Atomic resolution crystal structure of squid ganglion DFPase

Juergen Koepke,[a*] Eileen I. Scharff,[b] Christian Lücke,[b] Heinz Rüterjans[b] and Günter Fritzsch[a]

[a]Max-Planck-Institute of Biophysics, Department of Molecular Membrane Biology, Heinrich-Hoffmann-Strasse 7, D-60528 Frankfurt/Main, Germany, and [b]Institute of Biophysical Chemistry, Johann Wolfgang Goethe-University, Marie-Curie-Strasse 9, D-60439 Frankfurt/Main, Germany. E-mail: koepke@mpibp-frankfurt.mpg.de

Diisopropylfluorophosphatases (DFP-ases) are capable of detoxifying chemical warfare agents like diisopropyl fluorophosphate (DFP) by hydrolysis. The protein reported here was recombinantly expressed in *E. coli*. The X-ray cystal structure of this enzyme has been refined to a resolution of 0.85 Å and a crystallographic R value of 9.4%. Reversible flash-cooling improved both, mosaicity and resolution of the crystals considerably. The overall structure of this protein represents a six-bladed β-propeller with two calcium ions bound in a central water-filled tunnel. 496 water, 2 glycerol, 2 MES-buffer molecules and 18 PEG fragments of different lengths could be refined in the solvent region. The 208 most reliable residues, without disorder or reduced occupancy in their side-chains, were finally refined without restraints. A subsequent full-matrix refinement cycle for the positional parameters yielded estimated standard deviations (esds) by matrix inversion. The herewith calculated bond lengths and bond-esds were used to obtain averaged bond lengths, which have been compared to the restraints used in preceding refinement cycles.

Keywords: phosphotriesterase; reversible flash-cooling; nitrogen calcium coordination

1. Introduction

Phosphotriesterases (EC 3.1.8) hydrolyze organo-phosphorus triesters. Due to their large biochemical and sequence diversity they are subdivided into paraoxonases (EC 3.1.8.1) and diisopropylfluoro-phosphatases (DFPases; EC 3.1.8.2), which are further subdivided into Mazur-type DFPases (40-96 kDa) and squid-type DFPases (35-40 kDa) (Hoskin *et al*., 1984). The latter enzyme is capable to cleave the P-F bond in diisopropylfluorophosphate (DFP) about 2-20 times faster than the corresponding bond of Soman, releasing non-toxic diisopropylphosphate. Because of their irreversible inhibition of the acetylcholinesterase organo-phosphorus triesters act as nerve agents (Heath, 1961). Thus, the detoxification of these compounds emphasizes the great importance of DFPases as enzymatic decontaminants, especially when compared to currently discussed and applied decontamination strategies of chemical warfare agents (Cheng *et al*., 1999). Obviously, decontamination is not the physiological function of naturally evolved DFPases, since organophosphorus nerve agents have only been designed in the last century. However, the actual role of DFPases from squid still remains unclear.

Crystals of the squid-type DFPase from *Loligo vulgaris*, heterologously expressed in *E. coli*, were grown by the hanging-drop vapor diffusion technique (Scharff *et al*., 2001a). The structure was solved at a resolution of 2.2 Å by MAD phasing and refined in a previous study (Scharff *et al*., 2001b) to 1.8 Å resolution. A low- and a high-affinity calcium ion were observed in the central tunnel of the sixfold β-propeller structure (Fig. 1). The low-affinity calcium (Ca1) was found to be part of the active site, while the high-affinity calcium (Ca2) is completely embedded inside the protein within the central tunnel and is believed to stabilize the fold of the protein. It is octahedrally coordinated by two protein-bound oxygen atoms, three water molecules, and remarkably by a nitrogen atom belonging to the side chain of a histidine. Details about the calcium coordination may be found in Scharff *et al.* (2001b). The atomic resolution structure obtained in this study helped to confirm this unusual coordination.

