Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of MrsD, an FAD-binding protein of the HFCD family

Michael Blaesse,[a] Thomas Kupke,[b] Robert Huber[a] and Stefan Steinbacher[a]*†

[a]Max-Planck-Institut für Biochemie, Abteilung für Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany, and [b]Lehrstuhl für Mikrobielle Genetik, Universität Tübingen, Auf der Morgenstelle 15, Verfügungsgebäude, D-72076 Tübingen, Germany

† Present address: Division of Chemistry and Chemical Engineering, Mail Code 114-96, California Institute of Technology, Pasadena, CA 91125, USA.

Correspondence e-mail: steinbac@caltech.edu

Received 19 April 2003
Accepted 29 May 2003

PDB Reference: MrsD, 1p3y, r1p3ysf.

MrsD from *Bacillus* sp. HIL-Y85/54728 is a member of the HFCD (homo-oligomeric flavin-containing Cys decarboxylases) family of flavoproteins and is involved in the biosynthesis of the lantibiotic mersacidin. It catalyses the oxidative decarboxylation of the C-terminal cysteine residue of the MrsA precursor peptide of mersacidin, yielding a (*Z*)-enethiol intermediate as the first step in the formation of the unusual amino acid *S*-[(*Z*)-2-aminovinyl]-methyl-D-cysteine. Surprisingly, MrsD was found to bind FAD, in contrast to the three other characterized members of the HFCD family, which bind FMN. To determine the molecular discriminators of FAD binding within the HFCD family, the crystal structure of MrsD was analyzed at a resolution of 2.54 Å. Crystals of space group *F*432 contain one MrsD monomer in the asymmetric unit. However, a Patterson search with EpiD-derived models failed. Based on the consideration that the dodecameric MrsD particle of tetrahedral symmetry resembles the quaternary structure of EpiD, rotational and translational parameters were derived from the geometric consideration that the MrsD dodecamer is generated from a monomer by crystallographic symmetry around the position (1/4, 1/4, 1/4) of the unit cell. A structural comparison with the FMN-binding members of the HFCD family EpiD and AtHAL3a shows conserved sequence motifs in contact with the flavin's pyrimidine ring but divergent environments for the dimethylbenzene ring of the isoalloxazine moiety. The position of the ribityl chain differs in MrsD from that found in EpiD and AtHAL3a. However, the FMN-phosphate binding sites are also highly conserved in their exact positions. In all three cases, the flavin cofactor is bound to a structurally conserved region of the Rossmann-fold monomer, exposing its *Re* side for catalysis. The adenosyl phosphate of FAD is anchored in a well defined binding site and the adenosine moieties are oriented towards the interior of the hollow particle, where three of them pack against each other around the threefold axis of a trimeric facet.

1. Introduction

The lantibiotic mersacidin is active against methicillin-resistant *Staphylococcus aureus* strains (Chatterjee *et al.*, 1992) and inhibits the transglycosylation reaction in peptidoglycan biosynthesis by binding to the cell-wall precursor lipid II (Brötz *et al.*, 1997, 1998). Lantibiotics are ribosomally synthesized and post-translationally modified peptides that contain the thioether amino acid lanthionine as a characteristic building block (Schnell *et al.*, 1988). In mersacidin, epidermin, gallidermin and some of the mutacins, unsaturated thioether bridges are present (Allgaier *et al.*, 1985; Kellner *et al.*, 1988; Minami *et al.*, 1994; Mota-Meira *et al.*, 1997; Qi *et al.*, 1999, 2000; Smith *et al.*, 2000). Biosynthesis of the carboxy-

terminal *S*-[(*Z*)-2-aminovinyl]-3-methyl-D-cysteine residue of mersacidin occurs in two steps. The FAD-dependent flavoprotein MrsD catalyzes the oxidative decarboxylation of the COOH-terminal cysteine residue of the mersacidin precursor peptide MrsA to an aminoenethiol structure (Fig. 1) (Majer *et al.*, 2002). The unsaturated thioether is then formed by addition of the enethiol group to the didehydrobutyrine residue at position +15 of MrsA which is introduced by dehydration of Thr. The homologous enzyme EpiD, which is an FMN-dependent and not an FAD-dependent enzyme, catalyzes the oxidative decarboxylation of the peptidyl-cysteine precursor peptide EpiA (Kupke *et al.*, 1992, 1994, 1995). MrsD and EpiD belong to a new family of flavoproteins that was named HFCD (homo-oligomeric flavin-containing Cys decarboxylases; Blaesse *et al.*, 2000; Kupke, 2002*a*; Kupke *et al.*, 2000*b*). Other members of this flavoprotein family include the flavoenzymes Dfp from eubacteria (Kupke *et al.*, 2000*b*; Spitzer *et al.*, 1988; Spitzer & Weiss, 1985) and AtHAL3a from *Arabidopsis thaliana* (Espinosa-Ruiz *et al.*, 1999). Both catalyze the decarboxylation of (*R*)-4′-phospho-*N*-pantothenoylcysteine (*P*PC) to 4′-phosphopantetheine (*P*P) (Kupke, 2001; Kupke *et al.*, 2000*b*, 2001; Strauss & Begley, 2001), a reaction introducing the reactive cysteamine residue of coenzyme A. However, Dfp is a bifunctional enzyme: the COOH-terminal CoaB domain catalyzes the synthesis of *P*PC from 4′-phosphopantothenate and L-cysteine using cytidine-5′-triphosphate as the activating nucleotide (Kupke, 2002*b*; Strauss *et al.*, 2001), whereas the *P*PC decarboxylase activity resides in the amino-terminal CoaC domain (Kupke, 2001). The decarboxylation of *P*PC is a two-step reaction. In the first half-reaction *P*PC is oxidatively decarboxylated to the 4′-phosphopantothenoyl-aminoethene intermediate, which is reduced to 4′-phosphopantetheine in the second half-reaction (Hernández-Acosta *et al.*, 2002; Steinbacher *et al.*, 2003).

