Molecular Modeling of the Three-Dimensional Structure of Human Sphingomyelin Synthase

Zhang, Ya[a](张亚) Lin, Fu[b](林赋) Deng, Xiaodong[a](邓晓东)
Wang, Renxiao*[,b](王任小) Ye, Deyong*[,a](叶德泳)

[a] *School of Pharmacy, Fudan University, Shanghai 201203, China*
[b] *State Key Lab of Bioorganic Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China*

Sphingomyelin synthase (SMS) produces sphingomyelin and diacylglycerol from ceramide and phosphatidylcholine. It plays an important role in cell survival and apoptosis, inflammation, and lipid homeostasis, and therefore has been noticed in recent years as a novel potential drug target. In this study, we combined homology modeling, molecular docking, molecular dynamics simulation, and normal mode analysis to derive a three-dimensional structure of human sphingomyelin synthase (*h*SMS1) in complex with sphingomyelin. Our model provides a reasonable explanation on the catalytic mechanism of *h*SMS1. It can also explain the high selectivity of *h*SMS1 towards phosphocholine and sphingomyelin as well as some other known experimental results about *h*SMS1. Moreover, we also derived a complex model of D609, the only known small-molecule inhibitor of *h*SMS1 so far. Our *h*SMS1 model may serve as a reasonable structural basis for the discovery of more effective small-molecule inhibitors of *h*SMS1.

Keywords sphingomyelin synthase, molecular modeling, molecular dynamics

Introduction

Sphingomyelin synthase (SMS) is the enzyme that functions at the last step in the synthesis of sphingomyelin. It recognizes ceramide and phosphatidylcholine (PC) as substrates to produce sphingomyelin (SM) and diacylglycerol (DAG) (Figure 1).[1] Its activity influences the levels of SM, PC, ceramide, and DAG directly in living body, and is closely related with cell survival and apoptosis, inflammation, and atherosclerosis.[2-9]

SMS has two known subtypes, SMS1 and SMS2, which are classified by their cellular localizations. SMS1 is found merely in the *trans*-Golgi apparatus, and SMS2 is primarily found in the plasma membranes.[10,11] As most lipid phosphate phosphatase family, SMS catalyzes the choline phosphotransferase reaction possibly through a similar mechanism. First, a double-chain choline phospholipid (PC or SM) enters and binds to a single site of the enzyme. Then, a nucleophilic attack on the lipid-phosphate ester bond is executed by His328 in the assistance by Asp332. After the formation of a choline phosphohistidine intermediate and the release of DAG or ceramide, His285 acts as a nucleophile to attack on the carbon attached to the primary hydroxyl group on ceramide or DAG. Finally, the product (SM or PC) is released from the active site to allow the next round of catalysis.[12]

Due to its pharmaceutical implications, SMS has been noticed as a potential drug target in recent years. Huitema *et al*. reported the sequence of SMS using functional cloning strategy in yeast.[10] But the three-dimensional structure of SMS remains unresolved so far. Without such structural information, it remains as a challenge to understand the catalytic mechanism of SMS and discover potent inhibitors of SMS accordingly. In fact, very few small-molecule compounds that can regulate the biological function of SMS have been reported in literature. To the best of our knowledge, the only known SMS inhibitor so far is D609 (Figure 2), which was reported to have a weak inhibitory activity (IC_{50}＝500 μmol•L^{-1}) against SMS *in vitro*, but no activity *in vivo* due to its unstable chemical structure.[13-15] In this study, we combined homology modeling, molecular docking, and molecular dynamics simulation to derive a three-dimensional structural model of human sphingomyelin synthase 1 (*h*SMS1). Our model can reasonably explain some known experimental results regarding *h*SMS1. It can be applied to future structure-based discovery of novel small-molecule inhibitors of *h*SMS1.

* E-mail: dyye@shmu.edu.cn.; wangrx@mail.sioc.ac.cn.; Tel.: 0086-021-51980117, 0086-021-54925128; Fax: 0086-021-51980125
Received January 5, 2011; revised February 18, 2011; accepted Apil 28, 2011.
Project supported by the National Natural Science Foundation of China (Nos.30973641, 20902013), a special research fund for the Doctoral Program of Higher Education from the Chinese Ministry of Education (No. 20090071110054), and an open grant from the State Key Laboratory of Bio-organic and Natural Products Chemistry, Chinese Academy of Sciences.