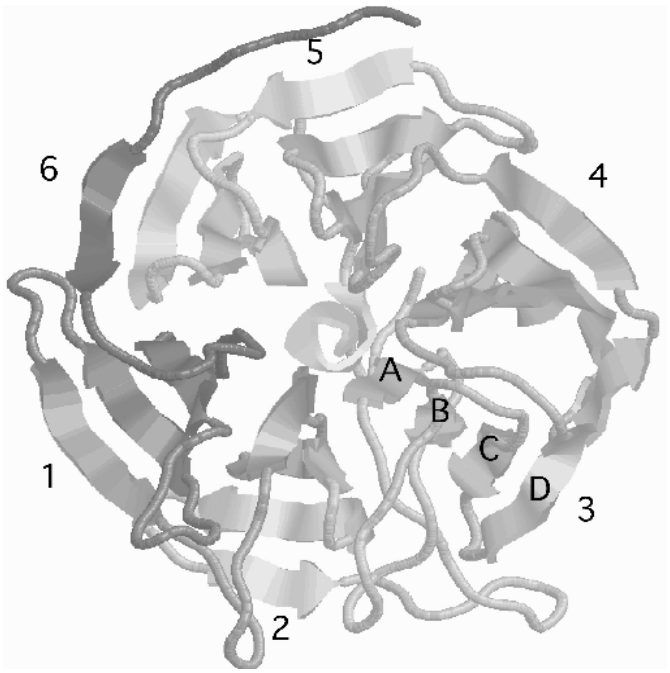

Figure 1 Overall structure of the squid-type DFPase. View of the molecule down the pseudo-sixfold axis. Each propeller blade labeled 1 to 6 consists of four antiparallel β-strands labeled from A to D.

2. Materials and methods

2.1. X-ray data collection and processing

Crystallization in the presence of PEG 6000 and 0.1 M MES-buffer at pH 6.5 was described previously (Scharff *et al*., 2001a). Prior to X-ray data collection the crystals were transferred to a cryoprotectant buffer identical to the reservoir solution, but supplemented with 20% (v/v) glycerol and annealed to 120 K in a stream of gaseous nitrogen from a cryocooling system. To improve the mosaicity of the crystals and to increase the resolution limit, the crystals were returned to ambient temperatures 3 to 5 times in a reversible flash-cooling process (Samygina *et al*., 2000). In each cycle, the crystal was transferred from the cold nitrogen-gas stream to the cryosolution and back. After each flash-cooling step a test measurement was made. Significant improvements in diffraction quality could be observed for several such steps. However, it is rather difficult to decide whether yet another step would destroy the crystal irreversibly. The limit for improvement in our case was at a maximal resolution of 0.82 Å.

Data were collected at the EMBL-Outstation Hamburg, beamline BW7B, during 5 days of beamtime using a wavelength of

0.8424 Å. Since the highest resolution reflections decayed after about 100 images, data from four different crystals were required to obtain a complete dataset. Data were processed with version 6.01 of Mosflm (Leslie, 1992) up to a resolution of 0.835 Å (Tab. 1). Structure factors were calculated from measured intensities by employing Truncate of the CCP4 package (Collaborative Computational Project Number 4, 1994).

2.2. Refinement

Starting with the model deposited in the Protein Data Bank under accession code 1E1A, the refinement was performed using Refmac5 (Murshudov *et al*., 1999). The geometry was restrained to the standard Engh & Huber (1991) values, data were used to a resolution of 0.835 Å, and anisotropic temperature factors were included. Subsequently, the model was refined against diffraction intensities rather than structure factor amplitudes using a parallelized version of Shelx-97 (Diederichs, 2000; Sheldrick & Schneider, 1997) The conjugate-gradient algorithm including anisotropic temperature factors and hydrogen atomes in the riding positions were used to a maximum resolution of 0.85 Å and a diffuse solvent correction was applied (Moewes & Kretsinger, 1975). Consequently, more details in the solvent region showed up. Water molecules were added by identification of peaks >3σ in the F_o-F_c difference density map with geometry suitable for hydrogen bonding. Between each refinement round, $2F_o$-F_c and F_o-F_c electron density maps were inspected using the graphics program XtalView (McRee, 1999). 45 residues and 17 water molecules have been refined with alternative orientations. In residues with alternative orientations or reduced occupancy no hydrogen atoms have been added to the model.