The crystal structures of EpiD, of the active-site mutant EpiD H67N with bound pentapeptide substrate DSYTC (Blaesse *et al.*, 2000), of AtHal3a (Albert *et al.*, 2000) and of the AtHAL3a mutant C175S with bound oxidatively decarboxylated intermediate pantothenoyl-aminoethenethiol (Steinbacher *et al.*, 2003) have recently been determined. These structures gave insight into the reaction mechanism, the substrate-binding mode and the substrate specificity of the HFCD proteins.

EpiD forms dodecamers with trimers disposed on the vertices of a tetrahedron. Each of the monomers consists of a single domain with a Rossmann-type fold. Oligomer formation is essential for binding of the flavin mononucleotide cofactor and the substrate, which is buried by an otherwise disordered substrate-recognition clamp. The FMN cofactor can be assigned to one subunit that provides the majority of the interactions and is buried at the center of the trimer side face contacting three subunits. The sequence motifs PASANT and P*X*MN*XX*MW which are characteristic for the HFCD proteins are involved in binding the cofactor (and the substrate). EpiD shares the Rossmann-type fold with flavodoxin-like proteins, although there is no sequence similarity. Superposition of flavodoxin and EpiD showed that the FMN cofactors bind to topologically similar positions. However, the ribityl moieties have different orientations and the FMN molecules expose different sides. The substrate-binding clamp of EpiD comprises residues Pro143–Met162 and forms a three-stranded β-sheet with the substrate peptide. The C-terminal cysteine residue of the substrate is fixed in the vicinity of the isoalloxazine ring of the FMN cofactor and its carboxylate group is hydrogen-bonded to Ser152 and Asn117 of the P*X*MN*XX*MW motif. Based on the geometry of the EpiD–substrate complex, we suggested that the side chain of the substrate cysteine residue is oxidized to a thioaldehyde structure. This thioaldehyde intermediate spontaneously decarboxylates, forming the enethiolate structure of the reaction product (Blaesse *et al.*, 2000).

Here, we present the structure of MrsD, the first characterized FAD-containing HFCD protein, and compare the known structures of HFCD proteins with respect to binding of the cofactor and the substrate. Modelling of an MrsD–substrate peptide complex explains the different substrate specificities of EpiD and MrsD. In total, the crystal structures of three HFCD proteins are now known, giving insight into the relationship between the function, sequence and structure of this new protein family. The work published here and in related papers shows how the side chain of cysteine residues is converted to a thioaldehyde structure, establishing new biochemistry of the thiol group.

2. Experimental

2.1. Expression, purification and crystallization

The MrsD protein was overexpressed in recombinant *Escherichia coli* M15 pREP4 harbouring the *mrsD* gene in the plasmid pQE12 as decribed previously (Majer *et al.*, 2002). Briefly, cells were grown at 310 K in 5 l Luria–Bertani medium to an OD_{600nm} of 0.6. Protein expression was induced by addition of 1 m*M* isopropyl-β-D-thiogalactopyranoside (IPTG). After 5 h, the cells were harvested by centrifugation,

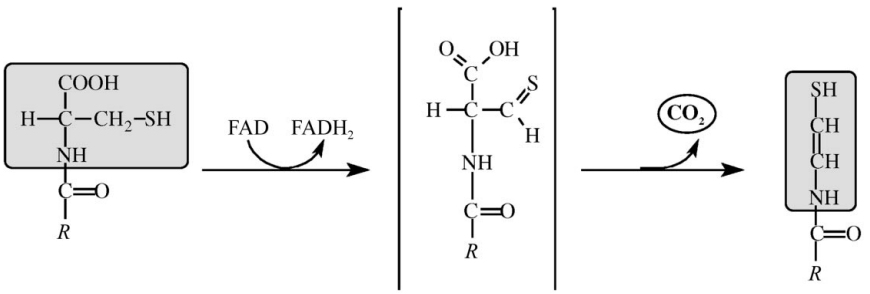

Figure 1
Reaction performed by MrsD. The C-terminal cysteine residue of the precursor peptide MrsA is oxidized by FAD to form a putative thioaldehyde intermediate that decarboxylates spontaneously. The resulting enethiol structure forms a thioether by adding to a double bond generated by dehydration of a threonine residue, which results in the unusual amino acid *S*-[(*Z*)-2-aminovinyl]-3-methyl-D-cysteine.

resuspended in 100 ml 10 m*M* Tris–HCl pH 8.0 and lysed by sonification.

After removal of the insoluble cell debris, the clear supernatant was applied to a Q-Sepharose anion-exchange column (100 ml) equilibrated with 20 m*M* NaCl in 10 m*M* Tris–HCl pH 8.0. After washing with the same buffer, MrsD was eluted with 60 m*M* NaCl in 10 m*M* Tris–HCl pH 8.0. The pooled protein was applied to a hydoxyapatite column (50 ml, MacroPrep ceramic hydroxyapatite 40 µm; BioRad) equilibrated with 20 m*M* potassium phosphate in 20 m*M* Tris–HCl, 150 m*M* NaCl pH 8.0, washed with the equilibration buffer and eluted in a linear gradient to 500 m*M* potassium phosphate in 20 m*M* Tris–HCl, 150 m*M* NaCl pH 8.0. The pooled fractions were dialysed against 20 m*M* Tris–HCl, 150 m*M* NaCl pH 8.0 and concentrated to 20 mg ml^{-1} by ultrafiltration. Aliquots of 0.3–0.5 ml were loaded onto a Superose 12 column (Pharmacia, HR 30/10) and eluted at a rate of 0.3–0.4 ml min^{-1}. The pooled fractions were dialysed against 50 m*M* NaCl in 20 m*M* Tris–HCl pH 8.0 and concentrated to a final concentration of 10 mg ml^{-1}.

Initial screening for crystallization conditions was performed with factorial solutions (Hampton Research Crystal Screens I, II and Cryo), which yielded two conditions for similar cubic crystal forms. Condition I (1.6 *M* $KNaHPO_4$ in 0.1 *M* HEPES pH 7.5) yielded crystals with a rather low FAD occupancy, presumably because the phosphate ions were competing with the FAD phosphates. Crystals from condition II (1.2 *M* potassium/sodium tartrate, 0.1 *M* Tris–HCl pH 9.1) could be grown in the presence of 2 m*M* FAD and showed a significantly higher FAD occupancy. Therefore, a crystal from the tartrate condition with space group *F*432 and unit-cell parameters $a = b = c = 191.8$ Å was used for structure determination.

