Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural and functional role of water molecules in bovine pancreatic phospholipase A_2: a data-mining approach

Shankar Prasad Kanaujia and Kanagaraj Sekar*

Bioinformatics Centre (Centre of Excellence in Structural Biology and Biocomputing), Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail: sekar@serc.iisc.ernet.in, sekar@physics.iisc.ernet.in

Received 29 August 2008
Accepted 22 November 2008

The water molecules in 25 (21 high-resolution and four atomic resolution) crystal structures of bovine pancreatic phospholipase A_2 have been analyzed in order to identify the invariant water molecules and their possible roles. A total of 24 water molecules have been identified that are invariant in all 25 crystal structures examined. These include the catalytic water molecule, which is directly involved in the enzyme mechanism, and the conserved structural water molecule, which stabilizes the extended hydrogen-bonding network of the active site. Furthermore, many other water molecules stabilize the structure, whilst a few have been found to maintain the active-site geometry and provide coordination to the functionally important calcium ion. The invariant water molecules have been carefully examined and their possible roles in the structure and/or function are discussed. Molecular-dynamics studies of all 25 crystal structures have also been carried out and the results provide a good explanation of and support the findings obtained from the crystal structures.

1. Introduction

Water molecules play a pivotal role in governing biomolecules, aiding in stabilizing the three-dimensional architecture, stability, dynamics and function (Halle, 2004; Eisenmesser *et al.*, 2005; Smolin *et al.*, 2005). It is also known that protein hydration plays an important role in biological processes (Otting *et al.*, 1991; Franks, 2002; Chaplin, 2006; Zhang *et al.*, 2007) and that hydration forces are responsible for the packing and stabilization of three-dimensional protein structure (Raschke, 2006). A notable involvement of water molecules is their participation in many hydrogen-bonding networks (Meyer, 1992). Therefore, water molecules are quantitatively considered to be an integral part of biomolecular systems and to be crucial in the protein-folding process (Cheung *et al.*, 2002; Papoian *et al.*, 2004). The common hydrophilic nature of the interfaces of protein–protein, protein–DNA and protein–ligand complexes and the abundance of water molecules at the interface suggest that water molecules are an indispensable component of biomolecular recognition and self-assembly (Tame *et al.*, 1996; Jayaram & Jain, 2004).

The location of many of these water molecules is conserved in identical or similar positions in the crystal structures of highly homologous proteins and their spatial conservation is common in active sites and metal coordination as well as in polar cavities. Furthermore, water molecules deeply buried in the core of the protein are considered to be important in the folded structure and make strong hydrogen bonds to polar groups; they are therefore believed to tighten the protein

molecules. It is well known that such ordered water molecules are best identified using crystallographic methods (X-ray crystallography or neutron diffraction) or in special cases by NMR spectroscopy (Otting, 1997). In general, for a water molecule to be described as ordered, it must make at least one contact (with a maximum distance of 3.5 Å) to the polar atoms of the protein molecule (Baker & Hubbard, 1984). With the availability of a large number of three-dimensional protein structures at higher resolutions, it is now possible to analyze and compare water structures. As indicated in the literature, studies have been performed on the water structures of T4 lysozyme (Zhang & Matthews, 1994), ribonuclease A (Kishan *et al.*, 1995), hen egg-white lysozyme (Biswal *et al.*, 2000), aspartic proteinases (Prasad & Suguna, 2002), legume lectins (Loris *et al.*, 1994) and serine proteases (Sreenivasan & Axelsen, 1992; Krem & Enrico, 1998). In these studies, highly homologous protein structures solved under varying experimental conditions with varying solvent contents and with minor mutations were studied and resulted in the identification of a number of water molecules that are invariant. These findings suggest that the positions of certain water molecules must be conserved for structural and/or functional reasons.

Therefore, we planned to carry out a similar study by analyzing the three-dimensional crystal structures of bovine pancreatic phospholipase A_2 (Fig. 1). The enzyme phospholipase A_2 (PLA_2; EC 3.1.1.4) catalyzes the hydrolysis of the *sn*-2 fatty-acid ester bond of phospholipids, producing a free fatty acid and a lysophospholipid in a calcium-dependent reaction (van Deenen & De Haas, 1964). The enzyme PLA_2 is widely distributed in snakes, lizards, bees and mammals. It is involved in a number of physiologically important cellular processes such as the liberation of arachidonic acid from membrane phospholipids (van den Berg *et al.*, 1995) and the subsequent conversion of arachidonic acid to leukotrienes and prostaglandins, which are involved in inflammatory diseases. The enzyme PLA_2 consists of 123 amino-acid residues (molecular weight of ∼14 kDa) and contains seven disulfide bonds. 32 recombinant bovine pancreatic PLA_2 structures are available. These include three crystal forms ($P2_12_12_1$, $P3_121$ and $C2$) and include native, native–inhibitor complex and mutant structures. The present study aims to better understand the role and the involvement of invariant water molecules in the three-dimensional architecture and function of the bovine pancreatic PLA_2 enzyme. To further strengthen our findings, molecular-dynamics studies have also been carried out on these structures and the results are compared with the crystal structures.

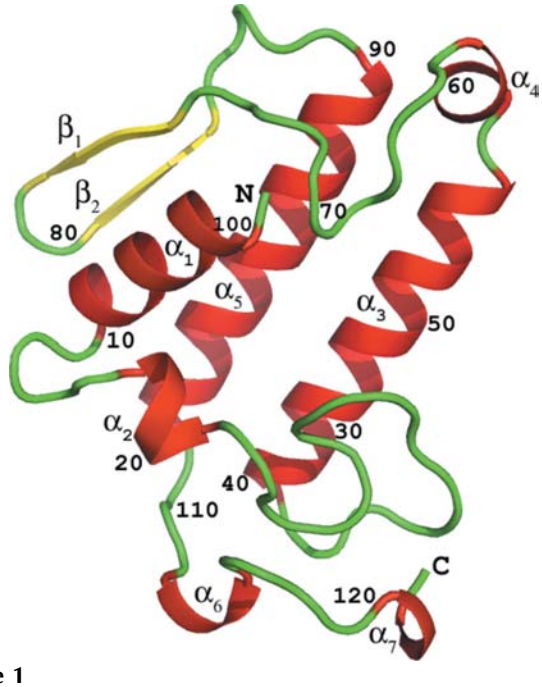

Figure 1
Overall three-dimensional structure of bovine pancreatic phospholipase A_2 (PDB code 1mkt). The seven disulfide bonds and the functionally important calcium ion are not shown.

2. Materials and methods

2.1. Data set

Seven of the 32 structures were excluded from the study either owing to the absence of water molecules in the structure (PDB codes 1bpq, 1bvm, 2bp2 and 2bpp) or low resolution (>2.0 Å; 1o2e and 3bp2) or because they were structures of pro-PLA_2 (4bp2). The three-dimensional atomic coordinates of the remaining 25 crystal structures (21 high-resolution and four atomic resolution) of bovine pancreatic PLA_2 were downloaded from the locally maintained PDB anonymous FTP server at the Bioinformatics Centre, Indian Institute of Science, Bangalore, India. The resolution of the investigated structures varies from 0.97 to 1.95 Å. In this resolution range, the positions of the solvent molecules are determined with high accuracy, particularly those in the first hydration shell. The crystal structure deposited with PDB code 1g4i (Steiner *et al.*, 2001) was taken as the reference or fixed molecule because it contained the highest number of water molecules. All the remaining structures were treated as mobile molecules and were superimposed on 1g4i to find invariant water molecules using the *3dSS* server (Sumathi *et al.*, 2006). The distance cutoff between pairs of superposed water molecules was taken to be 1.7 Å and a water molecule that had at least one common hydrogen bond to the protein atoms was considered to be invariant (Balamurugan *et al.*, 2007). While the hydrogen-bonding distance (maximum 3.4 Å) and angle (greater than 90°) criteria were generally followed, in some cases water molecules were considered to be equivalent if similar hydrogen bonds were observed even if the pairwise distance cutoff (1.7 Å) was not satisfied owing to variations in the side-chain conformation. Furthermore, in order to ascertain the water molecules, an investigation of the electron-density maps for the structures used in the present study was also carried out. The corresponding structure factors were downloaded from the PDB (Berman *et al.*, 2000) and the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) was used to generate mtz files. Subsequently, the modelling program *Coot* (Emsley & Cowtan, 2004) was used to visualize the difference electron-density maps. A water molecule was

Table 1
List of the three-dimensional crystal structures of bovine pancreatic phospholipase A_2 used in the present analysis.

