

A distant evolutionary relationship between bacterial sphingomyelinase and mammalian DNase I

YO MATSUO,^{1,3} ATSUKO YAMADA,² KIKUO TSUKAMOTO,² HIRO-OMI TAMURA,^{2,4}
HIROH IKEZAWA,² HARUKI NAKAMURA,¹ AND KEN NISHIKAWA^{1,5}

¹Protein Engineering Research Institute, 6-2-3 Furuedai, Saitama, Japan

²Department of Microbial Chemistry, Faculty of Pharmaceutical Sciences, Nagoya City University, Nagoya, Aichi 467, Japan

(RECEIVED July 23, 1996; ACCEPTED September 5, 1996)

Abstract

The three-dimensional structure of bacterial sphingomyelinase (SMase) was predicted using a protein fold recognition method; the search of a library of known structures showed that the SMase sequence is highly compatible with the mammalian DNase I structure, which suggested that SMase adopts a structure similar to that of DNase I. The amino acid sequence alignment based on the prediction revealed that, despite the lack of overall sequence similarity (less than 10% identity), those residues of DNase I that are involved in the hydrolysis of the phosphodiester bond, including two histidine residues (His 134 and His 252) of the active center, are conserved in SMase. In addition, a conserved pentapeptide sequence motif was found, which includes two catalytically critical residues, Asp 251 and His 252. A sequence database search showed that the motif is highly specific to mammalian DNase I and bacterial SMase. The functional roles of SMase residues identified by the sequence comparison were consistent with the results from mutant studies. Two *Bacillus cereus* SMase mutants (H134A and H252A) were constructed by site-directed mutagenesis. They completely abolished their catalytic activity. A model for the SMase–sphingomyelin complex structure was built to investigate how the SMase specifically recognizes its substrate. The model suggested that a set of residues conserved among bacterial SMases, including Trp 28 and Phe 55, might be important in the substrate recognition. The predicted structural similarity and the conservation of the functionally important residues strongly suggest a distant evolutionary relationship between bacterial SMase and mammalian DNase I. These two phosphodiesterases must have acquired the specificity for different substrates in the course of evolution.

Keywords: bacterial sphingomyelinase; DNase I; evolutionary relationship; site-directed mutagenesis; structure prediction

Bacterial sphingomyelinase (sphingomyelin cholinephosphohydro-lase; EC 3.1.4.12) catalyzes the hydrolysis of sphingomyelin into ceramide and phosphorylcholine. It is secreted from the bacterial cell and functions as a toxin that can damage the host cell membrane (Tomita et al., 1991). Recent studies revealed a novel signal transduction pathway in mammalian cells. In the pathway, mammalian SMases are activated in response to extracellular signals (tumor necrosis factor α , γ -interferon, etc.), and hydrolyze SM in

the membrane to produce ceramide; in turn, ceramide functions as a second messenger in mediating the effects of the extracellular signals on cell growth, differentiation, and apoptosis (for reviews, see Hannun & Bell, 1989; Kolesnick, 1992; Hannun, 1994). It is therefore possible that bacterial SMase, which is quite different in size and sequence from known mammalian SMases, may act as a cytotoxic exogenous agent to disturb the normal host cell regulation mechanisms (Okazaki et al., 1989; Tamura et al., 1994; Wiegmann et al., 1994).

The amino acid sequence of bacterial SMase shows no significant similarity with other known protein sequences (Yamada et al., 1988), and the three-dimensional structure is not known. The enzymatic activity is enhanced 20–30-fold in the presence of Mg^{2+} (Yamada et al., 1988). Chemical modification experiments suggested that acidic amino acid residues are involved in the catalysis (Tomita et al., 1993). Mutational analyses suggested that Asp 126 and Asp 156 are involved in the substrate recognition, and that Asp 295 is essential to the hydrolytic activity (Tamura et al., 1995). However, knowledge of the three-dimensional structure is desirable for a better understanding of the functional mechanism.

Reprint requests to: Ken Nishikawa, Center for Information Biology, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411, Japan; e-mail: knishika@genes.nig.ac.jp.

³Present address: Institute for Social Information Science, Fujitsu Laboratories Ltd., 1-9-3 Nakase, Mihamachi, Chiba 261, Japan.

⁴Present address: Kyoritsu College of Pharmacy, 1-5-30, Shiba-koen, Minato-ku, Tokyo 105, Japan.

⁵Present address: Center for Information Biology, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411, Japan.

Abbreviations: SM, sphingomyelin; SMase, sphingomyelinase; DNase I, deoxyribonuclease I; HNP, 2-hexadecanoylamino-4-nitrophenylphosphorylcholine; PC-PLC, phosphatidylcholine-hydrolyzing phospholipase C.

Protein fold recognition is a useful approach to three-dimensional structure prediction. In this approach, the amino acid sequence of a protein of unknown structure is compared with a library of known structures, and a likely structure is identified from the library. Recently, various groups developed fold recognition methods (e.g., Sippl, 1990; Bowie et al., 1991; Godzik et al., 1992; Jones et al., 1992; Maiorov & Crippen, 1992; Bryant & Lawrence, 1993; Ouzzounis et al., 1993; Wilmanns & Eisenberg, 1993; Kocher et al., 1994). We also developed a method that evaluates the sequence-structure compatibility in terms of side-chain packing, solvation, hydrogen bonding, and local conformation (Nishikawa & Matsuo, 1993; Matsuo & Nishikawa, 1994a, 1994b, 1995; Matsuo et al., 1995). In 1994, these methods were systematically assessed in a large-scale protein structure prediction contest (Moult et al., 1995). The results of the contest showed that, although there is much room for improvement, the methods can often recognize the correct fold successfully from a library of known structures, even in the absence of significant sequence similarity between the target and any protein of known structure (Lemer et al., 1995).

In the present work, the three-dimensional structure of bacterial SMase is predicted using our fold recognition method. According to the structure prediction, residues of the SMase that may be involved in catalysis and specific recognition of the substrate are proposed. Following the prediction, SMase mutants are created by site-directed mutagenesis, and their catalytic activities are measured. Finally, the evolutionary implications of the prediction are discussed.

