Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

X-ray structure of bovine pancreatic phospholipase A_2 at atomic resolution

Roberto A. Steiner,[a] Henriëtte J. Rozeboom,[a] André de Vries,[a] Kor H. Kalk,[a] Garib N. Murshudov,[b] Keith S. Wilson[b] and Bauke W. Dijkstra[a]*

[a]Laboratory of Biophysical Chemistry, Department of Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, and [b]York Structural Biology Laboratory, University of York, Heslington, York YO1 5DD, England

Correspondence e-mail: b.w.dijkstra@chem.rug.nl

Received 20 November 2000
Accepted 5 February 2001

PDB Reference: bovine pancreatic phospholipase A_2, 1g4i.

Using synchrotron radiation and a CCD camera, X-ray data have been collected from wild-type bovine pancreatic phospholipase A_2 at 100 K to 0.97 Å resolution allowing full anisotropic refinement. The final model has a conventional R factor of 9.44% for all reflections, with a mean standard uncertainty for the positional parameters of 0.031 Å as calculated from inversion of the full positional least-squares matrix. At 0.97 Å resolution, bovine pancreatic phospholipase A_2 reveals for the first time that its rigid scaffolding does not preclude flexibility, which probably plays an important role in the catalytic process. Functionally important regions (the interfacial binding site and calcium-binding loop) are located at the molecular surface, where conformational variability is more pronounced. A cluster of 2-methyl-2,4-pentanediol molecules is present at the entrance of the hydrophobic channel that leads to the catalytic site and mimics the fatty-acid chains of a substrate analogue. Bovine pancreatic phospholipase A_2 at atomic resolution is compared with previous crystallographic structures and with models derived from nuclear magnetic resonance studies. Given the high structural similarity among extracellular phospholipases A_2 observed so far at lower resolution, the results arising from this structural analysis are expected to be of general validity for this class of enzymes.

1. Introduction

Given a certain crystal quality, the level of detail obtainable from X-ray protein crystallography was until recently essentially limited by the strength of X-ray sources and by radiation-damage effects. The advent of synchrotron beamlines for macromolecular crystallography together with the development of cryotechniques have changed the scenario. The combined use of synchrotron radiation, efficient two-dimensional detectors and usually cryogenic conditions has increased the accuracy of data, prolonged the lifetime of crystals and extended the attainable resolution limits. In a so far exponentially growing number of cases, experimentalists are even able to collect intensities at or below atomic resolution (AR). According to a definition suggested by Sheldrick (1990), AR means that data should extend to at least 1.2 Å, with 50% or more of the theoretically measurable reflections in the outer shell having $\langle I \rangle \geq 2\sigma\langle I \rangle$. Advantages of AR are various: from a biochemical perspective, atomic resolution structures often allow the unravelling of structural details that are not visible at lower resolution and which may be of key mechanistic importance (Ferraroni *et al.*, 1999; Freitag *et al.*, 1999); methodologically, atomic resolution structures offer the possibility of a more meaningful statistical analysis of the

refined model with less bias from the applied restraints (Wang *et al.*, 1997; Dauter, Wilson *et al.*, 1997; Ridder *et al.*, 1999; Sevcik *et al.*, 1996; Longhi *et al.*, 1997). Availability of a statistically significant number of protein structures refined to AR may lead to the generation of a protein-based library of stereochemical target values to be generally used in macromolecular refinement. Needless to say, low- and medium-resolution structures, which still represent the vast majority of solved structures, will benefit most from such a dictionary. Reviews on the benefits of AR have been published recently (Longhi *et al.*, 1998; Dauter, Lamzin *et al.*, 1997; Dauter *et al.*, 1995).

In this account, we present the structure of soluble bovine pancreatic phospholipase A_2 (bpPLA$_2$) at atomic resolution (0.97 Å). Soluble extracellular phospholipases A_2 are 14 kDa enzymes that hydrolyse the *sn*-2 ester bond of phospholipids in a calcium-dependent manner (Pieterson, Volwerk *et al.*, 1974; Thunnissen *et al.*, 1990; White *et al.*, 1990). They have a digestive function and are found in large amounts in mammalian pancreatic juice and in snake and bee venoms. Activity towards organized lipids (micelles and vesicles) is often orders of magnitude greater than that displayed for dispersed substrates (interfacial activation) (Pieterson, Vidal *et al.*, 1974). Catalysis involves two kinetically and structurally independent steps (Gelb *et al.*, 1995): firstly, several residues arranged in a roughly flat exposed surface known as the interfacial recognition site (IRS) are responsible for the adsorption of soluble phospholipases A_2 to the phospholipid–water interface, whilst esterolysis occurs after productive binding of the substrate at the active site, which is located at the bottom of a hydrophobic channel. Several X-ray structures of extracellular phospholipases A_2 are available at different resolutions and from many different organisms and tissues, *e.g.* bovine pancreas (Dijkstra, Kalk *et al.*, 1981; Sekar & Sundaralingam, 1999), porcine pancreas (Dijkstra *et al.*, 1983), bee venom (Scott *et al.*, 1990), snake venom (White *et al.*, 1990) and human platelet (Wery *et al.*, 1991; Scott *et al.*, 1991). They all show striking structural similarities. The availability of well ordered highly diffracting crystals of orthorhombic bpPLA$_2$ and suitable cryoconditions allowed us to collect data at 100 K to 0.97 Å using synchrotron radiation. bpPLA$_2$ is the first phospholipase A_2 to be analysed in atomic detail.

2. Methods

2.1. Crystallization

Crystals of orthorhomic bpPLA$_2$ were obtained from conditions close to those previously described by Dijkstra *et al.* (1978). Freeze-dried wild-type bpPLA$_2$ was dissolved in Tris buffer (50 m*M* Tris, 5 m*M* $CaCl_2$ pH 7.6) to a final concentration of 10 mg ml^{-1}. 5 µl of the protein solution was then frozen at 253 K in melting-point capillaries, 5 µl of MPD was layered on top and the capillaries were sealed.

Crystals belonging to space group $P2_12_12_1$ grow at 293 K in 1–3 weeks to average dimensions of 0.4 × 0.4 × 0.6 mm. Crystals are stable in the crystallization solution, which also serves as a cryoprotectant.

Table 1
Data-processing statistics.

Values in parentheses refer to the highest resolution shell.

No. of crystals	1
Mosaicity (°)	0.33
Beamline	BM14
Wavelength (Å)	0.670
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	
a	46.53
b	63.56
c	37.62
No. of molecules in the asymmetric unit	1
Matthews coefficient (Å^3 Da^{-1})	2.00
Resolution range (Å)	20.0–0.97
No. of observations	469197
No. of unique reflections	66635
Completeness (%)	99.8 (97.4)
R_{sym} (%)	6.4 (17.2)
$\langle I\rangle/\langle\sigma\rangle$	33.4 (6.89)

2.2. X-ray data collection and processing

For data collection, a single orthorhombic crystal of dimensions 0.8 × 0.8 × 0.6 mm was transferred in a fibre loop to a stream of nitrogen gas at 100 K and flash-vitrified. Data were collected using synchrotron radiation at the ESRF beamline BM14 using a MAR CCD detector. Three data sets at different resolutions (0.97 Å, 0.5° oscillation; 1.30 Å, 1°; 2.20 Å, 1°) and different exposure times (15, 4 and 3 s) were collected in order to maximize the coverage of the dynamic range of the intensities. Owing to the small size of the CCD detector, rather hard X-rays (λ = 0.670 Å) were employed in order to record data to 0.97 Å resolution. Despite the short wavelength used, the total data-collection time was only about 6 h. Integration, scaling and merging were performed with the *HKL* suite (Otwinowski & Minor, 1997).

