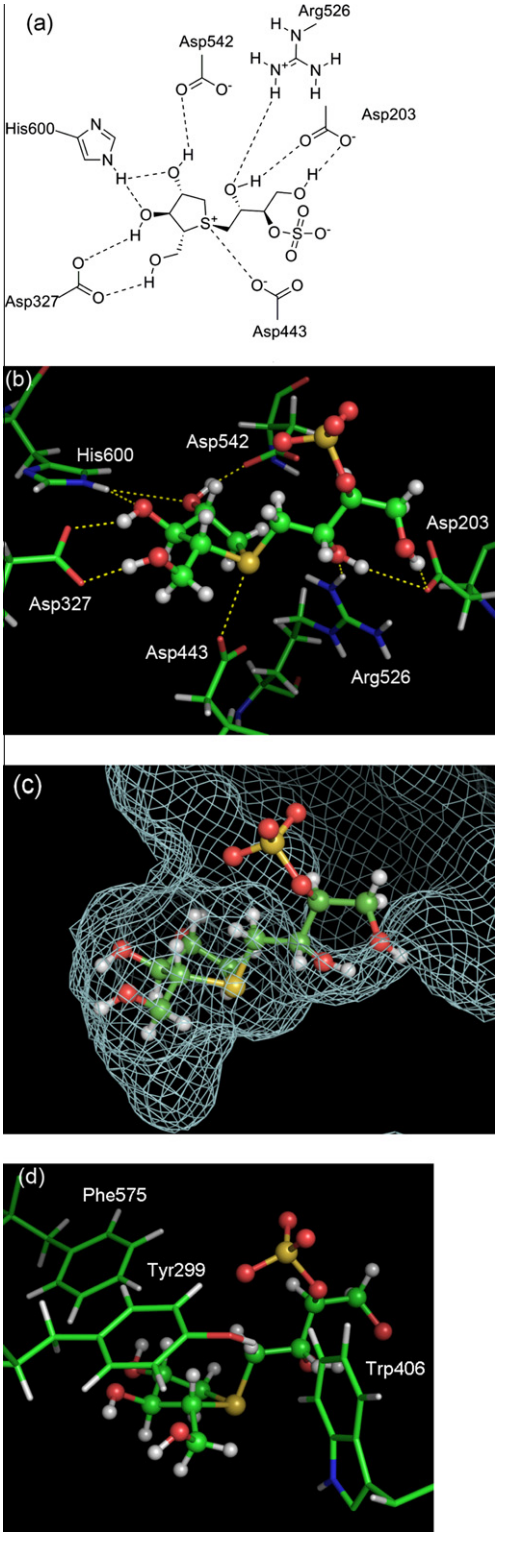


Figure 5. Predicted binding mode of salacinol. (a) Schematic interaction diagram. (b) The binding structure of salacinol (ball-and-stick) with neighboring amino acid residues. Hydrogen bonds are depicted by dashed lines. (c) Connoly surface of the binding pocket overlaid with docked salacinol. (d) The amino acid residues around the sulfate group at the entrance of the binding pocket.

casuarine; the distance between the cationic sulfur atom and the Asp443-Oδ atom is 2.8 Å. In the crystallographic analysis of the isolated salacinol molecule, it is reported that salacinol has an intramolecular salt bridge between the sulfonium cation and the sulfate anion in the crystal state. If the intramolecular salt bridge is broken during ligand binding, the salacinol molecule must lose large deformation energy to position itself stably. On that account, it is believed that salacinol does not possess an intramolecular salt bridge in a polar medium and both ionic parts interact with solvent molecules. Thus, the entire complex was stabilized by multiple hydrogen bonds and an intermolecular salt bridge.

The predicted binding mode could explain most of the structure–activity relationships of salacinol derivatives listed in Table 1. The R^1 hydroxy group at the C2′ position exists inside the binging pocket and is recognized by Arg526 and Asp203 through hydrogen bonds (Fig. 5a). Since this group contributes greatly to binding affinity, it is predicted that the lack of this hydroxy group weakens the affinity. Indeed, both **8** and **9**, which do not possess the R^1 hydroxy group, do not have inhibition activity. Also, the inversion of this group, as with **1**–**10** and **6**–**11**, reduces the pIC_{50} values. The narrow primary sugar binding pocket does not allow for the (*R*) configuration of the hydroxy group at the C2′ position.

