Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the prolyl-acyl carrier protein oxidase involved in the biosynthesis of the cyanotoxin anatoxin-a

Karine Moncoq,[a,b]* Leslie Regad,[b,c] Stéphane Mann,[d,e] Annick Méjean[b,d,e] and Olivier Ploux[d,e]*

[a]CNRS, UMR 7099, Institut de Biologie Physico-Chimique, 13 Rue Pierre et Marie Curie, 75005 Paris, France, [b]Université Paris Diderot-Paris 7, 75013 Paris, France, [c]INSERM, UMR-S973, MTi, 35 Rue Hélène Brion, 75205 Paris CEDEX 13, France, [d]Chimie ParisTech, ENSCP, Laboratoire Charles Friedel, 11 Rue Pierre et Marie Curie, 75231 Paris CEDEX 05, France, and [e]CNRS, UMR 7223, 75005 Paris, France

Correspondence e-mail: moncoq@ibpc.fr, olivier-ploux@chimie-paristech.fr

Received 18 April 2013
Accepted 5 August 2013

PDB Reference: AnaB, 4irn

Anatoxin-a and homoanatoxin-a are two potent cyanobacterial neurotoxins biosynthesized from L-proline by a short pathway involving polyketide synthases. Proline is first loaded onto AnaD, an acyl carrier protein, and prolyl-AnaD is then oxidized to 1-pyrroline-5-carboxyl-AnaD by a flavoprotein, AnaB. Three polyketide synthases then transform this imine into anatoxin-a or homoanatoxin-a. AnaB was crystallized in its holo form and its three-dimensional structure was determined by X-ray diffraction at 2.8 Å resolution. AnaB is a homotetramer and its fold is very similar to that of the acyl-CoA dehydrogenases (ACADs). The active-site base of AnaB, Glu244, superimposed very well with that of human isovaleryl-CoA dehydrogenase, confirming previous site-directed mutagenesis experiments and mechanistic proposals. The substrate-binding site of AnaB is small and is likely to be fitted for the pyrrolidine ring of proline. However, in contrast to ACADs, which use an electron-transport protein, AnaB uses molecular oxygen as the electron acceptor, as in acyl-CoA oxidases. Calculation of the solvent-accessible surface area around the FAD in AnaB and in several homologues showed that it is significantly larger in AnaB than in its homologues. A protonated histidine near the FAD in AnaB is likely to participate in oxygen activation. Furthermore, an array of water molecules detected in the AnaB structure suggests a possible path for molecular oxygen towards FAD. This is consistent with AnaB being an oxidase rather than a dehydrogenase. The structure of AnaB is the first to be described for a prolyl-ACP oxidase and it will contribute to defining the structural basis responsible for oxygen reactivity in flavoenzymes.

1. Introduction

Anatoxin-a and homoanatoxin-a (Fig. 1) are two potent neurotoxins produced by cyanobacteria. These alkaloids are agonists of the nicotinic acetylcholine receptor and provoke the rapid death of animals on ingestion (Wonnacott & Gallagher, 2006). Thus, the presence of these toxins in the environment represents a risk to animal and human health. We have recently identified the cluster of genes responsible for the biosynthesis of these toxins in cyanobacteria and shown that these genes could be used as genetic markers for the detection of cyanobacteria that produce these deadly neurotoxins, thus affording a simple way to monitor the presence of anatoxin-a-producing cyanobacteria in the environment (Méjean *et al.*, 2009; Cadel-Six *et al.*, 2009).

Based on sequence analysis of the *ana* cluster, we have proposed a biosynthetic route to anatoxin-a and homoanatoxin-a from proline (Fig. 1; Méjean *et al.*, 2009). In the first

steps of this pathway, proline is activated by an adenylation protein, AnaC, and is then loaded onto an acyl carrier protein (ACP), AnaD. AnaB, a flavoprotein, then oxidizes this prolyl-AnaD substrate to 1-pyrroline-5-carboxyl-AnaD (P5C-AnaD). This imine is thought to then be transformed into anatoxin-a and homoanatoxin-a by three consecutive polyketide synthases (PKSs), as shown in Fig. 1.

We have fully reconstituted the first steps of this pathway and characterized the two key enzymes AnaC and AnaB (Méjean *et al.*, 2010; Mann *et al.*, 2011). We showed that AnaC is highly specific for proline and that AnaB catalyzes an oxidation and an isomerization step. In a previous study (Mann *et al.*, 2011), we characterized the reaction mechanism of AnaB in detail by using deuterium-labelled substrates and fluoro and dehydro analogues of the substrate as well as site-directed mutagenesis. We showed that AnaB is related to acyl-CoA dehydrogenases (ACADs), although it uses oxygen as the second substrate rather than other electron acceptors such as 2,6-dichlorophenolindophenol and phenazine methosulfate (Méjean *et al.*, 2010). We also proposed a reaction mechanism that starts with a fast enolization step catalyzed by the active-site base Glu244 (Fig. 2). The conjugated imine is then formed by hydride transfer to the FAD cofactor and the active-site base catalyzes an aza-allylic isomerization. The product is then expelled from the active site and the FAD is re-oxidized by oxygen in a fast reaction.

The oxidation of proline has been well documented in bacterial or eukaryotic metabolism. In the primary metabolism of proline, this amino acid is oxidized to P5C by P5C reductase (EC 1.5.1.2), which uses NAD(P) as a co-substrate (Meng *et al.*, 2006). It is also oxidized to pyrroline-2-carboxylate (P2C) by P2C reductase (EC 1.5.1.1), which also uses NAD(P) as the co-substrate (Goto *et al.*, 2005). However, these two enzymes catalyze the reduction of the imine rather than the oxidation of proline. In other words, the equilibrium is in favour of reduction at physiological pH. In proline catabolism to glutamate, the proline dehydrogenase (EC 1.5.99.8) oxidizes free proline to P5C but uses a flavin cofactor instead of NAD(P) (Tanner, 2008). This enzyme is reoxidized by artificial or natural electron acceptors rather than by oxygen. Sequence alignments showed that AnaB is not related to any of these enzymes but rather to the ACADs (Méjean *et al.*, 2010).

Oxidation of proline linked to an ACP, that is prolyl-ACP, is also encountered in the biosynthesis of numerous bacterial secondary metabolites containing pyrrole rings (Méjean *et al.*, 2010; Walsh *et al.*, 2006). In these biosyntheses, a single enzyme is responsible for the four-electron oxidation to pyrrole-2-carboxyl-ACP, which in turn serves as a starter for PKSs or nonribosomal peptide synthases (NRPSs). Representatives of this class of enzymes have been studied *in vitro* and it has been shown that they contain an FAD cofactor and are re-oxidized by oxygen (Walsh *et al.*, 2006). Interestingly, AnaB is related to these enzymes, as revealed by sequence alignments, but it is unique because it catalyzes a two-electron oxidation rather than a four-electron oxidation reaction as in its homologues.

We have shown by sequence alignments that all of these prolyl-ACP oxidases, including AnaB, belong to the ACAD superfamily and that they share similarities to isovaleryl-CoA dehydrogenase (IVD), an enzyme involved in the degradation of leucine (Méjean *et al.*, 2010). Indeed, their active-site base, corresponding to Glu244 in AnaB, is fully conserved and its position in the sequence corresponds to that of the active-site base of human IVD and not to those of other ACADs such as the medium-chain acyl-CoA dehydrogenases (MCADs). However, they differ from the ACADs in using oxygen as the second substrate rather than an electron-transport protein. Thus, they are oxidases but are distantly related to the acyl-CoA oxidases (ACOXs) involved in peroxisomal fatty-acid degradation that use oxygen as the second substrate.

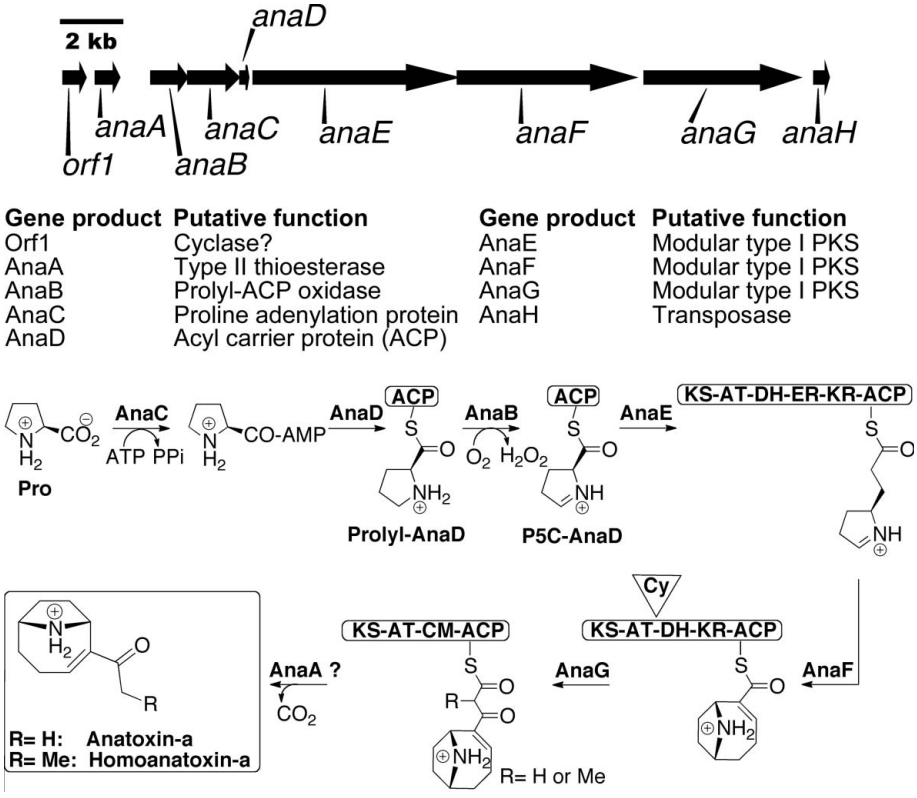

Figure 1
The *ana* gene cluster responsible for the biosynthesis of anatoxin-a and homoanatoxin-a in cyanobacteria and the postulated biosynthetic route leading to these neurotoxins. KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; ACP, acyl carrier protein; Cy, cyclase; CM, methyltransferase.