Data-collection and processing parameters are summarized in Table 1. The percentage completeness and R_{merge} of the data as a function of resolution are shown in Fig. 1. Data extending to 0.97 Å are complete at 99.8%. The overall *B* value from a Wilson plot is 6.4 Å^2.

2.3. Refinement

2.3.1. Overview. Full anisotropic refinement of bpPLA$_2$ was efficiently performed with the program *REFMAC* (Murshudov *et al.*, 1997, 1999) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) followed by the program *SHELX*97 (Sheldrick, 1997). The initial model for the refinement was the 291 K 1.5 Å resolution structure of orthorhombic bpPLA$_2$ (Sekar & Sundaralingam, 1999; PDB code 1une). R_{free} (Brünger, 1993), as a cross-validation measure of the progress of the refinement, was calculated using 5% of the reflections randomly excluded from the full data set. The automated refinement procedure *ARP* (Lamzin & Wilson, 1993) was adopted to model and update most of the

solvent molecules. As a general strategy, resolution and weighting of the X-ray terms were increased stepwise during the restrained refinement. At the end of the refinement, the program *SHELX*97 was employed to refine occupancies for alternate conformations and for all solvent molecules. Standard uncertainties (s.u.s) were calculated through inversion of the full positional least-squares matrix after removal of all restraints. Interpretation of electron-density maps and rebuilding of the atomic models when needed were carried out using the programs *XtalView* (McRee, 1999*b*) and *QUANTA* (Molecular Simulations Inc.).

2.3.2. Details. The complete 1.5 Å resolution model (123 amino acids, one calcium ion and 134 water molecules) was used as initial model for rigid-body refinement, which led in a few cycles to an *R* factor of 24.8% and an $R_{\rm free}$ of 24.5% with data in the resolution range 20.0–2.0 Å. Restrained refinement with *REFMAC* combined with *ARP* converged in five steps during which the resolution was extended stepwise to 0.97 Å. 108 new water molecules were added to the model, giving an *R* factor and $R_{\rm free}$ of 16.8 and 19.3%, respectively, for a working set of 63 381 reflections and an $R_{\rm free}$ set of 3372 reflections. Inspection at this stage of the 0.97 Å $2mF_o - DF_c$ and $mF_o - DF_c$ maps indicated that the side chains of Lys113, Lys116 and Leu118 needed correction. Moreover, clear signs of anisotropy were present and several residues showed possible alternate side-chain positions. Nevertheless, no attempt was made to model multiple conformations prior to the anisotropic treatment of the ADPs. Inclusion of protein H atoms at their calculated position decreased the *R* factor to 15.8% and the $R_{\rm free}$ to 18.3%, but only the anisotropic description of the ADPs led to a noticeable improvement of the quality of the electron-density maps. This was also reflected in the much lower values for the *R* factor and $R_{\rm free}$ (12.8 and 15.4%, respectively). At AR, decreases in *R* factor and $R_{\rm free}$ of about 3% seem to be common upon inclusion of anisotropy (Longhi *et al.*, 1997; Sevcik *et al.*, 1996). At this point an extensive but still conservative manual rebuilding was performed. Despite a rather strict selection criterion, 18 amino-acid residues (Asn6, Lys10, Pro14, Ser16, Glu17, Pro37, Gln54, Leu58, Ser60, Lys62, Val63, Asp66, Ser74, Ser78, Ser86, Lys108, Leu118 and Lys120) already showed unambiguous alternate positions of their side chains. Three molecules of MPD were also identified and a chloride ion was introduced as a replacement for a water molecule with a $B_{\rm eq}$ value suspiciously lower than the average figure for well defined water molecules. Occupancies for alternate residues were fixed at 0.5 throughout the refinement with *REFMAC*. Refinement of this extensively rebuilt model led to an *R* factor of 11.3% and an $R_{\rm free}$ of 13.4%. In a subsequent round of model rebuilding, 12 further residues were assigned multiple conformations (Trp3, Pro18, Leu19, Leu20, Phe22, Asn23, Asn24, Tyr25, Arg43, Val65, Pro68 and His115). At this stage, conformational variability was recognized to extend beyond the side chains. Accordingly, the main chain (encompassing residues Leu18–Leu20 and Phe22–Tyr25) was given two conformations. After minimization combined with solvent updating, the *R* factor and $R_{\rm free}$ became 10.4 and 12.1%, respectively. In the subsequent round of model rebuilding, two more MPD molecules were added to the structure and the side chains of Lys116 and Gln92 were also modelled in alternate conformations. After refinement the updated model had an *R* factor of 10.0% and an $R_{\rm free}$ of 11.8%. Throughout the refinement the carbonyl function of the calcium ligand Gly32 displayed a very elongated electron density in the $2F_o - F_c$ Fourier maps and strong positive peaks ($>5\sigma$) in the $F_o - F_c$ difference Fourier maps, indicating a possible similar but alternative coordination geometry. Anisotropy, when introduced, could not account for the extra density. In the later stages of the refinement, extra peaks for a flipped carbonyl function of the adjacent residue Leu31 appeared. Accordingly, the stretch of the calcium-binding loop Gly30–Gly33 was then modelled in two discrete conformations. A final solvent updating brought the *R* factor and $R_{\rm free}$ to 9.7 and 11.6%, respectively. To refine the occupancies and to obtain the s.u.s in the coordinates the refinement was continued with *SHELX*97 using the measured intensities instead of the structure-factor amplitudes as refinement targets. Throughout the *SHELX*97 refinement, the restraint on the similarity of the ADPs of adjacent atoms was decreased from the default 0.100 to 0.025 (tighter restraint), which proved to give a more symmetrical distribution of the anisotropy and lower *R* factors (Merritt, 1999). Refinement of the occupancies of alternate protein conformations, MPD molecules and of the chloride ion in 15 cycles of conjugate-gradient least-squares (CGLS) refinement lowered the *R* factor and $R_{\rm free}$ to 9.44 and 11.39%, respectively. A subsequent round of 15 cycles of CGLS refinement including all reflections only caused a slight increase in the conventional *R* factor to 9.46%. In the following round, all the occupancies of the solvent molecules were refined using blocked-matrix least-squares refinement (overlapping blocks of about 500 parameters). None of the water molecules refined to occupancies lower than 0.2 and those that had values higher than 0.95 were

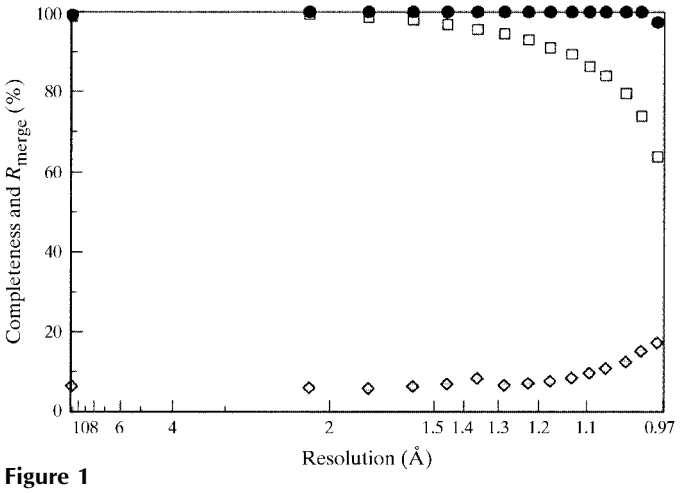

Figure 1
Percentage completeness of the data and the merging *R* for symmetry-related reflections as a function of resolution. Filled circles indicate the completeness for all reflections and open squares for reflections with $\langle I\rangle \geq 3\sigma\langle I\rangle$. Diamonds show the merging *R* factor for symmetry-related and multiply recorded reflections.

Table 2
Residues with alternate conformations.