Reference†	PDB code	Resolution (Å)	Space group	R_{work}/R_{free} (%)	No. of waters	Ligand/mutant‡
1	1g4i	0.97	$P2_12_12_1$	9.4/NA	247	
2	1vl9	0.97	$P2$	11.4/13.4	243	K53,56,121M
3	2bch	1.1	$P3_121$	10.4/13.3	228	K53,56,120,121M
4	2bax	1.1	$P3_121$	11.4/15.6	209	K53,56M
5	1vkq	1.6	$P3_121$	17.9/21.7	165	K53,56,120M
6	2zp3	1.9	$P3_121$	19.1/23.8	139	D49N
7	1une	1.5	$P2_12_12_1$	18.4/22.8	134	
8	2zp4	1.9	$P3_121$	17.8/20.1	132	H48N
9	2zp5	1.9	$P3_121$	19.7/23.5	127	D49K
10	1gh4	1.9	$P2$	19.6/25.9	125	K56,120,121M
11	2b96	1.7	$P3_121$	20.2/22.1	125	ANN/K53,56,121M
12	2bd1	1.9	$C2$	20.7/24.0	218§	K53,56,120,121M
13	1bp2	1.7	$P2_12_12_1$	17.1/NA	106	
14	1c74	1.9	$P3_121$	18.9/22.4	106	K53,56M
15	1mkt	1.72	$P3_121$	19.5/28.4	106	
16	1kvx	1.9	$P2_12_12_1$	20.0/31.3	98	D99A
17	1mkv	1.89	$P3_121$	18.0/NA	88	TSA
18	1fdk	1.91	$P3_121$	18.4/28.0	86	MJ33
19	1o3w	1.85	$P3_121$	19.3/23.2	85	K53,56,120M
20	1ceh	1.9	$P3_121$	18.5/NA	81	D99N
21	1irb	1.9	$P3_121$	19.2/NA	81	K120,121A
22	1mku	1.8	$P2_12_12_1$	19.6/NA	80	Y52,73F, D99N
23	1mks	1.9	$P3_121$	18.6/NA	77	Y52,73F, D99N
24	1kvy	1.9	$P3_121$	19.8/27.7	70	D49E
25	1kvw	1.95	$P3_121$	20.9/31.4	68	H48Q

† References: 1, Steiner *et al.* (2001); 2 and 4, Sekar *et al.* (2005); 3 and 12, Sekar, Yogavel *et al.* (2006) 5, Sekar *et al.* (2004); 6, 8 and 9, Kanaujia & Sekar (2008); 7, Sekar & Sundaralingam (1999); 10, Rajakannan *et al.* (2002); 11, Sekar, Gayathri *et al.* (2006); 13, Dijkstra *et al.* (1981); 14, Yu *et al.* (2000); 15, Sekar, Sekharudu *et al.* (1998); 16, 24 and 25, Sekar *et al.* (1999); 17, Sekar, Kumar *et al.* (1998); 18, Sekar, Eswaramoorthy *et al.* (1997); 19, Sekar *et al.* (2003); 20, Kumar *et al.* (1994); 21, Huang *et al.* (1996); 22 and 23, Sekar, Yu *et al.* (1997). ‡ ANN, 4-methoxybenzoic acid; TSA, 1-*O*-octyl-2-heptylphosphonyl-*sn*-glycero-3-phosphoethanolamine; MJ33, 1-decyl-3-trifluoroethyl-*sn*-glycero-2-phosphomethanol. § The number of water molecules given for PDB code 2bd1 is for both the molecules found in the asymmetric unit.

considered to be invariant if it was observed in all the structures except when the sites were occupied by other molecules such as 2-methyl-2,4-pentanediol (MPD) or tris(hydroxymethyl)aminomethane (Tris) (which were ingredients of the crystallization mixture) in some structures. The conformation with the higher occupancy was considered in the analysis in the case of alternate conformations. However, in the case of equal occupancies the first conformation was taken for the study. A similar approach was also followed for the water molecules. Hydrogen-bond interactions were calculated using the program *HBPLUS* (McDonald & Thornton, 1994). The solvent-accessible area of invariant water molecules was computed using the program *NACCESS* (Hubbard & Thornton, 1993) with a probe radius of 1.4 Å. Water molecules with an accessible surface area that was less than or equal to 2.5 $Å^2$ were considered to be internal or buried water molecules. The normalized B factor (B_i') for all invariant water molecules was calculated using the formula $B_i' = (B_i - \langle B \rangle)/\sigma(B)$, where B_i is the B factor of each atom, $\langle B \rangle$ is the mean B factor of the protein molecule and $\sigma(B)$ is the standard deviation of the B factors (Smith *et al.*, 2003).

2.2. Molecular-dynamics simulations

Molecular-dynamics (MD) simulations were performed using the *GROMACS* v.3.3 package (van der Spoel *et al.*, 2005) running on parallel processors with the OPLS-AA/L all-atom force field (Jorgensen *et al.*, 1996; Kaminski *et al.*, 2001). During molecular simulations, the crystallographic water molecules were removed from the protein models. However, the ligand, the molecules from the crystallization conditions and the calcium ions were retained during the MD simulations. A cubic box of dimensions 6.4 × 6.4 × 6.4 nm was generated using the *editconf* module of *GROMACS*. The necessary parameters and the topology file for MPD and Tris molecules were generated using the *PRODRG* server (Schüttelkopf & van Alten, 2004). Subsequently, the protein models were solvated with the SPC (simple point charge) water model using the *genbox* program available in the *GROMACS* suite. Energy minimization was performed using the conjugate-gradient method for 200 ps with the maximum force-field cutoff being 1 kJ mol^{-1}. Sodium and chloride ions were used to neutralize the overall charge of the system. The simulations utilized NPT ensembles with isotropic pressure coupling (τ_p = 0.5 ps) to 100 kPa and temperature coupling (τ_t = 0.1 ps) to 300 K. Parrinello–Rahman and Nose–Hoover coupling protocols were used for pressure and temperature, respectively. Long-range electrostatics were computed using the Particle Mesh Ewald (PME) method (Darden & York, 1993) and Lennard–Jones energies were cut off at 1.0 nm. Bond lengths were constrained with the *LINCS* algorithm (Hess *et al.*, 1997). MD simulations were performed for a time period of 3 ns for all the structures considered in the present study. Analyses were performed using the tools available in the *GROMACS* suite. The average structures used for comparison and analysis were calculated using the ensembles generated between 1 and 3 ns. The ensembles were computed every 2 ps and calculation of the residence frequency of the invariant water molecules was carried out using these ensembles. The interactions between protein atoms and solvent molecules were calculated with a hydrogen-bond distance of 3.4 Å. The residence of a particular invariant water molecule near the interacting residue(s) was assumed if at least one of the polar atoms made contact with any solvent molecule within a distance of 3.4 Å.

3. Results and discussion

3.1. All 24 invariant water molecules

The relevant details of all 25 crystal structures used in the present study are given in Table 1. A total of 24 invariant

Table 2
List of hydrogen-bond interactions calculated using the program *HBPLUS* (McDonald & Thornton, 1994) between invariant water molecules and protein atoms.

Water†	Invariant water	Main-chain protein atoms	Side-chain protein atoms	Waters†	Others
131	IW1		His48 $N^{\delta1}$, Asp49 $O^{\delta1}$	147	
132	IW2	Cys45 O	His48 $N^{\delta1}$, Asp49 $O^{\delta1}$	136	
133	IW3	Ala1 N, Pro68 O	Tyr52 OH, Asp99 $O^{\delta2}$		
134	IW4	Cys98 O			
136	IW5	Tyr28 O, Gly30 N		132	MPD126 O4
140	IW6	Ala93 O		145, 167	
141	IW7	Asp40 O		185	
143	IW8	Gly32 O		147, 276	
144	IW9	Leu41 N, Pro110 O		157	
145	IW10	Ser85 O	Asn88 $N^{\delta2}$	140, 207	
148	IW11		Glu17 $O^{\varepsilon2}$	358	
149	IW12	Leu19 N		316	
151	IW13		Arg100 NH1, Asn101 $N^{\delta2}$		
154	IW14	Glu81 O	Thr83 $O^{\gamma1}$	197, 230	
157	IW15	Asp40 N	Asp40 $O^{\delta2}$	144, 204	
159	IW16	Thr36 O, Asn122 O		163, 225	
162	IW17	Ser15 O	Asp21 $O^{\delta2}$	137	MPD129 O2
167	IW18		Asn97 $N^{\delta2}$	140, 173, 248	
168	IW19	Phe106 O, Val109 O		298, 328	
173	IW20	Cys84 N		167, 281	
183	IW21	Ser15 N	Ser15 O^{γ}		
208	IW22	Leu31 N	Asn23 $O^{\delta1}$	266, 302	
232	IW23		Ser107 O^{γ}	177, 255, 343	
236	IW24		Gln46 $N^{\varepsilon2}$, Thr47 $O^{\gamma1}$		

† The numbering scheme is taken from the reference structure (1g4i).

water molecules were identified and were further classified into three groups: Cluster1, Cluster2 and Cluster3 (Fig. 2). In addition, the details of the interaction(s) of these invariant water molecules are listed in Table 2. As indicated above, the crystal structure 1g4i (Steiner *et al.*, 2001) was taken as the reference structure throughout the discussion and unless otherwise mentioned the numbering scheme used for the residues and water molecules corresponds to that of 1g4i. Although the identification of a few well conserved water molecules was straightforward, difficulties were encountered with regard to the others. Such difficulties were anticipated even when analyzing the same protein in different crystal forms with different resolutions and space groups. Critical examination of protein–water interaction(s) and visual inspection methods were used to identify the invariant water molecules. A close examination of Table 2 reveals that the invariant water molecules are mainly hydrogen bonded to the main-chain polar atoms of the protein molecule. Furthermore, analysis of these interactions suggests that a significant number of water molecules are conserved in the vicinity of the active site and at the interfacial site of the enzyme. Interestingly, no invariant water molecules are observed near the surface loop (residues 60–70) except for the structural water molecule (hereafter referred to as IW3), which is hydrogen bonded to the backbone O atom of Pro68. The water-numbering scheme (first row), the normalized *B* factor (second row), solvent-accessible area (third row) and percentage residence frequency (fourth row) of these invariant water molecules are listed in Tables 3, 4 and 5.