In a final round of refinement, restraints for the 208 most reliable residues were removed, resulting in a marginally improved R_{cryst} of 11.1% and indicating that no major changes occured. The subsequent blocked full-matrix least-squares cycle for the positional parameters improved R_{free} to 12.8%. The matrix inversion yielded estimated standard deviations (esds) for the refined parameters from which the esds for bond lenghts and angles could be calculated. For important bond-length restraints used in Shelx-97, weighted mean values and weighted standard deviations were calculated from the bonds in the 208 unrestrained residues.

Table 1 Summary of crystallographic data

Unit cell:	a (Å)	43.1
	b (Å)	81.8
	c (Å)	86.5
	V (Å^3)	305,128
Space group		$P2_12_12_1$
Molecules per asymmetric unit		1
Molecular weight (Da)		35,079
Amino acids		314
V_m (Å^3/Da)		2.188
Solvent content (%)		43.8
B_{Wilson} (Å^2) [a]		5.9
Highest resolution (Å)		0.835
Unique reflections		264,417
Completeness: (%)	Overall	93.8 (7.7) [e]
	Highest resolution	77.8 (1.8) [e]
Atomes refined		5410
Solvent molecules:	Water	496
	Glycerol	2
	MES-buffer	2
	PEG fragments	18
R_{sym} (%) [b]		6.5
R_{cryst} (%) [c]		11.1 (9.4) [f]
R_{free} (%) [c,d]		12.8 (11.1) [f]

[a] B_{Wilson} is an average temperature factor

[b] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i-\langle I\rangle|/\Sigma\langle I\rangle$, where I_i is the intensity of the i^{th} measurement of reflection hkl and $\langle I\rangle$ is the average intensity of a reflection.

[c] for all reflections to a resolution of 0.85 Å.

[d] R_{free} is calculated from 1% of the measured unique data that were not used during refinement.

[e] I/σ(I) in parenthesis

[f] R values for Fo>4σ(Fo) in parenthesis

3. Results and discussion

Atomic resolution is defined as the data extending to at least 1.2 Å with at least 50% intensities higher than 2σ in the outer resolution shell (Sheldrick, 1990). At resolutions below 1.2 Å there is a dramatic increase in the quality and detail of the electron density observable (Longhi *et al*., 1998). Most notably, individual heavy atoms can be differentiated and electron density for hydrogen atoms can be discerned (Kuhn *et al*., 1998). Disorder of the DFPase side-chains is restricted to the outer two ß-strands, while about 50% of the inner ß-strands show densities for most of the hydrogen atoms in F_o-F_c electron density maps of hydrogen deficient structural models. Figure 2 shows the $2F_o$-F_c electron density map contoured at 4σ for residue His274 and Ca2. Electron density for C and N atoms in the aromatic ring of the histidine can be distinguished by their volume contured. The electron density for $N_{\delta1}$ and $N_{\varepsilon2}$ is slightly larger than the density for the adjacent C atoms $C_{\delta2}$ and $C_{\varepsilon1}$, but smaller than the oxygen atoms of the two water molecules and the carbonyl oxygen, clearly visualising the unusual coordination of the even bigger Ca2 by a nitrogen atom (Nayal & Di Cera, 1994). The bond-lengths in the imidazole ring of His274 (Figure 2, right side) are another indication for the ring orientation with respect to the C_β-C_γ bond. The bond C_γ-$N_{\delta1}$ (1.395(7) Å) is the longest in the histidine ring, whereas C_γ-$C_{\delta2}$ (1.323(9) Å) is significantly shorter. These bonds correspond to the restraints C5-NH1 (1.378 Å) and C5-CR1H (1.354 Å), respectively, in a fully protonated histidine. For some of the ring hydrogens, density is visible in Figure 2, indicating the protonation state of the imidazole ring. From the difference density close to the $N_{\varepsilon2}$ a protonation of this nitrogen atom instead of $N_{\delta1}$ seems to be more likely. Thus an unprotonated $N_{\delta1}$ would coordinate the Ca2, while the two bondlengths around $C_{\varepsilon1}$, identical with regard to their length of 1.335(8) and 1.335(9) Å, respectively, would suggest a doubly protonated histidine. The difference density close to the calcium ion seems to indicate deformations to a spherical atomic model due to bond effects. For the water molecule at the bottom of Figure 2 difference density for the two protons is visible.