Sc means that the entire side chain (from C^{β} onward) has been modelled in a double conformation, Residue means that the backbone is also in a double conformation. Atoms belonging to the regions Pro18–Leu20, Phe22–Tyr25 and Gly30–Gly33 were refined to a common occupancy. For the side chains of residues Lys56, Asn72, Asn79, Glu81 and Glu87 only one alternate conformation is visible, probably because of more pronounced disorder. Peptide bonds of the following residues have been modelled in alternative conformations: Pro18, Leu19, Leu20, Asp21, Phe22, Asn23, Asn24, Tyr25, Cys29, Gly30, Leu31, Gly32 and Gly33.

Residue	Occupancies	Atoms
Trp3	0.62, 0.38	Sc
Asn6	0.50, 0.50	Sc
Lys10	0.54, 0.46	Sc
Pro14	0.60, 0.40	C^{γ}
Ser16	0.58, 0.42	O^{γ}
Glu17	0.69, 0.31	Sc
Pro18	0.50, 0.50	C, O
Leu19	0.50, 0.50	Residue
Leu20	0.50, 0.50	Residue
Phe22	0.57, 0.43	N, C^{α}, C, O
Asn23	0.57, 0.43	Residue
Asn24	0.57, 0.43	Residue
Tyr25	0.57, 0.43	N, C^{α}, C, O
Gly30	0.69, 0.31	Residue
Leu31	0.69, 0.31	Residue
Gly32	0.69, 0.31	Residue
Gly33	0.69, 0.31	Residue
Pro37	0.64, 0.36	C^{γ}
Arg43	0.67, 0.33	C^{γ}, C^{δ}, N^{ε}, C^{ζ}, $N^{\eta1}$, $N^{\eta2}$
Gln54	0.54, 0.46	Sc
Lys56	0.48	C^{δ}, C^{ε}, N^{ζ}
Leu58	0.72, 0.28	C^{γ}, $C^{\delta1}$, $C^{\delta2}$
Ser60	0.65, 0.35	O^{γ}
Lys62	0.56, 0.44	C^{γ}, C^{δ}, C^{ε}, N^{ζ}
Val63	0.73, 0.27	$C^{\gamma1}$, $C^{\gamma2}$
Val65	0.71, 0.29	Sc
Asp66	0.64, 0.36	Sc and C^{α}
Pro68	0.51, 0.49	C^{γ}
Asn72	0.57	C^{γ}, $O^{\delta1}$, $N^{\delta2}$
Ser74	0.68, 0.32	O^{γ}
Ser78	0.68, 0.32	O^{γ}
Asn79	0.48	C^{γ}, $O^{\delta1}$, $N^{\delta2}$
Glu81	0.48	C^{γ}, C^{δ}, $O^{\varepsilon1}$, $N^{\varepsilon2}$
Ser86	0.56 0.44	Sc
Glu87	0.63	C^{γ}, C^{δ}, $O^{\varepsilon1}$, $N^{\varepsilon2}$
Glu92	0.61, 0.39	Sc
Lys108	0.50, 0.50	Sc
His115	0.60, 0.40	Sc
Lys116	0.58, 0.42	C^{γ}, C^{δ}, C^{ε}, N^{ζ}
Lys120	0.64, 0.36	Sc

given unitary occupancy. The R factor after this step was 9.44%. Variances and covariances of the parameters were calculated by inverting the full positional least-squares matrix after removal of all the stereochemical and shift-limiting restraints.

3. Description of the structure of bovine pancreatic phospholipase A_2 at 0.97 Å

3.1. Overall

The final model of bpPLA$_2$ at 0.97 Å resolution consists of 1409 non-H atom positions representing 123 protein residues, one Ca^{2+} and one Cl^{-} ion, five MPD molecules and 247 crystallographic water molecules (140 with full occupancy). The conventional R factor for the deposited structure is 9.44%. The last $R_{\rm free}$ value calculated before inclusion of all reflections was 11.39%. The final electron density is of excellent quality, with the vast majority of the molecule displaying well defined electron density in the $2mF_o - DF_c$ map when contoured at the 5σ level. A Ramachandran plot calculated for the final model with the program *PROCHECK* (Laskowski *et al.*, 1993) shows that 91.8% of the residues are in the most favoured region of the φ,ψ plane, with the remaining residues all falling into the additional allowed regions (if alternate backbone conformations with minor occupancy are considered, 90.9% of the structure falls in the most favoured region, 9.1% in the additional allowed regions and again none in the generously allowed or disallowed regions). The overall fold (Fig. 2) disclosed by previous crystallographic studies and the previously observed richness in structural features are confirmed at 0.97 Å. Five α-helices (Ala1–Ile13, Glu17–Phe22, Asp39–Leu58, Leu58–Leu64 and Asn89–Lys108), three 3_{10}-helices (Asn67–Asn71, Asn112–His115 and Asp119–Cys123) and one antiparallel β-sheet (Tyr75–Ser78, Glu81–Cys84) form the secondary-structure elements of bpPLA$_2$ (helical content = 52.8%, β-content = 6.5%). All 14 cysteines are engaged in disulfide bridges. In the 1.7 Å structure (Dijkstra, Kalk *et al.*, 1981) and in the recently published 1.5 Å structure by Sekar & Sundaralingam (1999) none of the amino-acid residues could be resolved in well defined alternate conformations. In the AR structure, 13 peptide bonds and 34 side chains have been modelled in two discrete conformations (Table 2). Side chains modelled in double conformation in general show all atoms clearly visible in the $2mF_o - DF_c$ map contoured at the 1σ level. Trp3 is not completely defined in its

Figure 2
Schematic drawing of the structure of bovine pancreatic phospholipase A_2 viewed in the direction of the plane of the interfacial activation site. The central hole loosely corresponds to the hydrophobic channel leading to the catalytic site. Strands are shown as arrows and helices as spirals. α-helices are labelled αA, αB, αC, αD and αE; 3_{10}-helices are labelled I, II and III; N and C represent the N-terminus and C-terminus. The seven disulfide bridges are also shown. Colouring is according to $B_{\rm eq}$: colour ramping is from orange (lower $B_{\rm eq}$) to green (higher $B_{\rm eq}$). The essential calcium ion is depicted as an ochre sphere. This figure was generated with the program *MOLSCRIPT* (Kraulis, 1991).

Table 3
Standard uncertainties for orthorhombic bpPLA$_2$.

S.u.s have been calculated with *SHELX*97 (Sheldrick, 1997) through inversion of the full unrestrained positional matrix.

Atoms	S.u. (Å)
All atoms (1409)	0.031
Protein atoms (1101)	0.026
Main-chain atoms (534)	0.019
Protein atoms in single conformation (796)	0.015
Main-chain atoms in single conformation (448)	0.013
All water sites (266)	0.050
Fully occupied waters (121)	0.034
MPD atoms (40)	0.043
Chloride ion (1)	0.0060
Calcium ion (1)	0.0023

minor conformation: density is only observed for the C^{β}, C^{γ}, C^{δ}, $N^{\varepsilon1}$ and $C^{\varepsilon2}$ atoms, which is, however, sufficient to define the orientation of the indole side chain. Lys56, Asn72, Asn79, Glu81 and Glu87 have their side chains refined to occupancies lower than 1.0 (0.48, 0.57, 0.48, 0.48 and 0.63, respectively), but no alternative conformations could be recognized for these residues, probably because of a more pronounced disorder. Lys116 has been modelled with its side chain in two conformations with occupancies of 0.58 and 0.42, but it is even possible to identify a third minor conformation (not included) for it which would correspond to that reported in the 1.5 Å model. Three major stretches of amino acids display flexibility of the main chain: Pro18–Leu20, Phe22–Gly26 and Gly30–Gly33. Five molecules of MPD could be included in the model with occupancies varying from 0.26 to 0.57. Interestingly, three of them form a cluster located at the entrance of the active-site channel in the plane of the interfacial recognition site. The plot of B_{eq} against residue number (Fig. 3) indicates that most of helix *C* and helix *E* are the best defined regions in the molecule, whilst disorder increases moving away from helix *E* towards the carboxy-terminus. Two prominent spikes are recognized in two strained regions of the molecule centred at residues 57 and 79. Leu57 constitutes the linking residue between the long helix *C* and the short helix *D*. Residues 78–80 belong to the hairpin joining the only two extended strands (Tyr75–Ser78 and Glu81–Cys84) with the hydrogen-bonding pattern of an antiparallel β-sheet.

