Molecular modeling of human alkaline sphingomyelinase

Panneer Selvam Suresh[1], Olujide Olubiyi[2], Chinnasamy Thirunavukkarasu[3], Birgit Strodel[2], Muthuvel Suresh Kumar[1]*

[1]Centre of Excellence in Bioinformatics, School of Life Sciences, Pondicherry University, Pondicherry, India; [2]Institute of Structural Biology and Biophysics / ISB-3, Research Centre Julich, 52425 Julich, Germany; [3]Department of Biochemistry and Molecular Biology, School of Life Sciences, Pondicherry University, Pondicherry, India; Muthuvel Suresh Kumar - Email: suresh.bic@pondiuni.edu.in, muthuvels@hotmail.com; Phone: +91 413 2655583; Fax: +91 413 2655211; *Corresponding author

Abstract:
Alkaline sphingomyelinase, which is expressed in the human intestine and hydrolyses sphingomyelin, is a component of the plasma and the lysosomal membranes. Hydrolase of sphingomyelin generates ceramide, sphingosine, and sphingosine 1-phosphate that have regulatory effects on vital cellular functions such as proliferation, differentiation, and apoptosis. The enzyme belongs to the Nucleotide Pyrophosphatase/Phosphodiesterase family and it differs in structural similarity with acidic and neutral sphingomyelinase. In the present study we modeled alkaline sphingomyelinase using homology modeling based on the structure of Nucleotide Pyrophosphatase/Phosphodiesterase from *Xanthomonas axonopodis* with which it shares 34% identity. Homology modeling was performed using Modeller9v7. We found that Cys78 and Cys394 form a disulphide bond. Further analysis shows that Ser76 may be important for the function of this enzyme, which is supported by the findings of Wu *et al*. (2005), that S76F abolishes the activity completely. We found that the residues bound to Zn^{2+} are conserved and geometrically similar with the template. Molecular Dynamics simulations were carried out for the modeled protein to observe the effect of Zinc metal ions. It was observed that the metal ion has little effect with regard to the stability but induces increased fluctuations in the protein. These analyses showed that Zinc ions play an important role in stabilizing the secondary structure and in maintaining the compactness of the active site.

Keywords: alkaline sphingomyelinase; homology modeling, NPP; gromacs

Background:
Sphingomyelin (SM) is a component of plasma and lysosomal membranes. Hydrolysis of SM generates ceramide, sphingosine, and sphingosine 1-phosphate that have regulatory effects on vital cellular functions such as proliferation, differentiation, and apoptosis [1-3]. Sphingomyelinase (SMase) hydrolyses SM, and so far five types of SMase have been identified. Alkaline sphingomyelinase (alk-SMase) is an old enzyme identified in the sphingomyelinase family. There is no structural similarity with another SMase and it belongs to the Nucleotide Pyrophosphatase/Phosphodiesterase family. It has specific properties such as bile salt dependency, trypsin resistance, high stability and tissue specific expression [4-5]. Alk-SMase is down regulated in colon cancer and supplementation of SM in the diet inhibits the formation of colonic aberrant crypt foci and colon tumors in animal studies [6]. However, hydrolysis of SM by alk-SMase generates antiproliferative messengers, such as ceramide and sphingosine, and proliferative messengers, such as sphingosine 1-phosphate, ceramide phosphate and lyso-sphingomyelin [7]. Alk-SMase can also hydrolyze lyso-phosphatidylcholine (lyso-PC) to monoacylglycerol, with phospholipase C activity. Lyso-PC is considered being mitogenic under the action of phospholipase D, which produces lysophosphatidic acid (LPA) from lyso-PC, leading to cell proliferation through different types of G protein coupled receptors. It is assumed that alk-SMase, by cleaving off the phosphocholine headgroup from lyso-PC, reduces the formation of LPA and

protects cells from tumorigenesis [8], [9] and [10]. Recent findings by Wu et al. (2005) stated that alk-SMase could hydrolyse platelet activating factor (PAF) and inactivate them. Alk-SMase may generate the proliferative signals by generating ceramide, reducing LPA formation, and inactivating PAF [11].

