http://www.jstage.jst.go.jp/browse/biophysics

BIOPHYSICS

Vol. 2, pp. 23–34 (2006)
doi: 10.2142/biophysics.2.23

Possible roles of S···O and S···N interactions in the functions and evolution of phospholipase A_2

Michio Iwaoka[1] and Noriyoshi Isozumi[1]

[1]*Department of Chemistry, School of Science, Tokai University, Kitakaname, Hiratsuka-shi, Kanagawa 259-1292, Japan*

Received 7 March, 2005; Accepted 30 January, 2006

To investigate possible roles of S···X (X=O, N, S) interactions in the functions and evolution of a protein, two types of database analyses were carried out for a vertebrate phospholipase A_2 (PLA_2) family. A comprehensive search for close S···X contacts in the structures retrieved from protein data bank (PDB) revealed that there are four common S···O interactions and one common S···N interaction for the PLA_2 domain group (PLA_2-DG), while an additional three S···O interactions were found for the snake PLA_2 domain group ($sPLA_2$-DG). On the other hand, a phylogenetic analysis on the conservation of the observed S···O and S···N interactions over various amino acid sequences of $sPLA_2$-DG demonstrated probable clustering of the interactions on the dendrogram. Most of the interactions characterized for PLA_2 were found to reside in the vicinity of the active site and to be able to tolerate the conformational changes due to the substrate binding. These observations suggested that the S···X interactions play some role in the functions and evolution of the PLA_2 family.

Key words: S···X interaction, disulfide bond, molecular evolution, dendrogram, protein data bank

Abbreviations: PLA_2, phospholipase A_2; PLA_2-DG, phospholipase A_2 domain group; $sPLA_2$-DG, snake phospholipase A_2 domain group; PDB, protein data bank; SS, disulfide; SSC, a type of the divalent sulfur atom involved in a half-cystine residue having an S–S–C linkage; CSC, a type of the divalent sulfur atom involved in a methionine residue having a C–S–C linkage; SCOP, Structural Classification of Protein; PDN, protein domain name; WT, wild type.

Corresponding author: Michio Iwaoka, Department of Chemistry, School of Science, Tokai University, Kitakaname, Hiratsuka-shi, Kanagawa 259-1292, Japan. e-mail: miwaoka@tokai.ac.jp.

Weak nonbonded interactions are important physicochemical forces that control the structure and function of proteins[1]. Ionic interactions, hydrogen bonds, and van der Waals forces are mainly considered part of this class of interactions, but some novel interaction patterns, such as C–H···O hydrogen bonds[2–4], cation-π interactions[5–7], and CH/π interactions[8,9], were recently characterized in folded protein structures and were claimed to be important for stability and function. The S···X interactions (Fig. 1)[10–12] that form between a cysteine (Cys) or methionine (Met) side chain and the nearby heteroatom X (X=O, N, S) belong to such non-conventional weak interactions that stabilize protein structures to some extent.

It has been well recognized in physical organic chemistry that a divalent sulfur atom[13–16], as well as a divalent selenium atom[17], frequently forms a weak hypervalent (or out-of-octet) interaction with a nearby oxygen, nitrogen, or sulfur atom (Fig. 1A). It was, therefore, assumed that similar nonbonded interactions would also occur in the protein architecture: S···X (X=O, N, S) interactions have recently been characterized in proteins by statistical analysis of the proximate atoms around Cys and Met sulfur atoms (SSC and CSC types, respectively) involved in 604 heterogeneous protein structures[10,11]. The features can be summarized as follows.

(1) Both SSC and CSC S atoms prefer to make a close contact with O, N, and S atoms in folded protein structures: out of a total of 790 Cys residues, 100 (13%), 33 (4%), and 15 (2%) have strong S···O, S···N, and S···S interactions, respectively, with less than the van der Waals atomic contact ($d \leq 0.0$ Å). Similarly, out of 2,124 Met residues, 153 (7%), 68 (3%), and 9 (0.4%) have strong S···O, S···N, and S···S interactions, respectively.

(2) According to ab initio calculations applying the MP2

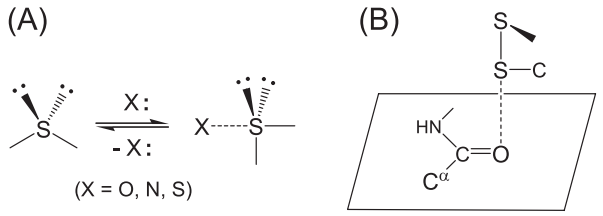

Figure 1 Formation and structural features of the S···X interaction. (A) The divalent S atom adopts the coordination from a heteroatom X, changing the valence state from a tetrahedral to a hypervalent trigonal bipyramidal state. (B) The most frequently observed S···O interactions in proteins. The S atom of a disulfide bond (SSC type) approaches a main-chain amide O atom in a direction vertical to the peptide plane maintaining the linearity of the three S–S···O atoms.

method, the S···X interactions are predominantly stabilized by dispersion and/or long-range electrostatic forces. However, the contribution from the orbital interactions is also important as evidenced by the observed directional preferences (*vide infra*).

(3) The S···X interactions *statistically* show various directionalities depending on the types of S and X atoms. For example, in the most frequent S(SSC)···O(amide) interaction (Fig. 1B), the O atom tends to approach the S atom in the direction of the antibonding σ^*_{SS} orbital, and the S atom tends to approach the O atom in the direction of the π_O orbital, thereby allowing the efficient $\pi_O \rightarrow \sigma^*_{SS}$ orbital (or charge-transfer) interaction. In other cases, the σ^*_{SC}, π_N, nonbonded n_O , n_N, or n_S orbital can be involved in the S···X interactions.

(4) The potential surfaces of the S···X interactions are relatively flat, probably due to the predominant contribution from isotropic dispersion force, but they have shallow, but distinct, potential holes in the location of the favorable orbital interactions. The stabilization energy was estimated to be up to 3.2 kcal/mol for the S(SSC)···O(amide) interaction (Fig. 1B) on the basis of ab initio calculations in vacuo. In addition, the S···O interaction distance becomes slightly shorter when the O atom simultaneously forms an NH···O hydrogen bond.

It was of great interest that the intrinsic directional preferences of the S···O interaction, which are originated from the favorable orbital interactions (*vide ante*), were *statistically* maintained in folded proteins despite the shallow potential holes as well as the presence of other nonbonded interactions, such as hydrogen bonds. Although the structural feature of an individual S···X interaction can vary significantly, this suggested not only that the S···O interaction is important to control protein structures but also that the interaction can be applied to the molecular engineering of three-dimensional protein structures[12].

S···X interactions in proteins have attracted the attention of various research groups. The first statistical analysis on S···O interactions[18] showed only marginal directional preferences of the interactions probably due to the use of a small number of protein structures. On the other hand, Pal and Chakrabarti[19] found similar directionality for the close S(CSC)···O contacts to that illustrated in Fig. 1B by using a larger number of protein structures. It was also reported for some enzymes, such as *S*-adenosylmethionine synthetase[20] and G-protein receptor[21], that the S···O interaction formed between the specific Cys or Met S atom and the carboxylate O atom plays important roles in the enzymatic functions. The existence of intermolecular nonbonded S···O interactions was suggested for the complexes between *N*-acetylglucosamine-thiazolin and β-hexosaminidase and between benzophenone and porcine odorant-binding protein[22]. Moreover, the importance of the possible S···O interaction between the Met10 and Thr317 residues of adenylosuccinate lyase for the stability of the protein structure was suggested by use of the M10L mutant[23]. Thus, the SSC and CSC S atoms, which are generally considered as merely hydrophobic groups in proteins, would have some positive structural and functional implications in the context of the amino acid sequence.

Phospholipase A_2 (PLA_2, EC 3.1.1.4)[24,25], a small globular protein containing about 130 amino acid residues, is an SS-rich enzyme that catalyzes hydrolysis of the 2-acyl ester bond of phosphoglycerides in the presence of a calcium ion. The active site is formed by the His48 and Asp99 residues[26,27], which are located close to each other by the support of the Cys44–Cys105 and Cys51–Cys98 SS linkages. Murakami and Kudo[28] recently reviewed the diverse functions of a mammalian PLA_2 family and pointed out that an understanding of the function as well as the expression and regulatory processes for individual PLA_2s is important for the treatment of various diseases. On the other hand, Ohno et al.[29] analyzed the molecular evolution of the myotoxic PLA_2s involved in the snake venom of *Trimeresurus flavoviridis* inhabiting the southern islands of Japan and suggested an accelerated evolution of snake PLA_2 during the acquisition of diversity in physiological activities. Mapping of the structural determinants for diverse functions of snake PLA_2 was also elaborated by means of sequence analysis and site-directed mutagenesis[30]. Thus, PLA_2 is an interesting protein in terms of both function and evolution.