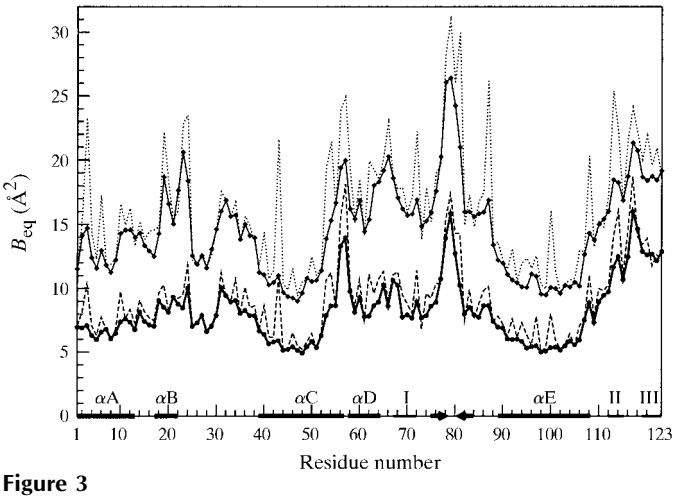

Figure 3
B_{eq} plot. Main-chain averaged (continuous line) and residue-averaged B_{eq} and *B* values for the current structure (lower) and for the room-temperature model at 1.5 Å (upper). On the *x* axis, thick lines represent the five α-helices, thin lines the three 3_{10}-helices and arrows the strands that form the antiparallel β-sheet. Residues within 3.5 Å from a symmetry-related molecule in the orthorhombic space group are Ala1, Leu2, Trp3, Lys10, Pro14, Ser15, Ser16, Glu17, Leu20, Asp21, Tyr28, Gly32, Ser34, Gly35, Asp40, Thr47, Asn50, Lys53, Glu54, Asp59, Lys62, Val65, Asn67, Tyr69, Ser74, Asn79, Ser86, Asn88, Asn89, Ala90, Asn97, Asn101, Cys105, Lys108, Val109, Lys113, Glu114, Lys116, Asp119, Lys121 and Cys123.

3.2. Assessment of standard uncertainties

At the end of an AR least-squares refinement a blocked-matrix or a full-matrix inversion is usually employed to estimate the standard uncertainties in the model. Which of the two approaches is chosen depends mainly on the number of atoms in the asymmetric unit. Blocked-matrix inversion requires less computational power but has the drawback of not properly estimating the errors owing to the omission of the sum parts of the matrix. A short discussion about the effect of the block size on the estimation of positional uncertainties can be found in McRee (1999*a*). In the present case of AR refinement the relatively small number of non-H atoms in the asymmetric unit (1409) of orthorhombic bpPLA$_2$ allowed the estimation of s.u.s in the positional parameters *via* the inversion of the full least-squares unrestrained positional matrix, keeping the anisotropic ADPs fixed. Table 3 presents the s.u.s for bpPLA$_2$ at the end of the refinement. In Fig. 4, a plot of the s.u.s for main-chain atoms and all atoms is given as a function of the residue number. The s.u. for protein atoms with a single

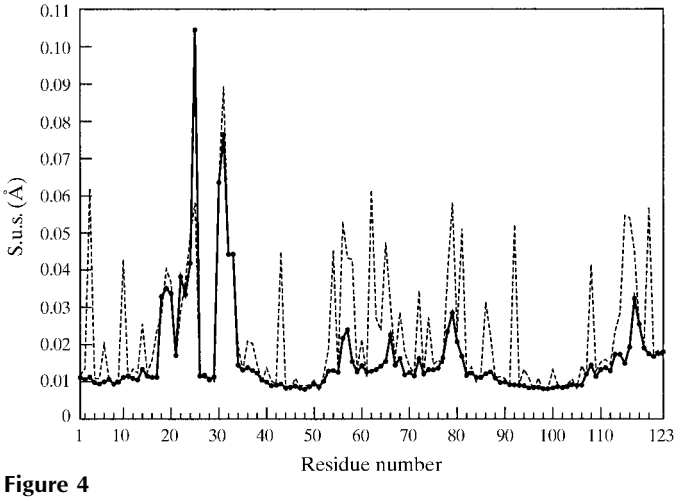

Figure 4
Local standard deviation plot. S.u.s calculated from inversion of the full unrestrained positional least-squares matrix are plotted as function of the residue number. The continuous line represents the averaged s.u.s for the main-chain atoms and the dashed line shows the residue-averaged s.u.s. Spikes in the continuous line with s.u.s above 0.03 Å describe the regions with alternate main-chain conformations.

Table 4
Interfacial recognition site (IRS).

List of residues having their side chain pointing to the substrate layer and therefore putative constituents of the IRS. For residues in single conformations, the average B_{eq} (Å^2) for the whole residue and side chain is given in parentheses. Average B_{eq} values considering only the conformation with the highest occupancy are 9.76 Å^2 (whole structure) and 11.00 Å^2 (side chains).

Residue	Location	Comment
Ala1	α-Helix *A*	Single conformation (7.36, 7.94)
Leu2	α-Helix *A*	Single conformation (8.45, 9.80)
Trp3	α-Helix *A*	Side chain in double conformation
Asn6	α-Helix *A*	Side chain in double conformation
Lys10	α-Helix *A*	Side chain in double conformation
Glu17	Initial residue of α-helix *B*	Side chain in double conformation
Leu19	α-Helix *B*	Whole residue in double conformation
Leu20	α-Helix *B*	Whole residue in double conformation
Asn23	β-Turn	Whole residue in double conformation
Asn24	β-Turn	Whole residue in double conformation
Leu31	Calcium-binding loop	Whole residue in double conformation
Lys56	Terminal part of α-helix *C*	Occupancy of side chain refined to 0.48 from C^{δ} onward
Val65	Loop	Side chain in double conformation
Asn67	Initial residue of 3_{10}-helix I	Single conformation (11.56, 12.04)
Tyr69	3_{10}-Helix I	Single conformation (9.94, 10.93)
Thr70	3_{10}-Helix I	Single conformation (9.08, 10.51)
Asn72	Loop	Occupancy of side chain refined to 0.56 from C^{γ} onward
Glu92	α-Helix *E*	Side chain in double conformation
Lys116	β-Turn	Side chain in double conformation from C^{γ} onward
Asn117	β-Turn	Single conformation (20.12, 23.40)
Asp119	3_{10}-Helix III	Single conformation (13.90, 14.39)
Lys120	3_{10}-Helix III	Side chain in double conformation
Lys121	3_{10}-Helix III	Single conformation (12.42, 11.75)
Lys123	3_{10}-Helix III	Single conformation (13.37, 13.97)

conformation is as low as 0.015 Å. The mean s.u. for the positional parameters of all atoms is 0.031 Å, which is six times lower than the value reported at 1.5 Å (0.19 Å) as derived from a Luzzati plot (Luzzati, 1952). As pointed out by Cruickshank (1999), a Luzzati plot of *R versus* $2\sin\theta/\lambda$ does not give final errors and therefore its use for that purpose should be abandoned. Several other methods for estimating the accuracy are available (Daopin, 1994).