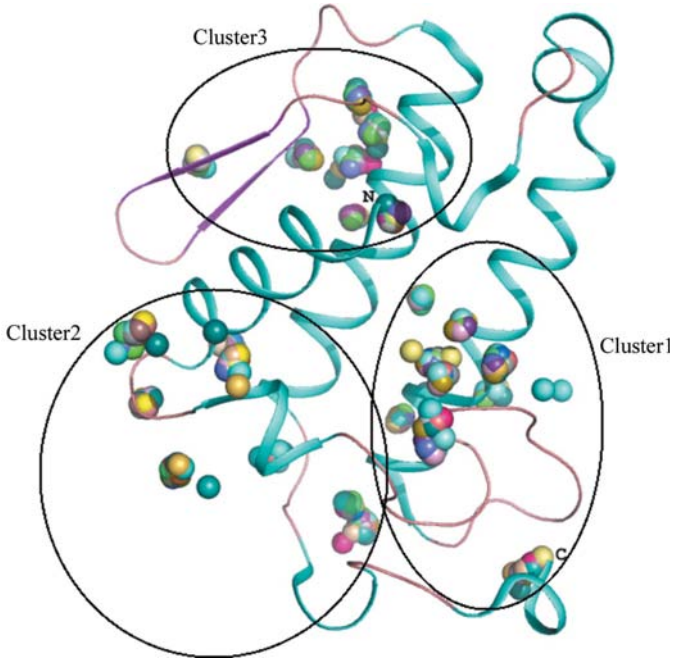

Figure 2
The invariant water molecules in all 25 crystal structures are shown in different colours and are grouped into three clusters based on their location in the three-dimensional structure. The protein model shown belongs to structure 1g4i.

3.1.1. Invariant water molecules in Cluster1. The nine invariant water molecules grouped into Cluster1 (Fig. 2) are mainly in the vicinity of the active site. Of these nine, a total of six invariant water molecules (IW1, IW2, IW4, IW5, IW8 and IW22) are very close to the active site of the enzyme (Fig. 3). Moreover, the water molecules IW1, IW2, IW4, IW5 and IW8 are buried and are highly stable with very low *B* factors (Table 3); they are considered to be internal water molecules. However, three internal water molecules (IW1, IW5 and IW8) are exposed with solvent-accessible areas of more than 5.0 Å^2 in some of the mutant structures (2zp4, 1kvx, 1o3w, 1ceh, 1mku and 1mks) and two native structures (1mkt and 1irb) owing to disturbances in the active-site hydrogen-bonding network.

The internal water molecule IW1, which is hydrogen bonded to His48 $N^{\delta1}$ and Asp49 $O^{\delta1}$, has previously been shown to be involved in the tautomerization of the catalytically important imidazole residue His48 (Sekar & Sundaralingam, 1999). Moreover, the analysis reveals that this invariant water molecule (IW1) may be involved in the stabilization of the functionally important residue Asp49 by donating a proton and is also hydrogen bonded to another water molecule (147) which stabilizes the surface-loop residue

Table 3
Invariant water molecules corresponding to the reference structure (1g4i) along with their normalized *B* factor in Å^2 (second row), accessible surface area in Å^2 (third row) and percentage residence frequency (fourth row) calculated from the ensembles generated using molecular-dynamics simulations for Cluster1.

M, mutant structure; D, electron density is present at 0.8σ; L, a ligand molecule is present at the spatial position of the water molecule; O, other atoms of the protein molecule; NSF, structure factors are not available in the PDB.

	IW1	IW2	IW4	IW5	IW7	IW8	IW16	IW22	IW24
1g4i	131	132	134	136	141	143	159	208	236
	−0.6	−0.8	−0.6	−0.7	−0.5	−0.2	0.0	0.6	0.1
	0.0	0.0	1.4	0.0	11.3	0.0	0.2	1.2	15.4
	100	100	98.8	100	94.4	100	99.6	100	100
1vl9	205	430	202	206	306	204	213	251	225
	−0.6	−0.8	−0.7	−0.8	−0.3	−0.7	−0.5	0.4	−0.1
	0.0	0.0	0.0	0.0	8.8	0.0	15.7	18.1	21.8
	100	100	94.2	100	38.4	100	72.9	100	100
2bch	240	217	206	231	204	297	227	236	215
	−0.1	−0.5	−0.8	−0.1	−0.6	0.4	0.1	1.0	−0.5
	0.0	0.0	2.4	0.0	9.5	0.0	1.5	3.0	15.1
	100	100	99.4	100	99.8	100	99.9	100	100
2bax	214	205	203	257	230	262	228	232	207
	−0.4	−0.6	−0.8	−0.5	−0.6	−0.2	0.2	0.8	−0.7
	0.0	0.0	2.2	0.0	11.2	0.0	1.2	8.7	15.8
	100	100	99.5	100	98.8	100	99.7	100	100
1vkq	7	2	3	17	13	20	28	40	8
	−0.5	−0.6	−0.8	−0.6	−0.7	−0.3	0.0	0.8	−0.6
	0.0	0.0	2.7	0.0	6.9	0.1	3.4	8.9	14.9
	100	100	93.2	100	99.8	100	94.5	100	100
2zp3	M	253	218	M	210	M	D	282	204
	—	0.7	−0.7	—	−0.6	—	—	0.9	−0.6
	—	1.1	2.3	—	0.0	—	—	25.3	6.4
	99.9	100	97.9	100	95.2	100	63.2	100	100
1une	266	268	201	263	203	264	213	225	216
	−0.5	−0.8	−1.0	−1.1	0.0	0.2	−0.1	2.3	0.9
	0.0	0.0	1.8	0.0	13.0	0.8	13.7	7.1	8.8
	100	100	95.7	100	97.3	100	99.9	100	100
2zp4	203	212	228	208	214	225	268	238	213
	−0.6	−0.9	−0.8	−0.8	−0.6	−0.2	−0.2	0.4	−0.5
	1.2	0.0	2.1	0.0	4.8	9.2	18.7	22.7	7.7
	100	100	89.7	100	95.0	100	93.1	100	100
2zp5	M	229	254	M	207	M	238	268	255
	—	0.4	−0.8	—	−0.9	—	0.3	1.6	−0.5
	—	0.0	2.7	—	4.3	—	12.5	0.0	5.9
	99.9	99.9	98.3	99.9	97.0	100	99.9	100	100
1gh4	206	218	209	201	321	202	212	230	233
	−0.4	−0.5	−0.4	−1.4	1.7	−0.8	0.0	0.9	0.9
	0.0	0.9	0.0	0.2	20.0	0.0	15.5	15.3	20.1
	100	100	97.4	100	96.7	100	98.5	100	100
2b96	L	204	212	L	216	203	225	259	275
	—	−0.1	−0.8	—	−0.1	−0.3	0.2	1.3	0.8
	—	0.0	2.5	—	10.8	0.5	4.5	6.3	4.9
	100	100	97.0	34.2	96.6	99.4	89.7	99.9	100
2bd1	204	231	210	202	300	208	213	356	D
	−0.4	−0.6	−0.8	−1.2	0.9	−0.8	−0.8	0.3	—
	0.0	0.3	0.0	0.0	20.2	0.2	13.5	13.0	—
	100	100	77.9	100	94.4	100	98.5	100	100
1bp2	7	6	O	5	16	12	13	47	O
	−0.5	−0.6	—	−0.6	−0.2	−0.3	−0.2	1.1	—
	2.2	0.0	—	2.3	6.7	1.4	13.1	10.3	—
	100	100	92.1	100	99.6	100	99.1	100	100
1c74	223	205	212	202	207	203	219	228	293
	−0.3	−0.5	−0.7	−0.8	−0.2	0.5	0.3	0.7	0.3
	0.1	1.2	2.5	0.0	12.9	2.1	15.4	20.5	4.8
	100	100	88.5	100	97.5	100	99.5	100	100

Tyr69 in the active site. As indicated in the literature, Tyr69 has been suggested to be involved in the binding of the pro-*S* nonbridging O atom of *sn*-3 phosphate (Scott & Sigler, 1994). Furthermore, observations from the crystal structure of PLA_2 with a transition-state analogue (PDB code 1mkv; Sekar, Kumar *et al.*, 1998) show that it may be involved in holding the substrate molecule during enzyme catalysis. In addition, the MD simulations also show the presence of this invariant water molecule (IW1) with 100% residence frequency (Table 3).