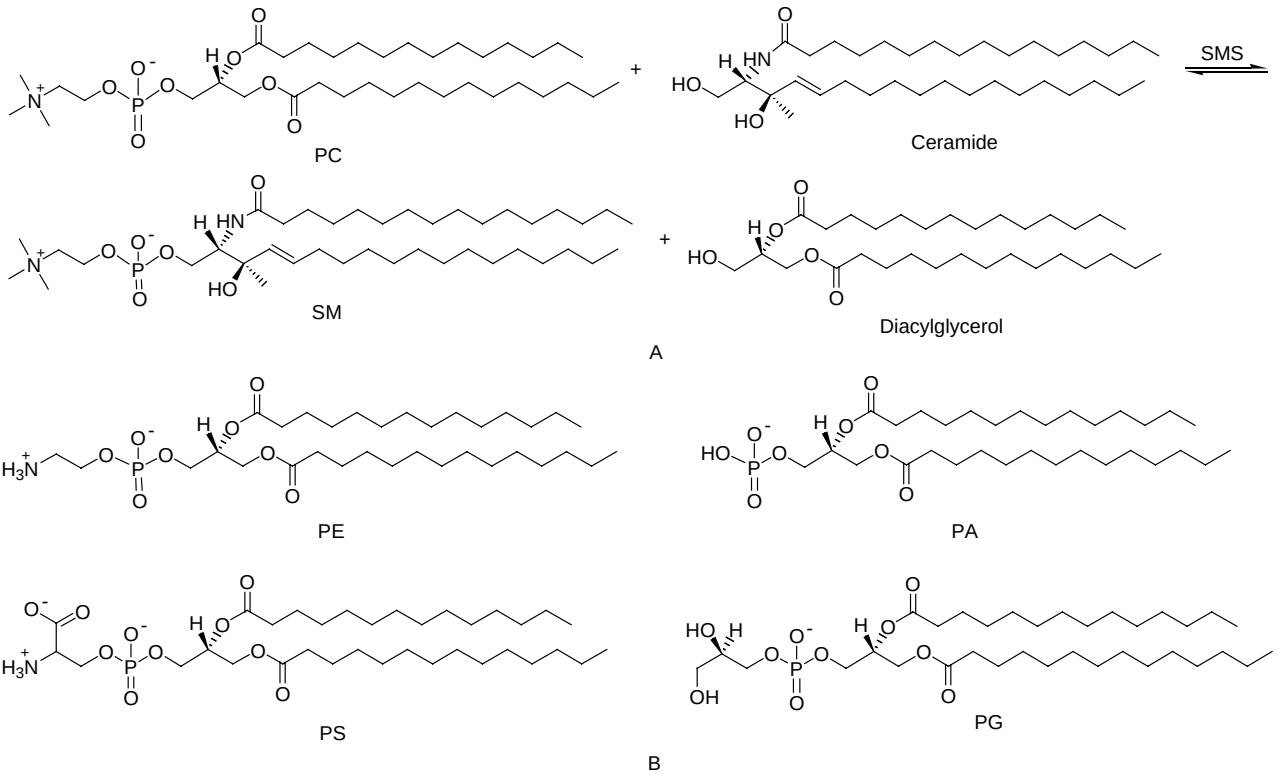

Figure 1 (A) *h*SMS1-catalyzed synthesis of sphingomyelin (SM) from phosphatidylcholine (PC). (B) Some other phosphatides related to ceramide, including phosphatidylethanolamine (PE), phosphatidic acid (PA), phosphatidylserine (PS) and phosphatidylglycerol (PG).

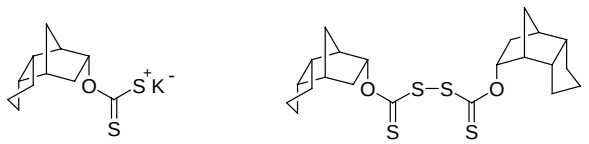

Figure 2 Chemical structures of D609 (IC_{50}=500 μmol•L^{-1} against SMS) and the corresponding dimer D609-dixanthogen (no inhibition activity against SMS).

Computational methods

Homology modeling of the *h*SMS1/lipid complex structure

The amino acid sequence of *h*SMS1 used in our study, which has 413 residues in total, was retrieved from PubMed (access ID=NP_671512). The *h*SMS1 has an *N*-terminal domain, a transmembrane domain, and a *C*-terminal domain.[16] Jiang *et al*. demonstrated recently that truncation of *N*-terminal and *C*-terminal of *h*SMS1 would not abort its activity.[17] Thus, only the core structure of *h*SMS1 (M130-Q353) was modeled in our study. Locations of the extracellular or intracellular loop and transmembrane domain on the sequence of *h*SMS1 were predicted by using the PSI-PRED Server.[18] The predicted extracellular or intracellular loops and transmembrane segments are given in Figure 3. This prediction is basically consistent with the results reported by Huitema *et al*. in a previous study.[10]

The transmembrane domain of *h*SMS1 is composed of six transmembrane helices (TMs). We employed 15 different computational methods to predict the locations of these TMs on the *h*SMS1 sequence.[19-33] Most of them produced consistent predictions (Table 1). In order to select an appropriate template for modeling the TMs of *h*SMS1, we retrieved a total of 61 entries with six TMs from the membrane structural proteins database PDBTM (http://pdbtm.enzim.hu).[34] After careful evaluations, which will be explained in the later Results and Discussion section, we selected the crystal structure of *Escherichia coli* GlpG (PDB entry: 2IC8) as the template (Table 2).[35] The Modeler function (as implemented in the Discovery Studio software suite)[36,37] was employed to generate a total of 30 structural models based on this template.

As for the extracellular or intracellular loop of *h*SMS1, loop 3 (residues 235—274) is relatively long (Figure 3). Therefore, it should be modeled based on an appropriate template. PDB entry 1BW0 (residues 144—181) (sequence identity=34.1%; sequence similarity=46.3%) was selected as the template for this purpose. This structure was selected throughout the entire PDB database according to the sequence similarity computed by the FASTA algorithm.[38,39] No qualified template was found for other loops of *h*SMS1 though. These loops were constructed from scratch by using the Modeler module in the Discovery Studio software suite. A total of 30 structural models were also generated for each

CHINESE JOURNAL OF CHEMISTRY

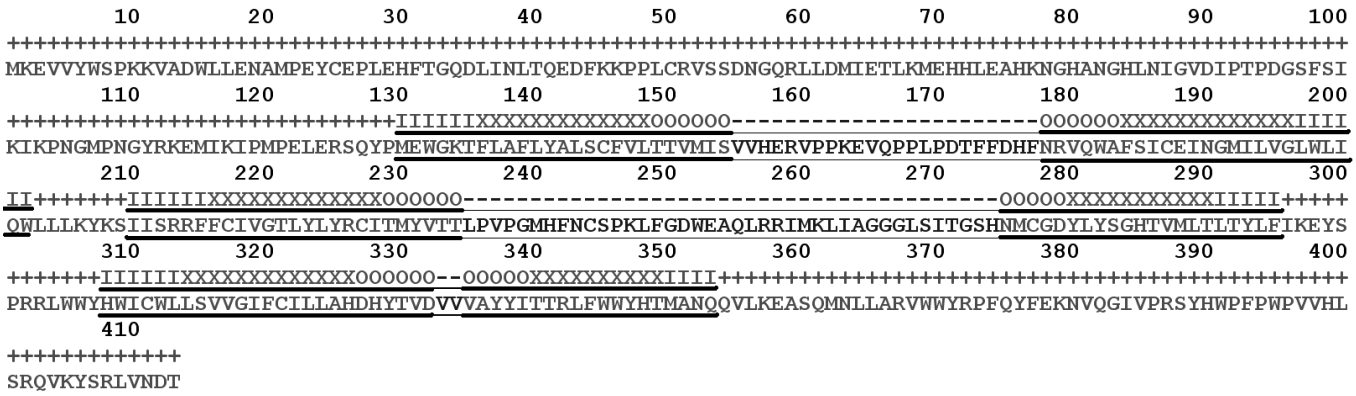

Figure 3 Transmembrane topology of *h*SMS1 predicted by MEMSAT3. Here, the characters with broad lines, with thin lines and without line stand for transmembrane domains, loop regions, and terminal domains respectively (+: intracellular domains; -: the extracellular loop; O : Outside helix cap; X : Central transmembrane helix segment; I: Inside helix cap).