In this paper, we focus on the functional and evolutional aspects of the S···O and S···N interactions in phospholipase A_2 (PLA_2). Two types of database analyses have been carried out: (1) a comprehensive search for close S···X (X=O, N, and S) contacts in the protein structures retrieved from protein data bank (PDB)[31] and (2) a phylogenetic analysis on the conservation of the S···O and S···N interactions over the various amino acid sequences. The first statistical data indicative of possible roles of S···O and S···N interactions in the evolution and functions of a protein are presented.

Materials and methods

Characterization of S···X interactions

A number of precise structures of PLA_2 have been registered in PDB[31] with a variety of amino acid sequences (including the wild-type proteins of several species and mutants). The structural data, which were determined by X-ray analysis with high resolution (resoln≤2.0 Å) and released before August 2003, were collected from PDB and sorted by the protein domain name (PDN) according to the SCOP[32] data. The structural data that were assigned to either PDN = phospholipase A_2 (PLA_2-DG) or snake phospholipase A_2 ($sPLA_2$-DG) were used in this study.

The high-resolution structural data of PLA_2 selected according to the above criteria were processed in an exhaustive search for close S···X (X=O, N, S) atomic contacts. According to a method described in the literature[11], proximate intramolecular and intermolecular heteroatoms (X) were sought for both SSC (Cys) and CSC (Met) S atoms, and the atomic distances between the S and X atoms ($r_{S\cdots X}$) were calculated. The directionality of the S···X contacts was also analyzed simultaneously. The values of $r_{S\cdots X}$ were then normalized to the relative distance ($d_{S\cdots X}$) by subtracting the van der Waals radii (*vdw*) of the corresponding atoms: $d_{S\cdots X}=r_{S\cdots X}-vdw(S)-vdw(X)$, where values of 1.80, 1.52, and 1.55 Å were used for S, O, and N atoms, respectively[33]. Close S···X contacts with $d_{S\cdots X}\leq 0.2$ Å were considered to mean that the two atoms form an S···X interaction. A similar criterion for S···X interactions was used in previous studies[10–12].

The location of an S···X interaction along the amino acid sequence was indicated in the following manner. When the interaction is formed between the S atom of cysteine 44 and the main-chain O atom of aspartic acid 40, it is denoted as S(C44)···O(D40). Similarly, the relative distance of the interaction is denoted as $d_{S(C44)\cdots O(D40)}$. When a side-chain heteroatom is involved in the interaction, the atom is labeled as Oγ, Nδ, etc. in order to distinguish it from the main-chain heteroatom, which is simply denoted by using an O or N label. The name of the polypeptide chain is indicated in the parentheses attached to a four-letter PDB code, e.g., 1FX9(A) and 1FX9(B) for chains A and B of 1FX9, respectively. The superimposed local structures around the interacting S atoms were drawn using the MOLDEN program[34].

Dendrogram

The dendrogram-based analysis for $sPLA_2$-DG was carried out using the FASTA-format data of the amino acid sequences, which can be obtained from the PDB data. The sequence data were first aligned on the basis of similarity by using the Clustal W program[35] and then converted to a dendrogram using the Tree View program[36]. All amino acid sequences of $sPLA_2$-DG selected from PDB were employed for the dendrogram analysis.

Results

Data on forty-two high-resolution crystal structures of PLA_2, which normally contained seven well-conserved SS linkages and one methionine (Met) residue, were collected from PDB. All of them belonged to a vertebrate PLA_2 family and could be classified into two groups according to the SCOP[32] data: the structures with PDN (protein domain name) = phospholipase A_2 and those with PDN = snake phospholipase A_2. Although in structure and amino acid sequence these groups (PLA_2-DG and $sPLA_2$-DG, respectively) were similar to each other, the two groups were analyzed separately.

Phospholipase A_2 domain group (PLA_2-DG)

Crystallographic information for the twenty-two high-resolution structures of PLA_2-DG selected from PDB is listed in Table 1 along with the data on relative atomic distances ($d_{S\cdots X}\leq 0.2$ Å) for the S···O and S···N interactions commonly found in this group. Bovine, porcine, and human PLA_2-DG were selected. Although a variety of amino acid sequences were found for bovine PLA_2, the seven SS bonds (C11–C77, C27–C123, C29–C45, C44–C105, C51–C98, C61–C91, and C84–C96) and the one Met residue (M8) were conserved. The same SS linkages and Met residue were found in porcine PLA_2 except for the C27–C123 linkage, which was assigned to the C27–C124 linkage in porcine PLA_2 due to the insertion of one amino acid residue. On the other hand, one additional SS bond (C49–C124 of 1KVO and C48–C122 of 1LE6) was present in human PLA_2, and one SS bond (C11–C77) was missing in 1KVO. The ligands and prosthetic groups were assigned to each polypeptide chain when the PDB data contained two crystallographically independent protein molecules. It is worth noting that all chains were crystallized with a calcium ion. This is of course related to the fact that a calcium ion is essential to enzymatic functions[24–27].

Various S···O and S···N interactions were detected in PLA_2-DG, but some of them were observed only in a few structures. Such rare nonbonded interactions were S(C11)···O(E81) in 1FX9(A) and 1FXF(B), S(C91)···Oδ(N71) in 1CEH, S(C91)···Nδ(N71) in 1MKU, etc. Although they might play some roles in the corresponding structures, they are ignored hereafter because more common S···X interactions would have more important implications for the function and evolution. No S···S interaction was found for PLA_2-DG.

As shown in Table 1, four S···O interactions and one S···N interaction were commonly observed for PLA_2-DG. The bold-style numbers of the relative distance $d_{S\cdots X}$ in the table indicate strong S···X (X=O and N) interactions with $d\leq 0.1$ Å. It should be noted that none of these common interactions were detected for 1GH4. This must be because the three lysine residues (K56, K120, and K121) have been mutated with three methionine residues, which would de-

Table 1 S···O and S···N interactions observed in PLA_2-DG

PDB code	species	resoln[a]	type	chain	ligands and prosthetic groups	S···O interactions				S···N interaction
						$d_{S(C44)\cdots O(D40)}$[b]	$d_{S(C61)\cdots O(A55)}$[c]	$d_{S(C84)\cdots O(C96)}$[d]	$d_{S(C98)\cdots O(F94)}$[e]	$d_{S(M8)\cdots N(R100)}$
1BP2	bovine	1.7	WT	–	Ca^{2+}, MPD[f]	**0.090**	**0.063**	0.165	**0.086**	**–0.025**
1BPQ	bovine	1.8	K56M	–	Ca^{2+}		**–0.262**	**0.015**	0.128	**–0.066**
1C74	bovine	1.9	K53M, K56M	A	Ca^{2+}	0.113	**–0.070**	0.130	0.179	**–0.020**
1CEH	bovine	1.9	D99N	–	Ca^{2+}	**0.039**	**–0.153**	**0.094**	**0.092**	**–0.025**
1FDK	bovine	1.96	WT	–	Ca^{2+}, GLE[g]		**–0.214**	**0.069**	0.156	**–0.009**
1G4I	bovine	0.97	WT	A	Ca^{2+}, Cl^-, MPD[f]	0.127	**–0.018**	0.117	**0.099**	**–0.034**
1GH4	bovine	1.9	K56M, K120M, K121M	A	Ca^{2+}, MPD[f]					
1IRB	bovine	1.7	K120A, K121A	–	Ca^{2+}	**0.022**	**–0.003**	0.172	**0.091**	0.186
1KVW	bovine	1.95	H48Q	–	Ca^{2+}	**0.006**	**0.052**		0.191	0.181
1KVX	bovine	1.9	D99A	–	Ca^{2+}	**0.024**	**0.058**	**–0.070**	**0.088**	0.120
1KVY	bovine	1.9	D49E	–	Ca^{2+}	**0.046**	**0.070**	0.115	**–0.059**	**0.096**
1MKS	bovine	1.9	Y52F, Y73F, D99N	–	Ca^{2+}	**0.086**	**–0.176**	**0.065**	0.184	
1MKT	bovine	1.72	WT	–	Ca^{2+}	**0.022**	**0.014**		0.136	**0.098**
1MKU	bovine	1.8	Y52F, Y73F, D99N	–	Ca^{2+}	**0.097**	0.183	**0.095**	0.169	
1MKV	bovine	1.89	WT	–	Ca^{2+}, GEL[h]	**0.006**	**–0.090**	**0.096**	0.167	0.148
1O3W	bovine	1.85	K53M, K56M, K120M	A	Ca^{2+}, MPD[f]	0.174	**–0.221**	0.115	0.134	**0.025**
2BPP	bovine	1.82	WT	–	Ca^{2+}		**–0.150**	**0.024**	0.134	**0.000**
1FX9	porcine	2.0	WT	A	Ca^{2+}, SO_4^{2-}, MJI[i]	0.137	0.197	**0.002**	**0.041**	0.130
			WT	B	Ca^{2+}, SO_4^{2-}, MJI[i]			**0.087**		**–0.256**
1FXF	porcine	1.85	WT	A	Ca^{2+}, PO_4^{3-}, MJI[i]	**0.086**		**–0.123**	**–0.016**	**0.060**
			WT	B	Ca^{2+}, PO_4^{3-}, MJI[i]	0.120	0.193	**0.031**		**–0.140**
1HN4	porcine	1.4	WT	A	Ca^{2+}, SO_4^{2-}, MJI[i]	0.162	**–0.237**	0.154	**0.024**	**–0.016**
			WT	B	Ca^{2+}, SO_4^{2-}, MJI[i]	0.134	**–0.020**	0.110	**0.027**	**–0.116**
1KVO	human	2.0	engineered	A	Ca^{2+}, OAP[j]	**0.053**		**0.048**	**–0.033**	
			engineered	B	Ca^{2+}, OAP[j]	**–0.165**			**0.060**	
1LE6	human	1.97	engineered	A	Ca^{2+}, MPD[f]	**0.089**	**0.003**	0.134	**–0.039**	k
			engineered	B	Ca^{2+}, MPD[f]	0.131	**0.036**	**–0.043**	**0.041**	k