In Table 2 the weighted means (x_m) and the respective weighted standard deviations (σ_m), calculated from the 208 most reliable residues of the refined DFPase structure are listed for the main-chain and C_α-C_β bond-restraints. The best experimental estimate of the parent standard deviation σ is given by the sample standard deviation (Bevington, 1969). According to Engh & Huber (1991) the standard deviation of the mean provides an estimate of the accuracy of this value, while the standard deviation of a parameter in the sample provides its force constant. Therefore we have listed the standard deviation of the sample in parentheses, and we have used this standard deviation to test the significance of differences relative to the restraints. As may be seen in the fourth column of Table 2, only two bond-types (CH2G-NH1 and CH1E-NH1) show differences of about $^1/_2\sigma$. All other bond-types show only minor differences to the Engh & Huber (1991) parameters in the fifth column, with respect to the standard deviations. The averaged bond lengths published for the 1.0 Å resolution structure of Cutinase (Longhi *et al*., 1998) also agree quite well with the values found in this study.

Table 2 Comparison of mean bond-lengths and restraints

bond-type[a]	bonds	mean[b] (Å)	Δ/σ_m[c]	restraint[d] (Å)	Cutinase[e]
C-O	284	1.235(20)	0.19	1.231(20)	1.234(17)
C-CH1E	227	1.520(19)	0.24	1.525(21)	1.518(15)
CH1E-NH1	208	1.451(16)	0.46	1.458(19)	1.457(16)
C-NH1	258	1.332(16)	0.15	1.329(14)	1.332(16)
CH1E-CH2E	159	1.533(22)	0.15	1.530(20)	1.530(20)
CH1E-N	19	1.460(33)	0.19	1.466(15)	1.471(16)
C-CH2G	31	1.515(19)	0.08	1.516(18)	1.506(19)
CH2G-NH1	31	1.444(15)	0.47	1.451(16)	1.447(15)
CH1E-CH1E	49	1.539(22)	0.06	1.540(27)	1.538(13)
CH2E-CH2E	86	1.512(38)	0.21	1.520(30)	1.505(32)[f]
CH1E-CH3E	90	1.520(21)	0.04	1.521(33)	1.524(20)

[a]The bond-types are labeled according to Engh & Huber (1991).
[b]The numbers in parenthesis refer to the last digits of the mean and represent the standard deviation of the sample.
[c]Δ is the absolute difference between restraint and the mean, σ_m is the standard deviation of the sample.
[d]Engh & Huber (1991)
[e]Longhi *et al*. (1998)
[f]Calculated from the values in table 6 of Longhi *et al*. (1998).