Table 1 Predicted locations of the trans-membrane helices on *h*SMS1

Method	N[a]	TM1	TM2	TM3	TM4	TM5	TM6
MEMSTAT3[19]	6	M130-S154 (25)[b]	N178-W202 (25)	I210-T234 (25)	N275-F295 (21)	I308-D332 (25)	V335-Q354 (20)
PSIPRED[20]	6	E131-H157 (27)	W182-L204 (23)	S209-T229 (21)	H285-E298 (14)	L304-A325 (22)	V331-A351 (21)
APSSP[21]	6	W132-H157 (26)	W182-K206 (25)	S209-I228 (20)	T286-F295 (10)	W305-A325 (21)	Y329-Q353 (25)
BetaTPred2[22]	6	W132-R159 (28)	F184-K206 (23)	I211-T234 (24)	T286-Y299 (14)	W306-A325 (20)	Y329-Q353 (25)
ConPred II[23]	5	F136-V156 (21)	S185-L205 (21)	F215-L235 (21)		D279-Y299 (21)	Y307-D327 (21)
PROF[24]	6	G133-V156 (24)	R179-K206 (28)	K208-T233 (26)	T286-K297 (12)	W306-L324 (19)	T330-M350 (21)
SSpro[25]	6	M130-H157 (28)	Q181-L204 (24)	S209-T229 (21)	Y280-Y299 (20)	W305-A325 (21)	V331-Q353 (23)
DPM[26]	6	M130-V160 (31)	Q181-K206 (26)	I210-T234 (25)	T286-E298 (13)	R302-H326 (25)	T330-Q353 (24)
DSC[27]	6	M130-H157 (28)	W182-L205 (24)	I210-T233 (24)	V287-K297 (11)	R302-L323 (22)	D332-Q353 (22)
GOR1[28]	6	M130-E158 (29)	R179-L205 (27)	S209-P238 (30)	M276-K297 (21)	I310-Y329 (20)	V331-Q353 (23)
GOR3[29]	6	M130-V156 (27)	I186-Y207 (22)	I210-T233 (24)	T286-R302 (17)	L304-A325 (22)	V331-Q353 (23)
MLRC[30]	6	W132-H157 (26)	I186-L205 (20)	S209-T234 (26)	Y280-K297 (18)	R303-A325 (23)	V331-Q353 (23)
PHD[31]	6	E131-R159 (29)	N178-L205 (28)	S209-L235 (27)	C277-K297 (21	R303-L324 (22)	T330-Q353 (24)
Predator[32]	6	T135-V156 (22)	F184-L205 (22)	S209-T234 (26)	T286-K297 (12)	W305-A325 (21)	T330-Q353 (24)
SOPM[33]	6	T135-R159 (25)	W182-L205 (24)	S209-T234 (26)	M276-K297 (22)	L304-324 (21)	T330-Q353 (24)

[a] Predicted number of TMs. [b] Numbers in brackets are the lengths of TMs.

Table 2 Comparison of the length of transmembrane helices on *h*SMS1 and GlpG (PDB entry: 2IC8)

Protein	TM number	TM1	TM2	TM3	TM4	TM5	TM6
*h*SMS1	6	25	25	25	21	25	20
GlpG	6	27	17	20	18	21	18

loop. All of the resulting models were then visually checked to exclude those having serious internal collisions. Moreover, as for the third extracellular loop (loop 5, residues 327—339), two highly conserved residues H328 and D332 on this loop formed hydrogen bond, which could be used as an additional criterion to select appropriate model for loop 5.

Then, all of the whole structural models of *h*SMS1 were visually inspected to exclude those containing crossing loops or serious internal steric collisions. Models on which the catalytic pocket were too small to accommodate PC or SM were excluded. Models on which the three key residues (H285, H328, and D332) could not form a reasonable arrangement in the catalytic pocket were also excluded. Among the remaining models, the one with the lowest probability density function (PDF) energy computed by Modeler was chosen for further refinement. In order to relax the steric repulsions between side chains, side chain refinement with restraints on backbone was performed. Each loop was refined by high-level optimizations by using Modeler. Side-chain refinement was performed again because the conformation of backbone could have been changed during the optimization performed at the previous step.

Molecular dynamics simulation of the *h*SMS1/SM complex structure

Based on the final representing model of *h*SMS1, SM was manually docked into the catalytic pocket in favor of the S_N2 nucleophilic substitution reaction. The

binding mode of SM was then optimized within the structural restraints of *h*SMS1 (Figure 4). Minimization was conducted using the adopted basis Newton-Raphson algorithm[40] with the CHARMm force field implemented in the Discovery Studio software.

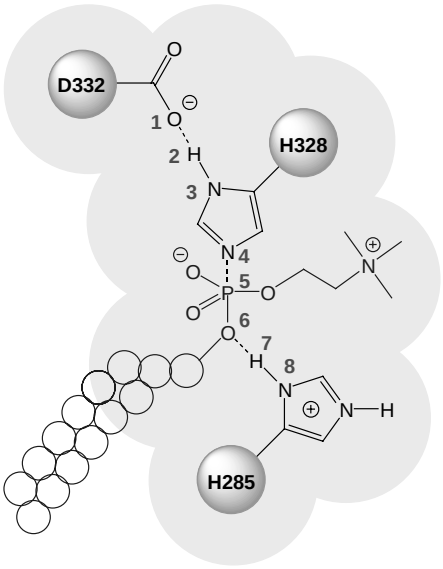

Figure 4 The binding mode of sphingomyelin to *h*SMS1. Three residues, *i.e*. H285, H328 and D332, in the catalytic site are critical for the catalytic mechanism of *h*SMS1.