2.2. Data collection, structure solution and refinement

For data collection under cryogenic conditions, the crystals from condition II were transferred to a buffer additionally containing 30%(*v*/*v*) glycerol. Data were collected at beamline BW6 at DESY in Hamburg with a MAR Research CCD detector from a crystal cooled to 100 K with an Oxford Cryostream. Data were integrated, scaled and merged with the *HKL* suite (Otwinowski & Minor, 1997). The crystals contained one molecule in the asymmetric unit and had a solvent content of about 67%. Patterson search methods with the EpiD protomer were unsuccessful despite its high structural similarity, as is sometimes encountered in high-symmetry space groups. The cubic space group *F*432 displays special positions of both octahedral and tetrahedral symmetry. As a tetrahedral particle was expected for MrsD, similar to EpiD, the EpiD dodecamer was oriented according to the crystallographic symmetry operators and its centre of mass translated to a special position with tetrahedral symmetry.

The polyalanine model derived from this manually placed EpiD monomer refined readily and could be completed by iterative rounds of model building with *MAIN* (Turk, 1992) and crystallographic refinement with *CNS* (Brünger *et al.*, 1998) using a test set of 5% of the reflections for cross-validation using the maximum-likelihood target using amplitudes. The structure displays good stereochemical parameters as estimated by the program *PROCHECK* (Laskowski *et al.*, 1993). 91% of the residues were found in the most favoured region of the Ramachandran plot, 9% in additional allowed regions and no residues in generously allowed or disallowed regions. Figures were prepared with *BOBSCRIPT* (Esnouf, 1997).

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

Data collection	
Unit-cell parameter (Å)	a = 191.8
Limiting resolution (Å)	2.54 (2.54-2.68)
Reflections	10384
$R_{\rm merge}$† (%)	4.8 (32.5)
$\langle I\rangle/\sigma(I)$	24.2 (5.4)
Redundancy	6.7 (6.8)
Completeness (%)	99.4 (99.5)
Refinement	
Resolution range (Å)	20–2.54
Reflections (working set)	9815 [93.8%]
Reflections (test set)	529 [5.1%]
$R_{\rm cryst}$‡ (%)	22.2
$R_{\rm free}$§ (%)	25.7
Non-H protein atoms	1345
FAD atoms	53
Solvent molecules	69
R.m.s.d. bond length (Å)	0.0105
R.m.s.d. bond angle (°)	1.19

† $R_{\rm merge} = \sum_{hkl}[(\sum_i |I_i - \langle I\rangle|)/\sum_i I_i]$. ‡ $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - k|F_{\rm calc}||/\sum_{hkl} |F_{\rm obs}|$ for hkl in the working set. § $R_{\rm free} = \sum ||F_{\rm obs} - k|F_{\rm calc}||/\sum |F_{\rm obs}|$ for hkl in the test set.

3. Results and discussion

3.1. Overall structure and oligomeric assembly

MrsD has been crystallized in the cubic space group *F*432 with one monomer in the asymmetric unit. As the dodecameric oligomer in these crystals is generated by crystallographic symmetry, the structure could be solved by placing the EpiD monomer based on geometric considerations. The crystal structure of MrsD has been refined at 2.54 Å resolution to a crystallographic R factor of 22.2% ($R_{\rm free}$ = 25.7%) (Table 1). The model consists of residues Ile3–Met155 and Arg168–Lys185. The disordered residues Ala156–Asn167 are part of the putative substrate-binding clamp that embraces the substrate in an extended conformation. Complete disorder of the binding clamp in the absence of substrate has also been observed for the HFCD proteins EpiD (Blaesse *et al.*, 2000) and AtHAL3a (Albert *et al.*, 2000; Steinbacher *et al.*, 2003). The oligomeric assembly of MrsD (Fig. 2) closely resembles that of EpiD, which has been described in detail (Blaesse *et al.*, 2000).

The monomer structure of MrsD is composed of classical mononucleotide-binding folds first observed in NADH-binding proteins (Rossmann *et al.*, 1974). It shares this fold with the FMN-binding homologues EpiD (Blaesse *et al.*, 2000) and AtHAL3a (Albert *et al.*, 2000; Steinbacher *et al.*, 2003)

from the HFCD family. Their structures are composed of six parallel β-strands interspersed by α-helices that appear on both sides of the central β-sheet. The monomer shows internal twofold topological symmetry and is built up from two halves known as Rossmann folds, $\beta1\alpha1\beta2\alpha2\beta3$ and $\beta4\alpha4\beta5\alpha5\beta6$, with a crossover α-helix ($\alpha3$) connecting $\beta3$ and $\beta4$. The α-helix $\alpha4$ is split into two characteristic helices that are involved in oligomerizaton in the HFCD family. Variations of this fold are found in FAD-binding proteins with respect to the second Rossmann fold or the crossover α-helix as well as in proteins that bind FMN or NADPH (Rao & Rossmann, 1973; Schulz & Schirmer, 1974; Dym & Eisenberg, 2001). In all cases the cofactor is bound at the C-terminal end of the central β-sheet.

All three known HFCD structures can be superimposed very well despite the rather low sequence homology of only 30.4% over 158 amino-acid residues between MrsD and EpiD and of 27.9% for 154 amino-acid residues between MrsD and AtHAL3a (Fig. 3). The substrate-binding clamps of the LanD proteins EpiD, MrsD (Pro151–Leu170) and MutD are four residues longer than that of the *P*PC decarboxylases AtHAL3a and Dfp (CoaC) (Blaesse *et al.*, 2000). Deviations are also observed for the irregular α-helix $\alpha2$ from Leu49 to Asp60 of MrsD. A major difference occurs around the loop from Glu65 to His73 in MrsD, which is involved in a dimer contact between trimers in dodecameric assemblies. This loop is longer in AtHAL3a to substitute for this dimer contact (Albert *et al.*, 2000; Steinbacher *et al.*, 2003). In MrsD, Glu65 and His66 of this loop contribute to the binding of the adenine moiety of FAD. In addition, the C-terminal helix that anchors the substrate-binding loop to the protein core shows certain variability. It is not surprising that the central β-sheets and adjacent isoalloxazine-binding loops show the highest degree of structural conservation displayed by the characteristic signature sequences PASANT ($\beta4$–$\alpha4$ loop) and P*X*MN*XX*MW ($\beta5$–$\alpha5$ loop) of the HFCD proteins.