3.3. Interfacial recognition site (IRS) and flexibility

The IRS is a flat external surface formed by a collar of hydrophobic residues arranged around the entrance to the catalytic site and by mainly cationic residues in more remote peripheral regions (Fig. 5). Using this surface, the enzyme adsorbs to micelles and other phospholipid aggregates. More than half of the residues originally proposed by Dijkstra, Drenth *et al.* (1981) to belong to the IRS (Table 4) have alternative conformations. They occur in the best defined regions of the molecule. In contrast, the residues with a single conformation are mostly in the more mobile regions as judged from their ADPs (residues 63–72 and C-terminus). Lys10 was not originally proposed to belong to the IRS, but its mutation to a glutamate resulted in a strong decrease in affinity for anionic phosphatidyl glycerol vesicles (Dua *et al.*, 1995). In our structure, we observe two conformations for Lys10, both of which have the side chain pointing in the direction of the substrate layer. Additional flexible residues and possible members of the IRS are Glu92 and Lys120. The flexible character of the IRS is not limited to its polar solvent-exposed side chains which, as observed by Smith *et al.* (1986), are the protein parts most likely to be found in double conformations, but also includes the inner ring of hydrophobic residues and stretches of protein backbone. A particularly complex region comprises residues Pro18–Tyr25. The main chain starts to diverge by a roughly 6.5° difference in the Pro18 φ angle and reconverges at the N atom of Asp21. It then deviates again from the main conformation at Phe22 ($\Delta\varphi$ = 26.5°) and returns to a single conformation at Tyr25. Asp21 is most likely to be kept in a single ordered conformation through hydrogen bonding of its side chain to the symmetry-related Asn101 and an MPD molecule. An additional stretch of main chain in double conformation is present at the calcium-binding loop Gly30-Leu31-Gly32-Gly33. Even after anisotropic refinement of the ADPs, positive peaks were still present in the Fourier synthesis, in particular near Gly32 O and Leu31 O. The latter peak indicated peptide-bond flipping. In the final model, the Gly32 C^{α} atom belonging to the chain with lower occupancy (30%) is the only atom with less than 1.0σ electron density in the $2F_o - F_c$ map, indicating that the minor conformation is not well defined (Fig. 6). Flexibility in the calcium-binding loop is not a consequence of partial occupancy for the cation, since its occupancy refines to values close to 1.0 and negative peaks are absent in the $F_o - F_c$ difference Fourier maps. From inspection of the electron density at AR we suggest that flexibility is an important factor for adsorption of bpPLA$_2$ to the phophospholipid–water interface. A recent and elegant EPR study carried out by Lin *et al.* (1998) on bee-venom PLA$_2$ shows that the protein sits on the membrane rather than digging into it and that the entrance to the active site faces the phospholipid aggregate without being firmly against it. It is therefore likely that the flexibility of the IRS residues assists in maintaining the contact between protein and phospholipid aggregate.

3.4. Calcium-binding site

Calcium is an essential cofactor for the catalytic activity of the enzyme. In Fig. 6, a representation of the calcium ion together with its seven ligands is given. The coordination of the calcium cation is octahedral, with an average distance of 2.39 Å. Both carboxylic O atoms of Asp49 contribute to the coordination sphere, sharing one of the six octahedral positions. The calcium ion has an uncertainty of 0.0021 Å in its atomic position and a B_{eq} of 6.59 Å^2. Compared with the 288 K model at 1.7 Å, the 100 K model at AR shows an average decrease of the calcium first coordination distance of about 0.073 Å (Table 5). The difference in temperature, resolution and the number of crystals used for data collection make a direct comparison not straightforward. Nevertheless,

Table 5
Distances of the calcium ion to its ligands in Å.

Ligand	1.7 Å model	1.5 Å model	0.97 Å model
28O	2.29	2.32	2.3537 (0.0055)
30O(a)	2.47	2.43	2.3159 (0.0097)
30O(b)			2.4552 (0.0244)
32O(a)	2.30	2.24	2.3198 (0.0084)
32O(b)			2.3494 (0.0161)
49O$^{\delta1}$	2.69	2.65	2.4724 (0.0046)
49O$^{\delta2}$	2.48	2.44	2.5001 (0.0058)
Wat	2.43	2.43	2.3593 (0.0071)
Wat	2.66	2.51	2.4097 (0.0060)
Average	2.47	2.43	2.3974 (0.0077)

we observe that the decrease in temperature does not cause a decrease for all calcium–ligand distances as, for example, there is an increase in the distance O28—Ca^{2+} from 2.32 Å (1.5 Å, room temperature) to 2.35 Å (0.97 Å, 100 K). We therefore conclude that the shorter average ligation distance of the main calcium coordination cage is a physical reality disclosed by the achievement of atomic resolution rather than an effect of data collection at cryotemperature. The average ligation distance of 2.397 Å is in perfect agreement with the average coordination distance extracted from a survey of the CSD for a calcium ion with a coordination number of seven and with one or more water molecules present in the coordination sphere (mean = 2.39 Å; Harding, 1999). One possible explanation for the longer average coordination distance observed at lower resolution is positional averaging of the carbonyl O atoms (O30 and O32) imposed by the impossibility of distinguishing their alternate positions. We have no explanation for the overly short distance of 2.24 Å observed for the O30—Ca^{2+} distance in the 1.5 Å resolution structure (2.30 Å at 1.7 Å and 2.32 Å here). Another detail observed at AR is a more symmetrical coordination of the carboxylic O atoms of Asp49. In the 1.5 Å resolution structure the degree of asymmetry (defined as the absolute value of the difference of the distances from each carbonyl O atoms to the metal centre) amounted to 0.21 Å, whilst in the present structure the two O atoms are located at 2.47 and 2.50 Å from the calcium ion, giving rise to an asymmetry of only 0.03 Å. *Bacillus lentus* subtilisin has a calcium coordination which is similar to that of bpPLA$_2$. The AR structure (0.78 Å; PDB entry 1gci; Kuhn *et al.*, 1998) at 100 K shows a slightly higher asymmetry (0.07 Å) for Asp41 and an identical average ligation distance of 2.39 Å.

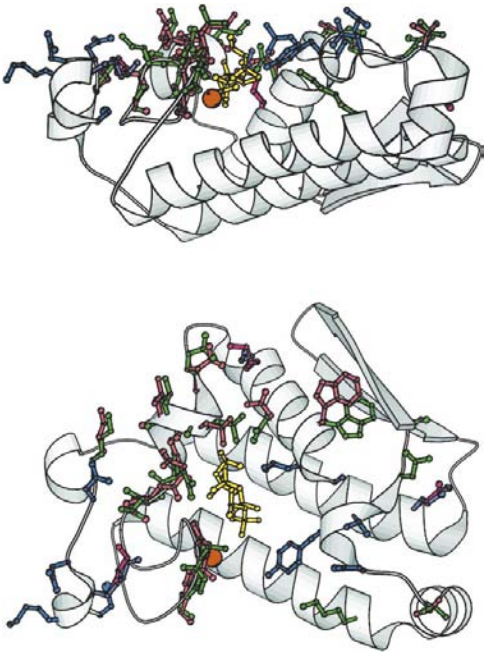

Figure 5
Side and top views of the interfacial recognition site (IRS). Residues proposed by Dijkstra, Drenth *et al.* (1981) to constitute the IRS are shown. Most of them are seen in double conformations (major in green, minor in pink) at AR. IRS residues in single conformation (brown) are located in the most flexible regions of the protein [surface 3_{10}-helix (67–72) and C-terminus]. In light and dark red newly proposed IRS residues (Lys10, Glu92 and Lys120) are shown. The essential calcium ion is shown as an ochre sphere. The MPD cluster (three MPD molecules with occupancies of 0.54, 0.51 and 0.26) seen near the calcium ion is displayed in yellow. This figure was generated with the program *MOLSCRIPT* (Kraulis, 1991).