In order to refine the structural model of the *h*SMS1/SM complex, it was further subjected to a long-time molecular dynamics (MD) simulation in which an implicit GB/SA model of bilayer membrane[41] was applied. Major parameters for setting the implicit membrane GB/SA model include: Grid spacing for lookup table (DGP)＝1.5 Å; Half membrane switching length (MSW)＝2.0 Å; Half smoothing length (SW)＝0.3 Å; Non-polar surface tension coefficient (SGAMMA)＝0.418 kJ/(mol•Å^2). Number of angular integration points 50 was used for volume integration in GB/SA calculation. To keep the catalytic site stable for S_N2 nucleophilic substitution reaction, three pairs of harmonic distance restraint were applied with harmonic force constant 418 (kJ/mol•Å^2): atom pairs 1—2, 4—5, and 6—7 (Figure 4). Firstly, *h*SMS1/SM was positioned in the center of the implicit membrane model. The planar membrane is perpendicular to the *z* axis and centered at *z*＝0 with a membrane thickness of 30 Å. Topology and coordinate files were generated with the Discovery Studio software. Minimization was performed using the program CHARMm (release c33b2).[42] A total of 5000 steps of steepest descent minimization and 5000 steps of conjugate gradient minimization were performed subsequently.

As the first step of MD simulation, the entire system was heated from 0 K to 310 K in 100 ps. Then, the whole system was equilibrated for 400 ps under a constant temperature of 310 K and a constant pressure of 1.0 MPa. After these preparative steps, a production of 10 ns long was conducted under the same temperature and pressure. During this process, the temperature was controlled with the Hoover method[43] whereas the pressure was controlled by coupling to a pressure bath using extended system algorithm.[44] The mass of the pressure piston was 1000 amu. Langevin piston collision frequency was set to 25.0 ps^{-1}. The dielectric constants of protein and water molecules were set to 1.0 and 80.0, respectively. Distance cutoff in generating the list of pairs was 14.0 Å. Switching function was used between 10.0 Å and 12.0 Å to treat non-bonding interactions. RMSD values relative to the initial structure were monitored as an indication of equilibrium along the MD trajectory. The last snapshot on the resulting MD trajectory was retrieved, minimized without restraint in couple with the implicit membrane model. This final refined model of the *h*SMS1/SM complex is shown in Figure 5. In order to evaluate this structural model, the PROCHECK program (version 3.5.4)[45] was applied to check its stereochemical quality. The Ramachandran plot of this *h*SMS1 structure produced by PROCHECK is shown in Figure 6.

Computational mutagenesis study on the *h*SMS1/lipid complex structure

We then performed computational mutagenesis on the *h*SMS1 complex structure to evaluate the importance of both the functional groups on the lipid substrate and several amino acid residues on *h*SMS1. Such results will provide additional proofs on the catalytic mechanism of *h*SMS1. Two types of mutations were performed accordingly: on the substrate side, SM was mutated in turn into PC, PE, PA, PS, and PG (Figure 1); whereas on the *h*SMS1 side, several important amino acid residues near the binding pocket were mutated in turn into alanine. Both types of mutations were done using the Discovery Studio software by deleting or adding some atoms while keeping original atoms as much as possible. After mutation, two rounds of minimization were applied to the *h*SMS1/lipid complex structure with the same force field parameters described above. Finally, binding energies between the substrate molecule and *h*SMS1 were calculated using an implicit GB/SA membrane model for each resulting *h*SMS1/lipid complex (Table 3 and Table 4).

Normal mode analysis and protein motion analysis of *h*SMS1

We conducted protein domain motion analysis for *h*SMS1 to study its major conformational motions with the DynDom on-line server (http://fizz.cmp.uea.ac.uk/dyndom/).[46] The initial and the last configuration of the *h*SMS1/SM complex in MD simulation were submitted as inputs. The DynDom server returned the information of the "hinge axes" (Figure 7). In order to confirm the outcomes of protein domain motion analysis, normal

CHINESE JOURNAL OF CHEMISTRY

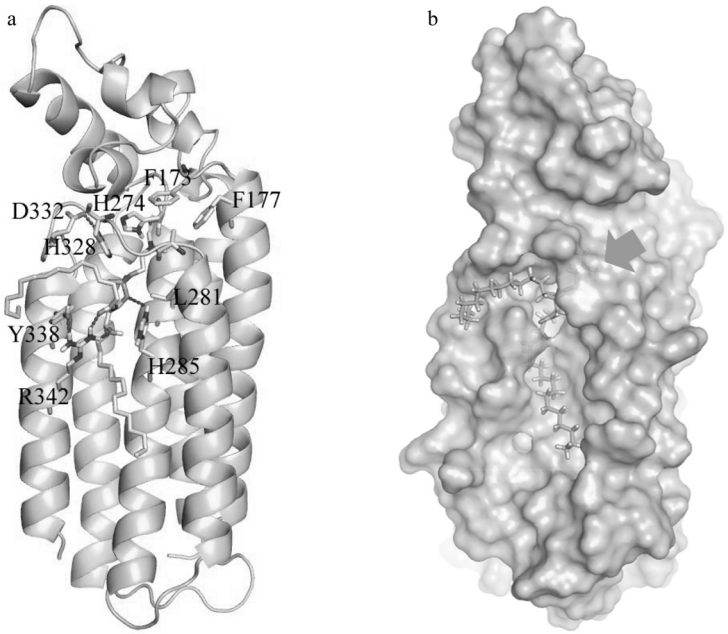

Figure 5 The *h*SMS1/SM complex structure as the last snapshot of 10 ns molecular dynamics simulation. (A) *h*SMS1 is rendered in ribbons. Several key residues around SM are shown in stick models. Dashed lines stand for hydrogen bonds. (B) *h*SMS1 is rendered in the solvent accessible surface. SM is rendered in stick model. The arrow indicates where the choline moiety on SM is buried inside.