Next, the substituted R^2 sulfate group and R^3 hydroxymethyl group at the C3′ position were located outside of the binding pocket and completely exposed to the solvent, excluding the hydroxy group (Fig. 5c). The R^3 hydroxymethyl group reaches the anionic Asp203 residue at the edge of the binding pocket and forms a hydrogen bond (Fig. 5b), thus explaining the large affinity difference between **1** and **7**.

On the other hand, the sulfate group did not have significant interaction with the protein, and it is positioned near the entrance of the binding pocket via weak van der Waals interaction with Tyr299, Trp406 and Phe575 (Fig. 5d). In other words, the sulfate group does not participate in specific recognition and it does not greatly affect binding affinity. This is consistent with the SAR data; the affinity of neosalacinol (**6**), which has the sulfate group replaced by a hydroxy group, is similar to that of salacinol. As the R^2 hydroxy group of **6** appeared not to interact with the protein significantly, we designed and synthesized C3′-non substituted compound **12**[12] in order to validate our speculation.

The affinity of **12** was about ten times weaker than that of **6**.[13] Though this result was reasonable for the docking pose, further calculation has been derived to investigate the change of the affinities by the R^2 substituent. The docked poses of **1**, **6** and **12** have been geometrically optimized by the MMFF94x forcefield with the generalized born/volume integral (GB/VI) solvation model[14] to consider the solvent effect and then the interaction for each compound has been calculated to examine the reason in detail. The order of the pIC_{50} values were reproduced by interaction energies (Table 2). In this case, the affinity difference among these compounds was sprung up from the balance of interaction. While **1** is inferior to **6** on electrostatic interaction by the less favorable interaction with the anionic surrounding environment, the bulky sulfate group of **1** gained more favorable steric interaction compared to **6**. The van der Waals interaction of **12** was almost the same as that for **6**, but the loss of polar interaction exceeded the advantage of desolvation penalty. This relative disadvantage of **12** compared to **6** causes further loss of affinity.

We have shown the validity of the predicted binding mode of salacinol and its derivatives to the α-glucosidase from the viewpoint of protein–ligand interaction. The predicted binding mode can satisfy the structure–activity relationship qualitatively. Furthermore, precise calculation with solvent effect suggested the affinity differences among three compounds, **1**, **6** and **12**, depended on the balance of the interaction energy components, such as van der Waals, electrostatic and desolvation.

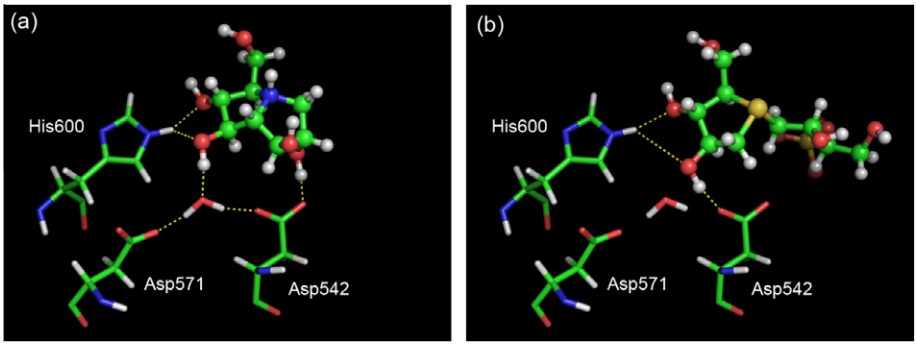


Figure 6. Intermediating water. (a) Experimental binding mode of casuarine. (b) Docking mode of salacinol with overlaid water molecule.

Table 2
The interaction energies (kcal/mol) of three compounds

	E_{int}[a]	E_{vdw}[b]	$E_{eff\text{-}ele}$[c]	E_{ele}[d]	G_{solv}[e]
1	−37.0	−19.6	−17.4	−317.4	300.0
6	−36.4	−14.5	−21.9	−479.3	457.4
12	−35.5	−14.7	−20.8	−472.1	451.3

[a] Total interaction ($E_{int} = E_{vdw} + E_{eff\text{-}ele}$) corresponds to binding free energy practically.
[b] van der Waals contribution for E_{int}.
[c] Effective electrostatic contribution for E_{int}. This term consists of coulombic interaction and desolvation penalty ($E_{eff\text{-}ele} = E_{ele} + G_{solv}$).
[d] Coulombic contribution for E_{int}.
[e] Desolvation penalty in binding calculated by the GB/VI model.

Acknowledgments

This work was supported by 'High-Tech Research Center Project': Matching Fund Subsidy for Private Universities from MEXT (Ministry of Education, Culture, Sports, Science and Technology).