Results and discussion

Prediction of SMase structure

The amino acid sequence of *Bacillus cereus* SMase (Yamada et al., 1988; NBRF-PIR code, S01130; 34 kDa) was compared with 325 known three-dimensional structures taken from the Protein Data

Bank (Bernstein et al., 1977) using the fold recognition method (Matsuo & Nishikawa, 1994b). Bovine pancreatic DNase I (EC 3.1.21.1, formerly EC 3.1.4.5; 30 kDa) structure (Oefner & Suck, 1986; PDB code, 3DNI) showed the best compatibility score (S_{tot} , -2.82) (Table 1). The difference in score between DNase I and the second best structure (bean pod mottle virus capsid protein) was -0.50. In our previous work (Matsuo & Nishikawa, 1994b), we found empirically that there is a good possibility that our fold recognition method gives a significant prediction if the top ranking structure shows a score S_{tot} better than approximately -3.0 and its difference in score from the other structures is larger than approximately -0.5. The prediction of the SMase structure is at the very limit of this empirical criterion.

The compatibility of other bacterial SMase sequences with the DNase I structure was also checked. The BLAST (Altschul et al., 1990) sequence database search detected four other bacterial SMase homologues from *B. cereus* strain GP-4 (PIR code, B32042; Gilmore et al., 1989), *B. cereus* strain SE-1 (S01950; Johansen et al., 1988), *Staphylococcus aureus* (S15766; Projan et al., 1989), and *Lepostospira interrogans* (S22634; Segers et al., 1990). Their sequence identities with the query sequence (S01130; SMase from *B. cereus* strain IAM1208) were 95.1, 93.5, 58.5, and 46.5%, respectively. When the four homologue sequences were compared with the 325 known structures, the positions of the DNase I structure (PDB code, 3DNI) in the score ranking were 1, 1, 1, and 10, respectively. The compatibility scores S_{tot} were -2.78, -3.01, -2.44, and -1.68, and the differences in score between DNase I and the best of the other structures were -0.53, -0.59, -0.02, and 0.48.

The secondary structures of the SMase homologues were also predicted using the method of Nishikawa and Noguchi (1991). As shown in Figure 1, the predictions agreed fairly well with the experimental structure of the DNase I.

According to the sequence alignment (Fig. 1) given by the prediction method, the SMase sequence was mounted on the DNase I

Table 1. Compatibility of the *B. cereus* SMase sequence with known structures^a

Rank	Structure	PDB ^b	NR ^c	%ID ^d	S_{tot}	S_{tot}^{res}
1	DNase I	3DNI	250	9.7	-2.82	-2.70
2	Bean pod mottle virus	1BMV2	300	8.8	-2.32	-1.08
3	Phthalate dioxygenase reductase	2PIA	294	7.8	-2.23	-1.08
4	Ovalbumin	1OVA	303	9.6	-2.04	-0.73
5	Rubisco	5RUBA	298	9.1	-2.04	-0.80
6	Aldose reductase	1ADS	285	7.2	-2.01	-0.97
7	Xylose isomerase	4XIS	298	6.9	-2.00	-0.76
8	D-Glyceraldehyde-3-phosphate dehydrogenase	1GD1O	280	10.7	-1.98	-1.02
9	Triosephosphate isomerase	5TIMA	245	11.7	-1.92	-1.62
10	Narbonin	1NAR	275	10.1	-1.91	-1.02

^aThe compatibility of the *B. cereus* SMase sequence (Yamada et al., 1988) with 325 known three-dimensional structures taken from the Protein Data Bank (PDB) (Bernstein et al., 1977) was evaluated using the program COMPASS (Matsuo & Nishikawa, 1994b). The 325 proteins had less than 30% sequence identities with one another. The structures were sorted in the order of their compatibility scores S_{tot} . The auxiliary scores, S_{tot}^{res} , which indicate the compatibility per residue, are also shown (see Matsuo & Nishikawa, 1994b, for details). The best ten structures are listed.

^bPDB code (the fifth character denotes the chain name).

^cNumber of residues of the SMase sequence mounted on the structure.

^d% Sequence identity.

Bovine DNase I (3DNI)*Bacillus cereus* SMase (S01130)*Staphylococcus aureus* SMase (S15766)*Leptospira interrogans* SMase (S22634)

Fig. 1. Alignment of the bovine DNase I sequence (PDB code, 3DNI; Oefner & Suck, 1986) and the SMase sequences from *B. cereus* (PIR code, S01130; Yamada et al., 1988), *S. aureus* (S15766; Projan et al., 1989) and *L. interrogans* (S22634; Segers et al., 1990). Residue conservation is denoted by an asterisk. The functionally important residues identified from the X-ray structure of DNase I by Weston et al. (1992) are boxed (E39, E78, H134, D168, N170, D212, D251, and H252 of DNase I). The two residue positions (D126 and D156 of *B. cereus* SMase) that may be involved in the substrate recognition (Tamura et al., 1995) are also boxed. The highly conserved pentapeptide sequence (residues 250–254 of DNase I) is underlined. Alignment of the *B. cereus* SMase and the bovine DNase I sequences was given by the structure prediction method (Matsuo & Nishikawa, 1994b), with slight manual modifications. Multiple alignment of the three SMase sequences was constructed using the program CLUSTAL W (Higgins et al., 1992). The residue 296 of *B. cereus* SMase was reported originally to be aspartate because of a sequencing error. In the present work, the correct residue, histidine, was substituted for aspartate (Tamura et al., 1995). The secondary structural residues of the DNase I were defined using the DSSP algorithm of Kabsch and Sander (1983) and shown in red (α -helices) and blue (β -strands). The secondary structures of the three SMase homologues were predicted using the method of Nishikawa and Noguchi (1991) and shown in the same manner.

structure (Fig. 2). The two cysteine residues of SMase, Cys 123 and Cys 159, are known to form a disulfide bond (Tomita et al., 1993). They were aligned with Cys 101 and Ala 142 of DNase I, respectively. The distance between the $C\beta$ atoms of the two residues is 15.0 Å, which seems too long for disulfide bond formation. However, Cys 123 is located on a long flexible loop between two β -strands (see Figs. 1, 2), and the loop is located next to the α -helix on which Cys 159 was mounted. The flexible loop may therefore change its conformation to form a disulfide bond in the actual structure of SMase. As described later, energy minimization and molecular dynamics calculations allowed the distance between the two cysteine residues to be reduced to 3.4 Å.

