Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction studies of phospholipase D from *Streptomyces* sp.

Ingar Leiros,[a] Edward Hough,[a]* Paola D'Arrigo,[b] Giacomo Carrea,[c] Giuseppe Pedrocchi-Fantoni,[b] Francesco Secundo[c] and Stefano Servi[b]

[a]Department of Chemistry, Faculty of Science, University of Tromsø, Norway, [b]CNR, Centro di Studio per le Sostanze Organiche Naturali, Dipartimento di Chimica Politecnico di Milano, Via Mancinelli 7, 20131 Milano, Italy, and [c]Instituto di Biocatalisi e Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano, Italy

Correspondence e-mail: edward@chem.uit.no

Received 20 September 1999
Accepted 18 January 2000

Crystals of purified phospholipase D (E.C. 3.1.4.4) from *Streptomyces* sp. strain PMF have been grown under two different crystallization conditions using vapour diffusion. Both conditions gave monoclinic crystals in space group $P2_1$. The unit-cell parameters were $a = 57.28$, $b = 57.42$, $c = 68.70$ Å, $\beta = 93.17^\circ$. The crystals diffract at 110 K to a resolution beyond 1.4 Å using synchrotron radiation.

1. Introduction

Phospholipids are amphiphilic molecules ubiquitous in nature. They are the basic components of natural membranes and cell walls, where they play a structural role and function as cofactors and activators of several membrane-associated enzymes. Hydrolysis of these compounds *in vivo* with the concomitant release of biologically active molecules, the so-called lipid second messengers diacyl glycerol (DAG), phosphatidic acid (PA) and inositol phosphate (IP), is mediated by a number of hydrolases whose effective physiological role is still the subject of intensive research and debate. One such enzyme, phospholipase D (PLD), is employed in the modification of natural phospholipids for industrial applications (D'Arrigo & Servi, 1997). PLD activity was first recognized 50 y ago in an enzyme from carrot roots and cabbage leaves (Hanahan & Chaikoff, 1947). Later, PLD was found in mammals (Kater *et al.*, 1976; Taki & Kanfer, 1981), plants (Kates & Sastry, 1969) and bacteria (Shimbo *et al.*, 1990). PLD catalyzes the hydrolysis of the phosphodiester bond of glycerophospholipids to generate phosphatidic acid and an alcohol. In addition to its hydrolytic activity, PLD also catalyzes transesterification (transphosphatidylation) when an alcohol is present as a nucleophilic donor. Using the transesterification reaction, it is possible to synthesize naturally occurring low-abundance phospholipids such as phosphatidylethanolamine, phosphatidylserine or phosphatidylglycerol from highly abundant ones such as phosphatidylcholine. These phospholipids have various applications in the pharmaceutical and food industries. It has been shown that PLDs from microbial species have a very wide transphosphatidylation activity compared with PLD from cabbage (Shimbo *et al.*, 1989, 1990).

Several recent publications (Hammond *et al.*, 1995; Ponting & Kerr, 1996; Sung *et al.*, 1997) have shown that PLD is a member of a superfamily of enzymes all sharing the ability to bind to or to perform catalytic action on a phosphodiester bond. The superfamily contains eight classes of enzymes, ranging from bacterial, plant and mammalian PLDs, phosphatidylserine synthases, cardiolipin synthases, endonucleases (nuc) and toxins, in addition to proteins having a so far unknown function: two *Poxviridae* viral proteins (p37K and K4) and a protein which is homologous to helicases (o338) and may encode an N-terminal helicase homologue with a C-terminal endonuclease domain (Ponting & Kerr, 1996). It is an interesting fact that many of the members of this PLD superfamily are proteins with two similar domains, *i.e.* they can to some extent be sequence aligned onto themselves. These enzymes are believed to function as bilobed monomers and to have evolved from one or more ancestors which at some stage in their evolution underwent a gene duplication and fusion event. It is also suggested that for each of these enzymes conserved histidine, lysine, aspartic acid and/or asparagine residues may be involved in a two-step ping-pong mechanism involving an enzyme–substrate covalently linked intermediate (Ponting & Kerr, 1996).