The internal water molecule IW2 (which is hydrogen bonded to the backbone O atom of Cys45, His48 $N^{\delta1}$ and

Table 3 (continued)

	IW1	IW2	IW4	IW5	IW7	IW8	IW16	IW22	IW24
1mkt	226	205	212	202	207	203	222	234	220
	0.6	0.3	−1.2	−0.2	−1.0	0.1	0.9	1.3	0.3
	0.0	0.2	2.8	0.0	22.0	5.3	13.9	17.1	10.5
	100	100	99.5	100	98.6	100	100	100	100
1kvx	259	228	201	227	202	226	208	214	NSF
	1.5	−0.3	−0.7	−1.0	−0.9	2.1	0.8	0.2	—
	5.7	2.5	2.2	8.9	12.0	1.8	14.3	14.1	—
	100	100	80.0	100	95.6	100	99.2	100	100
1mkv	L	L	208	L	204	L	D	226	214
	—	—	−1.1	—	−0.2	—	—	0.4	0.6
	—	—	2.8	—	9.5	—	—	24.2	6.4
	0.0	100	99.4	0.0	99.6	0.0	96.6	100	100
1fdk	L	L	206	L	202	D	209	L	222′
	—	—	−0.9	—	−0.4	—	0.0	—	—
	—	—	2.8	—	11.4	—	9.5	—	—
	100	0.0	99.4	0.0	99.7	100	100	100	100
1o3w	224	205	212	202	207	203	220	229	219
	0.1	−0.6	−1.0	−0.4	−0.4	0.3	0.6	0.9	0.9
	0.7	0.0	2.6	0.0	11.7	8.1	13.9	12.9	5.8
	100	100	98.4	100	99.3	100	100	100	00
1ceh	237	216	215	204	210	238	214	259	242
	0.3	−0.3	−0.7	−0.5	−0.2	1.2	−0.4	1.2	0.7
	0.6	0.0	2.5	1.7	10.0	11.7	11.8	27.8	11.8
	100	100	0.73	100	96.9	100	97.5	99.9	100
1irb	221	213	212	204	210	205	240	231	222
	0.5	−0.5	−0.6	1.1	−0.3	1.3	0.1	0.8	0.3
	5.5	2.9	2.9	16.6	7.5	15.1	15.5	16.6	10.5
	100	100	91.8	100	98.1	100	99.1	100	100
1mku	211	230	201	235	204	236	221	249	228
	−0.2	−0.2	−0.6	−0.3	−0.4	0.6	−0.1	1.3	0.5
	11.5	2.2	2.1	7.4	19.7	20.1	17.0	13.7	12.4
	100	100	97.0	100	96.2	100	97.0	100	100
1mks	238	205	210	202	207	201	215	253	214
	0.1	0.0	−1.2	0.4	−0.5	0.8	0.6	1.1	0.9
	6.5	0.5	2.5	8.2	9.8	23.3	17.4	23.2	6.3
	100	100	89.7	100	98.2	100	92.9	100	100
1kvy	250	251	208	M	204	M	252	NSF	215
	1.0	0.9	−0.9	—	−0.3	—	0.7	—	0.7
	0.5	0.6	3.0	—	21.9	—	8.3	—	11.5
	100	100	93.0	100	92.0	100	91.7	100	100
1kvw	M	204	208	202	245	243	NSF	NSF	215
	—	0.1	0.4	0.2	1.4	1.0	—	—	0.5
	—	0.5	3.8	0.4	23.9	0.2	—	—	17.2
	0.0	100	91.1	100	99.8	100	100	100	100
AVG†	−0.2	−0.5	−0.7	−0.4	−0.4	0.0	0.0	0.8	−0.4
	0.9	0.3	1.9	1.2	11.3	2.4	7.9	10.6	13.6
	100	100	95.9	98.5	90.1	100	94.4	100	100

† The average values of normalized *B* factor, solvent accessibility and residence frequency during MD simulations.

Table 4
Invariant water molecules corresponding to the reference structure (1g4i) along with their normalized *B* factor in Å^2 (second row), accessible surface area in Å^2 (third row) and percentage residence frequency (fourth row) calculated from the ensembles generated using molecular-dynamics simulations for Cluster2.

Symmetry-related water molecules are indicated with a prime; D, electron density is present at 0.8σ; O, other atoms of the protein molecule; NSF, structure factors are not available in the PDB.

	IW9	IW11	IW12	IW15	IW17	IW19	IW21	IW23
1g4i	144	148	149	157	162	168	183	232
	−0.4	−0.6	0.1	0.1	−0.3	0.6	0.9	1.20
	0.0	34.1	0.0	2.7	18.9	0.0	14.0	1.0
	100	100	79.0	100	100	100	79.4	100
1vl9	214	302	244	224	252	226	405	339
	−0.4	0.7	0.8	0.2	0.7	0.0	−0.8	2.1
	0.0	43.0	10.3	4.0	12.5	0.0	21.1	10.6
	100	100	64.7	100	100	100	76.9	99.9
2bch	211	203	202	420	219	212	238	220
	−0.4	−0.7	−0.7	0.0	−0.3	−0.3	0.3	0.1
	0.3	39.3	6.1	6.1	17.7	0.0	25.3	0.0
	100	100	53.8	100	100	100	76.3	99.9
2bax	210	209	202	225	218	271	242	222
	−0.5	−0.7	−0.7	−0.3	−0.3	−0.3	0.3	0.0
	0.0	47.0	0.1	0.5	36.7	0.2	13.0	0.0
	100	100	79.4	100	100	57.5	76.9	99.2
1vkq	10	18	4	23	37	27	53	22
	−0.6	−0.7	−0.7	−0.1	−0.1	−0.3	0.3	−0.1
	0.0	38.0	6.5	5.5	15.7	0.1	3.4	9.0
	100	100	82.1	100	100	98.0	89.8	99.4
2zp3	234	221′	215	227	237	229	223	216
	−0.4	—	−0.8	0.0	0.0	−0.4	0.0	−0.1
	0.0	—	7.5	19.5	31.5	0.0	25.5	9.9
	100	100	76.1	100	100	99.9	82.6	97.9
1une	208	215	209	207	223	228	226	327
	−0.4	−0.7	0.8	1.6	0.3	0.1	1.9	1.8
	0.0	47.2	10.3	4.1	27.7	9.3	24.3	17.5
	100	100	89.8	100	100	100	100	99.8
2zp4	226	233′	218	273	257	236	256	216
	−0.5	—	−0.6	0.2	0.0	−0.2	0.2	−0.4
	0.0	—	5.3	18.5	27.8	0.0	26.4	7.2
	100	100	57.6	100	100	100	96.8	98.8
2zp5	217	242	226	235	222	220	210	204
	−0.5	−0.8	−0.5	0.4	−0.6	−0.3	−0.2	−0.4
	0.0	43.5	0.0	16.5	22.3	3.5	14.5	8.8
	100	100	95.6	100	100	100	87.3	98.8
1gh4	229	282	264	222	320	234	268	232
	0.2	2.7	1.3	0.2	1.4	0.5	0.8	2.2
	0.0	22.6	32.3	19.2	26.4	10.0	23.9	22.4
	100	100	99.7	100	100	100	74.7	99.9
2b96	228	215	278	227	D	O	250	314
	−0.3	−0.5	0.3	−0.3	—	—	0.0	1.4
	0.0	47.7	2.1	29.4	—	—	20.8	9.7
	100	100	83.8	100	100	100	81.0	99.9
2bd1	242	O	414	224	264	240	274	338
	−0.4	—	1.3	0.5	0.3	0.2	0.5	32.0
	0.0	—	13.7	6.6	22.6	8.5	27.2	10.1
	100	100	97.1	100	100	100	93.8	99.9
1bp2	14	O	31	55	NSF	38	NSF	O
	−0.2	—	0.5	1.3	—	0.7	—	—
	0.0	—	5.2	15.4	—	2.8	—	—
	100	100	77.1	100	100	100	80.6	10.7
1c74	210	214	213	249	217	287	236	240
	−0.7	−0.7	−0.6	1.0	−0.1	0.3	1.1	0.6
	0.0	46.0	0.0	16.8	20.1	0.2	22.7	9.5
	100	100	77.6	100	100	100	73.7	99.8

Asp49 O$^{\delta1}$) is very important for the catalytic activity of the enzyme and is known to act as a nucleophile during enzyme hydrolysis (Steiner *et al.*, 2001; Sekar *et al.*, 2005). In fact, irrespective of the resolution, space group and biochemical properties of the enzyme, the water molecule IW2 is present in all the structures. Recently, the crystal structure of the active-site single mutant H48N of PLA_2 suggested the involvement and a possible role of the invariant water molecule IW2 in the low enzyme activity of the mutant enzyme (PDB code 2zp4; Kanaujia & Sekar, 2008). Interestingly, as expected, the catalytic water molecule (IW2) is present in all MD simula-