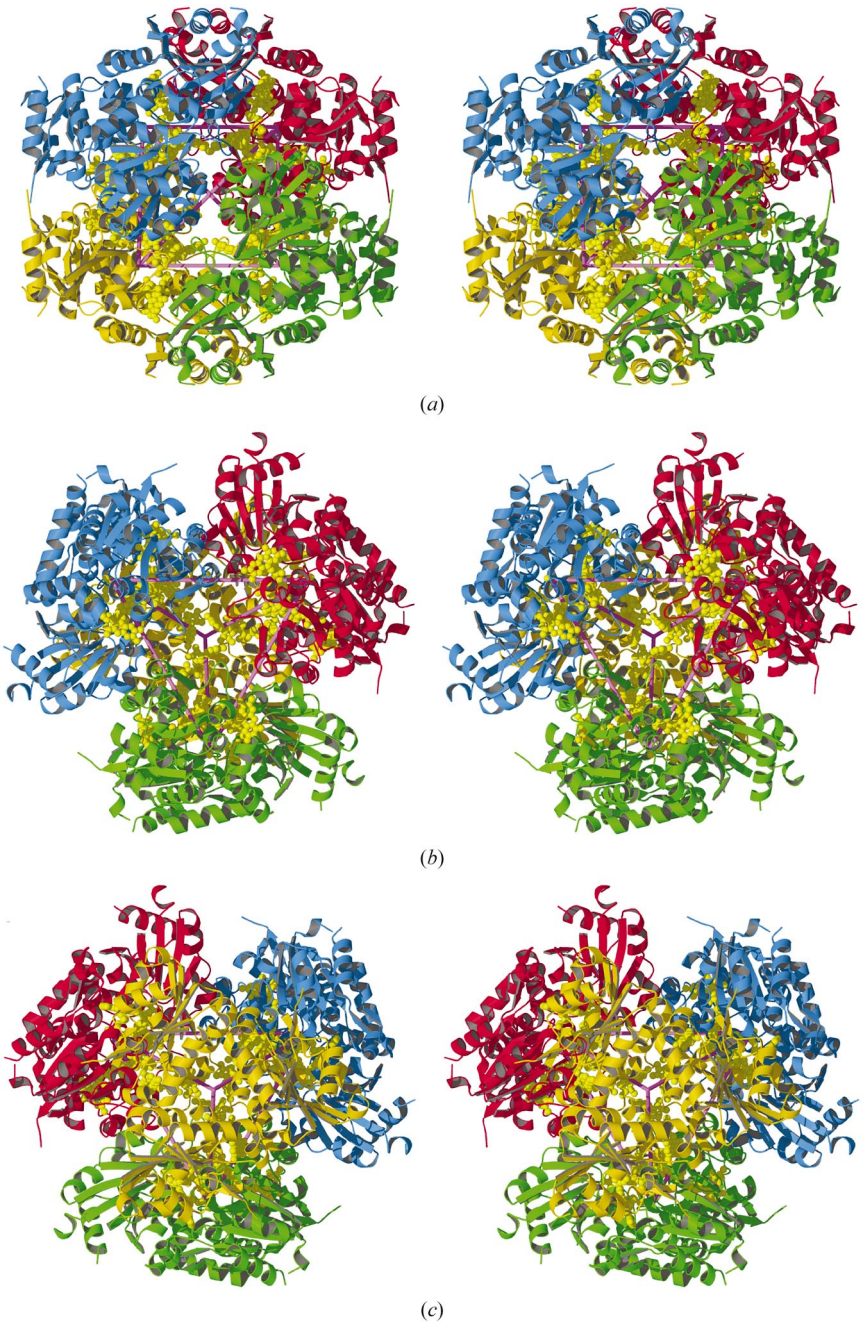

Figure 2
Three-dimensional structure of MrsD. Dodecamer structure viewed along the twofold (*a*) and both directions of the polar threefold (*c* and *d*) axes. The centres of gravity of each trimer are connected by pink rods to indicate the tetrahedral symmetry.

3.2. FAD binding site

The FAD cofactor of MrsD is completely defined by electron density (Fig. 4). It is buried between two neighbouring subunits of a trimeric facet in such a way that the adenosyl moiety points into the central cavity of the oligomer (Fig. 5). Surprisingly, FAD is not more tightly bound to MrsD than FMN to EpiD and the flavin cofactor is even partially lost during purification of MrsD. A major difference in the isoalloxazine-binding sites is the weak support for the dimethyzlbenzene ring in MrsD by Leu49 ($C6-C^{\delta1}$, 4.3 Å; $C7M-C^{\delta2}$, 3.8 Å) and Thr45 ($C9-C^{\gamma}$, 4.4 Å) compared with Phe43 in EpiD, and Phe59 and Trp78 in AtHAL3a, respectively. In EpiD and AtHAL3a these residues create a rigid platform for the dimethylbenzene moiety. This results in slightly different positions of the dimethylbenzene moiety in MrsD which is compensated for by a different confor-

mation of the ribityl chain (Fig. 5*a*), but also results in partial disorder of the dimethylbenzene moiety in MrsD as seen in the electron density (Fig. 4). It should be noted that for entropic reasons partial disorder is not automatically related to lower affinity (Forman-Kay, 1999). On the other hand, it may be argued that the weaker contacts between the isoalloxazine moiety of FAD and MrsD are partially compensated by favourable contacts of the adenosylphosphate moiety in order to result in an overall sufficient binding of the flavin cofactor. Unfortunately, the various contributions cannot be dissected in a simple way (Gohlke & Klebe, 2002). The residues supporting the dimethylbenzene ring are implicated in influencing the redox potential of the flavin cofactor (Fraaije & Mattevi, 2000). No data are available on the redox potential of any HFCD protein; however, all representatives are able to oxidize a thiol group. It cannot be ruled out that modulation of the redox potential by altering the environment of the dimethylbenzene ring also contributes to preventing the reduction of the enethiol intermediate in LanD proteins.