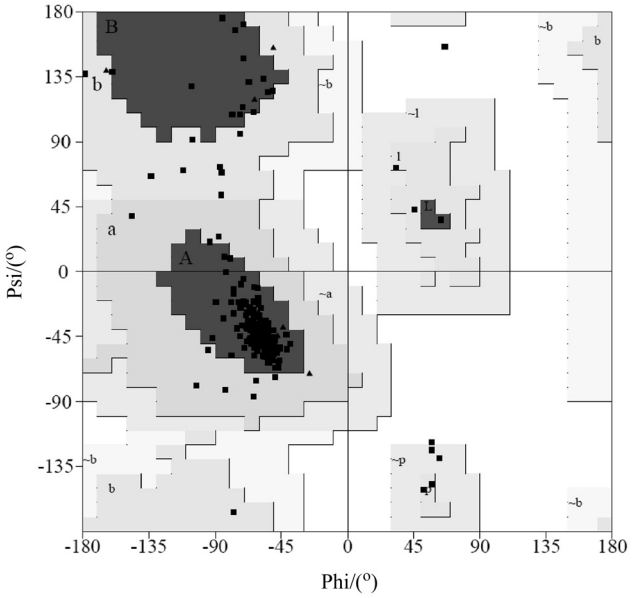

Figure 6 Ramachandran plot of the structural model of *h*SMS1 after molecular dynamics refinement. A: core alpha; a: allowed alpha; ~a: general alpha; B: core beta; b: allowed beta; ~b: general beta; L: core left-handed alpha; l: allowed left-handed alpha; ~l: general left-handed alpha; p: allowed epsilon; ~p: general epsilon. Glycines are shown as triangles. Here, 87.5% of residues are in the most favored regions (A, B, L), 10.5% of residues are in the additionally allowed regions (a, b, l, p), 1.0% of residues are in the generally allowed regions (~a, ~b, ~l, ~p), and only 1.0% of residues are in the disallowed regions.

mode analysis (NMA) was conducted to analyze the intrinsic motions of the *h*SMS1 structure. The elNémo on-line server (http://igs-server.cnrs-mrs.fr/elnemo/

Table 3 Computed binding energies between *h*SMS1 and some lipid substrates

Complex	Binding energy/(kJ•mol^{-1})
*h*SMS1/SM	0.00[a]
*h*SMS1/PC	−27.90
*h*SMS1/PE	62.80
*h*SMS1/PA	123.80
*h*SMS1/PS	63.76
*h*SMS1/PG	32.80

[a] The binding energy between *h*SMS1 and SM is taken as the energy reference.

Table 4 Computed binding energies between *h*SMS1 mutations and SM

*hh*SMS1 mutation	Binding energy/(kJ•mol^{-1})
Wild type	0.00[a]
D332A	−0.25
H285A	26.27
H328A	73.47
R342A	81.79
Y338A	21.08
F177A	7.57
H274A	18.95
L281A	14.60
F173A	17.61
F173A＋F177A＋H274A＋L281A	54.14

[a] The binding energy between SMS1 and SM is set as the reference.

index.html)[47] was employed in our study for this purpose. This server makes use of the Elastic Network Model, which provides a fast simple tool to compute, visualize, and analyze low-frequency normal modes of biological macromolecules. The structural model of *h*SMS1/SM was submitted for the NMA analysis. Key parameters used in computation included: DQMIN = −100, DQMAX = 100, DQSTEP = 20 and NRBL = "auto". The default cutoff of eight residues was used to identify elastic interaction ranges. A total of 100 normal modes with the lowest frequencies were computed. Essential features of the top ten low-frequency normal modes, including its frequency, are summarized in Table 5. The two lowest-frequency normal modes (*i.e.* modes 7 and 8 in Table 5) are illustrated in Figure 8.

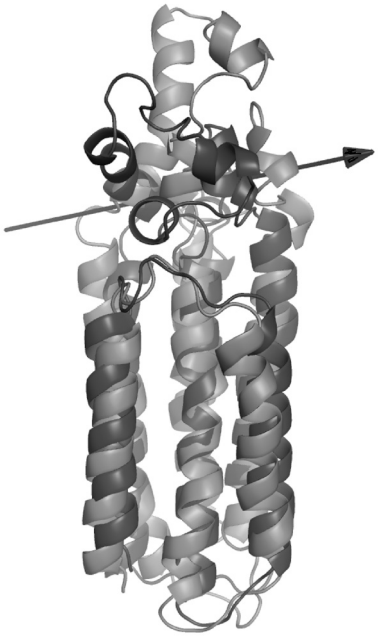

Figure 7 The two major states of the *h*SMS1 structure, *i.e.* an open state and a closed state, in which the extracellular loops adopt a different orientation from the transmembrane domain. The initial structure of *h*SMS1 before MD refinement is rendered in pale ribbons; the refined structure after 10 ns molecular dynamics simulation is rendered in dark ribbons. The arrow on the top stands for the hinge axe around which the extracellular loops rotate, which is given by the DynDom program.

Molecular docking of D609 to *h*SMS1

To understand how D609 affects the catalytic activity of *h*SMS1, the GOLD program (version 4.1, released by CCDC Inc.) was employed to perform automated molecular docking to explore the possible binding poses between *h*SMS1 and D609. The active site was defined as the residues within 10 Å from the reference ligand, *i.e.* SM. Other key parameters used in docking included: population size = 100, number of GA operations = 500000, mutation rate = 95%, crossover rate = 95% and scoring function = Chemscore. During docking process, the *h*SMS1 structure was kept fixed. A total of 20 top-ranked binding poses were output after docking was finished. These binding poses predicted by GOLD were quite similar to each other. Therefore, we chose the binding pose with the best binding score, and the corresponding *h*SMS1/D609 complex structure was minimized with the same force field parameters mentioned above in couple with the implicit membrane GB/SA model. The final model of the *h*SMS1/D609 complex is illustrated in Figure 9.

Table 5 Results of the normal mode analysis of the *h*SMS1 structure

Mode[a]	Frequency	Collectivity[b]
mode 7	1.00	0.6182
mode 8	1.07	0.6494
mode 9	1.23	0.6552
mode 10	1.85	0.6581
mode 11	1.87	0.6801
mode 12	2.07	0.5377
mode 13	2.16	0.6152
mode 14	2.65	0.6192
mode 15	2.70	0.5726
mode 16	2.75	0.3389

[a] Only the 10 normal modes with lowest frequencies are listed here; [b] Collectivity indicates the percentage of residues that are involved in a certain normal mode.