These results suggest that SMase may adopt a structure similar to that of DNase I. However, the significance of a prediction cannot be confirmed convincingly by the compatibility score alone (Matsuo & Nishikawa, 1994b, 1995). A prediction should be validated by its consistency with other known experimental data.

Below, we show how the predicted structural similarity between SMase and DNase I is supported by other data.

Conservation of functionally important residues

The substrates of SMase and DNase I are quite different; one is SM, and the other is DNA. The evolutionary relationship between the organisms, bacteria and mammals, from which the two enzymes in question are derived, is also very distant.

However, both bacterial SMase and mammalian DNase I catalyze the hydrolysis of the phosphodiester bonds of the substrates (Fig. 3), depending on the divalent metal cations. With this common feature as phosphodiesterases, together with the predicted structural similarity, the catalytic mechanisms were also expected to be similar between SMase and DNase I. If so, catalytically important residues should be conserved. To see residue conservation, the amino acid sequences of DNase I and bacterial SMase

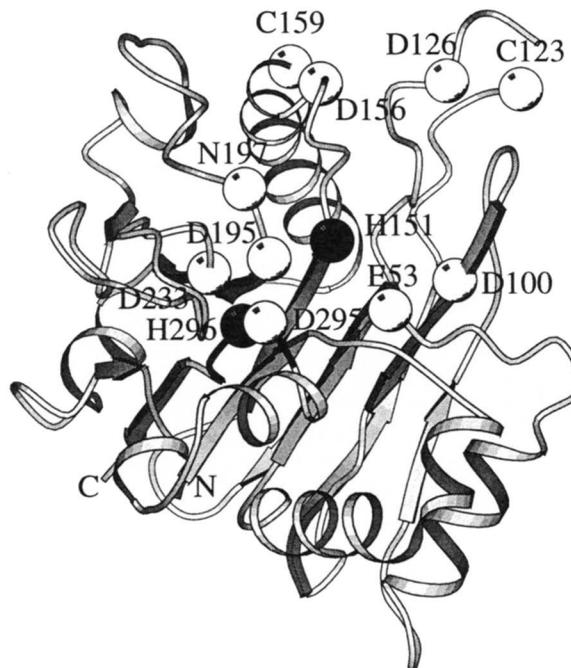


Fig. 2. *B. cereus* SMase sequence (Yamada et al., 1988) was mounted on the bovine DNase I structure (PDB code, 3DNI; Oefner & Suck, 1986). Predicted functionally important residues are shown in spheres: H151 (general acid in hydrolysis) and H296 (general base), in black; E53, D295 (stabilization of pentacovalent transition state), D100 (hydrogen bond to H151), D233 (hydrogen bond to H296), C123, C159 (S-S bond), D195, N197 (interaction with phosphate group), D126, and D156 (substrate recognition), in white. The highly conserved pentapeptide "SDHYP" (residues 294–298) is shown in black. The drawing was created using MOLSCRIPT (Kraulis, 1991).

were compared. The alignment of the sequences is shown in Figure 1.

From the X-ray structure of DNase I, Weston et al. (1992) proposed that, in DNase I, His 252 acts as a general base to take a proton from a water molecule, which then, as a nucleophile, attacks the phosphorus group and cleaves the P-O^{3'} bond; His 134

functions as a general acid to protonate the leaving O^{3'}; Asp 168 and Asn 170 interact with the scissile phosphate group; Glu 39 binds to the metal ion (Ca²⁺, Mg²⁺, or Mn²⁺), which is coordinated to the scissile phosphate bond; the metal ion, Glu 39, and Asp 251 are involved in the stabilization of the pentacovalent transition state. The sequence comparison revealed that these DNase I residues are completely conserved in the bacterial SMases (Fig. 1); the DNase I residues were aligned with His 296, His 151, Asp 195, Asn 197, Glu 53, and Asp 295 of the SMases, respectively (the residue numbering of SMase is based on the *B. cereus* mature SMase sequence). Asp 212 of DNase I, the side-chain carboxylate group of which is hydrogen bonded to N^{δ1} of His 252 (general base), was aligned with Asp 233 of the SMases. Glu 78 of DNase I, which is hydrogen bonded to His 134 (general acid), was aligned with Asp or Asn 100 of the SMases. No equivalent of Arg 9 of DNase I, which may play a role in stabilizing a pentacovalent transition state (Weston et al., 1992), was defined by the sequence alignment.

Suck and Oefner (1986) proposed another catalytic mechanism of DNase I from the three-dimensional structure of DNase I and the binding of Ca²⁺-thymidine 3',5'-diphosphate at the active site; that is, Glu 78 accepts a proton from His 134, and His 134 in turn accepts a proton from a water molecule, which then, as a nucleophile, attacks the phosphorus. However, from the structure of the DNA octanucleotide-DNase I complex, they later suggested the alternative mechanism involving the two histidine residues H252 and H134 (Suck et al., 1988). In fact, Glu 78 is replaced by Asn in *L. interrogans* SMase (Segers et al., 1990), as shown by the sequence comparison (Fig. 1).

The sequence database search (PIR release 43 was used, which contained 75,511 sequences, 22,468,834 residues) showed that the pentapeptide sequence "SDHYP" (residues 250–254 of DNase I) is conserved among all the known mammalian DNase I and bacterial SMase sequences (Fig. 1). This conserved motif includes the catalytically important residues Asp 251 and His 252 of DNase I. Among all other sequences in the database, the motif occurred only once in yeast dipeptidyl aminopeptidase B sequence (Roberts et al., 1989; PIR code, A30107). This indicates that the motif is highly specific to mammalian DNase I and bacterial SMase.