Only commonly observed interactions ($d_{S\cdots X} \leq 0.2$ Å) are listed in the table. The values of $d_{S\cdots X}$ are given in Å. Empty columns mean that the corresponding values of $d_{S\cdots X}$ are more than 0.2 Å. The numbers in bold correspond to strong S···X interactions with $d_{S\cdots X} \leq 0.1$ Å.

[a] Resolution in Å.

[b] $d_{S(C44)\cdots O(E40)}$ for pig, $d_{S(C43)\cdots O(A39)}$ for 1KVO, and $d_{S(C42)\cdots O(A38)}$ for 1LE6.

[c] $d_{S(C58)\cdots O(A53)}$ for 1LE6.

[d] $d_{S(C77)\cdots O(C88)}$ for 1KVO and $d_{S(C76)\cdots O(C88)}$ for 1LE6.

[e] $d_{S(C90)\cdots O(Q86)}$ for 1KVO and $d_{S(C90)\cdots O(L86)}$ for 1LE6.

[f] 2-Methyl-2,4-pentanediol.

[g] 1-Hexadecyl-3-trifluoroethyl-*sn*-glycero-2-phosphomethanol.

[h] 1-*O*-Octyl-2-heptylphosphonyl-*sn*-glycero-3-phosphoethanolamine.

[i] 1-Hexadecyl-3-trifluoroethyl-*sn*-glycero-2-phosphate methane.

[j] 4-(*S*)-[(1-Oxo-7-phenylheptyl)amino]-5-[4-(phenylmethyl)phenylthio]pentanoic acid.

[k] The corresponding Met residue is not present.

stabilize the three-dimensional structure significantly. Local structures around the five common S···X interactions are shown in Fig. 2.

The S(C44)···O(D40) interaction (Fig. 2A) did not have the structural features of a typical S–S···O interaction shown in Fig. 1B. The S(C105)–S(C44)···O(D40) angle ranged from 115 to 124°, and the S(C44) atom approached the O(D40) atom from a horizontal direction (i.e., the direction of the n_O orbital) rather than a vertical direction relative to the peptide plane. The S(C44)···O(D40) interaction, therefore, was assigned as a distorted S–S···O interaction. The structural distortion was in accord with the observation that the values of $d_{S(C44)\cdots O(D40)}$ are mostly positive in Table 1.

On the other hand, the S(C61)···O(A55) interaction (Fig. 2B) had almost the normal structural features of a typical S–S···O interaction. The S(C61) atom approached the O(A55) atom in a direction vertical to the peptide plane (i.e., the direction of the π_O orbital), and the O(A55) atom was located at the back of the S(C61)–S(C91) bond (i.e., the direction of the σ^*_{SS} orbital): the S(C91)–S(C61)···O(A55) angle was 143~169°. About half of the relative distance $d_{S(C61)\cdots O(A55)}$ was negative, indicating that the S(C61)···O(A55) interaction is significantly stronger than the S(C44)···O(D40) interaction (Fig. 2A). It can be seen in Fig. 2B that the position of the O(A55) atom relative to the S(C61) atom tends to converge compared with the neighboring atoms, supporting the notion that the S···O interaction stabilizes the local structure of PLA_2 to some extent.

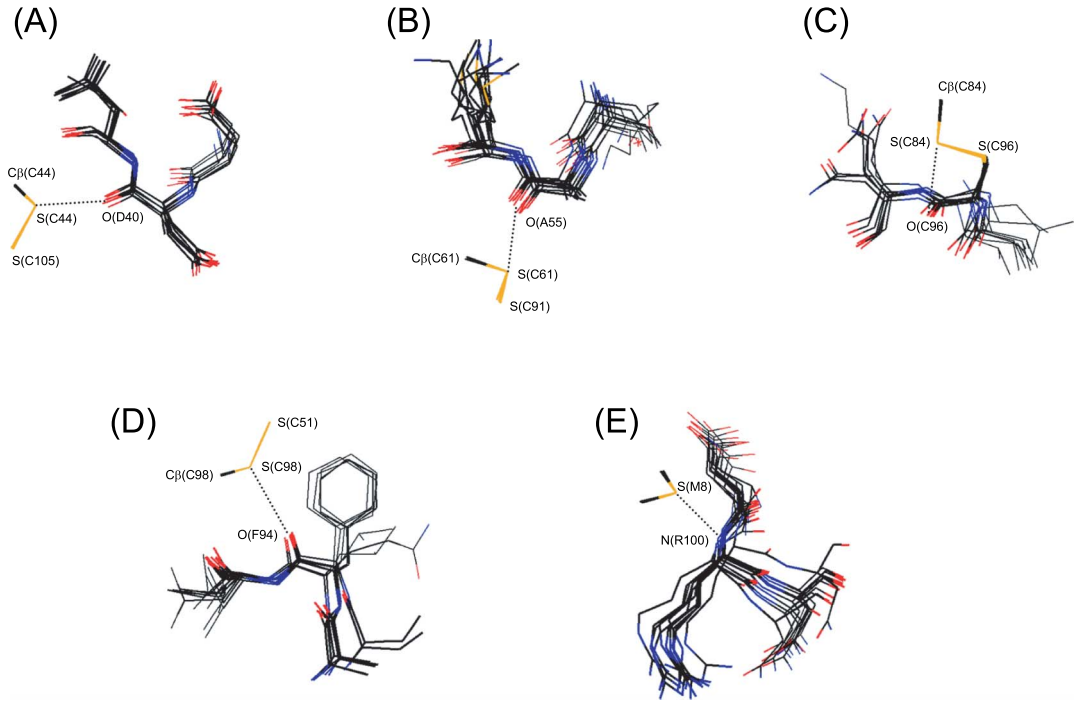

Figure 2 Superimposed local structures around the four S···O interactions and one S···N interaction observed in PLA_2-DG. Pictures were drawn using the structural data of PDB (Table 1) with $d_{S\cdots X} \leq 0.05$ Å. The sulfur, oxygen, nitrogen, and carbon atoms are shown in yellow, red, blue, and black, respectively. The residue numbers of bovine PLA_2 are indicated. For porcine and human PLA_2, the residue numbers may be different (see the footnotes of Table 1). (A) The S(C44)···O(D40) interaction. (B) The S(C61)···O(A55) interaction. The structures of human PLA_2 [1LE6(A) and 1LE6(B)] are omitted because they are significantly deviated. (C) The S(C84)···O(C96) interaction. (D) The S(C98)···O(F94) interaction. (E) The S(M8)···N(R100) interaction.

Similarly, the S(C84)···O(C96) interaction (Fig. 2C) was assigned as a normal C–S···O interaction with verticality with respect to the peptide plane and the linearity of the three C–S···O atoms. The S(C98)···O(F94) interaction (Fig. 2D) was a distorted S···O interaction with similar structural features to the S(C44)···O(D40) interaction (Fig. 2A). In this interaction, however, the spatial position of the O(F94) atom relative to the S(C98) atom fluctuated less than that of the surrounding atoms. The observation suggested that the S···O interaction contributes to the local stability of PLA_2.