The crystal structure of one PLD superfamily member, nuc from *Salmonella typhimurium* (Stuckey & Dixon, 1999), is so far the only structure known from this new class of enzymes. Nuc is an endonuclease which in addition to its endonuclease activity is able to cleave artificial phosphodiester substrates. Nuc is a protein consisting of 155 residues and is the smallest member of the PLD superfamily. The other proteins in the PLD superfamily are considerably larger, with the majority having about 500 amino-acid residues. The largest ones are plant and mammalian PLDs, with about 1000 amino-acid residues, which contain several additional domains thought to regulate their activity. Nuc crystallizes as a dimer

Table 1
Statistics from data processing for PLD from *Streptomyces* sp. strain PMF.

d_{min} (Å)	R_{sym} (I)†	R_{cum} (I)	$\langle I/\sigma(I)\rangle$‡	Unique reflections	Completeness (%)	Multiplicity	Cumulative percentages (%)			
							$<1\sigma(I)$	$<2\sigma(I)$	$<3\sigma(I)$	$>3\sigma(I)$
4.33	0.046	0.047	11.9	2702	89.5	3.0	0.9	1.6	1.9	98.1
3.10	0.045	0.045	13.9	5017	97.0	3.1	0.3	0.5	0.8	99.2
2.54	0.047	0.046	12.8	6591	99.6	3.2	0.7	1.3	1.8	98.2
2.21	0.048	0.046	12.5	7766	99.8	3.1	1.1	1.8	2.6	97.4
1.97	0.057	0.048	5.8	8780	99.8	3.1	1.3	2.7	4.3	95.7
1.80	0.076	0.050	8.0	9705	100.0	3.1	3.2	5.5	7.9	92.1
1.67	0.082	0.051	8.2	10544	100.2	3.0	4.5	8.5	12.8	87.2
1.56	0.103	0.053	6.8	11250	99.6	3.0	6.4	12.1	18.0	82.0
1.48	0.133	0.055	4.2	11958	99.6	3.0	9.8	18.0	25.8	74.2
1.40	0.204	0.058	3.4	12587	99.2	3.0	13.5	24.7	35.0	65.0
19.62–1.40	0.058	0.058	7.4	86900	99.2	3.0	5.4	9.9	14.3	85.7

† $R_{sym} = \sum_h \sum_i |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_i I(h)$, where $I_i(h)$ is the ith measurement of reflection h and $\langle I(h)\rangle$ is the weighted mean of all measurements of h. ‡ Intensity signal-to-noise ratio.

(Stuckey & Dixon, 1999) and is said to be multimeric in solution (Pohlman *et al.*, 1993). In the crystal structure of nuc, a single active site is made up of identical residues from two monomers (Stuckey & Dixon, 1999) and the reaction mechanism proposed from the native and tungstate-inhibited crystal structures is likely to be valid for most or even all members of the PLD superfamily. However, important differences between the proteins within the superfamily regarding their substrate specificity and active-site surroundings will undoubtedly be found when more members of the PLD superfamily are structurally characterized. This is necessarily the case as, although the members share the ability to bind a phosphodiester moiety, the true substrates for the various enzymes are very diverse, ranging from DNA for the endonucleases to phospholipids for the lipid-modifying members of the PLD superfamily.

In addition to the crystal structure of nuc, two preliminary crystallization notes regarding PLD superfamily members have been published; these concern PLD from *Streptomyces antibioticus* (Suzuki *et al.*, 1999), a protein of 512 amino-acid residues in the mature enzyme, and a murine toxin from *Yersinia pestis* (Rudolph *et al.*, 1999) having 531 amino-acid residues. Both these proteins have sequence similarities to the PLD reported here.

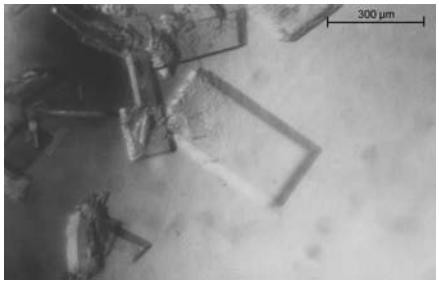

Figure 1
Crystals of PLD from *Streptomyces* sp. strain PMF. The crystals belong to space group $P2_1$ and grew in 2–5 d after microseeding.

2. Methods and results

2.1. Crystallization

Initial testing for crystallization conditions of PLD from *Streptomyces* sp. strain PMF, purified as previously described (Carrea *et al.*, 1995), was performed using the vapour-diffusion technique (McPherson, 1976) at room temperature (291 K). Conditions were surveyed using Hampton Research crystallization kits (Crystal Screen I & Crystal Screen II; Cudney *et al.*, 1994; Jancarik & Kim, 1991). The protein concentration used for initial screening was 12 mg ml^{-1} prior to mixing equal amounts of protein and reservoir solution. Droplets with an initial volume of 6 µl and reservoir volumes of 0.5 ml were used in the crystallization experiments. From the initial crystallization experiments, clusters of small needle-shaped crystals appeared after 2–3 weeks in a solution of 0.2 *M* NH_4Ac, 30% PEG 4000, 0.1 *M* sodium acetate buffer pH 4.6. This crystallization was easily reproduced but gave crystals which were unsuitable for X-ray diffraction studies. Therefore, a series of experiments around these conditions was initialized, varying the salts and buffers used and introducing microseeds into the droplets from the needle-shaped crystals obtained initially. The best crystals obtained were plates which were several hundred micrometres in two directions but were very thin. These plates were formed under two conditions, 0.2 *M* NH_4Ac, 30% PEG 4000 buffered with 0.1 *M* citrate–phosphate adjusted to pH 5.4 and 0.2 *M* $(NH_4)_2SO_4$, 30% PEG 4000 buffered with 0.1 *M* NaAc pH 5.4. They diffracted to 2 Å but were extremely fragile. Therefore, another set of microseeding experiments was initiated around these conditions with various additives. Well shaped crystals grew in trials where 20 m*M* $MnCl_2$ was added. They had a typical size of about 0.05 × 0.3 × 0.4 mm and grew in 2–5 d. Crystals from the two successful conditions were almost identical in shape and unit-cell parameters, but the first condition gave crystals diffracting to a higher resolution, probably owing to their increased thickness. These crystals were used in data collection. A representative view of some of the crystals obtained is shown in Fig. 1.