A direct product of SMase is ceramide, which has a variety of physiological functions including apoptosis, cell growth arrest, differentiation, cell senescence, cell migration and adhesion. On the other hand, downstream metabolites of ceramide are involved in a number of pathological states, including cancer, neurodegeneration, diabetes, microbial pathogenesis, obesity and inflammation. Inhibition of SMase activity can suppress pathological conditions including cancer. Recently, a model for alk-SMase is proposed by Duan *et al*.[12]. In the present study, the 3D structures of human alk-SMase have been predicted by homology modeling. We studied the effect of zinc metal ion on alk-SMase by molecular dynamics simulations.

Materials and methodology:
The sequence of human alk-SMase was obtained from the NCBI (National Center for Biotechnology Information) protein sequence database (GI: 33440070). The sequence was submitted to the fold recognition server Genesilico Metaserver [https://genesilico.pl/meta2/]. The obtained template was further submitted for template target alignment using ClustalW

BIOINFORMATION

[http://www.ebi.ac.uk/Tools/msa/clustalw2/], Multalign [http://multalin.toulou se.inra.fr/multalin/], T-Coffee [http://www.ebi.ac.uk/Tools/msa/tcoffee/] and Muscle [http://www.ebi.ac.uk/Tools/msa/muscle/]. Modeller was used for comparative modeling [18]. Loops were modeled using ModLoop [http://modbase.compbio.ucsf.edu/modloop/]. The resulting alk-SMase model was evaluated using the structure evaluation software Verify3D [http://nihserver.mbi.ucla.edu/Verify_3D/] and PROCHECK [http://www.ebi. ac.uk/thornton-srv/software/PROCHECK/]. Graphical presentations of the 3D model were prepared using Chimera [http://www.cgl.ucsf.edu/chimera/]. The model was submitted to the protein modeling database as PM0076830 [13].

Molecular dynamics simulations:
Two molecular dynamics simulations were performed for the alk-SMase model as obtained from homology modeling, one in the absence of Zn^{2+}, and the other one in the presence of Zn^{2+} ions. The setup and parameters were chosen to be identical for both systems. The protein coordinates were first stripped off the hydrogen atoms using the program Openbabel (http://openbabel.org). Since the united atom GROMOS force field (ffG43a2) was used we first removed all H atoms and then readded the polar H atoms with charge states of ionisable and terminal residues selected for a neutral pH. The GROMACS distribution 4.0.3 [14] was used for the simulations. A cubic box was set up and the molecules positioned at the centre of the box by defining a 10 Å distance between the protein and the box edge. A steepest descent minimization was carried out in vacuum. The systems were then solvated using the simple point charge (SPC) explicit water model and Na^+ and Cl^- ions were added to obtain charge neutrality and a near salt physiological concentration of 0.15 M. Two energy minimization methods were then applied to the solvated system, that is a steepest descent minimization followed by a conjugate gradient minimization. The molecular dynamics runs were carried out in the NPT ensemble for a total of 6 ns for each system. The first 1 ns involved restraining forces of 1000 $kJ.mol^{-1}.nm^{-2}$ applied to the heavy atoms of the protein in order to allow the solvent molecules to equilibrate around the protein and to remove bad contacts. During the equilibration dynamics the system was coupled to the Berensden thermostat to maintain a temperature of 300K and to the Berensden barostat to maintain a pressure of 1 bar. A spherical cut-off of 1.4 nm was used for the computation of short-range nonbonded interactions and the particle mesh-ewald (PME) method was used for computing long-range electrostatic interactions. After 1 ns, all restraints were removed from the protein heavy atoms and simulation data collected during the 5 ns production runs. Here, protein atoms and non-protein atoms were separately coupled to a Nose-Hoover thermostat at 300 K, while the Parrinello-Rahman barostat at 1 bar was used for pressure coupling. The collected data were then analyzed using functionalities of GROMACS.

Results and Discussion:
Homology modeling of alkaline sphingomyelinase:
The protein sequence of intestinal human alk-SMase was queried against the Protein Data Bank (PDB) using PSI-BLAST [http://www.ebi.ac.uk/ Tools/sss/psiblast/] and the result shows a 31% identity against the Nucleotide Pyrophosphatase/Phosphodiesterase (NPP) family of proteins. The Critical Assessment of Structure Prediction (CASP) experiments have lead to sufficient knowledge that allows us to build a reasonably good comparative model. There are two factors that affect the accuracy of template derived models: Firstly, the selection of the template, and secondly, target-template alignment errors in the distant comparative modeling [15]. We therefore concentrated on the template selection and target-template alignment by using different methods in order to obtain a reliable protein model for alk-SMase.