The S(M8)···N(R100) interaction (Fig. 2E) was significantly different from the four aforementioned S···O interactions in that it involved a methionine side chain (a CSC S atom), not a cysteine side chain (an SSC S atom). According to previous studies[10–12], CSC S atoms less frequently form S···X interactions in proteins than SSC S atoms. Therefore, it was a little surprising to find such a common and strong S···N interaction for the only methionine residue (M8) that is conserved in PLA_2-DG. Fig. 2E shows that for the S(M8)···N(R100) interaction, the S atom attacks the main-chain N atom perpendicularly while keeping the linear alignment of the three C–S···N atoms. Thus, the interaction was assigned as a typical C–S···N(π) interaction[11]. Moreover, the position of the N(R100) atom relative to the S(M8) atom was likely to converge. This suggested the importance of the interaction for the local stability of PLA_2. The S···N interaction was not detected in human PLA_2 because the local structure as well as the amino acid sequence around the interaction is different from that of bovine or porcine PLA_2.

Snake phospholipase A_2 domain group (sPLA$_2$-DG)

Twenty high-resolution structures were selected from PDB for sPLA$_2$-DG. The crystallographic information and the relative distances (d) of the eight specific S···X interactions are listed in Table 2.

The five structures obtained for Russell's viper (r. viper) possessed the same amino acid sequence but they were crystallized with different ligands and prosthetic groups. On the other hand, each of the amino acid sequences of the five structures obtained for Indian cobra (i. cobra) were different. Similarly, each structure for cottonmouth snake (c. snake) and sand viper (s. viper) had a different amino acid sequence. In the case of Chinese water moccasin (moccasin), the amino acid sequences of 1M8R and 1M8S were identical to each other but were different from that of 1BK9.

All snake PLA_2 contained seven SS bonds, of which six were conserved; C27–C126, C29–C45, C44–C105, C51–C98, C61–C91, and C84–C96 using the residue numbers for r. viper. The remaining bond was C50–C133 for the species from r. viper to s. viper in Table 2 or C11–C71, C11–C72, and C11–C77 for i. cobra, Taiwan cobra (t. cobra), and tiger

Table 2 S···O and S···N interactions observed in $sPLA_2$-DG

PDB code	species[a]	resoln[b]	chain	ligands and prosthetic groups	S···O interactions							S···N interaction
					$d_{S(C27)\cdots O(R36)}$[c]	$d_{S(C44)\cdots O(A40)}$[d]	$d_{S(C44)\cdots O(C105)}$[e]	$d_{S(C50)\cdots O(F46)}$[f]	$d_{S(C61)\cdots O(A55)}$[g]	$d_{S(C84)\cdots O(C96)}$[h]	$d_{S(C98)\cdots O(R94)}$[i]	$d_{S(M8)\cdots N(K100)}$[j]
1FB2	r. viper	1.95	A			**–0.030**		**0.017**	q			
			B			**0.094**	**0.096**	**–0.035**	q	0.194	**0.015**	0.158
1FV0	r. viper	1.7	A	9AR[k], AcO^-, SO_4^{2-}, glycerol		0.151	**0.039**	**–0.026**	q		0.180	0.188
			B	dioxane, SO_4^{2-}, glycerol		**0.031**	0.138	**–0.047**	q	0.141	**–0.101**	
1JQ8	r. viper	2.0	A	LAIYS[l], SO_4^{2-}, AcOH		**0.014**	0.119	0.118	q		0.106	
			B	AcOH		**0.083**	0.128	**–0.068**	q	0.182	**–0.134**	
1JQ9	r. viper	1.8	A	FLSYK[m], AcOH		**0.036**	0.138	0.117	q			
			B	AcOH		**0.084**	0.147	**–0.028**	q		**–0.029**	0.199
1KPM	r. viper	1.8	A	vitamin E, AcOH		**0.003**	0.145	**–0.002**	q		0.162	0.174
			B	AcOH		**0.075**	**0.075**	**–0.042**	q	0.181	**–0.026**	0.195
1MC2	h. snake	0.85	A	*i*-PrOH		**0.050**	**0.042**		q		**–0.047**	
1PPA	c. snake	2.0	–	cyclohexylamine	0.143	**–0.161**	**0.064**	**–0.050**	q	0.165	**–0.068**	
1VAP	c. snake	1.6	A			0.178	**–0.066**	0.101	q		0.120	p
			B		**0.079**	0.128	0.151	0.134	q		0.159	p
1BK9	moccasin	2.0	–	Ca^{2+}, PBP[n], 1,4-butanediol	**–0.053**	**0.043**	0.151	**–0.100**	q		0.160	p
1M8R	moccasin	1.9	A	Cd^{2+}, 1,4-butanediol	**0.074**	0.119	**0.068**	**0.066**	q		0.129	p
1M8S	moccasin	1.9	A	Cd^{2+}, 1,4-butanediol	**0.020**	**0.012**	0.117	**0.030**	q	0.181	**0.071**	p
1JLT	s. viper	1.4	A	MPD[o]	**0.041**	**–0.056**	0.115	0.137	q	**0.083**	0.151	
			B	MPD[o]	0.160	**–0.095**	0.175	**0.014**	q	**0.078**	**–0.089**	
1VPI	s. viper	1.72	–		**0.048**	**0.034**	0.149		q	**0.064**	**–0.027**	
1AE7	t. snake	2.0	–	SO_4^{2-}		**–0.000**	0.189	p		**–0.020**	0.169	p
1A3D	i. cobra	1.8	–	Na^+			0.144	p	0.181	**0.033**	**–0.188**	**0.000**
1LFF	i. cobra	1.5	A	Ca^{2+}, PO_4^{3-}, AcOH	0.159	0.125		p	0.118	0.107	**–0.101**	**0.083**
1LFJ	i. cobra	1.6	A	Ca^{2+}, PO_4^{3-}, AcOH		0.157	**0.014**	p	0.172	0.109	**–0.008**	**0.025**
			B	Ca^{2+}, PO_4^{3-}, AcOH			0.108	p	0.161	**0.054**	**–0.035**	**0.041**
1LN8	i. cobra	1.65	A	Ca^{2+}, PO_4^{3-}		0.180	0.115	p	0.106	0.127	**–0.117**	**0.094**
1MH7	i. cobra	2.0	A				0.126	p	0.156	**0.019**	**–0.049**	0.161
1POA	t. cobra	1.5	–	Ca^{2+}			0.106	p	0.114	**0.013**	**–0.030**	0.169

Only commonly observed interactions ($d_{S\cdots X} \leq 0.2$ Å) are listed in the table. The values of $d_{S\cdots X}$ are given in Å. Empty columns mean that the corresponding values of $d_{S\cdots X}$ are more than 0.2 Å. The numbers in bold correspond to strong S···X interactions with $d_{S\cdots X} \leq 0.1$ Å.

[a] r. viper = Russell's viper, h. snake = hundred-pace snake, c. snake = cottonmouth snake, moccasin = Chinese water moccasin, s. viper = sand viper, t. snake = tiger snake, i. cobra = Indian cobra, t. cobra = Taiwanese cobra.

[b] Resolution in Å.

[c] $d_{S(C26)\cdots O(R35)}$ for 1VAP, $d_{S(C27)\cdots O(Q36)}$ for 1PPA, and $d_{S(C27)\cdots O(T36)}$ for 1JLT(B) and 1LFF.

[d] $d_{S(C43)\cdots O(A39)}$ for 1VAP, $d_{S(C44)\cdots O(D40)}$ for 1LFJ and 1LN8, and $d_{S(C44)\cdots O(E40)}$ for 1AE7.

[e] $d_{S(C43)\cdots O(C95)}$ for 1VAP, $d_{S(C43)\cdots O(C99)}$ for 1A3D and 1POA, and $d_{S(C44)\cdots O(C100)}$ for 1LFF, 1LFJ, 1LN8, and 1MH7.

[f] $d_{S(C49)\cdots O(F45)}$ for 1VAP.

[g] $d_{S(C60)\cdots O(A54)}$ for 1A3D and 1POA.

[h] $d_{S(C78)\cdots O(C90)}$ for 1A3D and 1POA, and $d_{S(C79)\cdots O(C91)}$ for 1LFF, 1LFJ, 1LN8, and 1MH7.

[i] $d_{S(C88)\cdots O(Q84)}$ for 1VAP, $d_{S(C92)\cdots O(A88)}$ for 1POA, $d_{S(C92)\cdots O(S88)}$ for 1A3D, $d_{S(C93)\cdots O(S89)}$ for 1LFF, 1LFJ(B), and 1LN8, $d_{S(C93)\cdots O(T89)}$ for 1LFJ(A) and 1MH7, $d_{S(C98)\cdots O(A94)}$ for 1JLT(A) and 1VPI, $d_{S(C98)\cdots O(D94)}$ for 1JLT(B), $d_{S(C98)\cdots O(E94)}$ for 1BK9, 1M8R, 1M8S, 1PPA, and 1MC2, and $d_{S(C98)\cdots O(F94)}$ for 1AE7.