From the sequence comparison described above, we have proposed possible roles of the SMase residues in catalysis (Table 2).

Site-directed mutagenesis and catalytic activity of the mutant enzymes

The prediction of the functionally important residues of SMase (Table 2) is consistent with the results from mutant studies.

His 151 and His 296 of SMase were predicted to act as the general acid and base in the catalysis (Table 2); they were aligned with His 134 (general acid in catalysis) and His 252 (general base), respectively, in bovine DNase I. By site-directed mutagenesis, each of the two histidine residues of *B. cereus* SMase was substituted with alanine, and the catalytic activity of the two mutant enzymes H151A and H296A toward two phosphorylcholine-containing substrates SM and HNP were measured. HNP is a water-soluble synthetic substrate. As shown in Table 3, the two mutants completely lost their catalytic activity toward both the substrates. The result supports the prediction.

In a previous work, Tamura et al. (1995) analyzed the catalytic activities of the four SMase mutants D126G, D156G, D233G, and D295G. They all exhibited reduced or no activities. The activity of

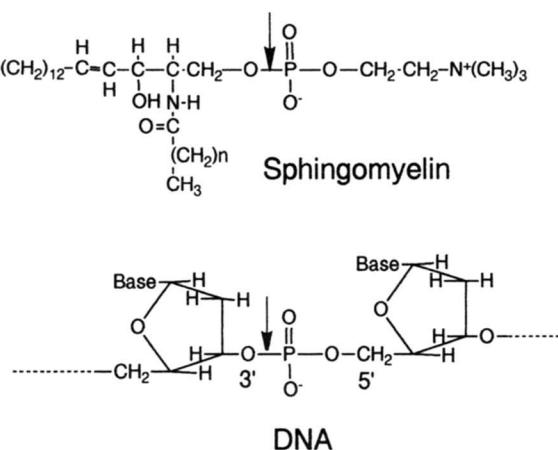


Fig. 3. SMase and DNase I both catalyze the hydrolysis of the phosphodiester bonds (indicated by arrows) of their substrates, SM and DNA.

Table 2. Predicted functional roles of SMase residues^a

SMase residue ^b	Predicted role	DNase I residue ^c
Glu 53	Metal ion binding; stabilization of pentacovalent transition state	Glu 39
Asp or Asn 100 ^d	Hydrogen bond to His 151	Glu 78
Asp 126	Substrate recognition ^e	Asp 107
His 151	General acid in hydrolysis	His 134
Asp 156	Substrate recognition ^e	Asp 139
Asp 195	Interaction with phosphate group	Asp 168
Asn 197	Interaction with phosphate group	Asn 170
Asp 233	Hydrogen bond to His 296	Asp 212
Asp 295	Stabilization of pentacovalent transition state	Asp 251
His 296	General base in hydrolysis	His 252

^aThe prediction is based on the sequence comparison with DNase I, whose functional residues have been identified using the X-ray structure by Weston et al. (1992).

^bResidue numbering is based on the *B. cereus* mature SMase sequence (Yamada et al., 1988).

^cEquivalent residues of DNase I.

^dAsp in *B. cereus* and *S. aureus* (Projan et al., 1989); Asn in *L. interrogans* (Segers et al., 1990).

^eAccording to the mutational analysis by Tamura et al. (1995).

D295G was completely abolished. This indicates that Asp 295, which was predicted to stabilize the pentacovalent transition state (Table 2), is essential for the catalysis. From the analysis of the hydrolytic activities of D126G and D156G toward four phosphorylcholine-containing substrates with different hydrophobicities—SM, HNP, lysophosphatidylcholine, and *p*-nitrophenylphosphorylcholine—Tamura et al. (1995) proposed that Asp 126 and Asp 156 may be involved in the substrate recognition rather than the hydrolysis. This is consistent with the structure prediction, where the two aspartate residues were located a little apart from the catalytic center (Fig. 2). The two aspartate residues are conserved in DNase I (Asp 107 and Asp 139; see Fig. 1), which suggests that the residues may also play some role in the substrate binding of DNase I.

A structural model for SMase-SM complex

As described above, the structure prediction, together with the mutant analysis, indicated that the bacterial SMase could adopt the

Table 3. Effect of amino acid substitutions at position 151 or 296 on the hydrolytic activity of SMase^a

SMase type	SM	HNP
Wild type	299	2.6
H151A	3.9	0.15
H296A	1.78	0.14
Vector	2.11	0.07

^aThe hydrolytic activity in the extract of sonically disrupted *E. coli* TG1 cells harboring the wild-type or mutant genes, or plasmid pUC119 (vector) was measured with SM or HNP as a substrate and the activity was expressed as units per mg of bacterial protein.

same catalytic mechanism with that of the mammalian DNase I. Another question is how the SMase recognizes its substrate, SM, specifically. To investigate the substrate specificity, we have built a three-dimensional structural model for the *B. cereus* SMase-SM complex based on the X-ray crystal structure of bovine DNase I-DNA complex.

First, a structure of the SMase was modeled with the DNase I structure as a template. As shown in Figure 1, the conserved active residues (boxed) are distributed over the sequence, which suggests that the folds of SMase and DNase I are globally similar. In the model building, the backbone conformations of those regions aligned with DNase I (Fig. 1) were borrowed from DNase I, and they were basically not changed. However, surface loop regions of SMase, including the loop containing Cys 123, may have conformations quite different from those of DNase I. The four insertions (residues 82–87, 92–95, 176–185, and 269–293) were not modeled, because they were apart from the active site and did not affect the discussion of possible SM binding residues. The residues Cys 123 and Cys 159, which are known to form a disulfide bond (Tomita et al., 1993), were initially 15.0 Å apart from each other. By energy minimization and molecular dynamics calculations, the flexible loop region containing Cys 123 was modified conformationally to allow the disulfide bond formation. The distance between the cysteine residues was 3.4 Å in the final model.