We now report the crystal structure of AnaB, the first for a prolyl-ACP oxidase, and provide a structural basis to support the reaction mechanism of this unusual enzyme. We describe several structural features that might explain the oxidase activity, rather than a dehydrogenase activity, of AnaB. Understanding the structural factors that govern the reactivity of flavoenzymes towards molecular oxygen remains a challenging issue and the AnaB structure will certainly provide further insights into this interesting question.

2. Materials and methods

2.1. General

Sequence alignments were performed using either the *Basic Local Alignment Search Tool* (*BLAST*) from the National Center for Biotechnology Information (Altschul *et al.*, 1997) or the *ClustalW* software (Thompson *et al.*, 1994). The dendograms were drawn using the sequence-alignment data and the *NJplot* software (Perrière & Gouy, 1996).

2.2. Protein purification and crystallization

AnaB was produced as a six-histidine-tagged protein as described previously (Mann *et al.*, 2011). The recombinant protein contains an extra N-terminal sequence (Supplementary Fig. S1[1]) that resulted from the cloning into the pET-28a expression vector (Novagen). However, the expressed protein lacked the first methionine residue as determined by mass spectrometry (Mann *et al.*, 2011). The protein was purified as described previously (Mann *et al.*, 2011) by metal-affinity chromatography followed by a size-exclusion chromatography step on a Superdex 200 10/30 column (GE Healthcare) using an ÄKTApurifier (GE Healthcare) in 20 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl, 10%(*v*/*v*) glycerol, 2 m*M* β-mercaptoethanol at 0.5 ml min^{-1} at 277 K. The column was calibrated by measuring the elution volumes of proteins of known molecular weight (ferritin, 440 kDa; aldolase, 158 kDa; conalbumin, 75 kDa; ovalbumin, 43 kDa; carbonic anhydrase, 29 kDa; ribonuclease, 13.7 kDa). We determined the void and total volumes by using blue dextran 2000 and acetone, respectively. The constants K_{av} were calculated and plotted against the logarithm of the molecular weight or against the Stokes radius (Supplementary Fig. S2). Prior to crystallization, AnaB was concentrated on an ultrafiltration device (Vivaspin 10 kDa MWCO, GE Healthcare) to 16 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl, 10%(*v*/*v*) glycerol, 2 m*M* β-mercaptoethanol, 1 m*M* FAD. Crystallization trials were performed by the sitting-drop vapour-diffusion method at room temperature using a sparse-matrix crystallization kit (Crystal Screen, Hampton Research) for initial screening. The best crystals were obtained by mixing 1 µl protein solution at 10 mg ml^{-1} with 0.5 µl reservoir solution consisting of 24% polyethylene glycol (PEG) 4000, 100 m*M* Tris–HCl pH 8.5, 100 m*M* $MgCl_2$. Yellow crystals grew after 2–4 d and were flash-cooled in mother liquor supplemented with 20%(*v*/*v*) glycerol.

2.3. Structure determination and refinement

Data collection was performed on beamline ID23-2 at the European Synchrotron Radiation Facility, Grenoble, France. Data were processed with *XDS* (Kabsch, 1993) and were scaled and reduced with *POINTLESS* (Evans, 2006) and programs from the *CCP*4 suite (Winn *et al.*, 2011). The crystals belonged to the monoclinic space group $P2_1$ and contained two tetramers per asymmetric unit. Data-collection and refinement statistics are given in Table 1. The structure of AnaB was solved by the molecular-replacement method with *Phaser* (McCoy *et al.*, 2007). The search-model ensemble was the tetramer structure of IVD and of short-chain acyl-CoA dehydrogenase (PDB entries 1ivh and 1jqi, respectively; Tiffany *et al.*, 1997; Battaile *et al.*, 2002) with unconserved residues pruned back to the β-atom using *CHAINSAW* from the *CCP*4 suite (Winn *et al.*, 2011). An initial round of rigid-body refinement followed by several cycles of refinement using *BUSTER* (Bricogne *et al.*, 2011) resulted in an R_{work} of 31.9% and an R_{free} of 34.3%. After automated building of the initial model using *Buccaneer* (Cowtan, 2006), further iterative cycles of model building and refinement were carried out using *Coot* (Emsley *et al.*, 2010) and *BUSTER* (Bricogne *et al.*, 2011) with LSSR NCS restraints (Smart *et al.*, 2012). During the later stages of refinement, the FAD molecules were modelled and water molecules were added automatically, followed by visual inspection. All residues from Ala4 to Leu381 of the monomer of wild-type AnaB were well defined in the electron-density map. However, the additional N-terminal tag (Supplementary Fig. S1) as well as Asp2 and Phe3 were not defined in the

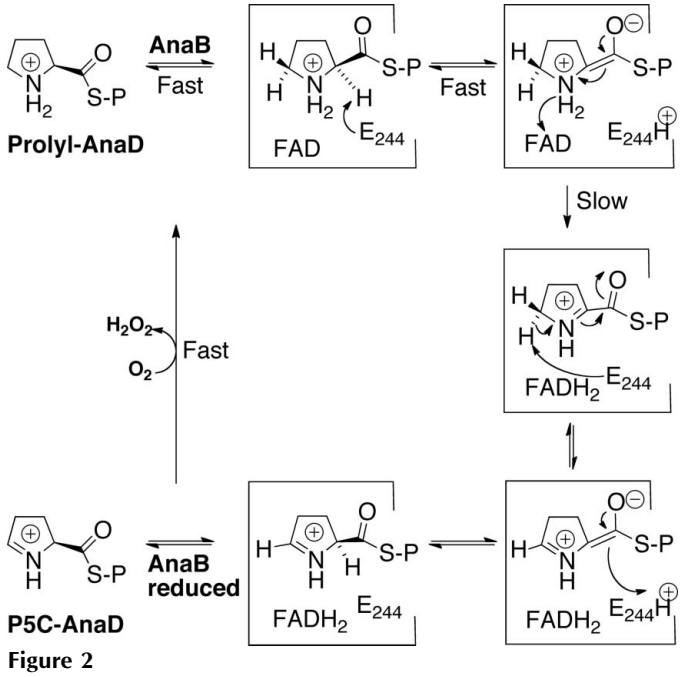

Figure 2
Proposed reaction mechanism for the oxidation of prolyl-AnaD catalyzed by AnaB (P represents the holo-AnaD protein).

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MH5096). Services for accessing this material are described at the back of the journal.

electron-density map. The final statistics of the AnaB structure refined to 2.8 Å resolution (R_{work} = 18.8%, R_{free} = 21.9%) are listed in Table 1. The coordinates and structure factors have been deposited in the Protein Data Bank (PDB) as entry 4irn.

2.4. Structure analysis

Structures were validated using the *MolProbity* web server (http://molprobity.biochem.duke.edu/; Chen *et al.*, 2010). Superposition of structures and root-mean-square deviation (r.m.s.d.) calculations were carried out with the *SSM* option in *Coot* (Emsley *et al.*, 2010). Figures were drawn using *PyMOL* (DeLano, 2002). The Stokes radius of AnaB was calculated using the *HYDROPRO* program (Ortega *et al.*, 2011). Interface areas were analyzed using the *PISA* software available at the European Bioinformatics Institute website (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html; Krissinel & Henrick, 2007). The water channel was visualized using the *HOLLOW* program (Ho & Gruswitz, 2008).

2.5. Structure calculations

All calculations were performed on the uncomplexed and homotetrameric form of the enzyme, which is the physiologically relevant oligomer. When the structure deposited corresponded to a complexed form, the ligand was removed prior to calculation (PDB entries 1ivh, 1buc and 1rx0) and when necessary (the presence of the dimer or the monomer in the deposited structure) the crystallographic symmetries were applied to the coordinates to rebuild the tetramer (PDB entries 1jqi, 1buc, 3mdd and 1siq). We determined the FAD-binding pocket in the enzymes as protein atoms situated 4 Å from any FAD atom. We estimated the substrate-binding pocket of the AnaB and IVD proteins using the *fpocket* software (Le Guilloux *et al.*, 2009), which allows pocket detection in uncomplexed protein structures. *fpocket*, which is based on protein decomposition into Voronoi polyhedra, allows surface inspection of the protein with spheres of varying diameters and estimates all pockets within the protein. Among the 33 pockets estimated by *fpocket* in IVD, we only retained the CoA persulfide binding pocket. In the case of AnaB, we first superimposed the AnaB and IVD structures using *PyMOL* (DeLano, 2002) and then selected the AnaB pocket overlapping the substrate-binding pocket of IVD. We calculated the accessible surface area (ASA) of the FAD molecule using *NACCESS* (Hubbard *et al.*, 1991). This software uses a sphere of 1.4 Å radius (the radius of a water molecule) that is rolled around the surface of the FAD and the path traced out by its centre is defined as the accessible surface (Lee & Richards, 1971).

3. Results and discussion

3.1. Crystal properties and oligomerization state

The recombinant AnaB containing a six-histidine tag at the N-terminal end (Supplementary Fig. S1) crystallized in the monoclinic space group $P2_1$ with two tetramers per asymmetric unit. The homotetramer is a dimer of dimers. In solution, as analyzed by gel filtration, AnaB behaved as a protein of 151 kDa (apparent Stokes radius R_S of 4.67 nm) and eluted as an apparent single and symmetrical peak (Supplementary Fig. S2). This experimental apparent molecular weight is actually lower than that of the homotetramer (4 × 46.4 = 185.6 kDa). The fact that the elution peak was symmetrical indicated that this situation was probably not the result of a rapid equilibrium between the dimers and tetramers; it is more likely to indicate that AnaB is a compact tetramer. We thus calculated the radius for the AnaB tetramer using the atomic coordinates and the *HYDROPRO* program (Ortega *et al.*, 2011) and found a value of R_S = 4.63 nm, in accordance with the gel-filtration data. Thus, AnaB is homotetrameric with a subunit of about 45 kDa, similar to the ACADs, but unlike the ACOXs, which are usually dimeric.

Table 1
Data-collection and refinement statistics for AnaB (PDB entry 4irn).