[j] $d_{S(M8)\cdots N(R94)}$ for 1A3D and 1POA, and $d_{S(M8)\cdots N(R95)}$ for 1LFF, 1LFJ, 1LN8, and 1MH7.

[k] 9-Hydroxy-8-methoxy-6-nitrophenanthrol[3,4D][1,3]dioxole-5-carboxylic acid.

[l] Leu-Ala-Ile-Tyr-Ser.

[m] Phe-Leu-Ser-Tyr-Lys.

[n] *p*-Bromophenacyl bromide.

[o] 2-Methyl-2,4-pentanediol.

[p] The corresponding Cys or Met residue is not present.

[q] The corresponding A55 residue and its neighbor are deleted in the amino acid sequence.

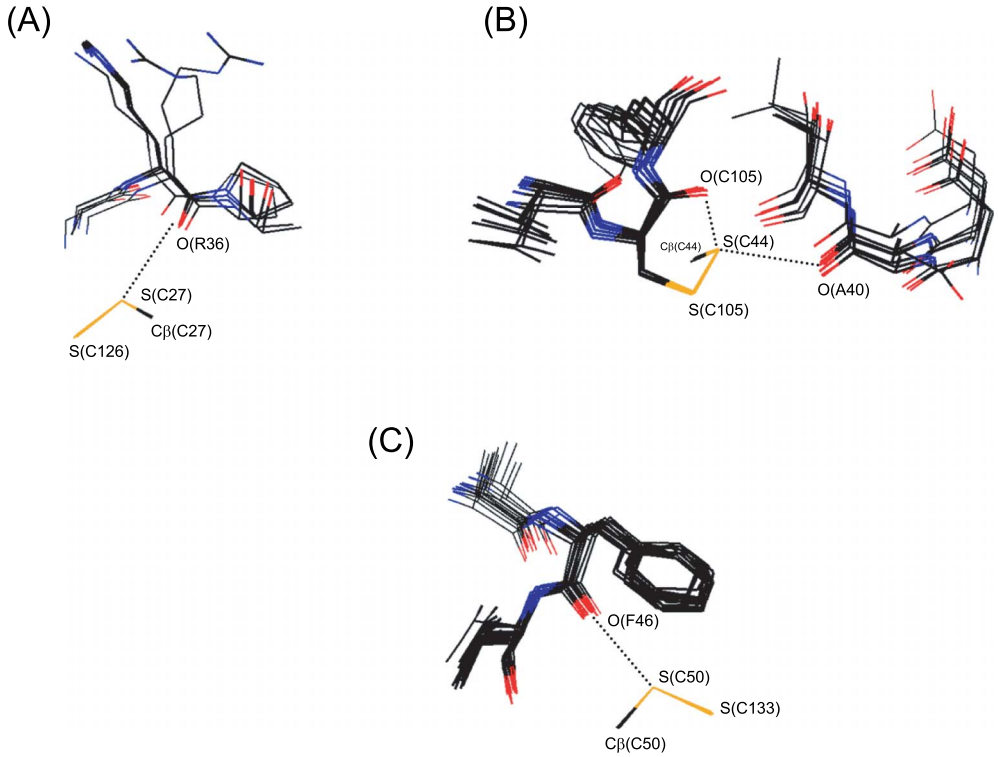

Figure 3 Superimposed local structures around the three S···O interactions additionally found in $sPLA_2$-DG. Pictures were drawn using the structural data of PDB (Table 2) with $d_{S\cdots O} \leq 0.10$ Å. The sulfur, oxygen, nitrogen, and carbon atoms are shown in yellow, red, blue, and black, respectively. The residue numbers of Russell's viper (r. viper) PLA_2 are indicated. For the other snake PLA_2, the residue numbers may be different (see the footnotes of Table 2). (A) The S(C27)···O(R36) interaction. (B) The S(C44)···O(C105) interaction shown with the bifurcated S(C44)···O(A40) interaction. (C) The S(C50)···O(F46) interaction.

snake (t. snake), respectively. Although the residue numbers were a little different between $sPLA_2$-DG and PLA_2-DG, the SS-linkage patterns of $sPLA_2$-DG were in common with those of PLA_2-DG except for the presence of an additional C50–C133 linkage for the species from r. viper to s. viper. The Met residue (M8) was also well conserved for the species of $sPLA_2$-DG, but it was replaced with another amino acid in c. snake (1VAP), moccasin (1BK9, 1M8R, 1M8S), and t. snake (1AE7).

Four of the S···O interactions [i.e., S(C44)···O(A40), S(C61)···O(A55), S(C84)···O(C96), and S(C98)···O(R94)] and one S···N interaction [S(M8)···N(K100)] summarized in Table 2 have also been observed in PLA_2-DG (see Table 1) although some of the amino acid residues are mutated in $sPLA_2$-DG. The local structures around these interactions were similar to those shown in Fig. 2. Therefore, we will not detail their structural features here. It should be noted, however, that S(C61)···O(A55) and S(M8)···N(K100) interactions cannot exist in some of the structures in $sPLA_2$-DG because the corresponding amino acids are missing or replaced. No S···S interaction was observed for $sPLA_2$-DG.

Fig. 3 shows superimposed local structures around the S(C27)···O(R36), S(C44)···O(C105), and S(C50)···O(F46) interactions that were found in $sPLA_2$-DG in addition to the five interactions common to PLA_2-DG.

The S(C27)···O(R36) interaction (Fig. 3A) showed an in-plane approach of the S(C27) atom to the amide O(R36) atom, while the three S(C126)–S(C27)···O(R36) atoms aligned almost linearly. The structural features indicated that the S···O interaction can be assigned as an S–S···O(n) interaction[12], in which the lone pair, instead of the π orbital, of the amide O atom would contribute to the stability.

In the S(C44)···O(C105) interaction (Fig. 3B), on the other hand, the S(C44) atom approached the O(C105) atom in the direction of the π orbital (the vertical direction), while the angles of S(C105)–S(C44)···O(C105) and Cβ(C44)–S(C44)···O(C105) were bent; 75~81° and 130~144°, respectively. It was interesting that the S(C44) atom simultaneously forms another S···O interaction with the O(A40) atom (the same S···O interaction to that shown in Fig. 1A). The bifurcated S···O interactions were unusual in that they are formed in the directions of the S lone pairs in the tetrahedral divalent state (see Fig. 1).

The S(C50)···O(F46) interaction (Fig. 3C) was assigned as a distorted S···O interaction without the verticality with respect to the peptide plane nor the linearity of the S–S···O atoms. Thus, various types of S···O interactions were detected in $sPLA_2$-DG.

The evolution of snake PLA_2 was subsequently analyzed. Fig. 4 shows a dendrogram of snake PLA_2 mapping the

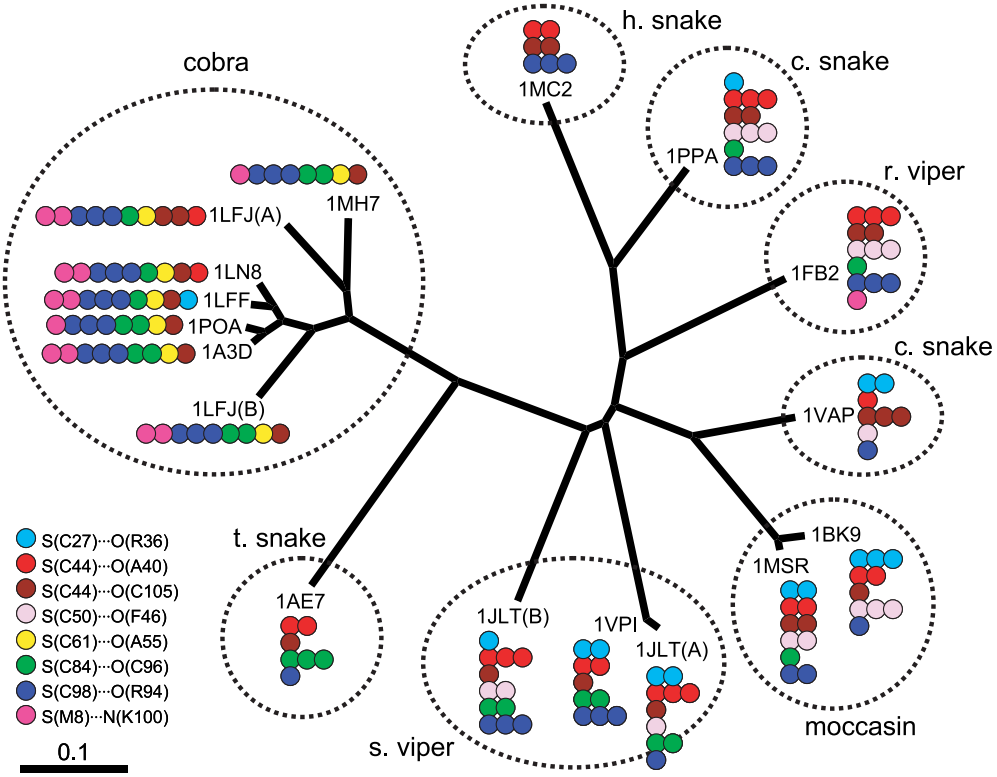

Figure 4 The dendrogram of snake PLA_2 ($sPLA_2$-DG). The eight commonly observed S···X interactions (see Table 2) are mapped on the figure by color circles, the number of which indicates the strength of the corresponding interaction; one circle for $0.1<d\leq0.2$ Å, two circles for $0.0<d\leq0.1$ Å, and three circles for $d\leq0.0$ Å. As for Russell's viper (r. viper), the smallest value of d was used for each interaction to determine the number of circles.

seven common S···O interactions and one common S···N.