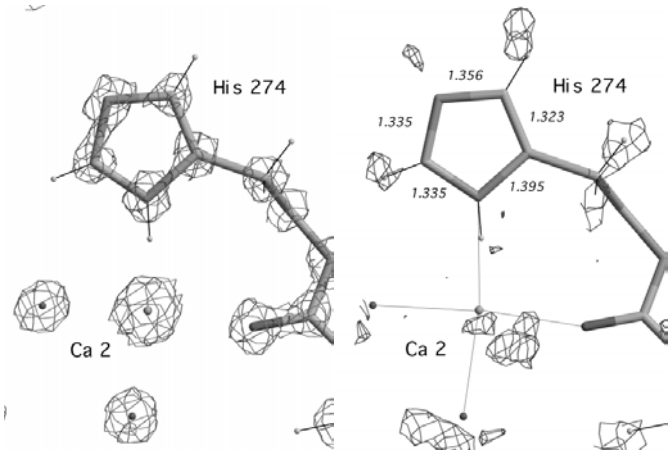

Figure 2 Electron density around His 274 and Ca2. On the left a $2F_o$-F_c map is contoured at the 4σ level. C, N, and O atoms as well as the Ca^{2+} ion can be distinguished by their size. On the right, with the same orientation, a F_o-F_c difference map of a hydrogen-deficient model is contoured at 2σ. Densities for hydrogen atoms attached to the imidazole ring of the histidine are visible. The bond lenghts in the ring are given in Å, and the coordination of Ca2 is indicated by thin grey lines. The protonation of some heavy atoms (C, N, O) is indicated by black lines connecting the riding hydrogen positions with the corresponding heavier atoms. The protonation state of histidines is preset in the used Shelx template file, why $N_{\delta 1}$ is protonated in this figure. Water molecules have no hydrogens in our calculations.

4. Conclusions

The standard deviations calculated for the averaged bond-lengths are in the range of the values derived from small molecule structures of the Cambridge Structural Database (CSD), which are commonly used as restraints. In the current study with data derived only from a single atomic resolution dataset, differences to these restraints are already detectable. In the Protein Data Bank (Bernstein *et al*., 1977) 135 protein and peptide entries are listed with a resolution better than 1.2 Å and 48 with a resolution below 1.0 Å. Hence, it will be useful to construct a protein-based library of stereochemical parameters based on a sufficiently large database of atomic resolution structures and to introduce corrections to the Engh & Huber (1991) set.

We wish to thank A. N. Popov for introducing us to the annealing method. We acknowledge help by the staff of EMBL Hamburg Outstation and the "European Community - Access to Research Infrastructure Action of the Improving Human Potential Programme to the EMBL Hamburg Outstation, contract number: HPRI-CT-1999-00017".

References

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanou-chi, T. & Tasumi, M. (1977). *J. Mol. Biol*. **112**, 535-542.
Bevington, P. R. (1969). *Data Reduction and Error Analysis for the Physical Sciences.* New York, USA: McGraw-Hill.
Cheng, T.-C., DeFrank, J. J. & Rastogi, K. V. (1999). *Chem. Biol. Interact*. **119-120**, 455-462.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst*. D**50**, 760-763.
Diederichs, K. (2000). *J. Appl. Cryst*. **33**, 1154-1161.
Engh, R. A. & Huber, R. (1991). *Acta Cryst*. A**47**, 392-400.
Heath, D.F. (1961). *Organophosphorus Poisons.*Oxford: Pergamon Press.
Hoskin, F. G. C., Kirkish, M. & Steinmann, K. (1984). *Fundam. Appl. Toxicol*. **4**, 165-172.
Kuhn, P., Knapp, M., Soltis, S. M., Ganshaw, G., Thoene, M. & Bott, R. (1998). *Biochemistry*, **37**, 13446-13452.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EAMCB Newsl. Protein Crystallogr*., **26**, 27-33.
Longhi, S., Czjzek, M. & Cambillau, C. (1998). *Curr. Opin. Struct. Biol.* **8**, 730-737.
McRee, D. E. (1999). *Practical Protein Crystallography*. San Diego, USA: Academic Press.
Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol*. **91**, 201-228.
Murshudov, G. N., Lebedev, A., Vagin, A. A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst*. D**55**, 247-255.
Nayal, M. & Di Cera, E. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 817-821.
Samygina, V. R., Antonyuk, S. V., Lamzin, V. S. & Popov, A. N. (2000). *Acta Cryst*. D**56**, 595-603.
Scharff, E. I., Lücke, C., Fritzsch, G., Koepke, J., Hartleib, J., Dierl, S. & Rüterjans, H. (2001a). *Acta Cryst*. D**57**, 148-149.
Scharff, E. I., Koepke, J., Fritzsch, G., Lücke, C. & Rüterjans, H. (2001b). *Structure*, **9**, 493-502.
Sheldrick, G. M. (1990). *Acta Cryst*. A**46**, 467-473.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol*. **277**, 319-343.