Values in parentheses are for the highest resolution shell.

Data collection	
Wavelength (Å)	0.87260
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 64.44, b = 191.18, c = 132.49, $\alpha = \gamma$ = 90, β = 98.94
Resolution (Å)	48.06–2.80 (2.95–2.80)
No. of unique reflections	76832 (11167)
R_{merge} (%)	0.118 (0.370)
$\langle I\rangle/\langle\sigma(I)\rangle$	6.3 (2.1)
Completeness (%)	99.0 (98.9)
Multiplicity	3.2 (3.1)
Refinement	
Resolution (Å)	30.00–2.80 (2.87–2.80)
No. of reflections	76749
R_{work} (%)	0.188 (0.213)
R_{free} (%)	0.219 (0.259)
No. of atoms	
Protein	23904
Ligand	424
Water	452
Average B (Å^2)	
Protein	29.79
Ligand	24.58
Water	20.01
R.m.s.d. from ideal	
Bond distances (Å)	0.008
Bond angles (°)	1.002
Ramachandran plot statistics† (%)	
Favoured regions	95.6
Allowed regions	4.2
Outliers	0.2

† A Ramachandran plot is supplied as Supplementary Fig. S10.

3.2. Homology to the ACAD and ACOX families of enzymes

ACADs are enzymes involved in the β-oxidation of fatty-acyl-CoA in mitochondria, while ACOXs catalyze the same reaction in peroxisomes (Ghisla & Thorpe, 2004; Kim & Miura, 2004). However, the ACOXs deliver the two electrons resulting from the oxidation to molecular oxygen, while dehydrogenases deliver them to an electron-transfer flavoprotein (ETF), which is coupled to the electron-transport chain of the mitochondrial membrane.

The ACAD family has been extensively studied because mutations in the genes coding for these enzymes lead to severe

Table 2
Sequence and structure comparison of AnaB and homologues taken from the PDB.

PDB code	Enzyme	Enzyme class	Source	Identity (%)	Similarity (%)	Monomer R.m.s.d. (Å)	Monomer No. of C^{α} atoms	Tetramer† R.m.s.d. (Å)	Tetramer† No. of C^{α} atoms	Reference
4irn	AnaB	Prolyl-ACP oxidase	*Oscillatoria*	100	100	—	378		1512	This study
1ivh	Isovaleryl-CoA dehydrogenase	IVD	*Homo sapiens*	35	53	1.53	368	1.77	1444	Tiffany *et al.* (1997)
1jqi	Short-chain acyl-CoA dehydrogenase	SCAD	*Rattus norvegicus*	35	55	1.29	351	1.40	1421	Battaile *et al.* (2002)
1buc	Butyryl-CoA dehydrogenase	BCAD	*Megasphaera elsdenii*	34	54	1.61	374	1.74	1485	Djordjevic *et al.* (1995)
3b96	Very long chain acyl-CoA dehydrogenase	VLCAD	*Homo sapiens*	33	52	1.60	351	ND‡		McAndrew *et al.* (2008)
2ix6	Short-chain acyl-CoA oxidase	ACOX4	*Arabidopsis thaliana*	31	48	1.87	352	2.60	1405	Mackenzie *et al.* (2006)
3mpi	Glutaryl-CoA dehydrogenase	GCD	*Desulfococcus multivorans*	31	51	1.57	370	1.65	1479	Wischgoll *et al.* (2010)
1rx0	Isobutyryl-CoA dehydrogenase	IBD	*Homo sapiens*	30	52	1.50	353	1.94	1413	Battaile *et al.* (2004)
3mdd	Medium-chain acyl-CoA dehydrogenase	MCAD	*Sus scrofa*	30	53	1.60	371	1.62	1483	Kim *et al.* (1993)
3d6b	Glutaryl-CoA dehydrogenase (apoenzyme)	GCD	*Burkholderia pseudomallei*	27	45	1.88	350	2.28	1367	Begley *et al.* (2011)
1siq	Glutaryl-CoA dehydrogenase	GCD	*Homo sapiens*	25	44	1.62	360	2.05	1423	Fu *et al.* (2004)
1r2j	Acyl-ACP dehydrogenase FkbI	ACAD	*Streptomyces hygroscopicus*	24	43	1.91	342	ND		Watanabe *et al.* (2003)
3djl	AidB	ACAD	*Escherichia coli*	24	42	2.17	329	ND		Bowles *et al.* (2008)
3d9g	Nitroalkane oxidase	NAOX	*Fusarium oxysporum*	23	45	1.90	354	ND		Héroux *et al.* (2009)
3mkh	Nitroalkane oxidase	NAOX	*Podospora anserina*	23	44	2.16	355	ND		Tormos *et al.* (2010)
1w07	Acyl-CoA oxidase 1	ACOX1	*Arabidopsis thaliana*	21	40	2.39	299	ND		Pedersen & Henriksen (2005)
2fon	Acyl-CoA oxidase 1	ACOX1	*Solanum lycopersicum*	21	41	2.35	319	ND		Powers *et al.* (2006)

† We only calculated the r.m.s.d. for homotetrameric enzymes and for the closest AnaB homologues. ‡ Not determined.

pathologies in humans: the so-called ACAD deficiencies. The ACAD family comprises several classes of enzymes depending on their substrate specificity: short-chain, medium-chain, long-chain and very long chain acyl-CoA dehydrogenases (SCAD, MCAD, LCAD and VLCAD, respectively) are involved in the degradation of straight-chain fatty-acyl-CoA, while isovaleryl-CoA, 2-methyl branched-chain acyl-CoA, isobutyryl-CoA and glutaryl-CoA dehydrogenases (IVD, SBCAD, IBD and GCD, respectively) are involved in amino-acid catabolism. The three-dimensional structures of several members of the ACAD and ACOX families have been described (Kim *et al.*, 1993; Djordjevic *et al.*, 1995; Tiffany *et al.*, 1997; Battaile *et al.*, 2002, 2004; Watanabe *et al.*, 2003; Fu *et al.*, 2004; Pedersen & Henriksen, 2005; Mackenzie *et al.*, 2006; Powers *et al.*, 2006; Bowles *et al.*, 2008; McAndrew *et al.*, 2008; Héroux *et al.*, 2009; Wischgoll *et al.*, 2010; Tormos *et al.*, 2010; Begley *et al.*, 2011) and Table 2 summarizes the results of a *BLAST* search performed using the AnaB sequence against sequences in the PDB. Supplementary Fig. S3 shows the corresponding multiple sequence alignment with its associated dendogram. Clearly, the closest AnaB homologues were ACADs (in particular SCAD and IVD) and ACOX4 from *Arabidopsis thaliana*, while other ACOXs and other related enzymes were more distant. Whereas AnaB is related to VLCAD, the latter enzyme is homodimeric and has an extra C-terminal domain of about 180 residues that is not present in other ACADs. ACOX4 from *A. thaliana* is also related to ACADs and to AnaB, although this plant enzyme has an extra N-terminal sequence that is likely to be involved in preventing electron transfer to ETF and thus facilitating electron transfer to oxygen (Mackenzie *et al.*, 2006). Interestingly, the acyl-ACP dehydrogenase from *Streptomyces hygroscopicus*, FkbI (Watanabe *et al.*, 2003), which is involved in the biosynthesis of ascomycin (FK520), an immunosuppressant, was not very similar to AnaB, although it uses an acyl-ACP substrate rather than an acyl-CoA substrate, like AnaB. Based on these sequence comparisons, we used the three-dimensional structure of IVD (Tiffany *et al.*, 1997) and that of SCAD (Battaile *et al.*, 2002) as models for phasing by molecular replacement.

3.3. Overall fold of AnaB and its similarities to the ACAD-family fold

The overall fold of the AnaB monomer is very similar to the ACAD-family fold, with three domains: an N-terminal domain consisting of six α-helices labelled A–F (residues 4–118), a core domain consisting of seven β-strands numbered 1–7 (residues 119–231) and a C-terminal domain consisting of five α-helices labelled G–K organized as a four-helix bundle (residues 232–381) (Fig. 3, Supplementary Fig. S4). Fig. 4 shows a structure-based alignment using the *ESPript* online tool (Gouet *et al.*, 2003) of AnaB with its closest homologues, showing these secondary-structure elements. The topology of the secondary-structure elements in the AnaB monomer is also conserved when compared with those of MCAD and IVD: the two α-helix domains are packed together and the β-strand domain forms two orthogonal superimposed β-sheets making contacts with helices E, F and G (Fig. 3*a*).

Cloning of the *anaB* gene into an expression vector (Mann *et al.*, 2011) resulted in the addition of 36 residues at the N-terminal end, which included a six-histidine tag (Supplementary Fig. S1). However, these residues as well as the

following Asp2 and Phe3 (wild-type AnaB numbering) were not identified in the electron-density map, likely because this portion of the polypeptide was too flexible. Given the location of the N-terminus in the tetramer, these extra disordered amino acids do not appear to disturb the structure.

The dimer is formed by two monomers interacting mainly by their C-terminal α-helices (Fig. 3*b*), with an interface area of 1785.8 Å^2. The interface between the monomers in AnaB is very similar to that of IVD and most of the hydrogen bonds between the two chains are conserved. Each monomer contains one FAD molecule with its ribityl chain in an extended conformation. The FAD molecule is buried between the β-sheet domain and the C-terminal domain of one monomer and its adenosine moiety is bound to the αG–αH and αI–αJ loops of the other monomer within the dimer.

The tetramer is formed by two dimers with an interface area of 3026.6 Å^2 and contacts through the helical C-terminal domains that are perpendicular to each other (Fig. 3*c*). The r.m.s.d.s between the monomers within the asymmetric unit were between 0.15 and 0.19 Å and that between the two tetramers in the asymmetric unit was 0.27 Å.