Table 4 (continued)

	IW9	IW11	IW12	IW15	IW17	IW19	IW21	IW23
1mkt	210	215	213	237	219	245	252	269
	−0.2	−0.1	−0.3	0.7	0.3	0.5	1.1	1.1
	0.0	34.4	0.0	10.0	24.0	1.1	26.5	9.6
	100	100	79.8	100	100	100	100	99.9
1kvx	206	250	207	233	213	NSF	241	O
	−1.0	−0.1	−0.7	1.1	−0.2	—	0.8	—
	0.0	42.0	0.0	7.9	25.5	—	26.6	—
	100	100	95.8	100	100	100	81.8	99.5
1mkv	207	288	209	229	213	211	232	246
	−0.2	−0.3	−0.4	0.4	0.3	0.0	1.3	0.9
	0.0	41.5	0.7	4.1	29.4	0.0	33.3	9.3
	99.9	100	95.4	100	100	100	80.8	99.7
1fdk	204	207	263	219	257	249	268	O
	−0.3	0.0	0.4	0.5	1.4	0.6	0.3	—
	0.0	47.5	0.1	13.3	19.2	1.4	40.0	—
	100	100	100	100	100	100	71.7	99.8
1o3w	210	214	213	230	218	235	238	243
	−0.3	−0.7	−0.6	0.9	0.1	0.4	0.7	0.8
	0.0	42.6	0.0	8.1	19.0	0.2	32.0	15.8
	100	100	74.7	100	100	100	80.6	99.7
1ceh	203	209	243	218	258	233	239	241
	−0.5	−0.4	0.3	−0.5	0.8	−0.2	0.6	2.1
	0.0	45.9	8.4	15.3	29.3	0.0	15.2	1.1
	100	100	76.7	100	100	99.9	62.5	95.6
1irb	203	209	D	214	243	218	D	242
	−0.2	0.0	—	0.6	0.0	0.1	—	1.9
	0.0	42.4	—	10.6	22.0	0.0	—	15.7
	100	100	92.7	100	100	100	92.0	99.2
1mku	210	226	212	208	246	257	252	O
	−0.5	−0.8	−0.5	0.4	0.3	1.2	1.1	—
	0.0	44.2	0.1	14.9	26.3	8.6	17.1	—
	100	100	39.7	100	100	100	74.1	100
1mks	209	233	211	255	236	213	D	256
	−0.2	0.0	−0.6	1.1	0.6	0.3	—	0.8
	0.0	49.6	0.3	11.4	22.3	2.7	—	14.1
	100	100	51.2	100	100	100	77.8	99.7
1kvy	207	210	209	225	214	212	NSF	243
	−0.4	−0.5	−0.6	0.9	−0.4	−0.4	—	1.2
	0.0	45.0	0.4	9.8	26.7	0.0	—	25.4
	100	100	30.2	100	100	100	78.6	99.4
1kvw	207	210	209	225	214	228	NSF	NSF
	−0.2	−0.1	−0.1	1.4	0.6	0.5	—	—
	0.0	55.9	0.1	13.0	20.6	0.2	—	—
	100	100	96.0	100	100	100	85.4	99.7
AVG†	−0.4	−0.3	−0.1	0.3	0.2	0.1	0.4	0.9
	0.0	38.0	4.3	8.1	21.7	1.2	18.4	7.4
	100	100	74.4	100	100	94.7	80.3	97.6

† The average values of normalized *B* factor, solvent accessibility and residence frequency during MD simulations.

Table 5
Invariant water molecules corresponding to the reference structure (1g4i) along with their normalized *B* factor in Å^2 (second row), accessible surface area in Å^2 (third row) and residence frequency (%) (fourth row) calculated from the ensembles generated using molecular-dynamics simulations for Cluster3.

SC, symmetry-related contact; NSF, structure factors are not available in the PDB; D, electron density is present at 0.8σ; O, other atoms of the protein molecule.

	IW3	IW6	IW10	IW13	IW14	IW18	IW20
1g4i	133	140	145	151	154	167	173
	−0.3	−0.3	−0.5	−0.3	0.5	0.3	0.3
	0.0	4.3	0.0	0.3	0.0	0.0	2.0
	100	88.7	92.4	100	99.6	100	97.0
1vl9	203	237	208	228	286	279	235
	−0.6	0.7	−0.4	0.1	1.2	0.7	0.1
	0.0	22.0	1.7	12.7	3.5	21.7	0.0
	100	89.4	98.0	100	79.8	100	98.6
2bch	208	226	209	419	207	424	235
	−0.6	−0.5	−0.7	−0.3	−0.7	0.3	0.2
	0.0	10.0	0.0	0.0	1.1	7.4	3.4
	54.1	87.6	98.5	100	99.3	100	96.3
2bax	211	215	208	313	206	231	221
	−0.5	−0.5	−0.6	−0.4	−0.6	−0.1	−0.5
	0.0	0.0	0.0	1.4	0.5	12.7	6.3
	52.3	70.0	99.5	100	93.9	100	99.2
1vkq	19	16	5	35	1	54	12
	−0.3	−0.5	−0.6	0.3	−0.7	0.2	−0.2
	0.0	14.0	2.6	0.3	0.4	15.0	11.7
	100	59.6	97.9	100	66.4	100	98.7
2zp3	224	231	203	212	208	273	205
	−0.4	−0.4	−0.9	0.0	−0.5	0.3	−0.9
	0.0	9.6	3.0	0.0	5.3	0.8	4.9
	100	93.7	99.1	100	99.9	100	92.4
1une	276	224	218	235	211	289	288
	−0.4	0.4	0.0	1.6	0.9	1.7	1.4
	0.0	7.7	0.4	0.2	0.7	9.2	12.2
	100	90.0	99.8	100	97.5	100	91.7
2zp4	227	253	201	219	207	210	205
	−0.5	−0.4	−0.7	−0.1	−0.5	0.0	−0.6
	0.0	14.1	3.0	7.9	4.7	16.9	21.1
	100	74.1	93.3	100	60.6	99.9	98.5
2zp5	219	215	225	218	213	246	216
	−0.6	−0.4	−0.7	−0.1	−0.6	−0.3	−0.7
	0.0	10.8	3.5	3.6	4.4	7.1	6.4
	100	59.3	96.8	100	83.2	99.9	98.4
1gh4	204	311	208	245	SC	265	316
	−1.3	1.3	0.2	0.6	—	1.7	1.3
	0.0	0.3	0.5	14.6	—	23.0	19.6
	100	93.8	98.2	100	98.9	100	98.3
2b96	202	258	236	255	234	322	235
	−0.5	0.1	−0.1	0.3	−0.4	2.1	−0.9
	0.0	24.7	0.0	5.5	6.6	3.7	18.7
	100	92.2	99.5	100	82.5	100	95.5
2bd1	214	296	244	259	SC	333	260
	−0.8	1.4	0.2	0.5	—	1.3	0.2
	0.0	25.2	0.4	0.0	—	12.9	15.8
	100	89.4	99.3	100	99.4	100	96.4
1bp2	11	23	18	36	32	NSF	50
	−0.3	0.2	−0.1	0.7	0.5	—	1.2
	0.0	4.9	0.0	2.1	2.7	—	25.3
	100	84.8	97.7	100	77.6	100	96.8
1c74	211	206	204	254	209	215	220
	−0.5	0.0	−0.4	0.6	0.5	0.7	0.2
	0.0	13.2	0.0	7.4	13.2	2.3	2.2
	100	75.3	96.5	100	99.8	100	92.1

tions with a residence frequency of 100% (Table 3), with the exception of 1fdk, in which one of the ligand atoms is hydrogen bonded to His48 N$^{\delta1}$. Surprisingly, in the case of 1mkv (in which the ligand is bound in the crystal structure), the catalytic water molecule (IW2) is also observed to be present during the MD simulations (Table 3). According to the proposed enzyme mechanism for the catalytic activity (Scott *et al.*, 1990), these observations correspond to the last step of the catalytic process, in which the substrate molecule is cleaved and the product is displaced from the active site. During this

Table 5 (continued)