We queried the alk-SMase protein sequence to Genesilico Metaserver for template identification which analyses the query using different servers. All servers predicted the Nucleotide Pyrophosphatase/Phosphodiesterase (NPP) protein of *Xanthomonas axonopodis* as target with reliable scores (**Table 1 see Supplementary material**). Three structures of the same NPP protein with PDB codes 2GSO (holoenzyme), 2GSU (Vanadate bound structure), and 2GSN (Adenosine monophosphate (AMP) bound structure) [16] were identified as a template by the servers. We selected 2GSO as the template for homology modeling based on good resolution. Target template alignment was carried out using ClustalW, T-Coffee, Multalign and FFAS. The sequence identity in the alignments between target and template varied from 31% to 34%. Individual alignments (Supplementary Figure 1 available with authors) were given as a input to Modeller to build 3D structures, and the resulting models were evaluated using Verify3D. Verify3D was used for the assessment of the reliability of the protein structures. This program analyzes the compatibility of an atomic model with its own amino acid sequence. Each residue is assessed as

a structural class (alpha, beta, loop, polar, apolar etc.) based on its location and environment. All the alignment models were found to have good Verify3D scores (**Table 2 see Supplementary material**). We chose the T-Coffee alignment model as final model for further studies because for this model 89.95% of the residues have a good Verify3D score. However, the loop region involving residues 243 to 264 showed a low Verify3D score for the T-Coffee model. We therefore used ModLoop to build the loop, and the resulting model yielded 91.75% of the residues with good Verify3D scores. The template structure 2GSO was evaluated as well using Verify3D and we found that 100% of the residues had an average score better than 0.2. From this analysis we can conclude that the modeled structure is a reliable model. As a further test, we subjected our final model to a Ramachandran plot test (PROCHECK) and found that 99.1% of the residues were in allowed regions, while only 0.9% of amino acid residues were in disallowed regions.

Structure of alkaline sphingomyelinase:
Alk-SMase is a novel member of the NPP family of proteins and catalyses the breakdown of SM via the phospholipase C activity. The enzyme shares ~30% similarity with the members of the NPP family but with no detectable NPP activity [5]. Furthermore, alk-SMase shares no similarity with either acidic or neutral SMases. Therefore, alk-SMase is also called NPP7. Alk-SMase was well studied in the work by Duan *et al*. (2006) [4]. However, there is still no structural information available for this enzyme. We therefore modeled alk-SMase by homology modeling. Out of the 458 amino acids in the sequence we modeled only the residues ranging from 29-415 due to insufficient coverage of the template. We were not able to model the hydrophobic transmembrane domain involving both the N and the C terminal of alk-SMase since the residues from 1-30 and 415-458 are predicted as disordered regions. It has been experimentally reported that N terminal results in a signal peptide, when it is cleaved in the mature enzyme, and that the C-terminal is a signal anchor, which hooks the enzyme to the plasma membrane [17]. The immunogold labeling experiments showed that the cellular location of alk-SMase is on the surface of microvillus membranes in the human intestinal tract [17]. It is noteworthy that both the membrane bound form and the secreted form of the enzyme are reported to be active [4]. The predicted glycosylation sites are shown in Fig. 1. It has been shown that the glycosylation not only affects the enzyme activity but also affects the transport of the enzyme to the plasma membrane [17]. The predicted model of alk-SMase is shown in **Figure 1**. The motif sequence TKTFPNH in the active site core of NPP family members is different in the template structure 2GSO [SLTFPNH, 88-94]. The first two amino acids were modified in the conserved motif of NPP. In alk-SMase the corresponding core motif [TMTSPCH, 73-79] is reported to be involved in Sphingomyelin hydrolysis [5].
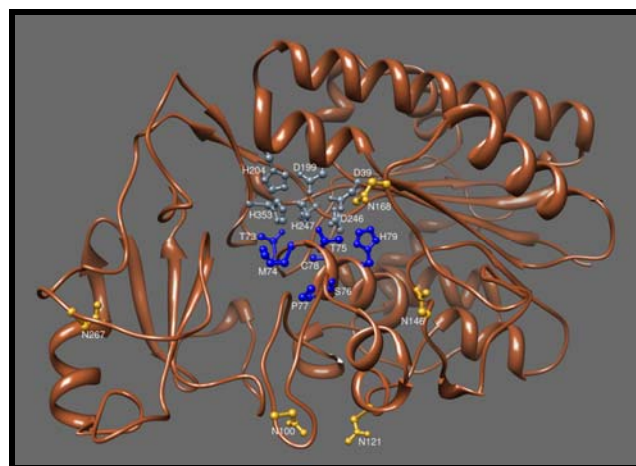