We have compared the three-dimensional structure of AnaB (monomeric or tetrameric) with those of its homologues and the data are compiled in Table 2. The main-chain C$^\alpha$ atoms of AnaB could be superimposed on those of closely related ACADs (sequence identity over 30%) in their free

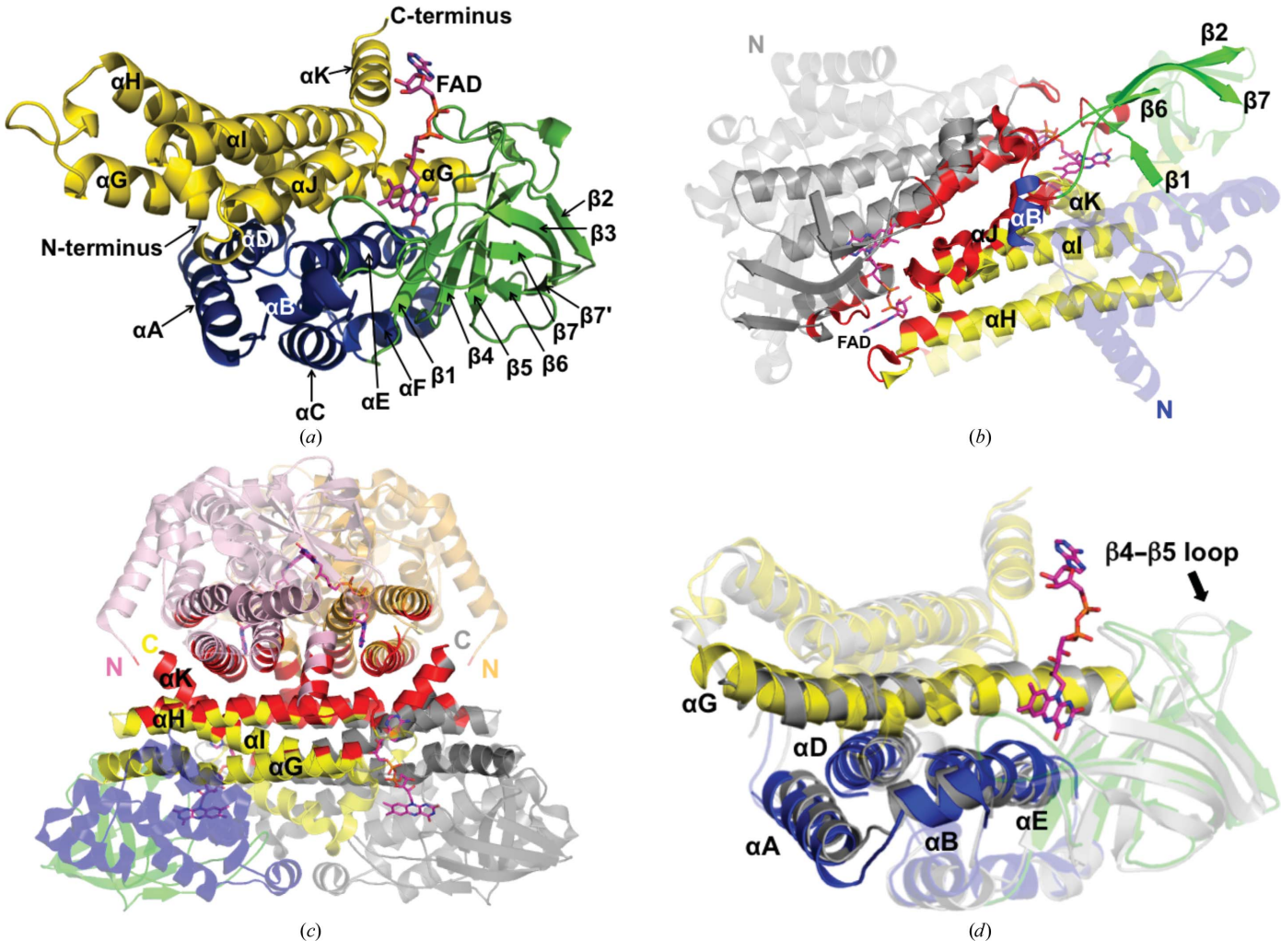

Figure 3
Overall fold of AnaB. (*a*) The AnaB monomer is shown as a ribbon structure. The N-terminal domain consisting of six α-helices (A–F) is coloured blue, the core domain consisting of seven β-strands forming two orthogonal superimposed sheets is coloured green and the C-terminal domain forming a four-helix bundle with α-helices G–K is coloured yellow. The bound FAD cofactor is represented as sticks. (*b*) The AnaB dimer is shown with one chain coloured as in the monomer and the other chain as a grey transparent cartoon. The structural elements involved in the dimer contacts are highlighted and labelled. The monomer–monomer interface area in the dimer is displayed in red. The two FAD molecules are coloured magenta. (*c*) The AnaB tetramer is represented as a transparent cartoon with one chain coloured as in the monomer and the other chain forming the dimer shown in grey, while the other dimer is shown with two monomers coloured orange and pink, respectively. The structural elements involved in the tetramer contacts are highlighted and labelled. The dimer–dimer interface area in the tetramer is displayed in red. (*d*) Superposition of AnaB and human IVD monomers. AnaB is coloured as in (*a*), the IVD chain is shown in grey and the FAD is coloured as in (*a*). Helices A, B, D, E and G, in which the main differences between AnaB and IVD fold are visible, are highlighted while the rest of the chains are shown as transparent cartoons. A stereoview of (*d*) is presented in Supplementary Fig. S5.

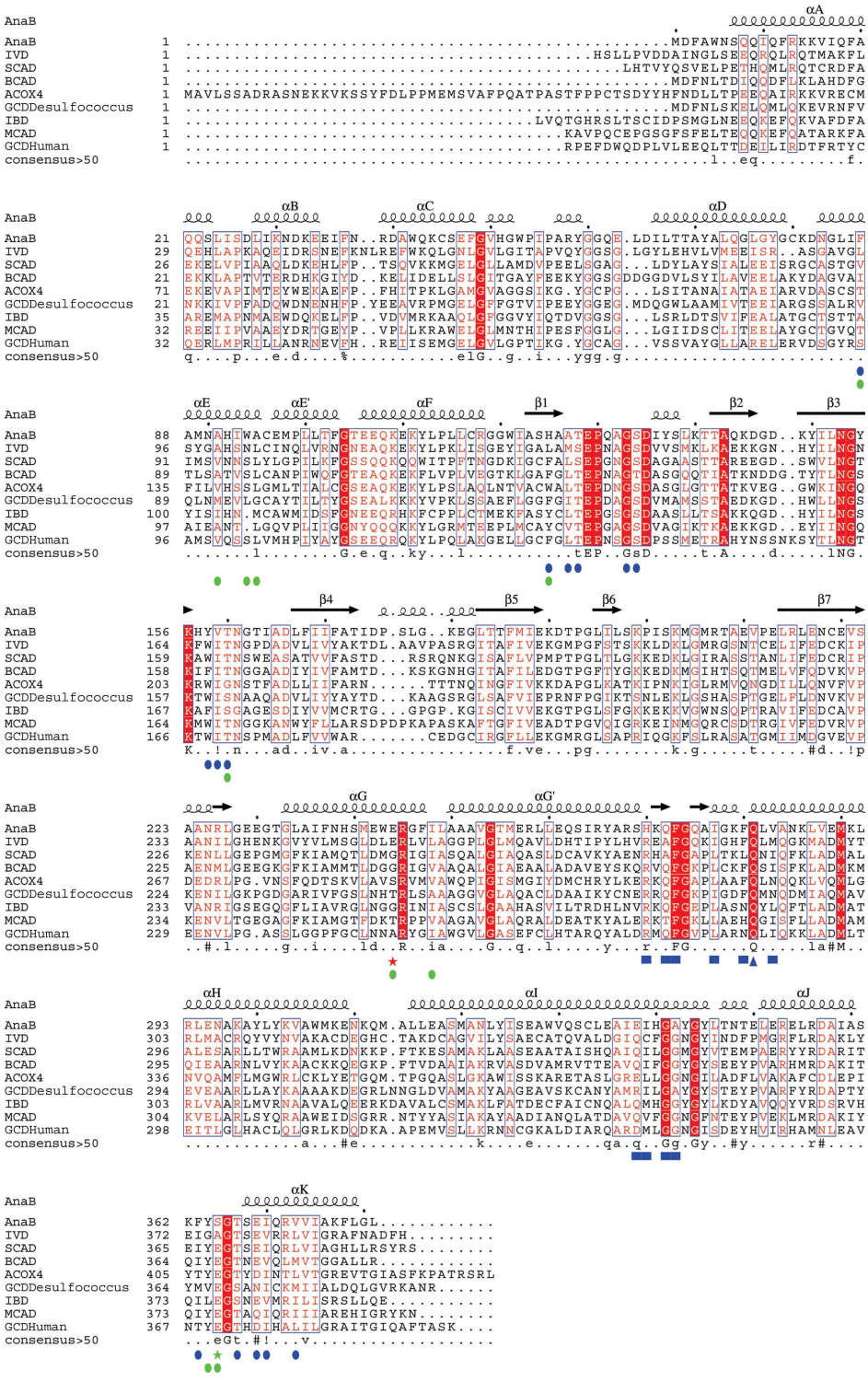

Figure 4
Structure-based multiple alignment of AnaB with its closest homologues. The alignment was performed using the *ESPript*2.2 software (Gouet *et al.*, 2003) based on a *ClustalW* (Thompson *et al.*, 1994) multiple alignment and on the structure of AnaB (chain *A*). The accession numbers for the sequences are AnaB, ZP_07113994.1; IVD, EAW92413; SCAD, AAA40669.1; BCAD, AAA03594.1; ACOX4, NP_190752.1; GCDDesufococcus, ACP50614.1; IBD NP_055199.1; MCAD AAW30430.1; GCDHuman, AAC52079.1. Strictly conserved residues are shown in white on a red background, similar residues are shown in red and blue frames indicate similarity in a group of consecutive residues. Consensus is indicated as follows: uppercase is identity, lowercase is consensus level higher than 50%, ! is I or V, $ is L or M, % is F or Y and # is any residue among N, D, Q and E. The secondary-structure elements of AnaB are shown on the top of the sequences and the α-helices and β-strands are labelled according to the structure of pig MCAD. The active-site bases of AnaB (Glu244) and IVD (Glu254), indicated by a red star, are located within α-helix G, while the active-site bases of the other enzymes, indicated by a green star, are located between α-helices J and K (Glu376 in pig MCAD). The residues in contact with FAD (hydrogen bonds or hydrophobic contacts), in AnaB, are labelled by a blue circle (chain *A*), a blue square (chain *B*) and a blue triangle (chain *D*) and those lining the substrate-binding pocket are labelled with a green circle at the bottom of the sequences.

uncomplexed or substrate-complexed forms (according to the available coordinates). The r.m.s.d.s ranged from 1.29 to 1.87 Å. This indicates that the overall structure and the active-site structure of the ACADs do not change significantly upon substrate binding. However, the closest homologues in terms of three-dimensional structure were IVD, SCAD and IBD, and this might be related to the shape and the size of their substrates: IVD, SCAD and IBD are specific for small or branched substrates and proline, the substrate of AnaB, is a small cyclic molecule.