Next, a structural model for SM was built according to standard bond lengths, angles, and dihedral angles. The SM structure was then rotated to superimpose its phosphate group onto the corresponding scissile phosphate group of the DNA, and the coordinate data of the SM structure were combined with those of the SMase structure. The initial structural model for the SMase-SM complex was thus derived. The model was then refined to remove van der Waals clashes between the SMase and the SM.

From the structural model for the complex (Fig. 4A,B), a set of SMase residues that may be involved in the binding of SM has been identified (Fig. 5). In the model, the hydrophobic tail of the sphingosine part of the SM is sandwiched between the aromatic rings of the side chains of Trp 28 and Phe 55. The hydroxyl group of the tail is hydrogen bonded to the side chain of Glu 53. The amide group of the tail is hydrogen bonded to the side chain of Gln 153. The carbonyl oxygen of the tail may be hydrogen bonded to the side chain of Glu 53 via a water. Alternatively, it can be hydrogen bonded to the side-chain amide group of Asn 130. There may be a salt bridge between the positive charge of the SM head group and the side chain of Asp 295. The side chain of Trp 232 is in van der Waals contact with the methyl groups of the SM head. These residues are all conserved in all known bacterial SMase homologues, although Glu 53 and Asp 295 are conserved in DNase I as well. The DNase I residues corresponding to Trp 28, Phe 55, Gln 153, and Trp 232 are replaced by Thr, Arg, Ala, and Tyr, respectively, and the residue corresponding to Asn 130 is deleted.

DNase I binds DNA mainly through hydrogen bonds and electrostatic interactions (Weston et al., 1992). When the bound DNA octamer is removed from the DNase I-DNA complex structure (1DNK; Weston et al., 1992), the following 32 residues of DNase I increase their water-accessible surface areas by more than 0.5 Å²: Arg 9, Thr 10, Gly 12, Glu 13, Thr 14, Glu 39, Arg 41, Asp 42, Ser 43, His 44, Gly 72, Arg 73, Asn 74, Ser 75, Tyr 76, Glu 78, Ser 110, Arg 111, His 134, Ala 136, Pro 137, Asp 168, Asn 170, Ser 174, Tyr 175, Thr 203, Thr 205, Ser 206, Thr 207, Tyr 211, Asp 251, and His 252. These are all hydrophilic, except for Gly 12, Gly 72, and Ala 136.

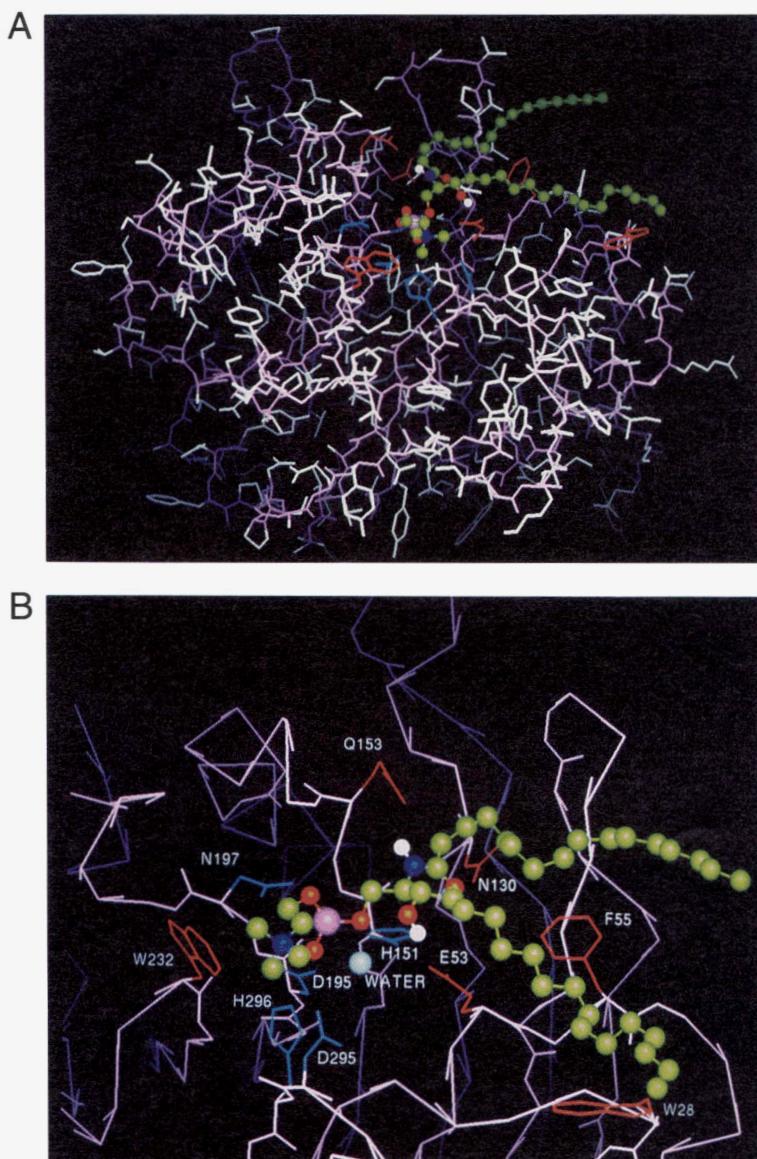


Fig. 4. **A:** Structural model for the *B. cereus* SMase–SM complex. The backbone of the SMase structure is shown in pink, and the side chains, in white. Side chains of residues that were predicted to be important in the catalysis and the substrate recognition are shown in navy and red, respectively. The SM structure is shown in ball-and-stick, where carbon atoms are shown in green; nitrogens, in blue; oxygens, in red; hydrogens, in white; and a phosphorus, in pink. **B:** An enlarged view of the substrate binding region in the model.