Figure 1: Predicted 3D Structure of alkaline Sphingomyelinase. Sphingomyelinase hydrolyzing motif (73-TMTSPCH-79) is shown in blue color, the metal binding residues are shown in grey color and the glycosylation residues are shown in yellow color. Labels are shown in single letter code with numbering. Amino acids are shown in ball & stick representations.

In **Figure 2** we superimposed the active site of alk-SMase, residues T73, M74, T75, S76, P77, C78 and H79, with the active site of NPP, residues S88, L89, T90, F91, P92, N93 and H94. We observed two main differences in the superimposed structures of template and target. Firstly, a disulphide bond in

alk-SMase is predicted between C78 and C394, which may have an essential role in the enzyme stability. Furthermore, C78 has an important role in substrate specificity. It is to be noted here that C78N mutation reduced the alk-SMase activity but it could not abolish activity completely [18]. The NPP structure, on the other hand, has a disulphide bond between residues C314 and C337, which does not involve in the active site residues. Secondly, there is a Serine at the 76th position of alk-SMase, which is different in the template structure having Phenylalanine instead. It has been reported by Gijsbers et al. (2001) that Phenylalanine in the motif [TKTFPNH] is responsible for the substrate specificity of the NPP family of proteins [19]. The corresponding Serine in alk-SMase might thus have a significant role in substrate specificity. This conclusion is supported by mutation studies of Wu *et al*. (2005), which showed that the S76F mutant abolishes the alk-SMase activity completely [18]. From the superposition of alk-SMase with the NPP template structure we found that the metal coordinating residues (Asp39, Asp199, Asp203, Asp246, His247, and His353) were conserved and also that the metal coordinating sites are geometrically similar. Mutations on the metal coordinating sites of alk-SMase abolish its enzymatic activity, suggesting that these metal binding sites play a vital role in the activity of alk-SMase. From this observation it was inferred that substrate binding sites may form a single pocket along with metal binding sites, because it has been shown that either binding of Zinc or mutations of this site abolish alk-SMase activity.
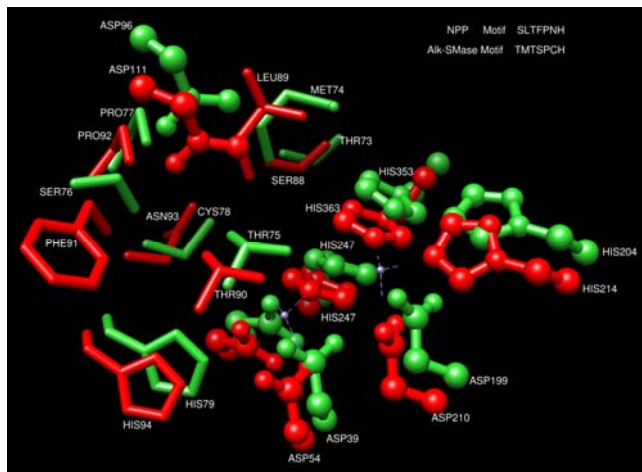

Figure 2: Superimposition of the 3-D models obtained from Chimera (green) compared with their template (PDBID 2GSO) (in red). Only the amino acids side chains are shown. Metal binding amino acid side chains are shown in ball and stick model and the active site motif [TMTSPCH for alk-SMase and SLTFPNH for template 2GSO] are shown in stick representation. Zinc binding to the residues are shown in blue color.