	IW3	IW6	IW10	IW13	IW14	IW18	IW20
1mkt	211	206	204	D	209	216	223
	−0.3	0.2	0.1	—	−0.5	1.3	1.2
	0.0	2.2	0.1	—	7.4	2.5	4.0
	100	94.3	99.4	100	90.0	100	94.1
1kvx	M	251	210	235	SC	217	222
	—	0.9	−0.2	−0.3	—	0.1	−0.6
	—	3.2	4.1	3.2	—	12.3	13.2
	100	86.2	99.2	100	98.3	100	98.3
1mkv	286	203	202	258	206	D	216
	−0.1	0.0	−0.7	0.5	−0.7	—	−0.7
	0.0	14.0	0.5	4.0	6.4	—	18.0
	0.0	90.2	98.2	100	97.2	100	98.4
1fdk	205	248	201	D	203	D	210
	−0.1	0.2	−0.3	—	−0.4	—	−0.7
	0.0	15.9	0.8	—	6.7	—	25.6
	100	93.1	99.6	100	98.4	100	98.9
1o3w	211	206	204	257	209	215	221
	−0.6	−0.1	−0.7	0.5	−0.3	0.3	−0.7
	0.3	1.0	4.4	8.4	4.7	13.6	7.2
	59.7	48.4	98.3	100	76.9	99.8	98.8
1ceh	O	220	206	226	205	222	234
	—	−0.5	−0.7	−0.2	−0.7	0.0	−0.4
	—	31.3	0.2	4.5	0.0	19.3	9.6
	0.0	70.2	96.7	100	64.0	99.9	97.4
1irb	201	215	206	235	225	262	219
	0.3	1.7	−0.8	−0.1	−0.1	0.9	−0.1
	0.6	23.5	2.5	7.5	11.9	8.4	5.5
	100	86.1	97.8	100	79.3	100	96.0
1mku	O	248	231	277	217	247	220
	—	0.3	−0.5	0.1	0.3	1.3	0.2
	—	4.3	0.3	6.1	1.7	13.5	10.8
	99.8	86.2	100	100	96.5	100	96.4
1mks	O	206	203	245	208	234	216
	—	−0.4	−0.6	−0.3	−0.7	0.6	0.4
	—	18.6	3.4	8.5	6.4	16.4	20.3
	0.0	86.1	95.2	100	87.8	100	99.4
1kvy	261	203	202	239	206	NSF	216
	0.0	−0.1	−0.9	0.6	−0.2	—	−0.1
	0.0	3.5	0.0	2.5	6.5	—	14.9
	100	87.0	98.1	100	72.8	100	97.9
1kvw	255	244	203	NSF	206	211	218
	0.8	0.6	−0.2	—	−0.1	0.4	1.0
	0.0	16.9	4.4	—	7.2	13.9	14.7
	100	89.1	99.8	100	82.6	100	96.6
AVG†	−0.4	0.1	−0.5	0.1	−0.1	0.5	0.0
	0.0	10.2	1.0	3.9	3.0	9.8	8.2
	84.7	83.2	97.6	100	89.7	100	97.3

† The average values of normalized *B* factor, solvent accessibility and residence frequency during MD simulations.

process, three water molecules (IW2, IW5 and IW8) occupy the void created by the substrate molecule in the active site.

The invariant water molecule IW4 seems to be structurally crucial in positioning the important residue Cys98. It is notable that the invariant water molecule IW4 is present in all of the structures, irrespective of crystal form and mutations, with the exception of 1bp2, in which one MPD molecule replaces this water molecule and is hydrogen bonded to the backbone O atom of Cys98 with a distance of 2.98 Å. Interestingly, water molecule IW4 is buried in a cavity that is relatively hydrophobic in nature. Furthermore, it is interesting to note that the disulfide bond (Cys98–Cys51) plays a key role in maintaining the two longest helices (α_3 and α_5) of the enzyme (Fig. 1). This water molecule is also present during the MD simulations with a residence frequency of 96%, except in one structure (1ceh; residence frequency of 1%). The low residence frequency of this water molecule in the single-mutant structure (D99N; PDB code 1ceh) may possibly be a consequence of the mutation at position 99. Furthermore, it is observed in the MD averaged structure that two residues (Thr47 and Asn101) have moved towards the spatial position of IW4, making it almost non-accessible to solvent molecules. Thus, water molecule IW4 can be treated as an integral part of the structure.

The invariant water molecules IW5 and IW8 provide coordination to the functionally important calcium ion. It has been shown that these two water molecules are replaced by the phosphonyl O atoms of the substrate (Scott *et al.*, 1990). Interestingly, the residence frequencies computed from MD simulations for these two calcium-coordination water molecules only reveal their absence in the case of ligand-bound structures, in which two of the ligand atoms are coordinated to the functionally important calcium ion (Table 3). Together with another water molecule (185), water molecule IW7 is involved in stabilizing the surface residue Asp40 (Fig. 3), which forms an ion pair with Arg43. Together with the disulfide bond between Cys27 and Cys123, the invariant water molecule IW16 basically connects the active-site loop and C-terminal residues of the enzyme. Another invariant water molecule, IW22, which is hydrogen bonded to the backbone N atom of Leu31 and the side-chain O atom of Asn23, is likely to anchor the active-site calcium-binding loop. Furthermore, the invariant water molecule IW24 seems to be involved in stabilizing the active-site residue His48 and is hydrogen bonded to Gln46 and Thr47. In addition, all these invariant water molecules (IW7, IW16, IW22 and IW24) are present during the MD simulations with more than 90% residence frequency and have very low normalized *B* factors, with the exception of IW22 (Table 3).

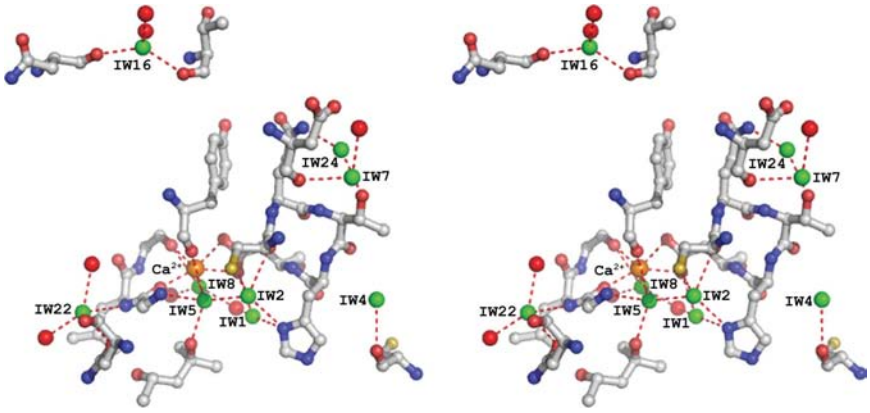

Figure 3
The nine invariant water molecules of Cluster1 are shown as green spheres together with their hydrogen-bonding interactions. Other water molecules that interact with the invariant water molecules are shown as red spheres. The catalytic calcium ion is shown in orange.

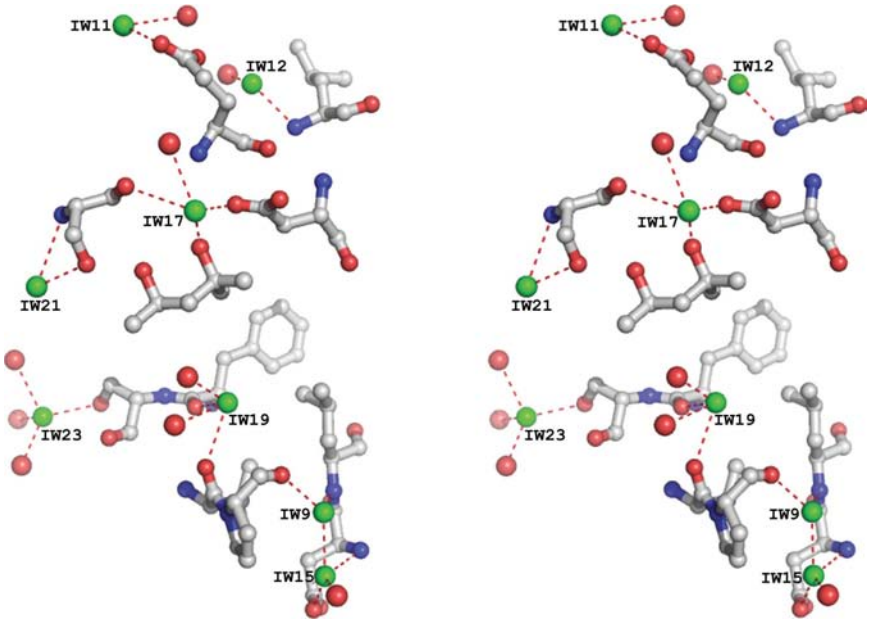

Figure 4
The eight invariant water molecules of Cluster2 are shown as green spheres together with their hydrogen-bonding interactions with the protein molecule. The red spheres are other water molecules that interact with the invariant water molecules.