2.2. Data collection and processing

A native data set from phospholipase D was collected at the Swiss–Norwegian beamline (SNBL) at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. A crystal was flash-cooled in liquid nitrogen and data were collected at 110 K. As the PEG concentration used in the crystallization experiments was relatively high, no additional cryoprotectant was necessary for the data collection. The intensity data were collected using a 30 cm MAR image-plate system (MAR Research), with a wavelength of 0.873 Å and a crystal-

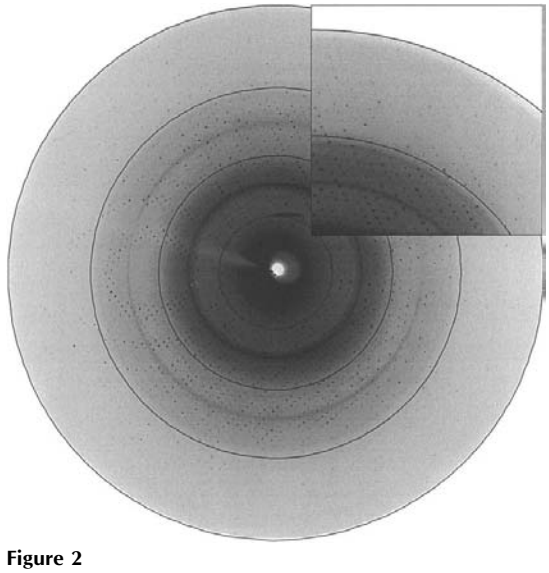

Figure 2
A 1° oscillation image of PLD from *Streptomyces* sp. strain PMF. The circles represent the resolution, with the innermost ring at 5.54 Å, the next at 2.75 Å, 1.85 Å and the outermost at 1.40 Å. In the upper right corner is a close-up of the same image, showing the diffraction to the higher resolution limit.

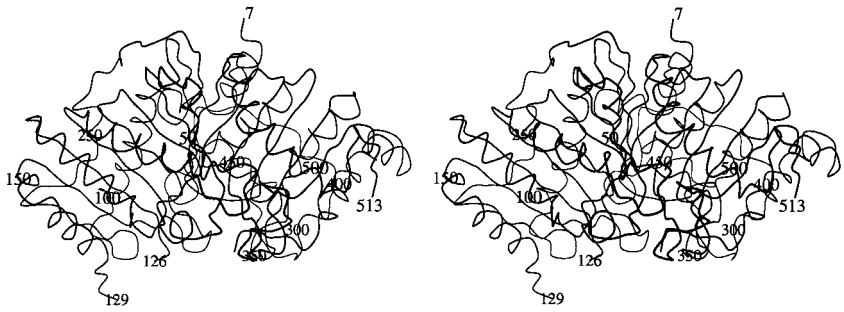

Figure 3
Stereo diagram showing the C^{α} trace of the temporary built model of PLD from *Streptomyces* sp. strain PMF.

to-detector distance of 20 cm. A total of 152 images were collected with 1° oscillations. An image is shown in Fig. 2.

The crystals diffracted to a resolution beyond 1.4 Å. The resulting unit-cell volume is 225 600 $Å^3$ and the molecular mass is 53.9 kDa (Carrea *et al.*, 1995), giving a Matthews coefficient, V_m, of 2.09 $Å^3$ Da^{-1} (Matthews, 1968), assuming one protein molecule in each asymmetric unit. This results in the reasonable water content of 41.2%. Data were indexed using the program *DENZO* (Otwinowski & Minor, 1997). The program output suggested the monoclinic space group $P2_1$, with unit-cell parameters a = 57.28, b = 57.42, c = 68.70 Å, β = 93.17°. This was confirmed by the systematic absences. Subsequent scaling and merging of intensities were carried out using the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Some statistics from the data processing are shown in Table 1.