On the other hand, in our SMase–SM complex model, the following 24 residues of SMase increase their accessible surface areas when the substrate is removed: Asn 23, Pro 26, Trp 28, Glu 53, Phe 55, Asp 56, Gly 91, Thr 97, Pro 98, Glu 99, Asp 100, Ser 129, Asn 130, His 151, Leu 152, Gln 153, Ala 154, Asp 195, Asn 197, Asn 202, Trp 232, Tyr 266, Asp 295, and His 296. These include four hydrophobic residues: Trp 28, Phe 55, Leu 152, and Trp 232. Leu 152 is not in van der Waals contact with the substrate SM, and the exposed surface area in its side chain is very small (1 \AA^2). The other three (Trp 28, Phe 55, and Trp 232) play important roles in the binding of SM in our model, as already described. The exposed areas (accessibility) of their side chains increase from 61 to 144 \AA^2 (28 to 66%), from 16 to 91 \AA^2 (9 to 52%), and from 37 to 76 \AA^2 (17 to 35%), respectively, on the removal of the SM. Here, the accessibility indicates the ratio of the water-accessible area to that in the tripeptide Gly-Xaa-Gly.

Among the 29 hydrophilic residues in the DNA-binding surface of DNase I listed above, the four residues Thr 14, Arg 41, Ser 174, and Tyr 211 are replaced by hydrophobic residues in SMase; the

replacements are Trp 28, Phe 55, Ile 201, and Trp 232, respectively (Fig. 1).

These changes in the surface region are consistent with the difference between the substrates. The substrate of the SMase, SM, is a component of the lipid bilayer of the host cell membrane. Unlike DNA, it is an amphipathic molecule. The existence of hydrophobic residues in the surface region would help SMase to bind the amphipathic substrate efficiently.

These results from the modeling of the complex structure should be useful for planning further experiments, such as site-directed mutagenesis, to identify residues important for the substrate recognition.

Evolutionary relationship between SMase and DNase I

The predicted structural similarity (Table 1; Fig. 2), the similarity in the enzymatic reaction (both SMase and DNase I are phosphodiesterases; Fig. 3), the conservation of functionally important residues (Fig. 1; Tables 2, 3), and the existence of a conserved

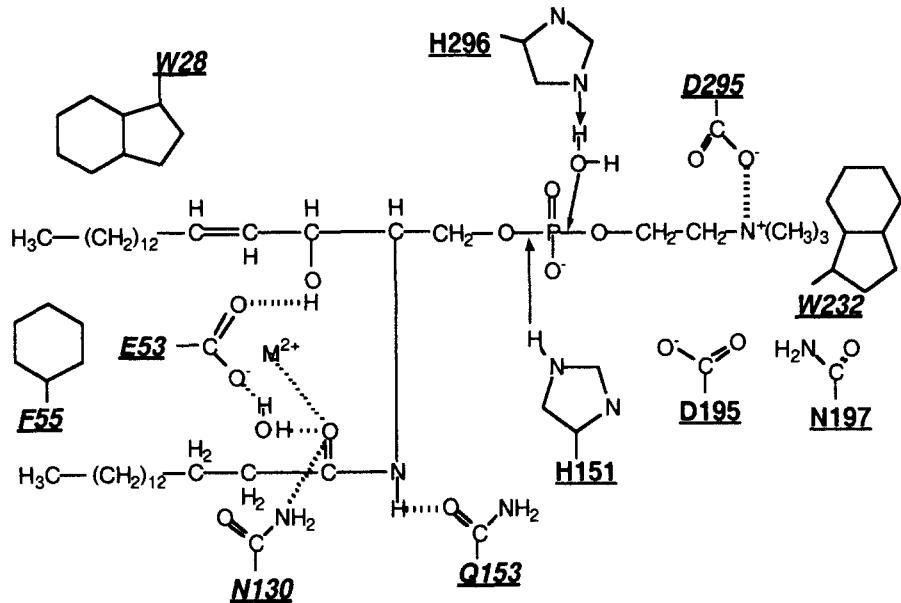


Fig. 5. A model for the interaction between SMase and SM. These residues are all conserved among known bacterial SMase homologues. The structural model for the SMase-SM complex suggested that the residues W28, E53, F55, N130, Q153, W232, and D295 (shown in italic) might be involved in the recognition of the substrate. The other residues, including E53 and D295, should be involved in the hydrolysis of the phosphodiester bond. M^{2+} denotes the metal ion that is essential in the catalysis.

pentapeptide sequence motif (Fig. 1) suggest a distant evolutionary relationship between bacterial SMase and mammalian DNase I. It is probable that they originated from a common ancestral phosphodiesterase and acquired the specificity for different substrates during evolution, although we can never deny the possibility that they have these similarities in sequence and function by chance or that they evolved these similarities convergently.

A similar example of evolutionary relationship was reported by Volbeda et al. (1991). They found a distant evolutionary relationship between two phosphodiesterases, *B. cereus* PC-PLC (Hough et al., 1989) and *Penicillium citrinum* P1 nuclease (Lahm et al., 1990; Volbeda et al., 1991). The two enzymes show no significant overall sequence similarity, but their known three-dimensional structures and catalytically important residues are well conserved (Volbeda et al., 1991).

If the relationship inferred in the present study between SMase and DNase I is really the case, we can see a parallelism of two pairs of phosphodiesterases, two phospholipases (SMase and PC-PLC) at one side and two nucleases (DNase I and P1 nuclease) at the other side. The X-ray structures of DNase I and P1 nuclease are totally different; DNase I contains a large amount of β -strands, whereas P1 nuclease consists dominantly of α -helices. Therefore, the two pairs of phosphodiesterases (SMase and DNase I, and PC-PLC and P1 nuclease) represent independent evolutionary events. The divergence of phosphodiesterases with specificity for phospholipids and nucleic acids from an ancestral phosphodiesterase may be a general evolutionary mechanism.

Materials and methods

Prediction of the SMase structure

The structure of the *B. cereus* SMase was predicted using the fold recognition method developed by us (Nishikawa & Matsuo, 1993; Matsuo & Nishikawa, 1994b; Matsuo et al., 1995). In this method, the compatibility of a sequence with a structure is evaluated using the following four functions: side-chain packing, solvation, hydrogen bonding, and local conformation functions. The details of the

evaluation functions are described elsewhere (Nishikawa & Matsuo, 1993; Matsuo et al., 1995). The fold recognition method has been implemented in the computer program named COMPASS (protein sequence-structure COMPAtibility Search System).