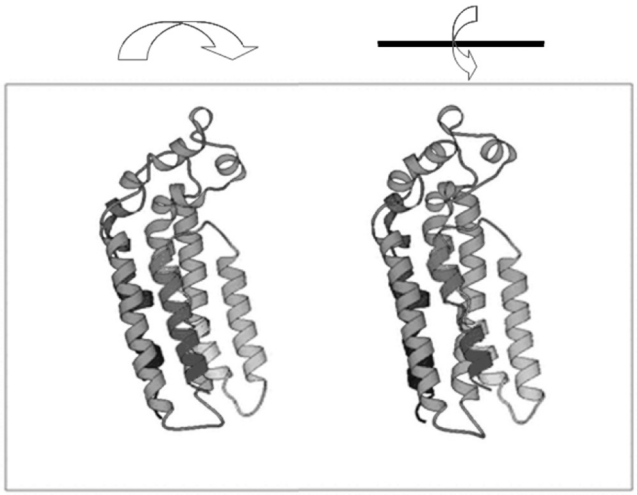

Figure 8 The two primary normal modes of *h*SMS1 with the lowest frequencies (Left: bending of the extracellular loops towards the transmembrane domain. Right: rotation of the extracellular loops on the top of the transmembrane domain).

Results and discussion

Homology modeling of the *h*SMS1/SM complex structure

In our study, it was of particular difficulty to select an appropriate template for homology modeling of the *h*SMS1 structure. In fact, no other membrane protein with known three-dimensional structure has a sequence

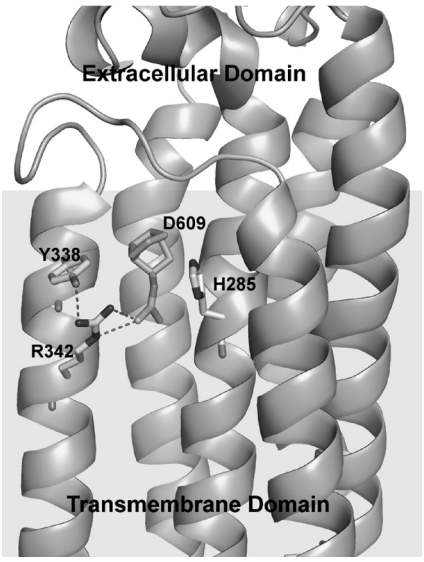

Figure 9 The binding mode of D609 with *h*SMS1 derived through molecular docking. *h*SMS1 is rendered in ribbons whereas D609 is rendered in stick model. The shaded region indicates the lipid bilayer.

similarity high enough with *h*SMS1, especially for the transmembrane domain. To tackle this problem, we retrieved all known membrane proteins with six TMs, 61 in total, from the membrane structural proteins database PDBTM.[34] Then, we examined them visually one by one with two criteria to select the most suitable one as template. Firstly, those largely different from *h*SMS1 in terms of biological function, such as aquaporin-like proteins, were excluded because they usually form a channel-like shape encircled by transmembrane helices. In addition, qualified candidates must be an enzyme. Secondly, for obvious reason, the TMs on the qualified candidates should be similar to those of *h*SMS1 in terms of length. After careful evaluation of all 61 entries, we chose the crystal structure of *Escherichia coli* GlpG (PDB entry: 2IC8) as the template. Indeed, the length of TMs in GlpG is comparable to the counterparts in *h*SMS1 (Table 2). In addition, there is an internal cavity on GlpG that harbors the Ser-His catalytic site, which is similar to *h*SMS1 that catalyzes PC into SM. An extracellular loop on GlpGare is also very similar to the counterpart on *h*SMS1 with two conserved amino acid residues (H328 and D332).

Our initial model of the *h*SMS1/SM complex structure was constructed through homology modeling and molecular docking. It was then subjected to a long-time MD simulation for refinement. An implicit membrane model, which is basically a GB/SA model, was employed in simulation to mimic the real environment round *h*SMS1. The final model of the *h*SMS1/SM complex structure is shown in Figure 5. The Ramachandran plot of this *h*SMS1 structure produced by PROCHECK (Figure 6) indicates that 87.5% of residues on our model are in the most favored regions (A, B, L), 10.5% of residues are in the additionally allowed regions (a, b, l, p), 1.0% of residues are in the generally allowed regions (~a, ~b, ~l, ~p), and only 1.0% of residues are in the disallowed regions. Thus, the quality of the *h*SMS1 structural model derived in our study is acceptable in this aspect.

According to our model (Figure 5A), residues H285, H328, and D332 are fully involved in the catalytic mechanism. D332 can form hydrogen bond with H328, which may act as hydrogen donor during catalysis. The choline group on the SM molecule stretches well into a hydrophobic patch formed by F173, F177, H274, and L281. Y338 and R342 can form hydrogen bond with SM. With these two kinds of interactions, the catalytic center on SM, *i.e.* the phosphorus atom, is well positioned inside the catalytic pocket of *h*SMS1 (Figure 5B). Our model (Figure 4) is generally consistent with the possible catalytic mechanism of *h*SMS1, which is given in the Introduction section of this article.

Mutagenesis study on *h*SMS1/lipid complexes

According to our computation results (Table 4), mutation of SM into other possible lipid substrates, including PG, PE, PA, and PS (Figure 1), leads to a decrease in binding energy of *h*SMS1/PG, *h*SMS1/PE, *h*SMS1/PA, and *h*SMS1/PS complexes as compared to that of the *h*SMS1/SM complex. However, mutation of SM into PC is quite different from lipid without choline above. The binding energy between *h*SMS1 and PC is a bit lower, *i.e.* more favorable, if compared to *h*SMS1/SM. Based on the computed binding energies, the choline group on SM is very important to binding between *h*SMS1 and SM. These results are consistent with Huitema's experimental proofs:[10] none of the non-choline phospholipids (PG, PE, PA, and PS) other than PC can trigger the production of SM. On the other hand, PC was efficiently recognized as a substrate of SMS and SM itself had a donor of the phosphorylcholine group, which can also be transformed back to PC. Therefore our model can, at least in a qualitatively manner, explain the high selectivity of *h*SMS1 towards PC and SM.