3.5. Solvent structure

A total of 247 discrete solvent molecules (266 sites) were identified in the crystal structure and interpreted as water molecules. Occupancies for all of them have been refined in *SHELXL* using all the available reflections, resulting in 140 sites with unitary occupancy. Water molecules whose occupancies refined to values higher than 0.95 were considered to be fully occupied. 19 solvent molecules were modelled with alternate positions with the sum of their occupancies fixed to 1.0 during the refinement. Summation of occupancies over all water sites gives a stoichiometric content of 197 solvent molecules. Fig. 7(*a*) shows a histogram of the derived distribution of occupancies. The solvent structure appears to be almost evenly divided into fully occupied water molecules (56.6%), which mainly belong to the first coordination sphere

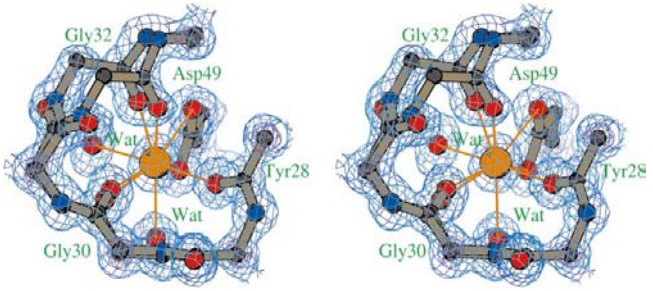

Figure 6
A stereoview of the catalytic calcium ion and its seven ligands. Ligands in double conformation are Gly30 and Gly32. An omit $2mF_o - DF_c$ electron-density map is shown at the 1.0σ level (blue) and at the 3.0σ level (purple) for the C$^{\alpha}$ atoms only. All C$^{\alpha}$ atoms of the ligands exhibit well defined spherical density when viewed at 3.0σ, except for the Gly32 C$^{\alpha}$ atom. Peptide flipping for Leu31 is also observed. This figure was generated with the program *BOBSCRIPT* (Esnouf, 1997).

of the protein, and water sites characterized by a slightly skewed distribution with the maximum located slightly below an occupancy of 0.5. The latter class represents solvent shells of higher order and also comprises the solvent engaged in hydrogen bonds with disordered protein residues. 63 more water molecules are included in the AR model compared with the 1.5 Å resolution structure. Assuming that the non-protein volume in the unit cell is entirely occupied by water, the modelled solvent accounts for 56% of the theoretically possible amount (about 350 water molecules). The AR model has therefore allowed an 18% expansion of the solvent description compared with the 1.5 Å model. To gain insight into the distribution of the defined water molecules, a histogram of the number of water molecules as a function of their distance to the closest protein N or O atom was produced (Fig. 7*b*). The first water shell is clearly observed at an average distance of 2.75 Å. A broader second hydration shell is present around 4.0 Å. Ordered fully occupied water molecules (dark bars) are located mainly in the first solvation shell, but they account only for part of it. 46 less than fully occupied water molecules and 80 ordered fully occupied water molecules have their closest contact within 3.1 Å of a fully occupied N or O atom. From these findings, it appears therefore that the solvent content in the orthorhombic form of $bpPLA_2$ possesses a consistent bulk character. In $bpPLA_2$ the first solvation shell around ordered amino-acid residues is only about 65% constituted of fully occupied ordered water molecules. This result points towards short residence times for many solvent molecules in the average position determined from X-ray diffraction.

3.6. MPD molecules

$bpPLA_2$ crystallizes from 50% MPD. Five molecules of precipitant are visible at atomic resolution, with occupancies varying from 0.26 to 0.57. Two molecules (MPD129 and MPD130) are located at the molecular surface near helix *B*, whilst the other three (MPD126, MPD127 and MPD128) form a cluster positioned at the entrance of the hydrophobic channel leading to the active site. One of the MPD molecules of the cluster (MPD126) was also observed in the 1.7 Å structure (Dijkstra, Kalk *et al.*, 1981) and in the 1.6 Å structure of bovine pro-phospholipase (Finzel *et al.*, 1991), but not in the 1.5 Å structure of the mature enzyme (Sekar & Sundaralingam, 1999) despite the virtually identical crystallization conditions (Dijkstra *et al.*, 1978). The MPD cluster lies in the flat surface of the IRS (Fig. 5). MPD strongly lowers the surface tension of water and, consistent with that, possesses a strong non-polar character (Pittz & Bello, 1971; Hammes & Schimmel, 1967) causing it to seek contact with non-polar residues in proteins (Kita *et al.*, 1994). A superposition of the trigonal form of $bpPLA_2$ complexed with the transition-state analogue L-1-*O*-octyl-2-heptylphosphonyl-*sn*-glycero-3-phosphoethanolamine (TSA; PDB code 1mkv; Sekar *et al.*, 1998) on the structure reported here shows that the arrangement of the fatty-acid chains of TSA matches very well the cluster of MPD molecules (Fig. 8). The MPD molecules have extensive interactions with the surrounding amino-acid residues mainly through hydrophobic contacts but also through hydrogen bonds to the protein and water molecules (Fig. 9). These contacts are very similar to those displayed by the fatty-acid chains of TSA in 1mkv and involve several conserved residues in the hydrophobic active-site channel: Leu2, Phe5, Asn6, Leu19, Phe22, Asn23, Cys29, Gly30 and Phe106. The finding that MPD mimics the fatty-acid part of bound phospholipids is supported by spin-labelling experiments, which show that the alkyl chains of a long-chain substrate molecule bound in the active site are in contact with the hydrophobic walls of the hydrophobic channel, as well as with solvent water and the interior of the bilayer (Lin *et al.*, 1998).

Table 6
$bpPLA_2$ helices.

The parameters listed are those calculated by the program *PROMOTIF* (Hutchinson & Thornton, 1996).

Helix	Type	Residues	Length (Å)	Residues per turn	Pitch (Å)	Deviation from ideal (°)
A	α	1–13	17.48	3.66	5.27	11.0
B	α	17–22	5.40	6.06	4.35	54.1
C	α	39–58	29.34	3.60	5.34	9.4
D	α	58–64	9.57	3.69	5.29	13.9
I	3_{10}	67–71	7.21	4.01	5.77	29.6
E	α	89–108	28.18	3.70	5.42	12.4
II	3_{10}	112–115	3.69	4.61	8.76	43.8
III	3_{10}	119–123	8.36	3.14	5.48	20.3

4. Comparison with NMR models with emphasis on debated structural regions

Recently, NMR experiments were carried out with $bpPLA_2$ at a functionally relevant pH (Yuan *et al.*, 1999). The NMR structure is generally well characterized, with the exception of the first three N-terminal residues, the calcium-binding loop (Tyr25–Thr36) and the surface loop (Val63–Asn72). Moreover, helix *B* (Glu17–Phe22) is absent, helix *D* is shorter (59–63 instead of 59–66) and the N-terminal hydrogen-bonding network is less defined. Fig. 10 reveals that the X-ray and NMR models overlap very well, with their polypeptide backbones showing similar secondary-structure elements and overall tertiary fold. The root-mean-square difference calculated between the X-ray model and the minimized average NMR model is 1.17 Å for all C^{α} atoms (123). The X-ray model agrees with the observation that the calcium-binding loop is disordered and that helix *D* is slightly shorter (Table 6). The disorder in the calcium-binding loop contributes to the flexibility of the IRS region surrounding the entrance to the active site. It also determines two almost identical calcium coordination geometries. The Val63–Asn72 region has an average B_{eq} of 10.45 $Å^2$, slightly higher than the average value for the whole molecule (9.76 $Å^2$). From the X-ray structure we additionally find that part of the region 63–72 (67–71) has the