The S(C27)···O(R36) interaction (sky blue circles) was mostly observed for c. snake, moccasin, and s. viper, which were located close to each other on the dendrogram. Similarly, the S(C44)···O(A40) interaction (red circles) was mostly observed for snake, viper, and moccasin but not observed for cobra. The S(C50)···O(F46) interaction (rose circles) was only observed for c. snake, r. viper, moccasin, and s. viper. The S(C84)···O(C96) interaction (green circles) was mostly observed for s. viper, t. snake, and cobra. On the other hand, the S(C61)···O(A55) interaction (yellow circles) and the S(M8)···N(K100) interaction (magenta circles) were mostly observed for cobra. These S···O and S···N interactions, except for S(C27)···O(R36) (sky blue circles), seem to make clusters on the dendrogram, assembling in one descent which can be separated from the dendrogram by cutting one branch, although there are a few exceptions as seen in Fig. 4. The ambiguity of the clustering is probably due to errors in the coordinates of the atomic positions in the protein crystals (*vide infra*). Meanwhile, the S(C44)···O(C105) interaction (brown circles) and the S(C98)···O(R94) interaction (blue circles) were found to be well conserved during the evolution of $sPLA_2$-DG.

Discussion

Structural and phylogenetic analyses of S···X interactions for PLA_2-DG and $sPLA_2$-DG have revealed the following: (1) there are many close contacts between the divalent S atoms and the nearby O or N atoms in both PLA_2-DG and $sPLA_2$-DG, (2) each S···X interaction has particular structural characteristics, but they are not always the same as those for the most stable and most frequent S···O interactions (Fig. 1B), and (3) the observed S···X interactions tend to cluster on the dendrogram of snake PLA_2. These observations are discussed below in relation to the conformational stability, molecular evolution, and functions of vertebrate PLA_2.

Choice of a model protein

To investigate possible roles of S···X interactions in the functions and evolution of proteins, we chose PLA_2 as a model not only because a number of precise structures have been registered in PDB for a variety of amino acid sequences but also because PLA_2 from snake venom has attracted interest in terms of its molecular evolution[29]. The appropriateness of the choice is confirmed by the result that a total of forty-two PDB entries with high resolution (resoln≤2.0 Å) were obtained for PLA_2-DG (Table 1) and

sPLA$_2$-DG (Table 2) with various amino acid sequences.

The accuracy of the structural data is essential for such a dendrogram as shown in Fig. 4, where the strength of the S···X interaction is classified as weak ($0.1<d\leq0.2$ Å), medium ($0.0<d\leq0.1$ Å), or strong ($d\leq0.0$ Å) depending on the relative atomic distance. The criterion for the resolution of PDB data was set at 2.0 Å. This criterion may be slightly low for discussing a difference of 0.1 Å in atomic distance in protein structure. For example, Dijkstra et al.[37] estimated the mean coordinate error of the bovine PLA$_2$ structure (1BP2) obtained at a resolution of 1.7 Å with an *R* factor of 17.1% to be 0.12 Å by applying Luzzati's diagram[38]. The magnitude of coordinate errors, which depends on both the *R* factor of the model and the resolution of the diffraction data[38], would be within a similar range for the PLA$_2$ structures listed in Tables 1 and 2. Therefore, absolute energies (or precise atomic distances) of the S···X interactions are not discussed below. Instead, we discuss only the tendency of the interactions.

Characteristics of S···X interactions in PLA$_2$

Many types of S···O and S···N interactions have been characterized in PLA$_2$ as shown in Figs. 2 and 3. Among them, five were commonly observed in bovine, porcine, and human PLA$_2$ (PLA$_2$-DG), while eight were commonly observed in snake PLA$_2$ (sPLA$_2$-DG). These interactions can be assigned to classes according to structural features. The S(C61)···O(A55) interaction (Fig. 2B) is a normal S–S···O(π) interaction with a similar partial structure to that in Fig. 1B. On the other hand, S(C84)···O(C96) (Fig. 2C), S(M8)···N(R100) (Fig. 2E), and S(C27)···O(R36) (Fig. 3A) are C–S···O(π), C–S···N(π), and S–S···O(n) interactions, respectively. The other four S···O interactions, i.e., S(C44)···O(D40) (Fig. 2A), S(C98)···O(F94) (Fig. 2D), S(C44)···O(C105) (Fig. 3B), and S(C50)···O(F46) (Fig. 3C), are considered to be distorted. It is interesting to note that three of the four distorted interactions, i.e., S(C44)···O(D40), S(C98)···O(F94), and S(C50)···O(F46), are involved in the α helices, suggesting that they support the stability of the α helices. The structural diversity would be due to the relatively shallow potential holes of S···X interactions[11] as summarized in the introduction section.

One characteristic feature of the observed S···X interactions was that most are formed with a main-chain amide O or N atom, not with a side-chain heteroatom. This feature was previously pointed out by statistical analyses[10,11] and has been rationalized by theoretical calculation[12]. Moreover, it was found that all the main-chain O atoms are simultaneously involved in a hydrogen bond with a proximate NH group. This is consistent with the result from ab initio calculations[10], which showed that the S···O interaction becomes shorter with the synergetic effect of a coexisting hydrogen bond.

Roles in local and global stability

The S···X interactions commonly detected in PLA$_2$ would play a role in local structural stability to some extent. Two possible roles can be considered.

One possibility is that the local structure around the S atom is controlled by the S···X interaction which acts in a similar manner to hydrogen bonds. If the S···X interaction plays a positive role in the local structure, the intrinsic directionality[11], i.e., linearity of S(or C)–S···X atoms and verticality with respect to the X atom, should be maintained. Indeed, some of the observed S···X interactions possess such structural characteristics as discussed above. Hence, they would act as structural determinants in PLA$_2$. The convergence of the positions of the O and N atoms relative to the S atom observed in some S···X interactions (for example, Figs. 2B and 2E) reasonably supports this possibility.

Another possibility is that the local structure around the S···X interaction is mainly determined by other nonbonded interactions and/or the steric congestion and the S···X interaction plays a minor role in the local stability. In this case, the structural features of the interaction should be largely distorted from the preferred structure shown in Fig. 1B[11]. As shown in Figs. 2 and 3, half of the eight S···X interactions are significantly distorted without linearity and verticality. Such interactions would stabilize the local structure only slightly. It should be noted that an isotropic dispersion (or long-range electrostatic) force is a major component of S···X interactions[11]. Therefore, the distorted S···X interactions still have the potential to stabilize the local structure.

The relationship between the presence of S···X interactions and the global stability, and hence the function, of a protein is an intriguing subject. In the literature, it was recently reported that the stability of adenylosuccinate lyase was reduced in the M10L mutant due probably to the deletion of the S···O interaction between the Met10 and Thr317 residues[23]. As for PLA$_2$, the roles of some of the individual SS linkages were studies by using several SS-engineered mutants[39]. However, the mutation of the Cys residues should erase the SS bond as well as the S···X interaction. On the other and, the roles of the S(M8)···N(R100) interaction in the global stability of PLA$_2$ can be directly analyzed by applying the M8X mutants. Yu et al.[40] recently reported an engineered M8,20L mutant of pig PLA$_2$ which showed a significant decrease in enzymatic activity. We suggest that the change is due largely to the loss of the nonbonded S(M8)···N(R100) interaction.

Roles in molecular evolution

The dendrogram shown in Fig. 4 demonstrated that most of the S···X interactions of sPLA$_2$-DG would make clusters in particular phylogenic regions though there are some exceptions. This in turn suggested that S···X interactions have played roles in the molecular evolution of snake PLA$_2$.