To better estimate the contribution of key residues to binding affinity between *h*SMS1 and SM, computational mutagenesis, *i.e.* alanine scanning, was conducted in our study on some residues near the catalytic site on *h*SMS1. For example, H285, H328, and D332 are known to be three highly conserved amino acids in SMS families, even in lipid phosphatases/phosphotransferases, and are key components participating in the catalytic reaction. Mutation of these three key residues on *h*SMS1 will lose catalytic activity completely.[17] In Table 4 and Figure 5A, one can see that mutation of D332A has little influence on the binding affinity between *h*SMS1 and SM. However, the carboxyl group on D332 functions as a hydrogen acceptor to stabilize the transition state during the catalysis process. As for H285 and H328, they are very important in both the binding process and the subsequent catalytic process. Mutation

of H285A and H328A will lead to a decrease in binding energy with value of 26.27, and 73.47 kJ/mol respectively. These results demonstrate the rationale of our model further.

Notably, our model also suggests some key residues which have not been well studied before. Mutation of R342A caused great loss of binding energy with value of 81.79 kJ/mol. This datum indicates that the hydrogen bonding interaction between SM and R342 is important for locking the conformation of SM to facilitate the catalyzed reaction. Mutation of Y338A also causes a decrease in binding energy of *ca*. 21.08 kJ/mol. One can see in Figure 5A that Y338 can form favorable hydrophobic interactions with one of the carbon chain on SM. In addition, Y338 can form a hydrogen bond with R342. In contrast, F173A, F177A, H274A, and L281A mutations do not have a major impact on the binding energy of SM (Table 4). However, the quadruple mutant F173A＋F177A＋H274A＋L281A leads to a large decrease of 54.14 kJ/mol in binding energy. It indicates that the hydrophobic interactions involving these residues are cooperative even though the contribution of each individual residue is trivial. These residues form a hydrophobic patch surrounding the choline group of SM, illustrating the importance of choline group in the PC and SM.[10]

Activation mechanism of *h*SMS1 based on NMA and DynDom analysis

Protein domain motion analysis indicated that the extracellular loops of *h*SMS1 rotate around the hinge axis shown in Figure 7. In order to further detect such conformational motions, we performed normal mode analysis (NMA) on the final snapshot on the MD trajectory of the *h*SMS1/SM complex. The two lowest-frequency normal modes (*i.e*. modes 7 and 8 in Table 3) are illustrated in Figure 8: one mode features bending of the loop towards the transmembrane domain whereas the other features rotating of the loop on top of the transmembrane domain. This motion prompts that there can be two major states existing in the physiological state: an open state and a closed state. These two states are connected by low-frequency conformational motions. This gating mechanism of the extracellular loops on *h*SMS1 illustrates how the reactant PC is locked after it enters the catalytic pocket and how the product SM is released afterwards. It also illustrates how an inhibitor may block the function of *h*SMS1. In particular, the choline moiety on PC can form favorable hydrophobic interactions with F173, F177, H274, and L281 in the closed state, which immobilizes the choline moiety and locks the shape of the catalytic center (Figure 5A). The lowest frequency motions that induce the movement of F173, F177, and H274 on the loops away from the binding site will disassemble the hydrophobic patch originally accommodating the choline moiety in the closed state. Consequently, SM will be released from the catalytic site in the open state.

The binding mode of a small-molecule inhibitor of *h*SMS1

D609 is essentially the only known small-molecule inhibitor of *h*SMS1 so far. Interestingly, dimerization of D609, *i.e*. D609 dixanthogen (Figure 2), was observed to have no inhibition activity at all.[13] Based on the structural model of the *h*SMS1/SM complex derived in our study, a binding mode of D609 was derived through molecular docking. Our model of the *h*SMS1/D609 complex (Figure 9) indicates that D609 occupies the catalytic pocket, which blocks the entry of substrate molecules, such as PC and SM. The key residues involved in the binding of D609 include R342 and H285. In particular, the negative charged sulfur atom on D609 forms a strong salt-bridge interaction with R342, which mimics the two critical hydrogen bonds between *h*SMS1 and SM. D609 dixanthogen does not have the negatively charged sulfur atom, leading to the loss of the salt-bridge with R342. In addition, the steric hindrance near R342 does not allow the second D609 moiety on D609 dixanthogen to go into the binding pocket deeply. Thus, our model provides a reasonable explanation on the different *h*SMS1 inhibition activities between D609 and D609 dixanthogen.

Conclusions

We have obtained a three-dimensional structural model of *h*SMS1 through homology modeling, molecular docking, and extensive molecular dynamics simulations. Our model reasonably explains how *h*SMS1 transforms a phosphocholine moiety on ceramide to produce sphingomyelin. It also explains the high selectivity of *h*SMS1 towards phosphocholine and sphingomyelin. Normal mode analysis on the low-frequency motions of *h*SMS1 suggests a gating mechanism of immobilizing phosphatidylcholine at the catalytic site and releasing sphingomyelin afterwards. Computational mutagenesis results explore the key residues for the binding of SM to *h*SMS1. Currently, very few small-molecule inhibitors of *h*SMS1 have been reported publicly in literature. The potential value of our structural model is that it can be applied to the discovery of effective *h*SMS1 inhibitors, for example, through virtual screening.