characteristics of a 3_{10}-helix. Helix *B*, which is not seen in solution, encompasses residues Glu17–Phe22. At AR we recognize this stretch as a short α-helix of about 5.4 Å but with parameters deviating highly from ideality (Table 6). Moreover, several residues in this region display discrete alternate conformations of the main chain. NMR studies on porcine (van den Berg, Tessari, de Haas *et al.*, 1995) and bovine (Yuan *et al.*, 1999) pancreatic PLA_2 have shown that the flexibility of the N-terminus in solution is higher than that observed in all crystal structures, including the present model where the first 13 amino acids form a very well defined α-helix. Interestingly, solution studies on porcine phospholipase A_2 (van den Berg, Tessari, Boelens *et al.*, 1995) also show that a more defined N-terminal helical structure is present in PLA_2 if it is bound to lipid micelles with a competitive inhibitor present in the active site. van den Berg, Tessari, de Haas *et al.* (1995) explain this difference by an ordering effect arising from substrate binding. Crystal packing and the presence of MPD molecules in the hydrophobic channel are in our opinion responsible for bringing $bpPLA_2$ into a stable active form in the crystal. In the orthorhombic space group, the side chains of symmetry-related Asp59 and Lys121 residues block a region which in several of the NMR models is seen occupied by Leu2 and Trp3. Moreover, two very clear intermolecular hydrogen bonds are observed between the Leu2 N atom and Asp59 $O^{\delta2}$ (2.83 Å) and between the Trp3 N and Asp59 $O^{\delta1}$ atoms (2.88 Å), which orient the peptide bonds such that the helical conformation is favoured. The MPD cluster observed at the entrance of the active site makes a number of apolar van der Waals interactions possible. In particular, one of the MPD molecules (MPD126) has interactions with Leu2 of the amino terminus. The fact that the geometry and interactions of the MPD cluster are virtually identical to those displayed by the fatty-acid chains of the TSA transition-state analogue in complex with $bpPLA_2$ offers an additional structural explanation for the ordering effect observed at the amino terminus in the native structure and indicates that MPD molecules might be acting as a pseudo-effector.

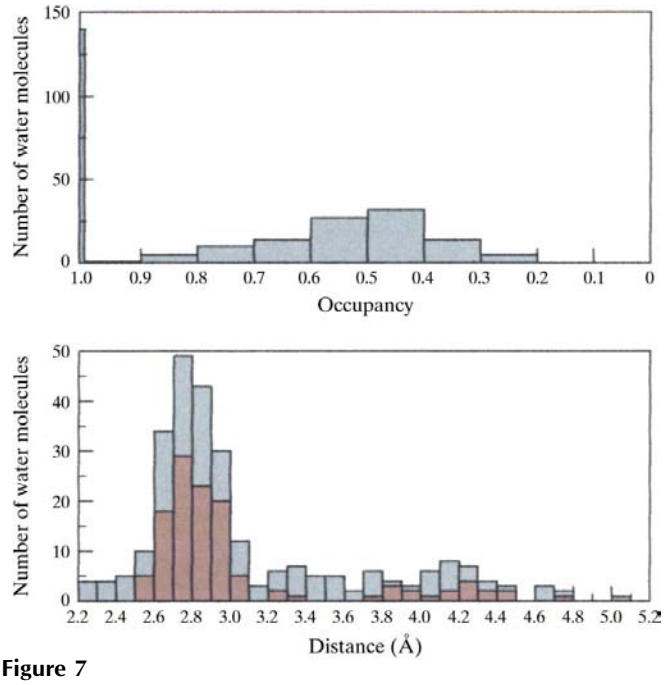

Figure 7
(*a*) Histogram of the occupancies of water molecules. (*b*) Histogram of the distribution of hydrogen-bond lengths for all water molecules (grey bars) and for ordered fully occupied waters. The distances from water molecules to the nearest N or O protein atom are plotted.

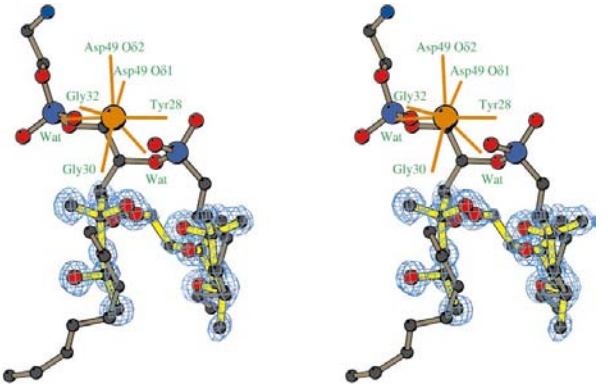

Figure 8
A stereoview of the superposition of the active-site MPD molecules found in our $bpPLA_2$ AR structure with the transition-state analogue L-1-*O*-octyl-2-heptylphosphonyl-*sn*-glycero-3-phosphoethanolamine (TSA) in complex with the trigonal form of $bpPLA_2$ (PDB entry 1mkv). MPD molecules are represented in yellow. The omit $2mF_o - DF_c$ map for the MPD molecules is shown at the 1.0σ level. The calcium ion and its ligands in the native structure are also displayed. This figure was generated with the program *BOBSCRIPT* (Esnouf, 1997).

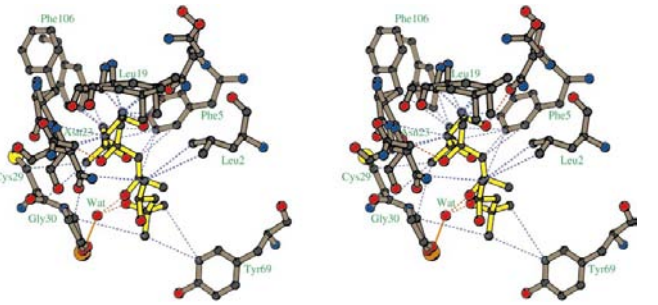

Figure 9
A stereoview representing the interactions of the active-site MPD molecules (in yellow) with the surrounding protein residues. Distances shorter than 4.0 Å are shown in violet. Hydrogen bonds are indicated in red. Two MPD residues are also in contact with one of the water molecules directly coordinating the calcium ion (ochre ball). Hydrogen bonds with other water molecules are not shown for clarity. This figure was generated with the program *MOLSCRIPT* (Kraulis, 1991).

5. Concluding remarks

Bovine pancreatic phospholipase A_2 has been anisotropically refined to 0.97 Å resolution, leading to a sixfold improvement

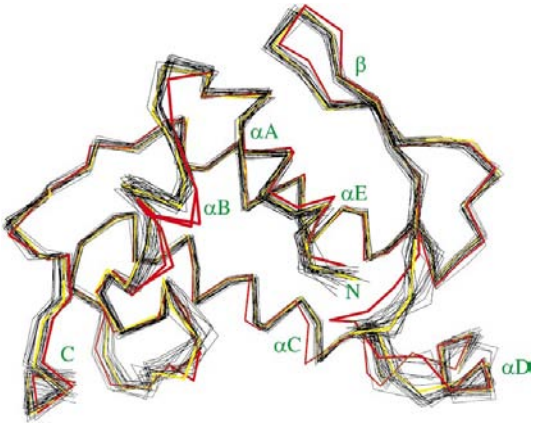

Figure 10
Picture showing the best least-squares fit (123 C^{α}) of the 20 NMR models (black) with the X-ray model (red). The fit of the best minimized average solution structure (yellow) to the crystallographic model is also shown.

in the accuracy of the model compared with the best structure to date (1.5 Å). The atomic resolution of the structure allowed for the first time a reliable modelling of the extensive flexibility of the molecules and the visualization of five molecules of the precipitant (MPD) used for crystallization. Flexibility of several side chains and main-chain stretches at the interfacial recognition site points to this being a possible important factor for enzymatic adsorption to the phospholipid–water interface. A cluster of MPD molecules located at the entrance of the hydrophobic channel leading to the active site is shown to mimic the fatty-acid chains of a substrate analogue bound in the active site. The crystallographic model of $bpPLA_2$ was compared with the ensemble of structures derived from recent NMR studies on the same system, showing a very good general agreement. From a functional point of view, probably the most important difference between the NMR and the X-ray structures is the enzyme's amino-terminus, which is observed in a helical conformation in the crystal but is flexible in solution. The conformation of the N-terminal region in the crystallographic model is in agreement with that of the enzyme in the presence of a bound substrate in the active site. The presence of the above-mentioned cluster of MPD molecules and crystal contacts are believed to be responsible. Further experiments are needed to specifically assess the geometry and role of the amino-terminus during adsorption and esterolysis.