3.4. Structural differences with respect to ACADs

Despite the similarity in the overall fold and topology of the AnaB monomer to those of ACADs, there are some key structural differences as illustrated by the overlay of the main-chain atoms of AnaB and IVD (Fig. 3*d*, Supplementary Fig. S5). The most notable differences occur in the N-terminal segment, in which helices A, B and D are slightly shifted away from the centre of the protein, yielding differences between AnaB and IVD of up to 3.3 Å in the position of the C^{α} traces of these helices. In addition, AnaB has fewer residues at the N-terminus involved in interactions that contribute to the tetramer interface. This might be owing to the presence of the flexible N-terminal 36-residue peptide, including a six-histidine tag, that was not excised from the protein.

The structural alignment (Fig. 4) shows that the homologous strands and helices in the structures of AnaB and ACADs have the same number of residues, with the exception of helix E, which is longer by one residue in AnaB (Trp94), IVD, SCAD, BCAD and GCD than in MCAD and IBD. This insertion in IVD, SCAD and BCAD restricts the space for the acyl chain of the substrate, a fact consistent with the specificity for small substrates of these enzymes (Djordjevic *et al.*, 1995). AnaB would thus have a small cavity suited to proline. Similarly, there is also a bulge in helix G located at position Ile248 of AnaB, causing helices E and G to be closer to one another and thus decreasing the width of the substrate-binding pocket (Supplementary Fig. S6).

The β-sheet domain of AnaB (shown in green in Fig. 3*a*) superimposed well with those of ACADs, although differences appear in the loop connecting β-strands 4 and 5. This loop is shorter in AnaB when compared with IVD and MCAD, but is similar in AnaB and SCAD or BCAD. It has been suggested that this loop together with loops β7–αG and αH–αI might be involved in the binding of the ACP substrate by FkbI (ACPD) of *S. hygroscopicus* (Watanabe *et al.*, 2003). However, there is no experimental structure to support this and these loops do not superimpose well in the structures of AnaB and FkbI, although both use acyl-ACP instead of acyl-CoA substrates.

The Cys318–Cys323 disulfide bridge in IVD in the loop connecting helices H and I is not conserved in AnaB and the corresponding residues are Lys308 and Met313.

3.5. The FAD-binding site

The FAD cofactor was readily identified in AnaB, and Fig. 5(*a*) shows its model built in the electron-density map. The interactions (hydrogen bonds and hydrophobic contacts) between the residues and FAD are summarized in Supplementary Fig. S7 as a *LIGPLOT* diagram. We superimposed the binding pocket of FAD, defined as the protein residues at a distance of 4 Å from the cofactor, in AnaB and IVD (Fig. 5). The amino acids lining the FAD-binding pockets are very similar, and of the 25 residues defining the FAD-binding pocket in AnaB and IVD ten are identical, 11 are chemically similar and only four are different (His125, Ala127, Gln272 and Ile339 of AnaB). However, Gln272 (Ala282 in IVD) is fairly well conserved in the other ACADs and Ile339 (Cys349 in IVD) is conserved or replaced by large hydrophobic residues (Met, Leu or Val) in the other ACADs (Fig. 4). In contrast, for the two other residues (His125 and Ala127 of AnaB) the replacements are likely to be significant. These two residues are located on the *Re* face of the isoalloxazine ring of FAD (Ala127 is actually almost in the plane of the FAD) and make backbone hydrogen bonds to the FAD N3 and O2 atoms, respectively. His125 of AnaB (Leu133 in IVD) is replaced by aromatic residues in the other ACADs, and Ala127 of AnaB (Met135 in IVD) is

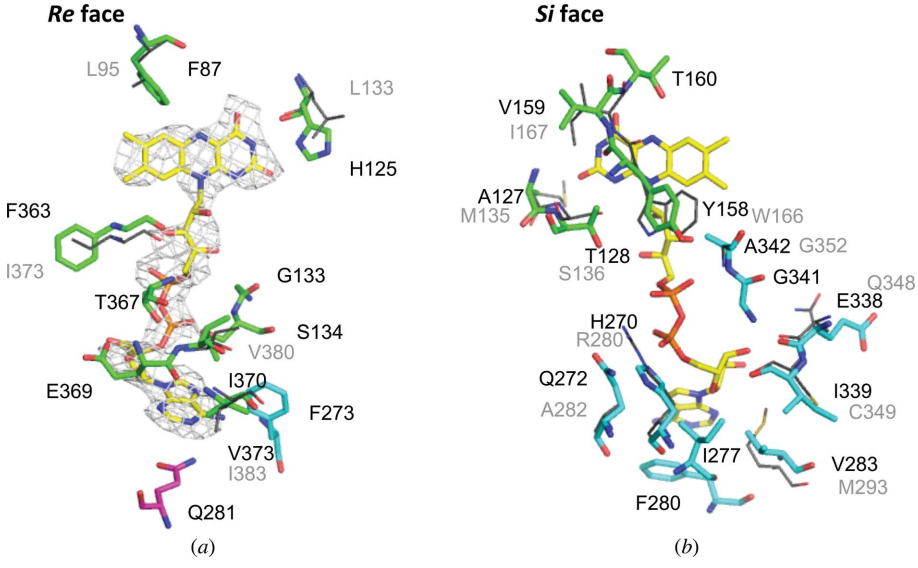

Figure 5
Overlay of the residues of AnaB and IVD involved in the FAD-binding pocket. The residues are displayed as sticks, with their backbone and side-chain C atoms in green, cyan and magenta for chains *A*, *B* and *D* of AnaB, respectively. The residues of IVD are not shown except when they differ from the corresponding AnaB residues; they are then shown as grey sticks. The FAD is coloured yellow and its OMIT $F_o - F_c$ electron-density map is shown contoured at the 2.5σ level. (*a*) The residues located on the *Re* face of FAD are displayed. (*b*) The residues located on the *Si* face of FAD are displayed.

replaced by large hydrophobic residues in the other ACADs (Fig. 4). As discussed below, we believe that these two positions are important for the oxidase activity of AnaB.

The *Si* face of the isoalloxazine moiety of the FAD cofactor lies on the loop connecting β3 and β4 and is surrounded by residues Thr128, Tyr158 and Thr160. Thr128 and Thr160 are highly conserved residues in the ACAD family and form hydrogen bonds to the N5 and O2 atoms of FAD, respectively. The dimethylbenzene ring of the flavin forms π-stacking interactions with the phenolic ring of Tyr158, a conserved aromatic residue in ACADs.

In ACADs there is a highly conserved hydrogen-bond network involving the diphosphate portion of FAD and a Ser and an Arg residue. These interactions are conserved in AnaB and they involve Ser134 and His270. The remainder of the FAD interacts extensively with conserved residues, including Phe273, Gln281, Glu338, Gly341, Ala342 and Glu369. Overall, the interactions involved in FAD binding in AnaB are very similar to those described in ACADs.

3.6. The active-site base

In the ACAD family the active-site base is always a glutamate residue (Ghisla & Thorpe, 2004; Kim & Miura, 2004). The majority of the enzymes have their active-site base located at the C-terminal end on the loop connecting α-helices J and K (Glu376 in pig MCAD), while IVD and LCAD have their active-site base within α-helix G (Glu254 in IVD and Glu261 in LCAD), approximately 120 residues upstream in the sequence. However, the active-site bases occupy the same position in the three-dimensional structure. Interestingly, permuting these positions in MCAD by site-directed mutagenesis gave active mutants (Nandy *et al.*, 1996; Lee *et al.*, 1996; Dakoji *et al.*, 1997). By sequence alignment, we identified Glu244 of AnaB as the active-site base and this was confirmed by site-directed mutagenesis (Mann *et al.*, 2011).

Superposition of the active-site structure of AnaB with those of IVD and MCAD clearly confirmed the identity of the catalytic base as Glu244, which has its side chain pointing towards N5 of FAD at a distance of 6.4 Å (Fig. 6). The side chain of Glu244 occupies the same position as that of Glu254 of IVD and is located in a position homologous to the catalytic residue of MCAD, Glu376. In addition, we have overlaid the CoA persulfide and octanoyl-CoA ligands of IVD and MCAD, respectively, into the substrate-binding pocket of AnaB (Fig. 6*b*). In this superposition, the C2—C3 bond of octanoyl-CoA is almost at the middle of the line between the carboxylate of Glu244 and the isoalloxazine ring of FAD. Also, the thioester O atom is sandwiched between the FAD and the catalytic base in a position allowing a potential hydrogen bond to the ribityl 2′-OH of FAD. It has been proposed that this interaction is important in the positioning and alignment of the flavin, substrate and catalytic base in MCAD for optimal catalysis (Kim & Miura, 2004). Therefore, it is likely that the substrate of AnaB is positioned in an analogous way as CoA persulfide in IVD and octanoyl-CoA in MCAD.