Molecular dynamics simulations:
RMSD, RMSF and SASA analysis:
We performed molecular dynamics simulations to investigate the effect of Zinc ions on the structure of alk-SMase. The main-chain root mean square deviations (RMSDs) were calculated for the trajectories of the two proteins, i.e., alk-SMase with and without Zn^{2+}, from the starting structures as a function of time. The resulting RMSD profiles are shown in **Figure 3a**. The 5-ns trajectory shows that the presence of Zn^{2+} ions has little or no stabilizing effect on the structure. The zinc bound protein structure has even a little higher RMSD compared to the unbound structure. For both structures a major structural change occurred during the initial few picoseconds leading to a RMSD to ~0.2 nm, followed by smaller structural deviations for the rest of both simulations. The final RMSD values are ~0.25-0.28 nm. The RMSDs reach a stable value in 2.5 ns for both structures. The main-chain root mean square fluctuations (RMSFs), calculated over the trajectories for the Zinc bound form and the unbound form, indicate that a large part of the residues fluctuates not more than 0.1 nm (**Figure 3b**). The two proteins display a different fluctuation pattern except for a few regions. In both structures, the N and C terminals have higher fluctuations than the other residues. The Zinc bound residues show similarly small fluctuations as the same residues in the protein without Zinc. However, the rigidity in the metal binding sites induces some degree of fluctuation in some other parts of the structure, leading to differences in the fluctuations observed for alk-SMase with and without Zn^{2+}. Major fluctuation differences are observed in four regions, namely for residues

29, 48-65, 250-265, and 285-310. We conclude from these findings that the Zinc ions induce increased fluctuations in some parts of the structure while not decreasing the overall stability of this structure as judged by the RMSD (see **Figure 3a**). The active site motif (TMTSPCH, residues 73-79) shows only small fluctuations in both structures, i.e., with and without zinc ions. This rigidity in the active site is due to the disulphide bond between residues C78 and C394. The solvent accessible surface area (SASA) analysis for the active site residues are shown in **Figure 3c**. The SASA is the surface area of a biomolecule that is accessible to a solvent and it can be related to the hydrophobic core. The results indicate that the hydrophobic core is more compact in the presence of bound zinc ions compared to the metal unbound form.
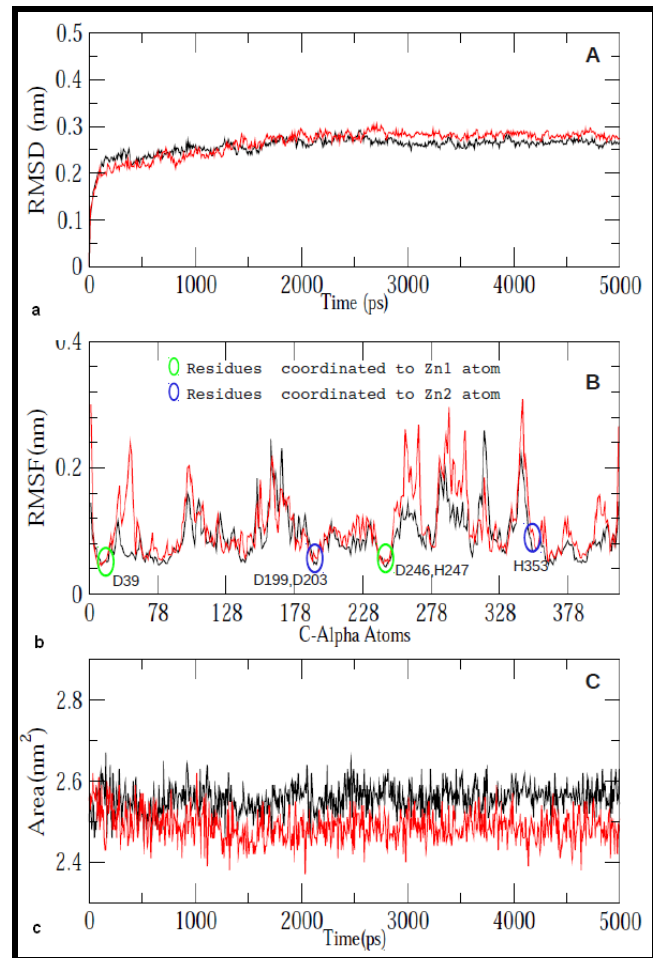