Using the program COMPASS, the *B. cereus* SMase sequence (Yamada et al., 1988) was compared with 325 known three-dimensional structures taken from the Protein Data Bank (Bernstein et al., 1977). The 325 proteins had less than 30% sequence identities with one another. The structures were sorted in the order of their compatibility scores S_{tot} (Table 1). The scores were calculated as described in Matsuo and Nishikawa (1994b). The structure that showed the best score was assumed to be the most similar to that of the SMase.

A structural model for the SMase-SM complex

The model for the *B. cereus* SMase structure was built according to standard model building procedures. The X-ray structure of bovine DNase I (Oefner & Suck, 1986, PDB code 3DNI; and Weston et al., 1992, PDB code 1DNK) was used as a template for the model. The alignment of the SMase and DNase I sequences was given by the program COMPASS. The structures of loop regions where residue insertions or deletions occurred were modeled by the conventional loop search method (Nakamura et al., 1991). The initial side-chain conformations were built using the dead-end elimination theorem (Desmet et al., 1992; Tanimura et al., 1994).

An initial structural model for SM was constructed using Insight II (Biosym, San Diego, California). The dihedral angles of the hydrocarbon chains of the tail of SM were initially set all *trans*. The SM structure was then combined with the SMase structure using the X-ray structure of the DNase I-DNA complex (Weston et al., 1992; PDB code, 1DNK) as a reference. The SM structure was rotated to superimpose its phosphate group onto the corresponding phosphate group of DNA using the method of Kabsch (1976).

The model was refined to remove van der Waals clashes. In the course of the model building, the model was also checked visually

using Insight II. The program PRESTO (Morikami et al., 1992) was used to perform energy minimization calculations and molecular dynamics simulations with the AMBER all atom force field (Weiner et al., 1986).

Site-directed mutagenesis and assays of hydrolyzing activities of SMase

The H151A and H296A mutants of *B. cereus* SMase were constructed by the oligonucleotide-directed PCR-overlap extension method (Maruta et al., 1991). Each mutant gene was sequenced to prove that only the mutation expected had occurred. *Escherichia coli* TG1 cells carrying a mutant gene on pUC119 were grown in 2× YT broth at 37 °C under aeration until the late logarithmic phase. Crude SMase was prepared by sonic disruption followed by centrifugation for 10 min at 15,000 × g and 4 °C. The supernatant was employed as a crude enzyme preparation. The hydrolytic activity of the wild and mutant enzymes toward bovine brain SM or HNP were measured as described (Tamura et al., 1995). One unit of the SMase activity was defined as the amount of enzyme that catalyzes the hydrolysis of 1 μmol substrate/min at pH 7.5 and 37 °C.