3.1.2. Invariant water molecules in Cluster2. The invariant water molecule IW9 is totally buried with a low normalized *B* factor (Table 4) and is hydrogen bonded to the backbone N atom of Leu41, the backbone O atom of Pro110 and another invariant water molecule IW15 (Fig. 4). Leu41 has been suggested to contribute to the back wall of the hydrophobic channel of the enzyme (Scott & Sigler, 1994). Thus, water molecule IW9 may play a role in holding the residue Leu41 and two helices (α_3 and α_5) in place by maintaining the overall tertiary structure. The MD analysis also confirms that water molecule IW9 has 100% residence frequency during the MD

simulations. Interestingly, water molecule IW11 is highly exposed but is stable with a very low normalized *B* factor (Table 4); it stabilizes the N-terminal capping of Glu17 in the α_2 helix (Fig. 4). The residence frequency computed from the MD simulations for water molecule IW11 is 100% (Table 4). A semi-accessible (solvent accessibility 4.3 Å^2) water molecule IW12 with a very low normalized *B* factor (Table 4) is hydrogen bonded to the backbone N atom of Leu19 and a water molecule (316). In most of the structures, water molecule 316 is replaced by one of the side-chain atoms of Asn6, which connects two α-helices (α_1 and α_2) through the invariant water molecule IW12. However, the residence frequency of this water molecule is only 74%, which may be a consequence of the hydrophobic nature of residue 19; it contributes to the interfacial adsorption surface (Scott & Sigler, 1994). Water molecule IW15 is hydrogen bonded to the backbone N atom and one of the side-chain carboxylate O atoms of Asp40. Together with another invariant water molecule (IW7; discussed earlier), water molecule IW15 may play a role in anchoring Asp40 (Fig. 4). The residence frequency computed from the MD simulations for this water molecule is 100% (Table 4).

It is interesting to note that invariant water molecule IW17 is hydrogen bonded to the backbone O atom of Ser15, Asp21 O$^{\delta2}$, an MPD molecule and a water molecule (137). The triangle formed by the backbone O atom of Ser15, IW17 and water molecule 137 is an approximate equilateral triangle (Fig. 4). In fact, another invariant water molecule (IW21) is also observed to be hydrogen bonded to the backbone N and O$^{\gamma}$ atoms of residue Ser15. Thus, it is very much possible that these two invariant water molecules stabilize residue Ser15, which is on the surface of the molecule and is located in a loop connecting two helices (α_1 and α_2). Interestingly, the PLA$_2$ enzymes from almost all groups consist of five or six short loops. However, the loop consisting of residues 14–16 is relatively rigid with a proline residue (Pro14) at the beginning and a characteristic type I β-turn. Furthermore, the short loop also provides structural support to the substrate-binding channel and it has been found that the amino-acid sequences and conformation of this loop are almost identical in all PLA$_2$ enzyme variants (Jabeen *et al.*, 2005). In addition, the invariant water molecule IW17 is present during the MD simulations with 100% residence frequency; the corresponding value for IW21 is 83%. However, the reason for the low residence frequency of IW21 is not clear. The buried invariant water molecule IW19 is involved in a hydrogen-bonding network that connects the backbone O atoms of Phe106 and Val109. In most of the crystal structures a Tris molecule is observed near IW19, preventing exposure of this water molecule. However, in three structures (1une, 1gh4 and 1mku), the water IW19 is more exposed because of the absence of the Tris molecule. The water molecule IW23 seems to stabilize the C-terminal capping surface residue Ser107 of helix α_5 (Fig. 1). Both these water molecules (IW19 and IW23) have a residence frequency of greater than 95% during the MD simulations.

3.1.3. Invariant water molecules in Cluster3. Water molecule IW3 has been suggested to maintain the catalytically important residue Asp99 (Kumar *et al.*, 1994; Sekar, Yu *et al.*, 1997; Sekar & Sundaralingam, 1999); it is totally buried and very stable. The invariant water molecule IW3 forms a hydrogen-bonding network with the surrounding residues (Ala1, Tyr52, Pro68 and Asp99) and has been proposed to serve as a link between the active site and the interfacial recognition site (Verheij *et al.*, 1980; Yuan & Gelb, 1988; Sekar & Sundaralingam, 1999). It is also noteworthy that the N-terminal residue Ala1 is believed to be involved in the interfacial catalysis. Ala1 performs the activation of the phospholipid, while the other end performs the hydrolysis of monomeric phospholipids (Sekar, Yu *et al.*, 1997). This suggests that this water molecule stabilizes the N-terminal residue Ala1 as well as Asp99. Analysis of the structures of single mutants at Asp99 (Kumar *et al.*, 1994; Sekar, Yu *et al.*, 1997; Sekar *et al.*, 1999) showed the absence of this water molecule at this position and it was observed that Ala1 occupies the void created by the missing structural water (IW3). MD analysis of all the crystal structures investigated reveals the presence of this water molecule at a similar position with a residence frequency of more than 84%. These observations suggest a requirement for the structural water at this position to anchor the N-terminal residue Ala1.

A further four invariant water molecules (IW6, IW10, IW18 and IW20) form an extended water bridge near the N-terminal

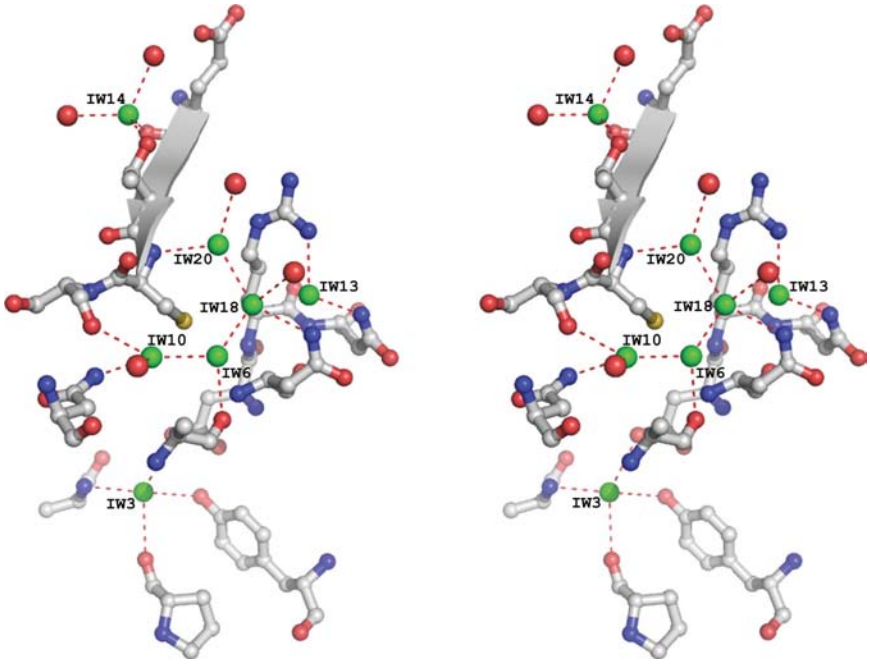

Figure 5
The seven invariant water molecules of Cluster3 are shown in green together with their hydrogen-bonding interactions. Other interacting water molecules are shown in red.

region (Fig. 5) and stabilize a stretch of residues (Cys84, Ser85, Asn88, Ala93 and Asn97). It is interesting to note that these four water molecules fill a small cavity that is observed on the surface of the enzyme. However, the importance of this cavity is not known. Moreover, IW20 seems to be involved in anchoring Cys84 (stabilizing the β-wing of the enzyme), which forms a disulfide bond between the β-sheet and one of the longest helices (α_5). In addition, IW10 (a buried water molecule) stabilizes the loop residue Ser85 and the N-terminal capping residue Asn88 of the α_5 helix. Furthermore, the second calcium-coordinating residues Asn71 and Glu92 have been suggested to be responsible for the binding of the enzyme to the membrane (Sekar, Yogavel *et al.*, 2006). It has also been observed that water molecules IW20 and IW6 are mediated through another invariant water molecule IW18, which is further hydrogen bonded to Asn97. Thus, it may be suggested that the water bridge formed by these four invariant water molecules is responsible for keeping the residues in place at the interfacial site (Fig. 5). Moreover, all four of these invariant water molecules are present during the MD simulations with more than 97% residence frequency, with the exception of IW6 (83%). The invariant water molecule IW13 is buried in a positively charged environment very close to the small cavity filled by an extended water bridge (IW6, IW10, IW18 and IW20). The invariant water molecule IW14, which is situated between two short β-strands (known as the neurotoxic region and located in the highly negatively charged cavity of the enzyme), is semi-accessible (solvent-accessible area = 3.0 Å^2) with a very low normalized *B* factor (Table 5). It may play a role in stabilizing the β_1 strand of the enzyme. The residence frequencies of these two invariant water molecules (IW13 and IW14) are greater than 90%.

3.2. Conclusion

There are 24 invariant water molecules that are conserved in all the structures of PLA_2. Of these, nine invariant water molecules (IW1, IW2, IW3, IW4, IW5, IW8, IW9, IW10 and IW19) are located in the core of the enzyme and are likely to be involved in the folding of the enzyme. The invariant water molecules IW1 and IW2 are also involved in the catalytic activity of the enzyme. In contrast, the invariant water molecules IW5 and IW8 are structurally essential, providing coordination to the functionally important active-site calcium ion. These invariant water molecules are also important to maintaining the correct active-site geometry. In addition, a few invariant water molecules are involved in mediating ion pairs that play an important role in stabilizing the tertiary structure. A set of water molecules also form a water bridge that stabilizes the functionally important residues. Approximately half of the invariant water molecules play a role in stabilizing the surface residues of the enzyme. Thus, it can be concluded that in addition to the structurally and functionally important water molecules, the present study helps to rationalize the water molecules that are significant to the folding and stability of the enzyme.