Figure 3: The starting structures of the protein with zinc (shown in red line) and without zinc (show in black line). (**3a**) Root Mean Square Deviations (RMSD); (**3b**) Root Mean Square Fluctuations (RMSF); (**3c**) Accessible Surface Area (ASA)

Effect of Zn^{2+} on the structure of alkaline sphingomyelinase:
To further analyze the effect of the zinc ions on alk-SMase, we superimposed the average structures of both alk-SMase with and without zinc. The results are shown in **Figure 4A & 4B**. Secondary structure changes are observed between the superimposed structure in the regions 43-47, 60-68, 99-104, 115-120, 155-158, 249-253, 282-286, and 348-352. The superimposed structures indicate that the presence of Zinc ions affects the secondary structure. In the absence of zinc ions, alpha helices are lost in the regions 43-47, and 282-286 whereas beta sheets are lost in the regions 60-64 and 155-158. On the other hand, alpha helices are gained in the regions 99-104, 115-120 and beta sheets are gained in the regions 64-68, 249-253, 348-352 and 380-83 in the zinc unbound structure.

In presence of the zinc ions, the ion binding helix region maintains its secondary structure whereas in the absence of these ions it loses the helix region partially, leading to a drift in the active site and a structural deviation of

about 4 Å (Supplementary Figure 2A available with authors). The mobility of the zinc binding residues in the absence of metal ions and the resulting changes of the active site conformation provides an explanation why the mutation on metal binding residues abolishes the enzyme activity. The metal binding residues and the active site motif form together the required conformational scaffold for sphingomyelin to bind, and mutations in either region abolish the enzyme activity. The active site motif helix region was stable during both simulations, i.e., in the presence and absence of the zinc ions, apart from a minor structural deviation of about 1.30 Å in the zinc unbound form (Supplementary Figure 2B available with authors). The results of our simulations show that the Zn ions are important for the overall stability of the active site conformation.
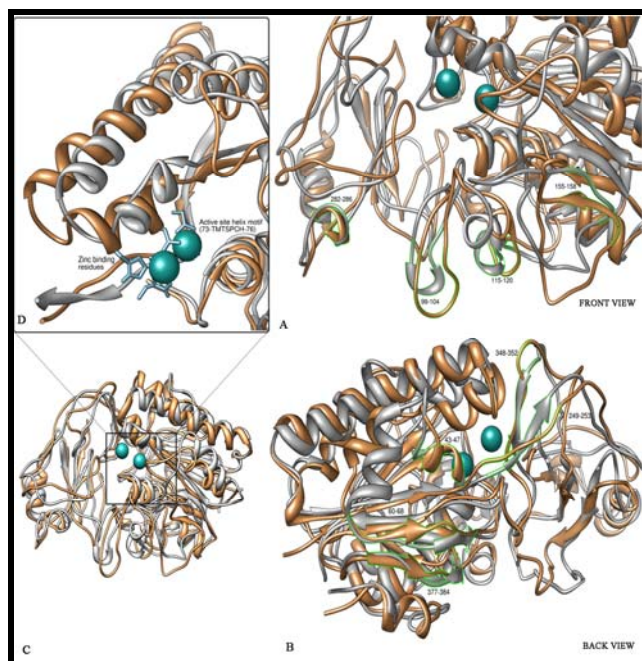

Figure 4: The superimposed average structure in presence of zinc (brown) and in absence of zinc (grey) is shown. Zinc ions are shown in spheres (cyan). Secondary structure changes observed regions are highlighted in green colour.

Conclusion:

This study was carried out to build the model structure of human alkaline Sphingomyelinase on the basis of the crystal structure of Nucleotide Pyrophosphatase/Phosphodiesterase (PDB code 2GSO). The resulting structure was evaluated using Verify3D and PROCHECK and was found to be good. Molecular dynamics simulations confirm that Zn^{2+} plays an important role in the activity of the enzyme. The Zinc metal ion has a crucial role in secondary structure and tertiary structure stabilization.