2.3. Structure solution

As no homologous protein structure had been determined at the time of this experiment, heavy-atom methods were used to solve the structure. A number of potential heavy-atom derivative data sets were collected, but most of these suffered from non-isomorphism, with unit-cell length changes of up to 2.5% compared with the native data set. A MAD experiment was performed at HASYLAB using a K_2PtCl_4 derivative prepared by transferring crystals into a new reservoir under the same conditions as used for crystal growth but with 10 m*M* K_2PtCl_4 added. The crystals were soaked for 48 h. However, phase information from the two platinum positions located from difference Patterson plots against the native data was insufficient to solve the structure. Data from a good tungstate derivative were collected at the Swiss–Norwegian Beamline (SNBL) at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. This derivative was prepared by modification of the already existing crystal-soaking conditions. Since the presence of either PO_4^{3-} or SO_4^{2-} ions was thought to hinder possible binding of WO_4^{2-} to the protein, crystals were transferred to a solution consisting of 0.2 *M* NH_4Ac, 30% PEG 4000 buffered with 0.1 *M* NaAc pH 5.4. Na_2WO_4 was then added to this condition to a final concentration of 10 m*M*. Difference Patterson plots against the native data showed a peak of about 12σ above the mean value. As this data set was collected at a wavelength of 0.873 Å, the anomalous differences could also be included. A second data set for the same derivative was collected with a threefold longer soaking time, but no changes in the heavy-atom occupancy or additional heavy-atom binding sites were observed. Using one of these derivatives made it possible to identify several heavy-atom positions in the difference Fourier plots of the other derivatives, but inclusion of one or more of these derivatives did not improve the lack-of-closure error or $R_{\rm Cullis}$ of the first derivative; therefore, the other derivatives were considered to be of no further use in the structure-determination process.

However, the singly incorporated WO_4^{2-} ion was highly suitable for MAD techniques. Initial phases obtained from a four-wavelength MAD experiment performed using station BM14 at the ESRF (Grenoble, France) produced an interpretable electron-density map of high quality; model building of the protein is in progress. A C^{α} trace of the current model is included in Fig. 3; the refined model of PLD from *Streptomyces* sp. strain PMF will be presented elsewhere.

The Norwegian Research Council and CNR, P. F. Biotecnologie (CNR Target Project on Biotechnology) is acknowledged for financial support.

References

Carrea, G., D'Arrigo, P., Piergianni, V., Roncaglio, S., Secundo, F. & Servi, S. (1995). *Biochim. Biophys. Acta*, **1255**, 273–279.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cudney, B., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.
D'Arrigo, P. & Servi, S. (1997). *Trends Biotechnol.* **15**, 90–96.
Hammond, S. M., Altshuller, Y. M., Sung, T. C., Rudge, S. A., Rose, K., Engebrecht, J., Morris, A. J. & Frohman, M. A. (1995). *J. Biol. Chem.* **270**, 29640–29643.
Hanahan, D. J. & Chaikoff, I. L. (1947). *J. Biol. Chem.* **168**, 233–240.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kater, L. A., Goetzl, E. J. & Austen, K. F. (1976). *J. Clin. Invest.* **57**, 1173–1180.
Kates, M. & Sastry, P. S. (1969). *Methods Enzymol.* **14**, 197–203.
McPherson, A. Jr (1976). *Methods Biochem. Anal.* **23**, 249–343.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pohlman, R. F., Liu, F., Wang, L., More, M. I. & Winans, S. C. (1993). *Nucleic Acids Res.* **21**, 4867–4872.
Ponting, C. P. & Kerr, I. D. (1996). *Protein Sci.* **5**, 914–922.
Rudolph, A. E., Stuckey, J. A., Zhao, Y., Matthews, H. R., Patton, W. A., Moss, J. & Dixon, J. E. (1999). *J. Biol. Chem.* **274**, 11824–11831.
Shimbo, K., Yano, H. & Miyamoto, Y. (1989). *Agric. Biol. Chem.* **53**, 3083–3085.
Shimbo, K., Yano, H. & Miyamoto, Y. (1990). *Agric. Biol. Chem.* **54**, 1189–1193.
Stuckey, J. A. & Dixon, J. E. (1999). *Nature Struct. Biol.* **6**, 278–284.
Sung, T. C., Roper, R. L., Zhang, Y., Rudge, S. A., Temel, R., Hammond, S. M., Morris, A. J., Moss, B., Engebrecht, J. & Frohman, M. A. (1997). *EMBO J.* **16**, 4519–4530.
Suzuki, A., Kakuno, K., Iwasaki, Y., Yamane, T. & Yamane, T. (1999). *Acta Cryst.* D**55**, 317–319.
Taki, T. & Kanfer, J. N. (1981). *Methods Enzymol.* **71**, 746–750.