The authors thank the Bioinformatics Centre, Interactive Graphics Based Molecular Modelling Facility and Supercomputer Education and Research Centre (SERC), Indian Institute of Science, Bangalore, India. This work was supported by a research grant awarded to KS by the Department of Science and Technology (DST), India.

References

Baker, E. N. & Hubbard, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.
Balamurugan, B. *et al.* (2007). *J. Appl. Cryst.* **40**, 773–777.
Berg, B. van den, Tessari, M., Boelens, R., Dijkman, R., de Haas, G. H., Kaptein, R. & Verheij, H. J. (1995). *Nature Struct. Biol.* **2**, 402–406.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Biswal, B. K., Sukumar, N. & Vijayan, M. (2000). *Acta Cryst.* D**56**, 1110–1119.
Chaplin, M. F. (2006). *Nature Rev. Mol. Cell Biol.* **7**, 861–866.
Cheung, M. S., Garcia, A. E. & Onuchic, J. N. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 685–690.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Darden, T. D. & York, P. L. (1993). *J. Chem. Phys.* **98**, 10089–10092.
Deenen, L. L. van & de Haas, G. H. (1964). *Adv. Lipid Res.* **2**, 167–234.
Dijkstra, B. W., Kalk, K. H., Hol, W. G. & Drenth, J. (1981). *J. Mol. Biol.* **147**, 97–123.
Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E. & Kern, D. (2005). *Nature (London)*, **438**, 36–37.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Franks, F. (2002). *Biophys. Chem.* **96**, 117–127.
Halle, B. (2004). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359**, 1207–1224.
Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. (1997). *J. Comput. Chem.* **18**, 1463–1472.
Huang, B., Yu, B. Z., Rogers, J., Byeon, I. J., Sekar, K., Chen, X., Sundaralingam, M., Tsai, M. D. & Jain, M. K. (1996). *Biochemistry*, **35**, 12164–12174.
Hubbard, S. J. & Thornton, J. M. (1993). *NACCESS Computer Program.* Department of Biochemistry and Molecular Biology, University College, London.
Jabeen, T., Singh, N., Singh, R. K., Ethayathulla, A. S., Sharma, S., Srinivasan, A. & Singh, T. P. (2005). *Toxicon*, **46**, 865–875.
Jayaram, B. & Jain, T. (2004). *Annu. Rev. Biophys. Biomol. Struct.* **33**, 343–361.
Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. (1996). *J. Am. Chem. Soc.* **118**, 11225–11236.
Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. (2001). *J. Phys. Chem. B*, **105**, 6474–6487.
Kanaujia, S. P. & Sekar, K. (2008). *Acta Cryst.* D**64**, 1003–1011.
Kishan, R. V. R., Chandra, N. R., Sudarsanakumar, C., Suguna, K. & Vijayan, M. (1995). *Acta Cryst.* D**51**, 703–710.
Krem, M. M. & Enrico, D. C. (1998). *Proteins*, **30**, 34–42.
Kumar, A., Sekharudu, C., Ramakrishnan, B., Dupureur, C. M., Zhu, H., Tsai, M. D. & Sundaralingam, M. (1994). *Protein Sci.* **3**, 2082–2088.
Loris, R., Stas, P. P. G. & Wyns, L. (1994). *J. Biol. Chem.* **269**, 26722–26733.
McDonald, I. K. & Thornton, J. M. (1994). *J. Mol. Biol.* **238**, 777–793.
Meyer, E. (1992). *Protein Sci.* **1**, 1543–1562.
Otting, G. (1997). *Prog. Nucl. Magn. Res. Spectrosc.* **31**, 259–285.
Otting, G., Liepinsh, E. & Wuthrich, K. (1991). *Science*, **254**, 974–980.

Papoian, G. A., Ulander, J., Eastwood, M. P., Luthey-Schulten, Z. & Wolynes, P. G. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 3352–3357.
Prasad, B. V. L. S. & Suguna, K. (2002). *Acta Cryst.* D**58**, 250–259.
Rajakannan, V., Yogavel, M., Poi, M. J., Jeyaprakash, A. A., Jeyakanthan, J., Velmurugan, D., Tsai, M. D. & Sekar, K. (2002). *J. Mol. Biol.* **324**, 755–762.
Raschke, T. M. (2006). *Curr. Opin. Struct. Biol.* **16**, 152–159.
Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.
Scott, L. D. & Sigler, P. B. (1994). *Adv. Protein Chem.* **45**, 53–88.
Scott, L. D., White, S. P., Otwinowski, Z., Yuan, W., Gelb, H. M. & Sigler, P. B. (1990). *Science*, **250**, 1541–1546.
Sekar, K., Biswas, R., Li, Y., Tsai, M.-D. & Sundaralingam, M. (1999). *Acta Cryst.* D**55**, 443–447.
Sekar, K., Eswaramoorthy, S., Jain, M. K. & Sundaralingam, M. (1997). *Biochemistry*, **36**, 14186–14191.
Sekar, K., Gayathri, D., Velmurugan, D., Jeyakanthan, J., Yamane, T., Poi, M.-J. & Tsai, M.-D. (2006). *Acta Cryst.* D**62**, 392–397.
Sekar, K., Kumar, A., Liu, X., Tsai, M. D., Gelb, M. H. & Sundaralingam, M. (1998). *Acta Cryst.* D**54**, 334–341.
Sekar, K., Rajakannan, V., Gayathri, D., Velmurugan, D., Poi, M.-J., Dauter, M., Dauter, Z. & Tsai, M.-D. (2005). *Acta Cryst.* F**61**, 3–7.
Sekar, K., Rajakannan, V., Velmurugan, D., Yamane, T., Thirumurugan, R., Dauter, M. & Dauter, Z. (2004). *Acta Cryst.* D**60**, 1586–1590.
Sekar, K., Sekharudu, C., Tsai, M. D. & Sundaralingam, M. (1998). *Acta Cryst.* D**54**, 342–346.
Sekar, K. & Sundaralingam, M. (1999). *Acta Cryst.* D**55**, 46–50.
Sekar, K., Vaijayanthi Mala, S., Yogavel, M., Velmurugan, D., Poi, M. J., Vishwanath, B. S., Gowda, T. V., Jeyaprakash, A. A. & Tsai, M. D. (2003). *J. Mol. Biol.* **333**, 367–376.
Sekar, K., Yogavel, M., Kanaujia, S. P., Sharma, A., Velmurugan, D., Poi, M.-J., Dauter, Z. & Tsai, M.-D. (2006). *Acta Cryst.* D**62**, 717–724.
Sekar, K., Yu, B. Z., Rogers, J., Lutton, J., Liu, X., Chen, X., Tsai, M. D., Jain, M. K. & Sundaralingam, M. (1997). *Biochemistry*, **36**, 3104–3114.
Smith, D. K., Radivojac, P., Obradovic, Z., Dunker, A. K. & Zhu, G. (2003). *Protein Sci.* **12**, 1060–1072.
Smolin, N., Oleinikova, A., Brovchenko, I., Geiger, A. & Winter, R. (2005). *J. Phys. Chem. B*, **109**, 10995–11005.
Spoel, D. van der, Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. C. (2005). *J. Comput. Chem.* **26**, 1701–1718.
Sreenivasan, U. & Axelsen, P. H. (1992). *Biochemistry*, **31**, 12785–12791.
Steiner, R. A., Rozeboom, H. J., de Vries, A., Kalk, K. H., Murshudov, G. N., Wilson, K. S. & Dijkstra, B. W. (2001). *Acta Cryst.* D**57**, 516–526.
Sumathi, K., Ananthalakshmi, P., Roshan, Md M. N. A. & Sekar, K. (2006). *Nucleic Acids Res.* **34**, W128–W138.
Tame, J. R. H., Sleigh, S. H., Wilkinson, A. J. & Ladbury, J. E. (1996). *Nature Struct. Biol.* **3**, 998–1001.
Verheij, H. M., Volwerk, J. J., Jansen, E. H., Puyk, W. C., Dijkstra, B. W., Drenth, J. & de Haas, G. H. (1980). *Biochemistry*, **19**, 743–750.
Yu, B. Z., Poi, M. J., Ramagopal, U. A., Jain, R., Ramakumar, S., Berg, O. G., Tsai, M. D., Sekar, K. & Jain, M. K. (2000). *Biochemistry*, **39**, 12312–12323.
Yuan, W. & Gelb, M. H. (1988). *J. Am. Chem. Soc.* **110**, 2665–2666.
Zhang, N. Z. & Matthews, B. W. (1994). *Protein Sci.* **3**, 1031–1039.
Zhang, L., Wang, L., Kao, Y.-T., Qiu, W., Yang, Y., Okobiah, O. & Zhong, D. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 18461–18466.