Acknowledgments

We thank Dr. Kosuke Morikawa for valuable suggestions and discussions, and Drs. Ryo Taguchi and Masahiro Tomita for critical comments. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410.
- Bernstein FC, Koetzle TF, Williams GTB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol* 112:535–542.
- Bowie JU, Lüthy R, Eisenberg D. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253:164–170.
- Bryant SH, Lawrence CE. 1993. An empirical energy function for threading protein sequence through the folding motif. *Proteins Struct Funct Genet* 16:92–112.
- Desmet J, De Maeyer M, Hazes B, Lasters I. 1992. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature* 356:539–542.
- Gilmore MS, Cruz-Rodz AL, Leimeister-Wachter M, Kreft J, Goebel W. 1989. A *Bacillus cereus* cytolytic determinant, cereolysin AB, which comprises the phospholipase C and sphingomyelinase genes: Nucleotide sequence and genetic linkage. *J Bacteriol* 171:744–753.
- Godzik A, Kolinski A, Skolnick J. 1992. Topology fingerprint approach to the inverse protein folding problem. *J Mol Biol* 227:227–238.
- Hannun YA. 1994. The sphingomyelin cycle and the second messenger function of ceramide. *J Biol Chem* 269:3125–3128.
- Hannun YA, Bell RM. 1989. Functions of sphingolipids and sphingolipid breakdown products in cellular regulation. *Science* 243:500–507.
- Higgins DG, Bleasby AJ, Fuchs R. 1992. CLUSTAL V: Improved software for multiple sequence alignment. *CABIOS* 8:189–191.
- Hough E, Hansen LK, Birknes B, Jyng K, Hansen S, Hordvik A, Little C, Dodson E, Derewenda Z. 1989. High-resolution (1.5 Å) crystal structure of phospholipase C from *Bacillus cereus*. *Nature* 338:357–360.
- Johansen T, Haugli FB, Ikezawa H, Little C. 1988. *Bacillus cereus* strain SE-1: Nucleotide sequence of the sphingomyelinase C gene. *Nucleic Acids Res* 16:10370.
- Jones DT, Taylor WR, Thornton JM. 1992. A new approach to protein fold recognition. *Nature* 358:86–89.
- Kabsch W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr A* 32:922–923.
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.
- Kocher JPA, Rooman MJ, Wodak SJ. 1994. Factors influencing the ability of knowledge-based potentials to identify native sequence–structure matches. *J Mol Biol* 235:1598–1613.
- Kolesnick R. 1992. Ceramide: A novel second messenger. *Trends Cell Biol* 2:232–236.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 24:946–950.
- Lahm A, Volbeda A, Suck D. 1990. Crystallisation and preliminary crystallographic analysis of P₁ nuclease from *Penicillium citrinum*. *J Mol Biol* 215:207–210.
- Lemer CMR, Rooman MJ, Wodak SJ. 1995. Protein structure prediction by threading methods: Evaluation of current techniques. *Proteins Struct Funct Genet* 23:337–355.
- Maiorov VN, Crippen GM. 1992. Contact potential that recognizes the correct folding of globular proteins. *J Mol Biol* 227:876–888.
- Maruta H, Holden J, Sizeland A, D'Abaco G. 1991. The residues of Ras and Rap proteins that determine their GAP specificities. *J Biol Chem* 266:11661–11668.
- Matsuo Y, Nakamura H, Nishikawa K. 1995. Detection of protein 3D-1D compatibility characterized by the evaluation of side-chain packing and electrostatic interactions. *J Biochem* 118:137–148.
- Matsuo Y, Nishikawa K. 1994a. Prediction of the structural similarity between spermidine/putrescine-binding protein and maltose-binding protein. *FEBS Lett* 245:23–26.
- Matsuo Y, Nishikawa K. 1994b. Protein structural similarities detected by a sequence–structure compatibility method. *Protein Sci* 3:2055–2063.
- Matsuo Y, Nishikawa K. 1995. Assessment of a protein fold recognition method that takes into account four physicochemical properties: Side-chain packing, solvation, hydrogen-bonding, and local conformation. *Proteins Struct Funct Genet* 23:370–375.
- Morikami K, Nakai T, Kidera A, Saito M, Nakamura H. 1992. PRESTO (PROtein Engineering SimulaTOr): A vectorized molecular mechanics program for biopolymers. *Computers Chem* 16:243–248.
- Moult J, Pedersen JT, Judson R, Fidelis K. 1995. A large-scale experiment to assess protein structure prediction methods. *Proteins Struct Funct Genet* 23:ii–iv.
- Nakamura H, Katayanagi K, Morikawa K, Ikehara M. 1991. Structural models of ribonuclease H domain in reverse transcriptases from retroviruses. *Nucleic Acids Res* 19:1817–1823.
- Nishikawa K, Matsuo Y. 1993. Development of pseudoenergy potentials for assessing protein 3D-1D compatibility and detecting weak homologies. *Protein Eng* 6:811–820.
- Nishikawa K, Noguchi T. 1991. Predicting protein secondary structure based on amino acid sequence. *Methods Enzymol* 202:31–44.
- Oefner C, Suck D. 1986. Crystallographic refinement and structure of DNase I at 2 Å resolution. *J Mol Biol* 192:605–632.
- Okazaki T, Bell RM, Hannun YA. 1989. Sphingomyelin turnover induced by vitamin D₃ in HL-60 cells: Role in cell differentiation. *J Biol Chem* 264:19076–19080.
- Ouzounis C, Sander C, Scharf M, Schneider R. 1993. Prediction of protein structure by evaluation of sequence–structure fitness: Aligning sequences to contact profiles derived from three-dimensional structures. *J Mol Biol* 232:805–825.
- Projan SJ, Kornblum J, Kreiswirth B, Moghazeh SL, Eisner W, Novick RP. 1989. Nucleotide sequence: The β-hemolysin gene of *Staphylococcus aureus*. *Nucleic Acids Res* 17:3305.
- Roberts CJ, Pohlig G, Rothman JH, Stevens TH. 1989. Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an integral membrane glycoprotein of the yeast vacuole. *J Cell Biol* 108:1363–1373.
- Segers RPAM, van der Drift A, de Nijs A, Corcione P, van der Zeijst BAM, Gaastraa W. 1990. Molecular analysis of a sphingomyelinase C gene from *Leptospira interrogans serovar hardjo*. *Infect Immun* 58:2177–2185.
- Sippl MJ. 1990. Calculation of conformational ensembles from potentials of mean force: An approach to the knowledge-based prediction of local structures in globular proteins. *J Mol Biol* 213:859–883.
- Suck D, Lahm A, Oefner C. 1988. Structure refined to 2 Å of a nicked DNA octanucleotide complex with DNase I. *Nature* 332:464–468.
- Suck D, Oefner C. 1986. Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature* 321:620–625.
- Tamura H, Noto M, Kinoshita K, Ohkuma S, Ikezawa H. 1994. Inhibition of NGF-induced neurite outgrowth of PC12 cells by *Bacillus cereus* sphingomyelinase, a bacterial hemolysin. *Toxicol* 32:629–633.
- Tamura H, Tameishi K, Yamada A, Tomita M, Matsuo Y, Nishikawa K, Ikezawa H. 1995. Mutation in aspartic acids modifies catalytic and hemolytic activities of *Bacillus cereus* sphingomyelinase. *Biochem J* 309:757–764.
- Tanimura R, Kidera A, Nakamura H. 1994. Determinants of protein side-chain packing. *Protein Sci* 3:2358–2365.
- Tomita M, Taguchi R, Ikezawa H. 1991. Sphingomyelinase of *Bacillus cereus* as a bacterial hemolysin. *J Toxicol-Toxin Rev* 10:169–207.

- Tomita M, Ueda Y, Tamura H, Taguchi R, Ikezawa H. 1993. The role of acidic amino-acid residues in catalytic and adsorptive sites of *Bacillus cereus* sphingomyelinase. *Biochim Biophys Acta* 1203:85–92.
- Volbeda A, Lahm A, Sakiyama F, Suck D. 1991. Crystal structure of *Penicillium* P1 nuclease at 2.8 Å resolution. *EMBO J* 10:1607–1618.
- Weiner SJ, Kollman PA, Nguyen DT, Case DA. 1986. An all atom force field for simulations of proteins and nucleic acids. *J Comput Chem* 7:230–252.
- Weston SA, Lahm A, Suck D. 1992. X-ray structure of the DNase I-d(GGTATACC)₂ complex at 2.3 Å resolution. *J Mol Biol* 226:1237–1256.
- Wiegmann K, Schütze S, Machleidt T, Witte D, Krönke M. 1994. Functional dichotomy of neutral and acidic sphingomyelinases in tumor necrosis factor signaling. *Cell* 78:1005–1015.
- Wilmanns M, Eisenberg D. 1993. Three-dimensional profiles from residue-pair preferences: Identification of sequences with β/α -barrel fold. *Proc Natl Acad Sci USA* 90:1379–1383.
- Yamada A, Tsukagoshi N, Udaka S, Sasaki T, Makino S, Nakamura S, Little C, Tomita M, Ikezawa H. 1988. Nucleotide sequence and expression in *Escherichia coli* of the gene coding for sphingomyelinase of *Bacillus cereus*. *Eur J Biochem* 175:213–220.