References and notes

1. (a) Yoshikawa, M.; Murakami, T.; Shimada, H.; Matsuda, H.; Yamahara, J.; Tanabe, G.; Muraoka, O. *Tetrahedron Lett.* **1997**, *38*, 8367; (b) Yoshikawa, M.; Morikawa, T.; Matsuda, H.; Tanabe, G.; Muraoka, O. *Bioorg. Med. Chem.* **2002**, *10*, 1547.
2. Yoshikawa, M.; Murakami, T.; Yashiro, K.; Matsuda, H. *Chem. Pharm. Bull.* **1998**, *46*, 1339.
3. Sim, L.; Quezada-Calvillo, R.; Sterchi, E. E.; Nichols, B. L.; Rose, D. R. *J. Mol. Biol.* **2008**, *375*, 782.
4. Cardona, F.; Parmeggiani, C.; Faggi, E.; Bonaccini, C.; Gratteri, P.; Sim, L.; Gloster, T. M.; Roberts, S.; Davies, G. J.; Rose, D. R.; Goti, A. *Chem. Eur. J.* **2009**, *15*, 1627.
5. MOE Ver 2008.10, Chemical Computing Group Inc., Montreal, Canada.
6. PYMOL Ver 0.99rc6, DeLano Scientific LLC. San Carlos, California.
7. Goto, J.; Kataoka, R.; Muta, H.; Hirayama, N. *J. Chem. Inf. Model.* **2008**, *48*, 583.
8. Labute, P. *Proteins* **2009**, *75*, 187.
9. (a) Muraoka, O.; Yoshikai, K.; Takahashi, H.; Minematsu, T.; Lu, G.; Tanabe, G.; Matsuda, H.; Yoshikawa, M. *Bioorg. Med. Chem.* **2006**, *14*, 500; (b) 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; (c) Yoshikawa, M.; Xu, F.; Nakamura, S.; Wang, T.; Matsuda, H.; Tanabe, G.; Muraoka, O. *Heterocycles* **2008**, *75*, 1397.
10. Ghavami, A.; Johnston, B. D.; Maddess, M. D.; Chinapoo, S. M.; Jensen, M. T.; Svensson, B.; Pinto, B. M. *Can. J. Chem.* **2002**, *80*, 937.
11. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 587. modified slightly in MOE.
12. Compound **12**: colorless oil. $[\alpha]_D^{25}$ −147.5 (*c* 0.16, CH_3OH). IR (neat): 3356, 1653, 1647, 1055, 1023 cm^{-1}. ^1H NMR (700 MHz, CD_3OD) δ: 1.80 (2H, dt, *J* = 6.2, 6.0 Hz, H-3′a and H-3b′), 3.61 (1H, dd, *J* = 13.0, 9.6 Hz, H-1′a), 3.70 (1H, dt, *J* = 11.0, 6.0 Hz, H-4′a), 3.72 (1H, dt, *J* = 11.0, 6.0 Hz, H-4′b), 3.77 (1H, dd, *J* = 13.0, 3.0 Hz, H-1′b), 3.82 (1H, dd, *J* = 12.6, 3.4 Hz, H-1a), 3.86 (1H, dd, *J* = 12.6, 1.4 Hz, H-1b), 3.92 (1H, dd, *J* = 10.6, 9.0 Hz, H-5a), 4.02 (1H, br dd-like, *J* = 9.0, 5.2 Hz, H-4), 4.04 (1H, dd, *J* = 10.6, 5.2 Hz, H-5b), 4.24 (1H, dtd, *J* = 9.6, 6.2, 3.0 Hz, H-2′), 4.36 (1H, br d-like, *J* = ca. 1.4 Hz, H-3), 4.61 (1H, br ddd-like, *J* = ca. 3.4, 1.4, 1.4 Hz, H-2). ^{13}C NMR (175 MHz, CD_3OD) δ: 40.1 (C-3′), 52.0 (C-1), 54.7 (C-1′), 59.0 (C-4′), 66.9 (C-2′), 61.0 (C-5), 73.4 (C-4), 79.45 (C-2), 79.50 (C-3). FABMS *m/z*: 239 $[M]^+$ (pos.). FABHRMS *m/z*: 239.0949 ($C_9H_{19}O_5S$ requires 239.0953).
13. The pIC_{50} value of compound **12** was determined by the same method with other compounds.
14. Labute, P. *J. Comput. Chem.* **2008**, *29*, 1693.