The FMN phosphate-binding sites of all three available structures coincide almost exactly when superimposing the entire structures without special emphasis on the flavin-binding sites (Fig. 5*a*). It is well documented that phosphate binding in Rossmann-fold proteins is associated with conserved sequence motifs (Dym & Eisenberg, 2001), but in other proteins such as TIM-barrel proteins it is associated with the so-called standard phosphate-binding motif (SPB; Nagano *et al.*, 2002). In HFCD proteins the FMN phosphate is bound by the PASANT motif, which includes Thr91 and Asn93 in MrsD. The side chains of Thr91 and Asn93 bind to one phosphate oxygen (Fig. 5*a*). The adenosyl-phosphate of FAD is tightly bound to both the backbone carbonyl and the side-chain hydroxyl of Thr45. Interestingly, Ser39 of EpiD and Ser55 of AtHAL3a are present at the same position. The side-chain hydroxyl of Ser39 is shifted by 1.3 Å, whereas that of Ser55 is found at an almost identical position. A second well defined contact is formed to the side chain of Asn105, which is Asn97 in EpiD and Asn120 in AtHAL3a and is shifted by 1.3 and 1.6 Å, respectively. Therefore, no significant sequence difference can be seen between the adenosyl-phosphate-binding site of MrsD and the corresponding positions in EpiD and AtHAL3a. A major source of stabilization of FAD binding can be seen in the tight packing of three neighbouring FAD molecules along the symmetry axis of the molecule including hydrogen bonds between N7 (3.1 Å) and N6 (2.9 Å) of the adenosine and O2 of the neighbouring pyranose. In addition, Met104* (where * denotes residues from the neighbouring subunit) packs against the pyranose ring, whereas Lys44 packs against the adenine moiety. Met104* is unique to MrsD, as EpiD has Asp96* and AtHAL3a has Asp119* at this position. Lys44 of MrsD is also found in AtHAL3a (Lys54), whereas it is replaced by Pro38 in EpiD. However, in all cases a tight van der Waals packing with the adenosine moiety seems possible. A major difference in backbone position and sequence is seen around Glu65*–His66*. Glu65* contacts O3 of the pyranose and His66* is involved in a van der Waals contact (4.6 Å) to the adenine moiety. The sequences are Asp74*–Glu75* in EpiD and Glu59–Ile60* in AtHAL3a. In summary, the surprising finding that MrsD binds FAD and not FMN as do EpiD, AtHAL3a or Dfp cannot be explained by a single structural reason. Instead, it has to be assumed that a larger number of small contributions add up

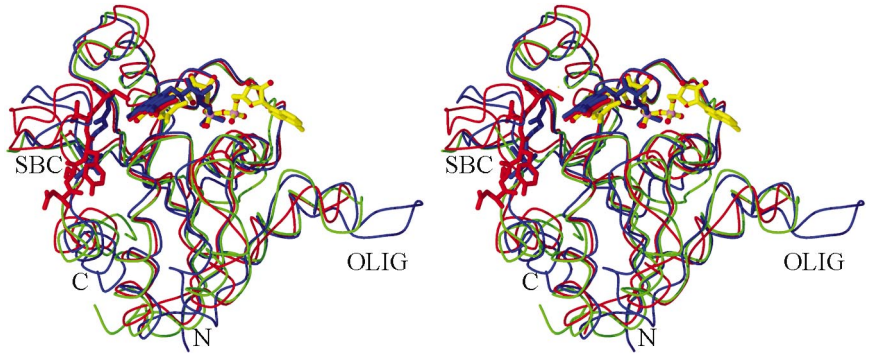

Figure 3
Superposition of HFCD monomers. MrsD is shown in green and FAD bound to it as a yellow ball-and-stick model. EpiD, the substrate pentapeptide and FMN are shown in red. AtHAL3a, the reaction intermediate pantothenoyl-aminoethenethiol bound to it and $FMNH_2$ are depicted in blue. SBC denotes the substrate-binding clamp and OLIG the loop in AtHAL3a that substitutes for a neighbouring subunit from another trimer in the dodecameric assemblies of EpiD and MrsD.

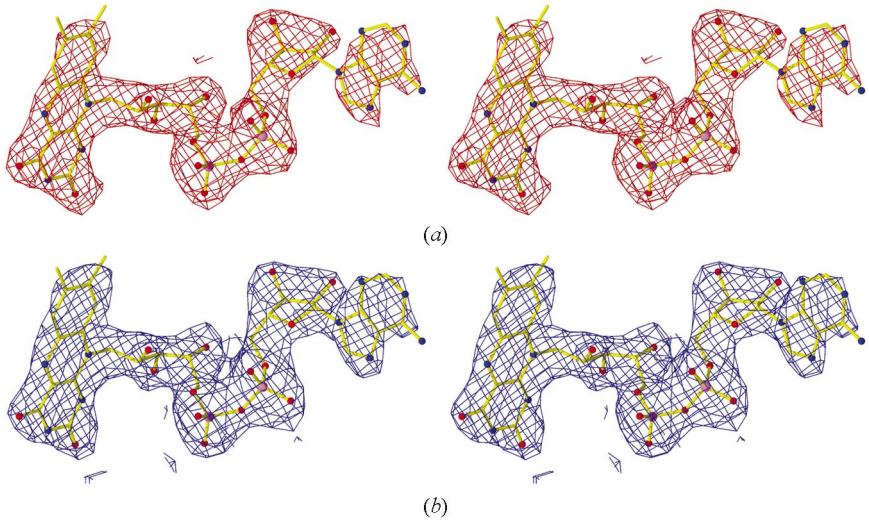

Figure 4
Electron density of FAD bound to MrsD at 2.54 Å resolution. (*a*) $F_o - F_c$ simulated-annealing omit electron density contoured at 2.5σ. (*b*) Final $2F_o - F_c$ electron density contoured at 1σ.

to a substantial increase in affinity for FAD compared with FMN.