We thank Dr Andrew Thompson for his assistance with the data collection. Dr Chunhua Yuan is gratefully acknowledged for providing us with the coordinates of the minimized average NMR structure. RAS thanks the EC (contract number CT94-0690) and the BBSRC (grant number 87/SB09829) for financial support during his stays at the York Structural Biology Laboratory. RAS is grateful to Eleanor J. Dodson for kind hospitality. These investigations were supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organization for Scientific Research (NWO). We thank the EU for supporting the work at the ESRF, Grenoble through the HCMP Access to Large Installations Programme.

References

Berg, B. van den, Tessari, M., Boelens, R., Dijkman, R., Kaptein, R., de Haas, G. H. & Verheij, H. M. (1995). *J. Biomol. NMR*, **5**, 110–121.
Berg, B. van den, Tessari, M., de Haas, G. H., Verheij, H. M., Boelens, R. & Kaptein, R. (1995). *EMBO J.* **14**, 4123–4131.
Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.
Daopin, S., Davies, D. R., Schlunegger, M. P. & Grütter, M. J. (1994). *Acta Cryst.* D**50**, 85–92.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). *Curr. Opin. Struct. Biol.* **5**, 784–790.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.
Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J. & Moulis, J. M. (1997). *Biochemistry*, **36**, 16065–16073.
Dijkstra, B. W., Drenth, J. & Kalk, K. H. (1981). *Nature (London)*, **289**, 604–606.
Dijkstra, B. W., Drenth, J., Kalk, K. H. & Vandermaelen, P. H. (1978). *J. Mol. Biol.* **124**, 53–60.
Dijkstra, B. W., Kalk, K. H., Hol, W. G. J. & Drenth, J. (1981). *J. Mol. Biol.* **147**, 97–123.
Dijkstra, B. W., Renetseder, R., Kalk, K. H., Hol, W. G. J. & Drenth, J. (1983). *J. Mol. Biol.* **168**, 163–179.
Dua, R., Wu, S. K. & Cho, W. (1995). *J. Biol. Chem.* **270**, 263–268.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 133–138.
Ferraroni, M., Rypniewski, W., Wilson, K. S., Viezzoli, M. S., Banci, L., Bertini, I. & Mangani, S. (1999). *J. Mol. Biol.* **285**, 413–426.
Finzel, B. C., Ohlendorf, D. H., Weber, P. C. & Salemme, F. R. (1991). *Acta Cryst.* B**47**, 558–559.
Freitag, S., Le Trong, I., Klumb, L. A., Stayton, P. S. & Stenkamp, R. E. (1999). *Acta Cryst.* A**55**, 1118–1126.
Gelb, M. H., Jain, M. K., Hanel, A. M. & Berg, O. G. (1995). *Annu. Rev. Biochem.* **64**, 653–688.
Hammes, G. G. & Schimmel, P. R. (1967). *J. Am. Chem. Soc.* **89**, 442–446.
Harding, M. M. (1999). *Acta Cryst.* D**55**, 1432–1443.
Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.
Kita, Y., Arawaka, T., Lin, T. Y. & Timasheff, S. N. (1994). *Biochemistry*, **33**, 15178–15189.
Kraulis, P. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kuhn, P., Knapp, S., Soltis, S. M., Ganshaw, G., Thoene, M. & Bott, R. (1998). *Biochemistry*, **37**, 13446–13452.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lin, Y., Nielsen, R., Murray, D., Hubbell, W. L., Mailer, C., Robinson, B. H. & Gelb, M. H. (1998). *Science*, **279**, 1926–1929.
Longhi, S., Czjzek, M. & Cambillau, C. (1998). *Curr. Opin. Struct. Biol.* **8**, 730–738.
Longhi, S., Czjzek, M., Lamzin, V. S., Nicolas, A. & Cambillau, C. (1997). *J. Mol. Biol.* **268**, 779–799.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
McRee, D. E. (1999*a*). *Practical Protein Crystallography*, 2nd ed. New York: Academic Press.
McRee, D. E. (1999*b*). *J. Struct. Biol.* **125**, 156–165.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.

Murshudov, G. N., Vagin, A. A. & Dodson, E. A. (1997). *Acta Cryst.* D**53**, 240–255.
Murshudov, G. N., Vagin, A. A. & Dodson, E. A. (1999). *Acta Cryst.* D**55**, 247–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pieterson, W. A., Vidal, J. C., Volwerk, J. J. & de Haas, G. H. (1974). *Biochemistry*, **13**, 1455–1460.
Pieterson, W. A., Volwerk, J. J. & de Haas, G. H. (1974). *Biochemistry*, **13**, 1439–1445.
Pittz, E. P. & Bello, J. (1971). *Arch. Biochem. Biophys.* **146**, 513–524.
Ridder, I. S., Rozeboom, H. J. & Dijkstra, B. W. (1999). *Acta Cryst.* D**55**, 1273–1290.
Scott, D. L., Otwinowski, Z., Gelb, M. H. & Sigler, P. B. (1990). *Science*, **250**, 1563–1566.
Scott, D. L., White, S. P., Browning, J. L., Rosa, J. J., Gelb, M. H. & Sigler, P. B. (1991). *Science*, **254**, 1007–1010.
Sekar, K., Kumar, A., Liu, X., Tsai, M. D., Gelb, M. H. & Sundaralingam, M. (1998). *Acta Cryst.* D**54**, 334–341.
Sekar, K. & Sundaralingam, M. (1999). *Acta Cryst.* D**55**, 46–50.
Sevcik, J., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1996). *Acta Cryst.* D**52**, 327–344.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. (1997). *SHELX*97 *Manual*. University of Göttingen, Germany.
Smith, J. L., Hendrickson, W. A., Honzatko, R. B. & Sheriff, S. (1986). *Biochemistry*, **25**, 5018–5027.
Thunnissen, M. G. M. A. B. E., Kalk, K. H., Drenth, J., Dijkstra, B. W., Kuipers, O. P., Dijkman, R., de Haas, G. H. & Verheij, H. M. (1990). *Nature (London)*, **347**, 689–691.
Wang, Z., Luecke, H., Yao, N. & Quiocho, F. A. (1997). *Nature Struct. Biol.* **4**, 519–522.
Wery, J. P., Schevitz, R. W., Clawson, D. K., Bobbitt, J. L., Dow, E. R., Gamboa, G., Goodson, T. Jr, Hermann, R., Kramer, R. M., McClure, D. B., Mihelich, E. D., Putman, J. E., Sharp, J. D., Stark, D. H., Teater, C., Warrick, M. W. & Jones, N. D. (1991). *Nature (London)*, **352**, 79–82.
White, S. P., Scott, D. L., Otwinowski, Z., Gelb, M. H. & Sigler, P. B. (1990). *Science*, **250**, 1560–1563.
Yuan, C., Byeon, I. J. L., Li, Y & Tsai, M. D. (1999). *Biochemistry*, **38**, 2909–2918.