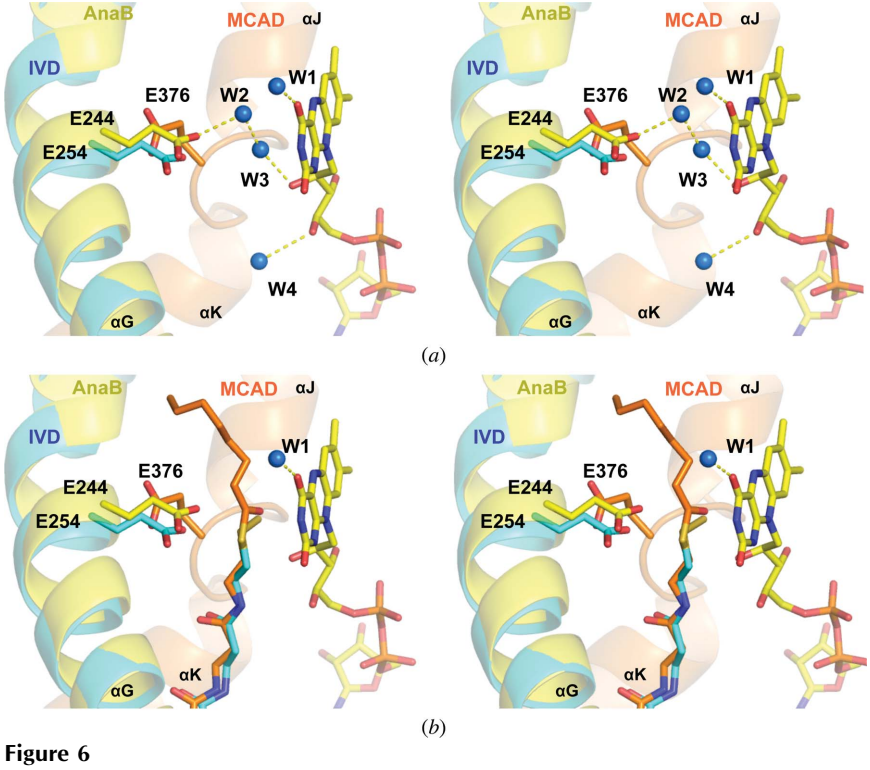

Figure 6
Stereoview of the superposition of the active site of AnaB with those of IVD and MCAD. The FAD cofactor of AnaB as well as the catalytic bases are shown in stick representation. Helices G of AnaB and IVD, which contain the catalytic base (Glu244 and Glu254), are represented in yellow and cyan, respectively. Helices J and K of MCAD and their connecting loop, which contains the catalytic base Glu376, are shown in orange. Hydrogen bonds are shown as dotted lines. (*a*) The water molecules present in the substrate cavity of crystalline AnaB are indicated as blue spheres. W1 is common to all known ACAD structures. (*b*) The ligands of the IVD and the MCAD structures (PDB entries 1ivh and 3mde), CoA persulfide and octanoyl-CoA, respectively, were overlaid into the active site of AnaB.

In AnaB, in the absence of ligand the active site is occupied by well defined water molecules (Fig. 6*a*). These water molecules have also been found at equivalent positions in the ligand-free structures of MCAD, IBD and human GCD, but they are displaced upon substrate binding. However, a water molecule, W1, is the only well ordered conserved water molecule found near the active site of all other ACAD structures, free or in complex. Finally, this structure of AnaB has conclusively identified the catalytic base and confirmed our previous site-directed mutagenesis experiments and mechanistic proposals (Mann *et al.*, 2011).

3.7. The substrate-binding pocket

In ACADs, the fatty-acyl chain of the thioester substrate is bound at the *Re* face of the flavin ring and is buried inside the

monomer between helices G and E and the loop between helices J and K. The size and shape of the substrate-binding pocket is of major importance for ACAD substrate specificity (Ghisla & Thorpe, 2004; Kim & Miura, 2004). We have estimated the substrate-binding pockets of AnaB and IVD using the *fpocket* program (Le Guilloux *et al.*, 2009) and they are represented as surfaces in Fig. 7 for comparison. The ten residues lining the binding pocket are mostly hydrophobic (marked with green circles in Fig. 4), with six residues likely to be mainly responsible for shaping the binding cavity of AnaB: Phe87, Trp94, Ala95, His125, Ile248 and Tyr364 (Fig. 4). Among these residues, Phe87, Trp94, Ile248 and Tyr364 seemed to particularly restrict the binding cavity. Phe87 is bulkier than the corresponding residues Leu95, Ile88 and Val90 in IVD, BCAD and SCAD, respectively. This residue is located within the bulge of helix E, which together with helix G limits the depth of the binding cavity. Trp94, which is bulkier than the corresponding residue in the other ACADs (Fig. 4), seems to block the bottom of the substrate-binding cavity. Ile248, which is located on helix G, has its side chain pointing into the binding cavity, as do the corresponding residues Leu258 and Ile251 in IVD and SCAD, respectively, and thus contributes to making the substrate-binding pocket shallower.

In contrast, the small Ala95 and His125 of AnaB in comparison to the corresponding Leu or Met residues in other ACADs probably make the binding cavity slightly wider at this position and could thus eliminate potential steric repulsions with the pyrrolidine ring of the prolyl-ACP substrate.

Finally, Tyr364 of AnaB is conserved in all known ACADs except for IVD and IBD, which have a glycine (Gly374) or a leucine (Leu375) at this position. The presence of less bulky residues at this position has been shown to allow space for branched-chain substrates.

Overall, the substrate-binding pocket of AnaB is small and is likely fitted to accept the pyrrolidine ring of proline.

In the absence of a three-dimensional structure of AnaB in complex with its prolyl-ACP substrate, it is difficult to predict the binding site for the ACP portion of the substrate. Nevertheless, we have superimposed the structure of AnaB with that of IVD in complex with CoA persulfide (Supplementary Fig. S8). This superposition shows that several residues of IVD involved in the binding of the ribose 3′-phosphate of the CoASSH ligand are not conserved in AnaB. These AnaB residues seem to make steric or charge repulsions with the adenosine of CoASSH (see Supplementary Fig. S8 for details), a part of the substrate that is not present in ACP.

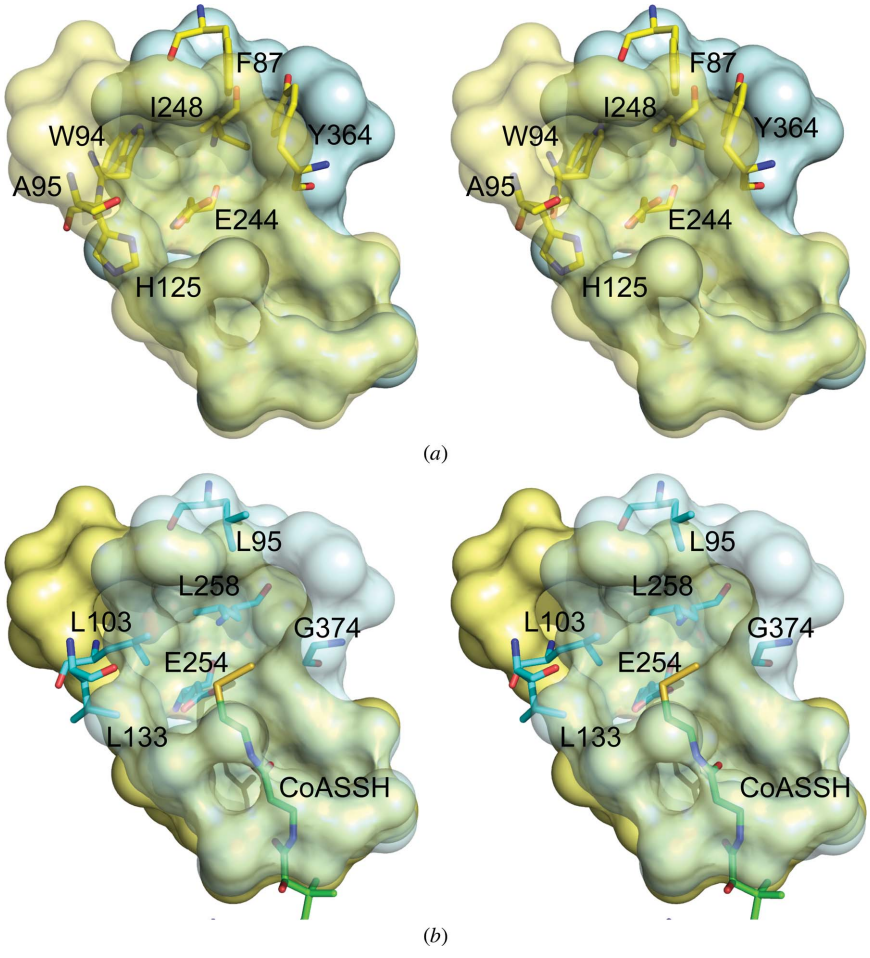

Figure 7
Graphical representation (stereoview) of the superimposed substrate-binding pockets of AnaB and IVD. The pockets, estimated using the *fpocket* software (Le Guilloux *et al.*, 2009), are displayed as surfaces and are coloured yellow for AnaB and cyan for IVD. Residues important for the two pockets are displayed as sticks with their backbone and side-chain C atoms coloured yellow for AnaB and cyan for IVD. The ligand (CoA persulfide, CoASSH) is displayed as sticks with its C atoms coloured green. (*a*) The substrate-binding pocket of AnaB and its important residues are represented in the foreground and the substrate-binding pocket of IVD is in the background. (*b*) The substrate-binding pocket of IVD and its important residues are represented in the foreground and the substrate-binding pocket of AnaB is in the background.

3.8. Oxygen accessibility and oxidase activity

The reactivity of molecular oxygen with flavoenzymes is a fascinating research field and recent reviews have summarized the latest published data in this area (McDonald *et al.*, 2011; Chaiyen *et al.*, 2012; Gadda, 2012). In fact, despite numerous studies there is no clear understanding of the structural factors that govern the different reactivity of flavoenzymes towards molecular oxygen, and in particular in ACADs and ACOXs (DuPlessis *et al.*, 1998). It is well established that mammalian ACADs show relatively high oxygen reactivity in their ligand-free form compared with their product-bound form and it has been proposed that the strong binding of the product leads to nearly complete desolva-

tion of the active site, preventing oxygen diffusion and contributing to the suppression of any oxidase activity. This protective effect against oxygen reactivity by ligand binding is physiologically crucial to favour flavin re-oxidation by ETF.