3.3. Comparison with other FAD-binding proteins

A recent survey of FAD-binding proteins identified four family folds, which are exemplified by glutathione reductase (GR), ferredoxin reductase (FR), *p*-cresol methylhydroxylase (PCMH) and pyruvate oxidase (PO) (Dym & Eisenberg, 2001). This study also demonstrated variability in FAD binding, with no single protein 'pharmacophore' and the involvement of highly conserved sequence motifs in pyrophosphate binding that allow the identification of phosphate-binding proteins. Searching the PDB with the program *DALI* (Dietmann *et al.*, 2001) shows MrsD has closest similarity to the pyruvate oxidase family of the four FAD-binding families. The FAD-binding domain of pyruvate oxidase (PDB code 1pox; Muller *et al.*, 1994) can be superimposed onto MrsD with an r.m.s. deviation of 3.5 Å for 113 residues and with a *Z* score of 4.5 (Fig. 6). Interestingly, the FAD cofactor is bound with opposite directionality, *e.g.* the MrsD isoalloxazine ring binds approximately where the adenosine moiety is located in pyruvate oxidase. Both proteins share a phosphate-binding site with only 1.3 Å displacement of the phosphate atoms. However, this site is occupied by the flavin phosphate in MrsD, whereas in pyruvate oxidase the adenosyl-phosphate is present. Common properties of this conserved phosphate-binding site include the backbone amide of Gly17 (MrsD) and Ile221 (PO), and the side chains of Thr43 (MrsD) and Thr244 (PO). Lower similarities are detected for ferredoxin reductase (*Z* score 3.9, r.m.s.d. 3.0 Å for 88 residues) and glutathione reductase (*Z* score 3.0, r.m.s.d. 4.1 Å for 100 residues). As can be expected from the utterly different folding topology, no representatives of the PCMH family showed up in the search. This lower similarity to FAD-binding Rossmann-fold proteins is not surprising, as the HFCD family shows a higher structural similarity to FMN-binding proteins such as flavodoxin (Blaesse *et al.*, 2000).

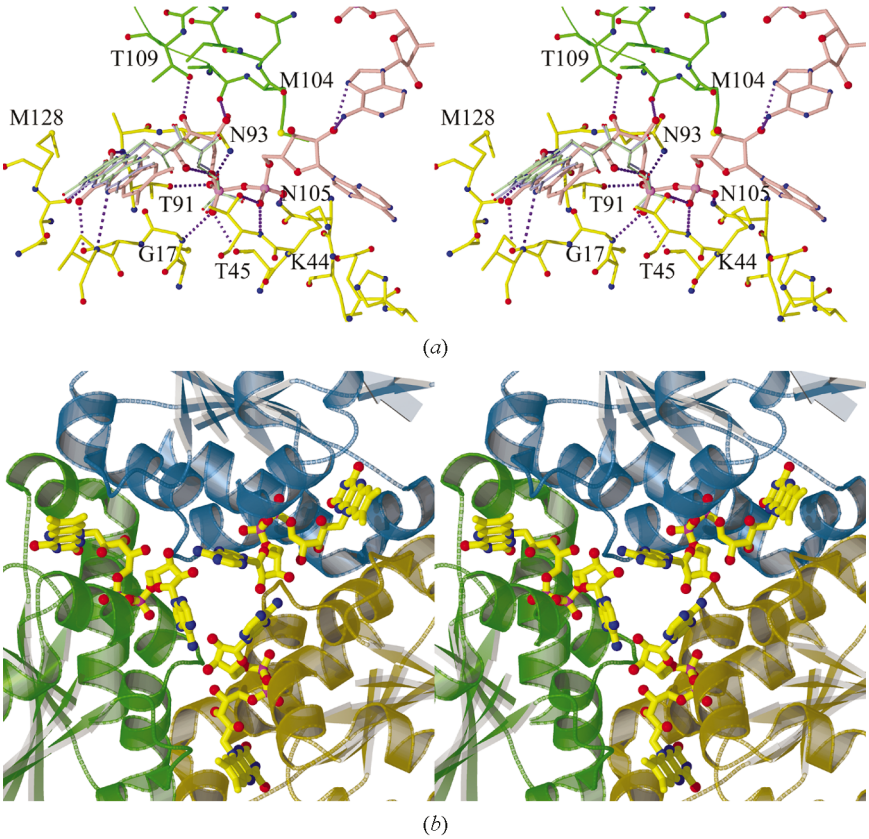

Figure 5
Interactions between MrsD and FAD. (*a*) FAD is shown in pink. Thr91 and Asn93 are part of the conserved PASANT motif; Met128 belongs to the P*X*MN*XX*MW motif. FMN bound to EpiD is in light blue and AtHAL3a is in turquoise. The orientation is identical to Fig. 3. (*b*) View from the centre of the dodecamer along the threefold axis.

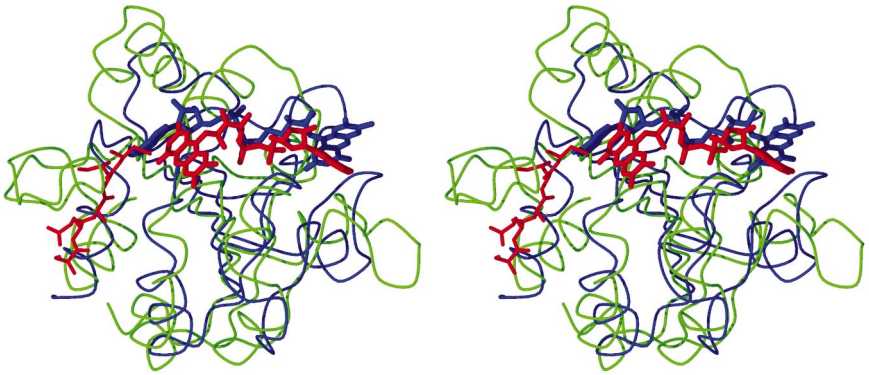

Figure 6
Superposition of MrsD and the FAD-binding domain of pyruvate oxidase (residues 213–340). The FAD cofactors are bound with opposite directionality in extended conformations. Pyruvate oxidase in shown in blue and MrsD in green, including the modelled substrate-binding clamp and the substrate peptide (thin ball-and-stick model in red). The MrsD FAD is shown in red and the pyruvate oxidase FAD in blue. Both structures share a conserved phosphate-binding site.