For snake PLA$_2$, the dendrogram (Fig. 4) represents the molecular evolution that has occurred during division of the

species, which would have been controlled by the environment and source of food[29]. Therefore, the clustering of S···X interactions on the dendrogram strongly suggested that the PLA_2 from snake venom of specific species has acquired (or lost) specific S···X interactions during evolution. For example, the PLA_2 group with the S(C50)···O(F46) interaction (rose circles) can be partitioned by a border between s. viper and t. snake. The groups with either S(C44)···O(A40) (red circles), S(C61)···O(A55) (yellow circles), or S(M8)··· N(K100) (magenta circles) interaction can be partitioned by a border between t. snake and cobra. Similarly, for the S(C84)···O(C96) interaction (green circles), the border may be drawn between moccasin and s. viper.

However, the observed clustering of the S···X interactions does not necessary mean that the S···X interactions have positively controlled the molecular evolution. As shown in Table 2, the formation of some S···X interactions is inhibited for some species due to replacement or deletion of the corresponding amino acid residues. This possibly indicates that the clustering of the S···X interactions resulted from other more essential factors.

The dendrogram shown in Fig. 4 provides the first evidence of distinct coupling between the presence of S···X interactions and the molecular evolution of a protein. However, the role of the S···X interactions in the molecular evolution of $sPLA_2$-DG is not conclusive at the moment and should be a subject of further study.

Roles in function

The active site of PLA_2 is assigned to the residues His48 and Asp99[26,27]. For bovine PLA_2, several mutants (e.g., H48Q[41] and D99A[41]) have been designed to characterize the importance of these residues, and some of the structures have been determined by X-ray crystallography at high resolution as shown in Table 1. Fig. 5 illustrates the structure around the active site of wild-type bovine PLA_2. The active site is fixed by two SS linkages (C44–C105 and C51–C98) and supported by two other SS linkages (C29–C45 and C84–C96). It is of significant interest to us that four of the five common S···X interactions (i.e., S(C44)···O(D40), S(C84)···O(C96), S(C98)···O(F94), and S(M8)···N(R100)) have been found in this region. The observation suggested that the S···X interactions also play roles in the enzymatic functions. The significant decrease in enzymatic activity reported recently for the M8,20L mutant of porcine PLA_2[40] reasonably supports this consideration, because such a mutant cannot form the S(M8)···N(R100) interaction near the active site.

The calcium binding site of PLA_2 is assigned to Tyr28, Gly30, Gly32, and Asp49[37]. Since all wild-type bovine PLA_2s in Table 1 (i.e., 1BP2, 1FDK, 1G4I, 1MKT, 1MKV, and 2BPP) were crystallized with a calcium ion, it would be interesting to examine the relationship between the presence of a ligand other than a calcium ion and the strengths of the five common S···X interactions. However, the S···X interactions of 1MKT and 2BPP, which contained only a calcium ion as a ligand, showed no significant difference in strength from those of the other wild-type bovine PLA_2s, which contained various ligands, such as MPD and GLE (see the footnotes of Table 1 for the abbreviations). Thus, the five common S···X interactions tolerate the conformational changes caused by the substrate binding.

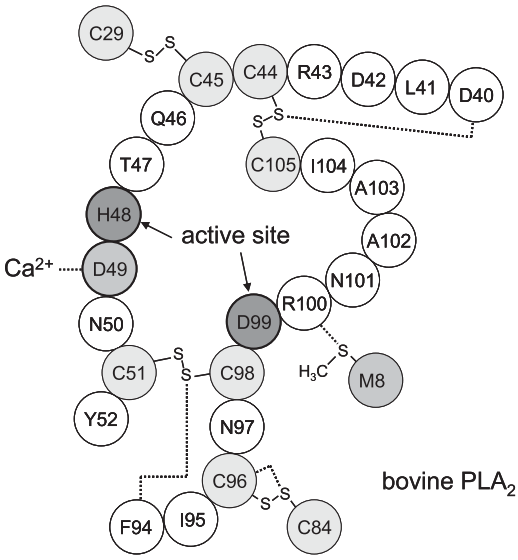

Figure 5 The active site of bovine PLA_2. The dotted lines indicate the S···O and S···N interactions present in the vicinity and the binding of Asp49 to a calcium ion.

A similar feature can be seen in Table 2 for the S···X interactions of Russell's viper PLA_2 with various ligands and prosthetic groups (i.e., chains A and B of 1FB2, 1FV0, 1JQ8, 1JQ9, and 1KPM). The structures without ligands (i.e., 1FB2) maintained strong S(C44)···O(A40) and S(C50)··· O(F46) interactions, medium S(C44)···O(C105) and S(C98)··· O(R94) interactions, and weak S(C84)···O(C96) and S(M8)··· N(K100) interactions. The strength did not seem to change on the binding of various substrates. Thus, the S···X interactions of Russell's viper PLA_2 also tolerate the conformational changes due to the substrate binding.

PLA_2 is a conformationally tough protein due to the presence of numerous SS linkages (normally seven) along the relatively short peptide chain comprising about 130 amino acid residues. However, the conformational changes relevant to the recognition of a substrate would not be small: the root mean square deviation for all the backbone atoms was 0.32 Å between the complex and inhibitor-free PLA_2[42], and the structural changes around the active-site were much more significant[43]. Therefore, the tolerance observed for the S···X interactions of bovine and Russell's viper PLA_2 would suggest that these interactions contribute to the local conformational stability around the active site. If the interactions are not important, they would readily disappear in the complexes with the substrate.

Structural flexibility is an important feature of proteins.

Some nonbonded interactions would be broken by the conformational changes caused by the substrate binding but others would not. The S···X interactions observed in PLA_2 can be classified as the latter type of nonbonded interactions.

Implications for protein force fields

If S···X interactions play important roles in the structure and functions of proteins, their incorporation into force fields is necessary to obtain accurate results in the molecular simulation of proteins. It is beyond the scope of this paper to argue whether conventional force fields, such as AMBER[44] and CHARMM[45], implicitly include the effects of S···X interactions or not. S···X interactions are normally formed near the van der Waals contact and are predominantly stabilized by dispersion and/or long-range electrostatic forces[11]. Therefore, it may be that the interactions are included implicitly in the Lennard-Jones and Coulombic terms of the force field. However, as for the directional tendency of the interactions, such as shown in Fig. 1B, it is not clear if conventional force fields can reproduce the tendency properly because the directionality is related to orbital interaction, which is not explicitly counted in the Lennard-Jones and Coulombic terms. To examine these points, molecular simulations of model polypeptides or proteins will be needed.

Conclusions

Many types of S···O and S···N interactions have been characterized in PLA_2. The dendrogram for $sPLA_2$-DG revealed a possible clustering of the S···X interactions. Most of the S···X interactions were found to reside in the vicinity of the active site and to tolerate the conformational changes caused by the binding of the substrate in the cases of bovine and Russell's viper PLA_2. Thus, the possibility that the S···X interactions play roles in the functions and evolution of PLA_2 has been suggested.

The present study is the first to show the possible coupling of S···X interactions with the evolution of a protein. Nevertheless, the conclusions are fairly putative at the moment. Similar dendrogram-based analyses should be carried out for other proteins as well. However, according to our preliminary search in PDB, there are at present no suitable model proteins, other than PLA_2, for which the evolution of the weak nonbonded S···X interactions can be systematically analyzed by using a set of precise structures with a variety of amino acid sequences.

The three-dimensional structure, and hence the function, of a protein is controlled by the interplay of weak nonbonded interactions, such as hydrogen bonds and van der Waals forces. Based on the present study as well as previous ones[10–12], we strongly suggest that S···X interactions are new members of such weak interactions. Indeed, a significant loss of the conformational stability or enzymatic activity of a protein on the elimination of a possible S···X interaction[23,40] has recently been reported.

Acknowledgments

We thank Dr. K. Yura of the Quantum Bioinformatics Group at the Japan Atomic Energy Research Institute for assistance with the software for constructing the dendrogram and valuable discussions of the results. This work was supported by Grants-in-aid for Scientific Research on Priority Areas (C) "Genomic Information Science" (Nos. 15014211 and 16014223) and Scientific Research (B) (No. 16350092) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. It was also supported in part by a Fellowship to Researchers from the Association for the Progress of New Chemistry of Japan and by the Research and Study Program of the Tokai University Educational System General Research Organization.