References

1 Merrill, A. H. Jr.; Jones, D. D. *Biochim. Biophys. Acta* **1990**, *1044*, 1.
2 Nagao, K.; Takahashi, K.; Hanada, K.; Kioka, N.; Matsuo, M.; Ueda, K. *J. Biol. Chem.* **2007**, *282*, 14868.
3 Sano, O.; Kobayashi, A.; Nagao, K.; Kumagai, K.; Kioka, N.; Hanada, K.; Ueda, K.; Matsuo, M. *J. Lipid Res.* **2007**, *48*, 2377.
4 Schutze, S.; Potthoff, K.; Machleidt, T.; Berkovic, D.; Wiegmann, K.; Kronke, M. *Cell* **1992**, *71*, 765.
5 Luberto, C.; Yoo, D. S.; Suidan, H. S.; Bartoli, G. M.;

Hannun, Y. A. *J. Biol. Chem.* **2000**, *275*, 14760.
6 Hailemariam, T. K.; Huan, C.; Liu, J.; Li, Z.; Roman, C.; Kalbfeisch, M.; Bui, H. H.; Peake, D. A.; Kuo, M. S.; Cao, G. *Arterioscler. Thromb. Vasc. Biol.* **2008**, *28*, 1519.
7 Miyaji, M.; Jin, Z. X.; Yamaoka, S.; Amakawa, R.; Fukuhara, S.; Sato, S. B.; Kobayashi, T.; Domae, N.; Mimori, T.; Bloom, E. T.; Okazaki, T.; Umehara, H. *J. Exp. Med.* **2005**, *202*, 249.
8 Van der Luit, A. H.; Budde, M.; Zerp, S.; Caan, W.; Klarenbeek, J. B.; Verheij, M.; Van Blitterswijk, W. J. *Biochem. J.* **2007**, *401*, 541.
9 Ding, T.; Li, Z.; Hailemariam, T.; Mukherjee, S.; Maxfield, F. R.; Wu, M. P.; Jiang, X. C. *J. Lipid Res.* **2008**, *49*, 376.
10 Huitema, K.; van den Dikkenberg, J.; Brouwers, J. F.; Holthuis, J. C. *Embo. J.* **2004**, *23*, 33.
11 Yamaoka, S.; Miyaji, M.; Kitano, T.; Umehara, H.; Okazaki, T. *J. Biol. Chem.* **2004**, *279*, 18688.
12 Tafesse, F. G.; Ternes, P.; Holthuis, J. C. *J. Biol. Chem.* **2006**, *281*, 29421.
13 Meng, A.; Luberto, C.; Meier, P.; Bai, A.; Yang, X.; Hannun, Y. A.; Zhou, D. *Exp. Cell. Res.* **2004**, *292*, 385.
14 Bai, A.; Meier, G. P.; Wang, Y.; Luberto, C.; Hannun, Y. A.; Zhou, D. *J. Pharmacol. Exp. Ther.* **2004**, *309*, 1051.
15 Dong, J. B.; Liu, J.; Lou, B.; Li, Z. Q.; Wu, M. P.; Jiang, X. C. *J. Lipid Res.* **2006**, *47*, 1307.
16 Sigal, Y. J.; McDermott, M. I.; Morris, A. J. *Biochem. J.* **2005**, *387*, 281.
17 Yeang, C.; Varshney, S.; Wang, R.; Zhang, Y.; Ye, D. Y.; Jiang, X. C. *Biochim. Biophys. Acta* **2008**, *1781*, 610.
18 Bryson, K.; McGuffin, L. J.; Marsden, R. L.; Ward, J. J.; Sodhi, J. S.; Jones, D. T. *Nucleic Acids Res.* **2005**, *33*, W36.
19 Jones, D. T. *Bioinformatics* **2007**, *23*, 538.
20 Jones, D. T. *J. Mol. Biol.* **1999**, *292*, 195.
21 Raghava, G. P. S. *CASP4* **2000**, 75.
22 Kaur, H.; Raghava, G. P. *Protein Sci.* **2003**, *12*, 627.
23 Arai, M.; Mitsuke, H.; Ikeda, M.; Xia, J. X.; Kikuchi, T.; Satake, M.; Shimizu, T. *Nucleic Acids Res.* **2004**, *32*, W390.
24 Ouali, M.; King, R. D. *Protein Sci.* **2000**, *9*, 1162.
25 Cheng, J.; Randall, A. Z.; Sweredoski, M. J.; Baldi, P. *Nucleic Acids Res.* **2005**, *33*, W72.
26 Deleage, G.; Roux, B. *Protein Eng.* **1987**, *1*, 289.
27 King, R. D.; Sternberg, M. J. *Protein Sci.* **1996**, *5*, 2298.
28 Garnier, J.; Osguthorpe, D. J.; Robson, B. *J. Mol. Biol.* **1978**, *120*, 97.
29 Gibrat, J. F.; Garnier, J.; Robson, B. *J. Mol. Biol.* **1987**, *198*, 425.
30 Guermeur, Y.; Geourjon, C.; Gallinari, P.; Deleage, G. *Bioinformatics* **1999**, *15*, 413.
31 Rost, B.; Sander, C. *J. Mol. Biol.* **1993**, *232*, 584.
32 Frishman, D.; Argos, P. *Protein Eng.* **1996**, *9*, 133.
33 Geourjon, C.; Deleage, G. *Protein Eng.* **1994**, *7*, 157.
34 Tusnady, G. E.; Dosztanyi, Z.; Simon, I. *Bioinformatics* **2004**, *20*, 2964.
35 Wang, Y.; Zhang, Y.; Ha, Y. *Nature* **2006**, *444*, 179.
36 Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.
37 *Discovery Studio Software*, Version 1. 7, Accelyrs Inc., San Diego, California, U. S. A. **2007**.
38 Bhat, T. N.; Bourne, P.; Feng, Z. *Nucleic Acid Res.* **2001**, *29*, 214.
39 Pearson, W. R. *Methods Enzymol.* **1990**, *183*, 63.
40 Momany, F. A.; Rone, R. *J. Comp. Chem.* **1992**, *13*, 888.
41 Im, W.; Feig, M.; Brooks, C. L. *Biophys. J.* **2003**, *85*, 2900.
42 Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
43 Hoover, W. G. *Phys. Rev.* **1985**, *31*, 1695.
44 Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384
45 Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
46 Lee, R. A.; Razaz, M.; Hayward, S. *Bioinformatics* **2003**, *19*, 1290.
47 Suhre, K.; Sanejouand, Y. H. *Nucleic Acid Res.* **2004**, *32*, W610.

(E1101051 Cheng, F.; Fan, Y.)