3.4. Substrate binding

HFCD proteins act on terminal cysteine residues with free carboxylate groups. Two reactions have been characterized so far. Either a decarboxylation without overall net change of the redox state occurs (*P*PC decarboxylases; Kupke *et al.*, 2000*a*, 2001) or the substrate is oxidatively decarboxylated, resulting in an enethiol group (LanD proteins; Kempter *et al.*, 1996; Kupke *et al.*, 1994). In both cases, the reaction is likely to start with the oxidation of the thiol group to a thioaldehyde which spontaneously decarboxylates to the enethiolate structure. This enethiol intermediate is then reduced

by the *P*PC decarboxylases but not the LanD enzymes, reoxidizing the flavin cofactor which was reduced in the first step (Hernández-Acosta *et al.*, 2002). The ability of *P*PC decaboxylases to reduce the enethiol intermediate is ascribed to the precise positioning of the intermediate with respect to the enzyme and the presence of a cysteine residue that acts as a proton source during the reduction (Steinbacher *et al.*, 2003). The LanD proteins EpiD, MrsD and MutD share the peptidyl-cysteine nature of their substrates and have highly homologous substrate-binding clamps of identical lengths. Therefore, it is reasonable to assume that substrate recognition is also quite similar, which allows modeling of the MrsD complex based on the EpiD–substrate complex. The substrate specificity of EpiD has been probed extensively (Kupke *et al.*, 1994, 1995; Majer *et al.*, 2002; Schmid *et al.*, 2002), yielding a consensus sequence [(Cys)/Val/Ile/Leu/(Met)/Phe/Tyr/Trp]-[Ala/Ser/Val/Thr/Cys/(Ile/Leu)]-Cys for the three C-terminal positions. EpiD and MrsD have different substrate specificities and the substrate peptides EpiA and MrsA were not interchangeable (Majer *et al.*, 2002). A major determinant of EpiD substrates is a large hydrophobic residue in the penultimate position. In the native substrate EpiA a tyrosine residue is present at this position, whereas in MrsA this residue is cysteine. It was recently shown that, in principle, EpiD also accepts a cysteine residue at this position, but the substrate peptide SFNSCCC was only very inefficiently oxidatively decarboxylated (Schmid *et al.*, 2002).

Based on the EpiD–substrate complex (Blaesse *et al.*, 2000), a model for MrsD substrate recognition can be obtained (Fig. 7). The C-terminus of the MrsD substrate MrsA has the sequence Ser-Glu-Cys-Ile-Cys. Compared with the peptide used in the EpiD complex (Asp-Ser-Tyr-Thr-Cys), a major difference is the requirement for a large hydrophobic residue (Tyr) at the −3 position of the EpiD substrate that is replaced by Cys in MrsA. The residue at −2 is Ile in MrsA but is preferentially a small and/or hydrophilic residue in an EpiD substrate. The larger side chain of Ile in MrsA is compensated for by the exchange of Ser148 (EpiD) for Ala156 in the substrate-recognition clamp. However, Ile can also be accepted by EpiD at that position with lower efficiency. A major difference is the replacement of the large hydrophobic residue at position −3 in EpiA by the significantly smaller Cys in MrsA. The theoretical model predicts that this residue is recognized in a hydrophobic pocket formed by Ile19, Phe157 from the substrate-binding clamp and Phe58** of MrsD (** denotes residues related by a twofold axis from a neighbouring trimer). Phe157 is predicted to contact Leu29**. It is a striking feature of substrate recognition in HFCD proteins that this specificity pocket is not static and preformed, as often observed in proteases, but that this binding site is generated during the dynamic process of substrate binding. The specificity pocket of MrsD is smaller than that of EpiD owing to the replacement of Val23 (EpiD) by Leu29 (MrsD).

Both structures have an asparagine residue in common that is involved in orienting the C-terminal carboxylate group (Asn117 in EpiD and Asn125 in MrsD) and a hydrophobic residue (Ile68 in EpiD and Val76 in MrsD) that packs against the thiol methylene groups of the terminal cysteine residue. The *P*PC decarboxylase AtHAL3a complexed to the enethiol reaction intermediate also shows these active-site characteristics (Ile91 and Asn142; Steinbacher *et al.*, 2003).

3.5. Conclusions

MrsD represents the third structure of the HFCD family of flavoproteins to be described. The members of this novel family of flavoproteins share a trimeric building block composed of Rossmann-fold type monomers that bind their flavin cofactor at the interface of two subunits. In the case of the plant *P*PC decarboxylase AtHAL3a from *A. thaliana*, these trimers represent the functional oligomeric state. However, four of the trimers can further assemble to form dodecamers as observed for EpiD and MrsD. Similar dodecamers probably represent the core of the *E. coli* Dfp protein which is composed of bifunctional monomers. In Dfp, the *P*PC decarboxylase activity resides in the N-terminal domain (CoaC) homologous to HFCD proteins, whereas the unrelated C-terminal domain (CoaB) harbours the *P*PC synthase activity.

The binding of FAD to MrsD cannot be ascribed to a single structural feature when comparing it with the FMN-binding proteins EpiD and AtHAL3a. It is assumed to be the result of a larger number of rather small contributions. It is, however, an unusual feature of FAD binding to MrsD that the adenine moiety has no well defined binding site provided by the protein. Major contacts are formed by the tight packing of three adenosine moieties along the threefold axis of the trimeric facets of the dodecamer.