References

1. Dobson, C. M., Sali, A. & Karplus, M. Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.* **37**, 869–893 (1998).
2. Derewenda, Z. S., Lee, L. & Derewenda, U. The occurrence of C–H···O hydrogen bonds in proteins. *J. Mol. Biol.* **252**, 248–262 (1995).
3. Bella, J. & Berman, H. M. Crystallographic evidence for C^{α}–H···O=C hydrogen bonds in a collagen triple helix. *J. Mol. Biol.* **264**, 734–742 (1996).
4. Vargas, R., Garza, J., Dixon, D. A. & Hay, B. P. How strong is the C^{α}–H···O=C hydrogen bond? *J. Am. Chem. Soc.* **122**, 4750–4755 (2000).
5. Flocco, M. M. & Mowbray, S. L. Planar stacking interactions of arginine and aromatic side-chains in proteins. *J. Mol. Biol.* **235**, 709–717 (1994).
6. Gallivan, J. P. & Dougherty, D. A. Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci.* **96**, 9459–9464 (1999).
7. Gallivan, J. P. & Dougherty, D. A. A computational study of cation-π interactions vs salt bridges in aqueous media: Implications for protein engineering. *J. Am. Chem. Soc.* **122**, 870–874 (2000).
8. Umezawa, Y. & Nishio, M. CH/π interactions as demonstrated in the crystal structure of guanine-nucleotide binding proteins, Src homology-2 domains and human growth hormone in complex with their specific ligands. *Bioorg. Med. Chem.* **6**, 493–504 (1998).
9. Brandl, M., Weiss, M. S., Jabs, A., Sühnel, J. & Hilgenfeld, R. C–H···π-interactions in proteins. *J. Mol. Biol.* **307**, 357–377 (2001).
10. Iwaoka, M., Takemoto, S., Okada, M. & Tomoda, S. Statistical characterization of nonbonded S···O interactions in proteins. *Chem. Lett.* 132–133 (2001).
11. Iwaoka, M., Takemoto, S., Okada, M. & Tomoda, S. Weak nonbonded S···X (X=O, N, and S) interactions in proteins. Statistical and theoretical studies. *Bull. Chem. Soc. Jpn.* **75**, 1611–1625 (2002).
12. Iwaoka, M., Takemoto, S. & Tomoda, S. Statistical and theoretical investigations on the directionality of nonbonded S···O interactions. Implications for molecular design and protein engineering. *J. Am. Chem. Soc.* **124**, 10613–10620 (2002).
13. Ciuffarin, E. & Guaraldi, G. Nucleophilic substitution at dicoordinated sulfur. Effect of the leaving group on the reaction between triphenylmethyl sulfenyl derivatives and *n*-

butylamine. *J. Org. Chem.* **35**, 2006–2010 (1970).
14. Rosenfield, R. E. Jr., Parthasarathy, R. & Dunitz, J. D. Directional preferences of nonbonded atomic contacts with divalent sulfur. 1. Electrophiles and nucleophiles. *J. Am. Chem. Soc.* **99**, 4860–4862 (1977).
15. Row, T. N. G. & Parthasarathy, R. Directional preferences of nonbonded atomic contacts with divalent sulfur in terms of its orbital orientations. 2. S···S interactions and nonspherical shape of sulfur in crystals. *J. Am. Chem. Soc.* **103**, 477–479 (1981).
16. Desiraju, G. R. & Nalini, V. Database analysis of crystal-structure-determining interactions involving sulphur: Implications for the design of organic metals. *J. Mater. Chem.* **1**, 201–203 (1991).
17. Iwaoka, M., Komatsu, H., Katsuda, T. & Tomoda, S. Nature of nonbonded Se···O interactions characterized by ^{17}O NMR spectroscopy and NBO and AIM analyses. *J. Am. Chem. Soc.* **126**, 5309–5317 (2004).
18. Carugo, O. Stereochemistry of the interaction between methionine sulfur and the protein core. *Biol. Chem.* **380**, 495–498 (1999).
19. Pal, D. & Chakrabarti, P. Non-hydrogen bond interactions involving the methionine sulfur atom. *J. Biomol. Struct. Dynam.* **19**, 115–128 (2001).
20. Taylor, J. C. & Markham, G. D. The bifunctional active site of *S*-adenosylmethionine synthetase — Roles of the active site aspartates. *J. Biol. Chem.* **274**, 32909–32914 (1999).
21. Brandt, W., Golbraikh, A., Täger, M. & Lendeckel, U. A molecular mechanism for the cleavage of a disulfide bond as the primary function of agonist binding to G-protein-coupled receptors based on theoretical calculations supported by experiments. *Eur. J. Biochem.* **261**, 89–97 (1999).
22. Nagao, Y., Honjo, T., Iimori, H., Goto, S., Sano, S., Shiro, M., Yamaguchi, K. & Sei, Y. Intramolecular nonbonded S···O interaction in acetazolamide and thiadiazolinethione molecules in their dimeric crystalline structures and complex crystalline structures with enzymes. *Tetrahedron Lett.* **45**, 8757–8761 (2004).
23. Palenchar, J. B., Crocco, J. M. & Colman, R. F. The characterization of mutant *Bacillus subtilis* adenylosuccinate lyases corresponding to severe human adenylosuccinate lyase deficiencies. *Protein Sci.* **12**, 1694–1705 (2003).
24. Dennis, E. A. Diversity of group types, regulation, and function of phospholipase A_2. *J. Biol. Chem.* **269**, 13057–13060 (1994).
25. Jain, M. K., Gelb, M. H., Rogers, J. & Berg, O. G. Kinetic basis for interfacial catalysis by phospholipase A_2. *Methods in Enzymology* **249**, 567–614 (1995)
26. Dijkstra, B. W., Drenth, J. & Kalk, K. H. Active site and catalytic mechanism of phospholipase A_2. *Nature* **289**, 604–606 (1981).
27. Sekar, K., Yu, B.-Z., Rogers, J., Lutton, J., Liu, X., Chen, X., Tsai, M.-D., Jain, M. K. & Sundaralingam, M. Phospholipase A_2 engineering. Structural and functional roles of the highly conserved active site residue aspartate-99. *Biochemistry* **36**, 3104–3114 (1997).
28. Murakami, M. & Kudo, I. Diversity of phospholipase A_2 enzymes. *Biol. Pharm. Bull.* **27**, 1158–1164 (2004).
29. Ohno, M., Chijiwa, T., Oda-Ueda, N., Ogawa, T. & Hattori, S. Molecular evolution of myotoxic phospholipase A_2 from snake venom. *Toxicon* **42**, 841–854 (2003).
30. Chioato, L. & Ward, R. J. Mapping structural determinants of biological activities in snake venom phospholipase A_2 by sequence analysis and site directed mutagenesis. *Toxicon* **42**, 869–883 (2003).
31. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. The protein data bank. *Nucl. Acids Res.* **28**, 235–242 (2000).
32. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540 (1995).
33. Bondi, A. van der Waals volumes and radii. *J. Phys. Chem.* **68**, 441–451 (1964).
34. Schaftenaar, G. & Noordik, J. H. Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Design* **14**, 123–134 (2000).
35. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680 (1994).
36. Page, R. D. M. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357–358 (1996).
37. Dijkstra, B. W., Kalk, K. H., Hol, W. G. J. & Drenth, J. Structure of bovine pancreatic phospholipase A2 at 1.7 Å resolution. *J. Mol. Biol.* **147**, 97–123 (1981).
38. Luzzati, P. V. Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr.* **5**, 802–810 (1952).
39. Janssen, M. J. W., Verheij, H. M., Slotboom, A. J. & Egmond, M. R. Engineering the disulphide bond patterns of secretory phospholipases A2 into porcine pancreatic isozyme — The effects on folding, stability and enzymatic properties. *Eur. J. Biochem.* **261**, 197–207 (1999).
40. Yu, B.-Z., Pan, Y. H., Janssen, M. J. W., Bahnson, B. J. & Jain, M. K. Kinetic and structural properties of disulfide engineered phospholipase A_2: Insight into the role of disulfide bonding patterns. *Biochemistry* **44**, 3369–3379 (2005).
41. Sekar, K., Li, Y., Tsai, M.-D. & Sundaralingam, M. Structures of the catalytic site mutants D99A and H48Q and the calcium-loop mutant D49E of phospholipase A_2. *Acta Crystallogr.* **D55**, 443–447 (1999).
42. Sekar, K., Eswaramoorthy, S., Jain, M. K. & Sundaralingam, M. Crystal structure of the complex of bovine pancreatic phospholipase A_2 with the inhibitor 1-hexadecyl-3-(trifluoroethyl)-*sn*-glycero-2-phosphomethanol. *Biochemistry* **36**, 14186–14191 (1997).
43. Sekar, K., Kumar, A., Liu, X., Tsai, M.-D., Gelb, M. H. & Sundaralingam, M. Structure of the complex of bovine pancreatic phospholipase A_2 with a transition-state analogue. *Acta Crystallogr.* **D54**, 334–341 (1998).
44. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7**, 230–252 (1986).
45. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Computat. Chem.* **4**, 187–217 (1983).