References:

[1]   Zeisel SH *et al. J Nutr*. 1986 **116**: 50 [PMID: 3944656]
[2]   Hannun YA & Linardic CM. *Biochim Biophys Acta*. 1993 **1154**: 223 [PMID: 8280742]
[3]   Merrill AH *et al. Toxicol Appl Pharmacol*. 1997 **142**: 208 [PMID: 9007051]
[4]   Duan RD. *Biochim Biophys Acta*. 2006 **1761**: 281 [PMID: 16631405]
[5]   Duan RD *et al. J Biol Chem*. 2003 **278**: 38528 [PMID: 12885774]
[6]   Hertervig E *et al. Cancer* 1997 **79**: 448 [PMID: 9028353]
[7]   Meyer zu Heringdorf D *et al. Biochim Biophys Acta*. 2002 **1582**: 178 [PMID: 12069827]
[8]   Lin DA & Boyce JA. *J Immunol*. 2005 **175**: 5430 [PMID: 16210650]
[9]   Yang M *et al. Proc Natl Acad Sci U S A*. 2005 **102**: 6027 [PMID: 15837931]
[10]   Saatian B *et al. Biochem J*. 2006 **393**: 657 [PMID: 16197369]
[11]   Wu J *et al. Biochem J*. 2006 **394**: 299 [PMID: 16255717]
[12]   Duan J *et al. Biochemistry* 2010 **49**: 9096 [PMID: 20839774]
[13]   Castrignano T *et al. Nucleic Acids Res*. 2006 **34**: D306 [PMID: 16381873]
[14]   Van Der Spoel D *et al. J Comput Chem*. 2005 **26**: 1701 [PMID: 16211538]
[15]   Venclovas C & Margelevicius M. *Proteins* 2005 **61** Suppl 7: 99 [PMID: 16187350]
[16]   Zalatan JG *et al. Biochemistry* 2006 **45**: 9788 [PMID: 16893180]
[17]   Duan RD *et al. J Lipid Res*. 2003 **44**: 1241 [PMID: 12671034]
[18]   Wu J *et al. Biochem J*. 2005 **386**: 153 [PMID: 15458386]
[19]   Gijsbers R *et al. J Biol Chem*. 2001 **276**: 1361 [PMID: 11027689]

BIOINFORMATION

open access

Supplementary material:

Table 1: Summary of the fold recognition analysis for alkaline sphingomyelinase

S. No	Threading server	PDBID	Protein name	Reliability*
1	pdbblast	2GSOA	Nucleotide Pyrophosphatase	Reliable
2	blastp	2GSOA	Nucleotide Pyrophosphatase	Reliable
3	ffas	2GSNA	Nucleotide Pyrophosphatase	Reliable
4	hhsearch	2GSOA	Nucleotide Pyrophosphatase	Reliable
5	phyre	2GSUB	Nucleotide Pyrophosphatase	Reliable
6	pcons5	2GSUA	Nucleotide Pyrophosphatase	Reliable

*Each individual servers calculates their own assessment of the reliability of the results whether the fold-recognition alignment is reliable or not. The scores from the different servers rely on different scales. Reliability threshold values are given in **Table 3**.

Table 2: Protein Validation scores from Verify 3d and Procheck

S.No	Alignment tool	Verify3d	Ramachandran plot			
			Core	Allowed	Generously allowed	Disallowed
1	Ffas	81.40%	88%	9.0%	2.4%	0.6%
2	Clustalw	86.08%%	86.0%	11.0%	1.5%	1.5%
3	Tcoffee	91.75%	86.6%	11.0%	1.5%	0.9%
4	Muscle	78.81%	88.4%	9.6%	1.2%	0.9%
5	Mulalign	78.81%	88.4%	9.6%	1.2%	0.9%

Table 3: Reliability score values

S. No	Threading Server	Scores for Alk-SMase	Reliable	Unreliable
1	pdbblast	1e-154	< 2e-06	> 0.023
2	blastp	3e-50	< 2e-06	> 0.023
3	ffas	-99	<-34.5	> -8.5
4	hhsearch	E value is 0	E value < 1	E value < 1
5	phyre	5.1e-27	<0.085	> 0.27
6	pcons5	6.8845	>2.17	< 1.03