However, a clearer picture has emerged from several recent studies on diverse flavoenzymes: the oxygen reactivity is likely to be modulated by a combination of several factors such as the charge distribution and solvent accessibility to the reactive flavin N5–C4a locus (McDonald *et al.*, 2011; Chaiyen *et al.*, 2012; Gadda, 2012). Particularly, the oxygen reactivity seems to be enhanced by the presence of a positively charged residue in the vicinity of the N5–C4a locus that should stabilize the caged radical pair superoxide-FAD semiquinone.

Using the *NACCESS* program (Hubbard *et al.*, 1991), we calculated and compared the solvent-accessible surface area (ASA) of the FAD cofactor and of its isoalloxazine ring in AnaB and homologous ACADs and in some ACOXs (Figs. 8*a* and 8*b*). With an average ASA of 62 Å^2, the whole FAD molecule of AnaB is clearly more accessible to the solvent than those of other ACADs, which have an average ASA around 40 Å^2. The FAD cofactor of GCD from *Desulfococcus multivorans* is, however, an exception, with an average ASA of 65 Å^2. This high ASA value for FAD in this enzyme might be explained by the fact that *D. multivorans* is an obligate anaerobe and thus GCD had no evolutionary pressure to protect itself from oxidation by molecular oxygen (Wischgoll *et al.*, 2010). On the other hand, ACOX4, a plant short-chain acyl-CoA oxidase from *A. thaliana*, has its FAD cofactor quite buried, with an average ASA of 36 Å^2. This strikingly low ASA for an oxidase has previously been noted (Mackenzie *et al.*, 2006) and the authors proposed that the oxidase activity of this enzyme was owing to a blockade of electron-transfer reactions to acceptors by an extra N-terminal domain and to a small cavity for molecular oxygen near the N5 of FAD. Nevertheless, this oxidase seems to be an exception among other oxidases. We also found a high solvent accessibility for the isoalloxazine ring of FAD in AnaB when compared with those of homologous enzymes (Fig. 8*b*). As a comparison, we calculated the FAD ASAs in the rat ACOX II (rACOXII; PDB entry 1is2; Nakajima *et al.*, 2002) and found high values as expected for an oxidase (Fig. 8).

As pointed out earlier (Figs. 4 and 5), the isoalloxazine ring of FAD in AnaB is surrounded by smaller residues than those found in its homologues (His125, Ala127 and Tyr158 in AnaB). For instance, all known ACADs except for IBD and BCAD have a tryptophan (Trp166 of IVD) that protects the dimethylbenzene ring of the flavin at its *Si* face. The presence of a smaller residue, Phe or Tyr, at this corresponding position makes the flavin more exposed to solvent and thus more reactive towards oxygen, as was shown for BCAD (Djordjevic

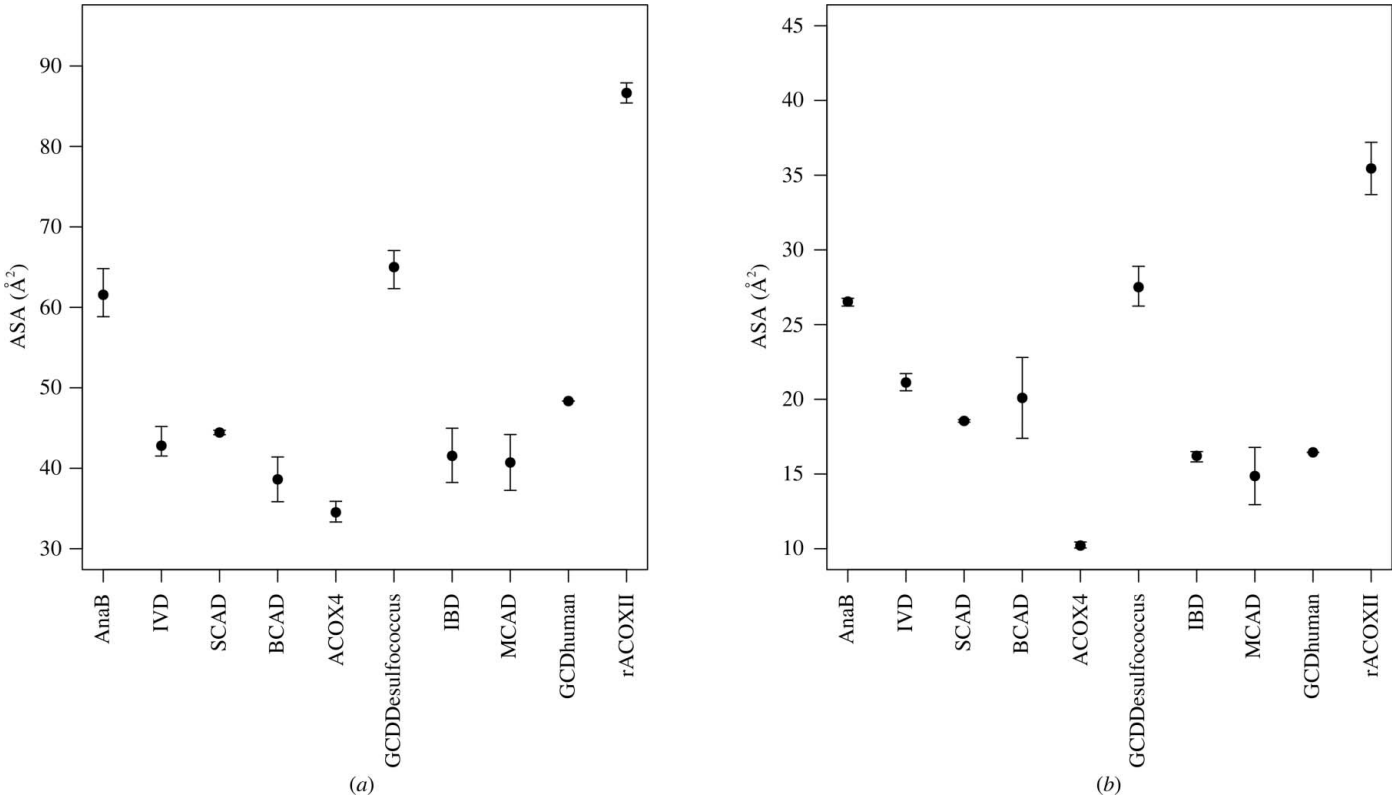

Figure 8
Accessible surface area (ASA) of FAD calculated for AnaB and for some of its homologues. The ASAs were calculated using the *NACCESS* program (Hubbard *et al.*, 1991) using the tetrameric and ligand-free structures. There are thus several ASA values per enzyme because the monomers are not strictly equivalent. The data points represent the average ASA value and the bars span from the highest to the lowest ASA value. The graphs were drawn using the *R* Program for Statistical Computing and Graphics (R Development Core Team, 2012). (*a*) Plot of the ASA of the whole FAD molecule. (*b*) Plot of the ASA of the isoalloxazine ring of the FAD molecule.

et al., 1995). Overall, the high solvent accessibility of the AnaB flavin suggests that oxygen might easily reach the flavin for electron transfer.

Secondly, we found a network of seven well defined water molecules that link the bulk solvent to Tyr158 at the *Si* face of the FAD cofactor in the crystal structure of AnaB. Fig. 9 shows the volume of this water channel in AnaB. In IVD there are also well defined water molecules but the channel (Fig. 9) is smaller and narrower, particularly around Arg280. Quite interestingly, His270 of AnaB replaces this strictly conserved Arg residue in ACADs (Fig. 4) and it seems that this smaller residue leaves space for formation of the water channel in AnaB. Also, it is quite clear from Fig. 9 that Tyr158 of AnaB, being smaller than Trp166 of IVD, does not block the channel, likely allowing oxygen to reach the flavin. In MCAD, it was shown that a possible binding site for ETF is actually exactly around this position and is in contact with Trp166 for electron transfer (Toogood *et al.*, 2004). We thus believe that we have identified a possible channel for water and oxygen access to the flavin in AnaB and for release of H_2O_2.

A third important point is the presence of a charged residue, His125, close to the N5–C4a locus of the FAD of AnaB. It has been suggested in recent studies that the presence of charged residues (Lys and His) near the flavin in oxidases is an important factor for oxygen activation in these enzymes (Chaiyen *et al.*, 2012; Gadda, 2012). For instance, in glucose oxidase two protonated histidine residues at positions 559 and 516 in the active site seem to be involved in oxygen activation. It has been proposed that these positively charged side chains are involved in stabilizing and caging the superoxide-flavin semiquinone radical pair (Leskovac *et al.*, 2005). We superimposed the structure of AnaB and glucose oxidase and, quite interestingly, His125 of AnaB is at almost the same position as His559 of glucose oxidase (Supplementary Fig. S9). Although located 5.7 Å from the flavin C4a atom (4.5 Å in the glucose oxidase structure), His125 might contribute to oxygen activation. In addition, the presence of the protonated imine of the product of AnaB (P5C-AnaD; Fig. 2) might increase the polarity of the active site, again favouring activation.

Altogether, the data reported here, *i.e.* (i) a high solvent accessibility to FAD, (ii) the presence of a solvent-accessible channel near the *Si* face of FAD and (iii) a potential positive charge environment in the active site, provide a structural basis for the fact that AnaB is an oxidase rather than a dehydrogenase.

4. Conclusion

Here, we report the three-dimensional structure of AnaB, a prolyl-ACP oxidase involved in the biosynthesis of the cyanobacterial neurotoxins anatoxin-a and homoanatoxin-a. The structure is strikingly similar to those of ACADs and in particular that of human IVD. The overall fold is conserved, the FAD-binding site is very similar and the active-site base is conserved. However, we found that the substrate-binding pocket is smaller in AnaB when compared with ACADs and is likely to be adapted to the five-membered cyclic moiety of prolyl-ACP. We have also found that the flavin in AnaB is more accessible to solvent than in homologous ACADs, which is a consequence of the presence of smaller residues around FAD. Furthermore, we have identified a water channel in AnaB that is likely to be the preferred path for molecular oxygen in the oxidative reaction of this oxidase and a charged residue (His125) that most likely facilitates oxygen activation. These differences between ACADs and AnaB are probably sufficient to explain the fact that this enzyme is an oxidase rather than a dehydrogenase and provide valuable information for better understanding the oxygen reactivity of flavoenzymes.

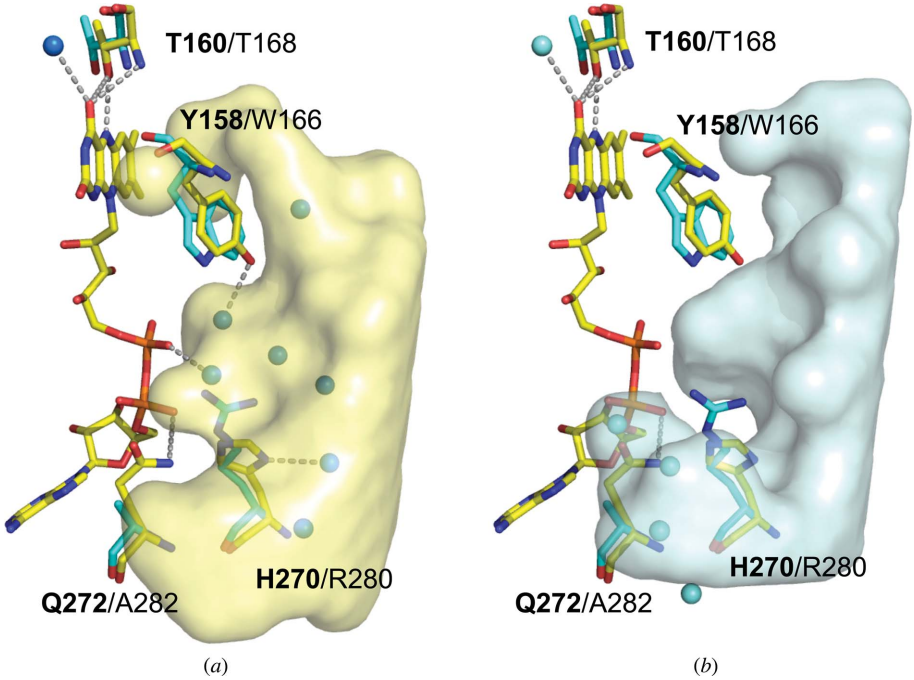

Figure 9
Comparison of water accessibility to FAD in AnaB and IVD. The channel surfaces in AnaB (*a*) and IVD (*b*) created using the program *HOLLOW* (Ho & Gruswitz, 2008) are represented in yellow and cyan, respectively. The FAD cofactor is shown as yellow sticks. Residues of AnaB potentially involved in oxygen accessibility and the corresponding residues in IVD are shown in yellow and cyan, respectively. Water molecules present in the AnaB and IVD structures are represented as blue and cyan spheres, respectively. Hydrogen bonds are shown as dotted lines; for clarity, hydrogen bonds involving water molecules are not represented.

OP, AM and KM planned and designed the experiments; KM performed the crystallizations, the

diffraction experiments and solved the structure; LR and KM performed the calculations; SM purified the protein; OP, KM, LR and AM wrote the manuscript; OP, AM and KM afforded financial support. All authors have given approval to the final version of the manuscript. This work was supported by the Université Paris Diderot-Paris 7, the CNRS and the ENSCP. We acknowledge the European Synchrotron Radiation Facility for providing access to beamline ID23-2. We thank Ahmed Haouz and Patrick Weber from the Pasteur Institute crystallization facility (Paris, France) for initial crystallization screening.

References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Battaile, K. P., Molin-Case, J., Paschke, R., Wang, M., Bennett, D., Vockley, J. & Kim, J.-J. P. (2002). *J. Biol. Chem.* **277**, 12200–12207.
Battaile, K. P., Nguyen, T. V., Vockley, J. & Kim, J.-J. P. (2004). *J. Biol. Chem.* **279**, 16526–16534.
Begley, D. W., Davies, D. R., Hartley, R. C., Hewitt, S. N., Rychel, A. L., Myler, P. J., Van Voorhis, W. C., Staker, B. L. & Stewart, L. J. (2011). *Acta Cryst.* F**67**, 1060–1069.
Bowles, T., Metz, A. H., O'Quin, J., Wawrzak, Z. & Eichman, B. F. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 15299–15304.
Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. & Womack, T. O. (2011). *BUSTER*. Cambridge: Global Phasing Ltd.
Cadel-Six, S., Iteman, I., Peyraud-Thomas, C., Mann, S., Ploux, O. & Méjean, A. (2009). *Appl. Environ. Microbiol.* **75**, 4909–4912.
Chaiyen, P., Fraaije, M. W. & Mattevi, A. (2012). *Trends Biochem. Sci.* **37**, 373–380.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Dakoji, S., Shin, I., Battaile, K. P., Vockley, J. & Liu, H. (1997). *Bioorg. Med. Chem.* **5**, 2157–2164.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Djordjevic, S., Pace, C. P., Stankovich, M. T. & Kim, J.-J. P. (1995). *Biochemistry*, **34**, 2163–2171.
DuPlessis, E. R., Pellett, J., Stankovich, M. T. & Thorpe, C. (1998). *Biochemistry*, **37**, 10469–10477.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fu, Z., Wang, M., Paschke, R., Rao, K. S., Frerman, F. E. & Kim, J.-J. P. (2004). *Biochemistry*, **43**, 9674–9684.
Gadda, G. (2012). *Biochemistry*, **51**, 2662–2669.
Ghisla, S. & Thorpe, C. (2004). *Eur. J. Biochem.* **271**, 494–508.
Goto, M., Muramatsu, H., Mihara, H., Kurihara, T., Esaki, N., Omi, R., Miyahara, I. & Hirotsu, K. (2005). *J. Biol. Chem.* **280**, 40875–40884.
Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.
Héroux, A., Bozinovski, D. M., Valley, M. P., Fitzpatrick, P. F. & Orville, A. M. (2009). *Biochemistry*, **48**, 3407–3416.
Ho, B. K. & Gruswitz, F. (2008). *BMC Struct. Biol.* **8**, 49.
Hubbard, S. J., Campbell, S. F. & Thornton, J. M. (1991). *J. Mol. Biol.* **220**, 507–530.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kim, J.-J. P. & Miura, R. (2004). *Eur. J. Biochem.* **271**, 483–493.
Kim, J.-J. P., Wang, M. & Paschke, R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 7523–7527.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.
Lee, H.-J. K., Wang, M., Paschke, R., Nandy, A., Ghisla, S. & Kim, J.-J. P. (1996). *Biochemistry*, **35**, 12412–12420.
Le Guilloux, V., Schmidtke, P. & Tuffery, P. (2009). *BMC Bioinformatics*, **10**, 168.
Leskovac, V., Trivić, S., Wohlfahrt, G., Kandrac, J. & Pericin, D. (2005). *Int. J. Biochem. Cell Biol.* **37**, 731–750.
Mackenzie, J., Pedersen, L., Arent, S. & Henriksen, A. (2006). *J. Biol. Chem.* **281**, 31012–31020.
Mann, S., Lombard, B., Loew, D., Méjean, A. & Ploux, O. (2011). *Biochemistry*, **50**, 7184–7197.
McAndrew, R. P., Wang, Y., Mohsen, A. W., He, M., Vockley, J. & Kim, J.-J. P. (2008). *J. Biol. Chem.* **283**, 9435–9443.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McDonald, C. A., Fagan, R. L., Collard, F., Monnier, V. M. & Palfey, B. A. (2011). *J. Am. Chem. Soc.* **133**, 16809–16811.
Méjean, A., Mann, S., Maldiney, T., Vassiliadis, G., Lequin, O. & Ploux, O. (2009). *J. Am. Chem. Soc.* **131**, 7512–7513.
Méjean, A., Mann, S., Vassiliadis, G., Lombard, B., Loew, D. & Ploux, O. (2010). *Biochemistry*, **49**, 103–113.
Meng, Z., Lou, Z., Liu, Z., Li, M., Zhao, X., Bartlam, M. & Rao, Z. (2006). *J. Mol. Biol.* **359**, 1364–1377.
Nakajima, Y., Miyahara, I., Hirotsu, K., Nishina, Y., Shiga, K., Setoyama, C., Tamaoki, H. & Miura, R. (2002). *J. Biochem.* **131**, 365–374.
Nandy, A., Kieweg, V., Kräutle, F. G., Vock, P., Küchler, B., Bross, P., Kim, J.-J. P., Rasched, I. & Ghisla, S. (1996). *Biochemistry*, **35**, 12402–12411.
Ortega, A., Amorós, D. & García de la Torre, J. (2011). *Biophys. J.* **101**, 892–898.
Pedersen, L. & Henriksen, A. (2005). *J. Mol. Biol.* **345**, 487–500.
Perrière, G. & Gouy, M. (1996). *Biochimie*, **78**, 364–369.
Powers, R. A., Rife, C. L., Schilmiller, A. L., Howe, G. A. & Garavito, R. M. (2006). *Acta Cryst.* D**62**, 683–686.
R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. http://www.r-project.org.
Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C. & Bricogne, G. (2012). *Acta Cryst.* D**68**, 368–380.
Tanner, J. J. (2008). *Amino Acids*, **35**, 719–730.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W., Vockley, J. & Kim, J.-J. P. (1997). *Biochemistry*, **36**, 8455–8464.
Toogood, H. S., van Thiel, A., Basran, J., Sutcliffe, M. J., Scrutton, N. S. & Leys, D. (2004). *J. Biol. Chem.* **279**, 32904–32912.
Tormos, J. R., Taylor, A. B., Daubner, S. C., Hart, P. J. & Fitzpatrick, P. F. (2010). *Biochemistry*, **49**, 5035–5041.
Walsh, C. T., Garneau-Tsodikova, S. & Howard-Jones, A. R. (2006). *Nat. Prod. Rep.* **23**, 517–531.
Watanabe, K., Khosla, C., Stroud, R. M. & Tsai, S.-C. (2003). *J. Mol. Biol.* **334**, 435–444.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Wischgoll, S., Demmer, U., Warkentin, E., Günther, R., Boll, M. & Ermler, U. (2010). *Biochemistry*, **49**, 5350–5357.
Wonnacott, S. & Gallagher, T. (2006). *Mar. Drugs*, **4**, 228–254.