The HFCD family of flavoproteins can be seen as paradigm of variability within homologous proteins. From sequence

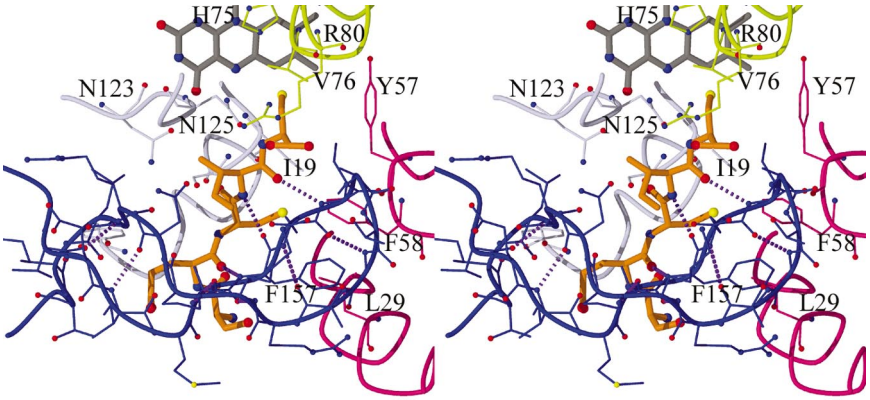

Figure 7
Theoretical model of the substrate complex. The peptide NH_2-Ser1-Glu2-Cys3-Ile4-Cys5-COOH which represents the COOH-terminal pentapeptide of MrsA and the substrate-binding clamp (dark blue) were modelled as determined in the EpiD pentapeptide complex (Blaesse *et al.*, 2000). The specificity pocket for the cysteine residue Cys3 is formed by Phe58, Ile19 and Phe157. Asn125 and Val76 are predicted to orient the C-terminal Cys5 residue of the substrate.

information alone, neither the substrate nor the actually bound flavin cofactor, FMN or FAD, can be predicted unambiguously. Perhaps the unusual binding mode of FAD facilitated by intermolecular interactions between cofactors in neighbouring molecules contributes to this uncertainty. The biochemical function and properties have to be tested and verified experimentally in every single case. For example, the function of the SIS2 protein of *Saccharomyces cerevisiae* (Ferrando *et al.*, 1995) with yet another composition of the substrate-binding clamp (Blaesse *et al.*, 2000) is still unknown.

The *P*PC decarboxylase reaction essential for coenzyme A biosynthesis is likely to be the original function of the HFCD family. Changes within the substrate-binding clamp of the *P*PC decarboxylases resulted in loss of the second half-reaction and enabled the oxidative decarboxylation of peptidyl-cysteines by LanD enzymes, used by some lantibiotic-synthesizing systems. The resulting enethiol structure retained in LanD-catalyzed reactions and present in the unsaturated thioether rings of lantibioticts such as epidermin and mersacidin contribute to rigidity of these peptide antibiotics. Simultaneously, the terminal carboxylate group of these lantibiotics is removed, changing the net charge of these compounds. At the moment it is not known if and how these properties (the unsaturated thioether ring and the missing carboxylate group) contribute to the mode of action of epidermin and mersacidin. It will be interesting to learn whether nature has found yet another application of the simple principle of oxidizing terminal cysteines in a flavin-dependent reaction to make use of the resulting chemistry.

The work was supported by Deutsche Forschungsgemeinschaft SFB 413 (StS) and Grant KU869/4 (TK). We thank Gabi Bierbaum for helpful discussion and Gleb Bourenkov for support with synchrotron data collection.

References

Albert, A., Martínez-Ripoll, M., Espinosa-Ruiz, A., Yenush, L., Culiáñez-Macià, F. A. & Serrano, R. (2000). *Structure*, **8**, 961–969.

Allgaier, H., Jung, G., Werner, R. G., Schneider, U. & Zähner, H. (1985). *Angew. Chem. Int. Ed. Engl.* **24**, 1051–1053.

Blaesse, M., Kupke, T., Huber, R. & Steinbacher, S. (2000). *EMBO J.* **19**, 6299–6310.

Brötz, H., Bierbaum, G., Reynolds, P. E. & Sahl, H.-G. (1997). *Eur. J. Biochem.* **246**, 193–199.

Brötz, H., Josten, M., Wiedmann, I., Schneider, U., Götz, F., Bierbaum, G. & Sahl, H.-G. (1998). *Mol. Microbiol.* **30**, 317–327.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Chatterjee, S., Chatterjee, D. K., Jani, R. H., Blumbach, J., Ganguli, B. N., Klesel, N., Limbert, M. & Seibert, G. (1992). *J. Antibiot.* **45**, 839–845.

Dietmann, S., Park, J., Notredame, C., Heger, A., Lappe, M. & Holm, L. (2001). *Nucleic Acids Res.* **29**, 55–57.

Dym, O. & Eisenberg, D. (2001). *Protein Sci.* **10**, 1712–1728.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.

Espinosa-Ruiz, A., Bellés, J. M., Serrano, R. & Culiáñez-Macià, F. A. (1999). *Plant J.* **20**, 529–539.

Ferrando, A., Kron, S. J., Rios, G., Fink, G. R. & Serrano, R. (1995). *Mol. Cell Biol.* **15**, 5470–5481.

Forman-Kay, J. D. (1999). *Nature Struct. Biol.* **6**, 1086–1087.

Fraaije, M. W. & Mattevi, A. (2000). *Trends Biochem. Sci.* **25**, 126–132.

Gohlke, H. & Klebe, G. (2002). *Angew. Chem. Int. Ed. Engl.* **41**, 2645–2676.

Hernández-Acosta, P., Schmid, D. G., Jung, G., Culiáñez-Macià, F. A. & Kupke, T. (2002). *J. Biol. Chem.* **277**, 20490–20498.

Kellner, R., Jung, G., Hörner, T., Zähner, H., Schnell, N., Entian, K.-D. & Götz, F. (1988). *Eur. J. Biochem.* **177**, 53–59.

Kempter, C., Kupke, T., Kaiser, D., Metzger, J. W. & Jung, G. (1996). *Angew. Chem. Int. Ed. Engl.* **35**, 2104–2107.

Kupke, T. (2001). *J. Biol. Chem.* **276**, 27597–27604.

Kupke, T. (2002*a*). *Flavins and Flavoproteins 2002*, edited by S. Chapman, R. Perham & N. Scrutton, pp. 951–962. Berlin: Rudolf Weber.

Kupke, T. (2002*b*). *J. Biol. Chem.* **277**, 36137–36145.

Kupke, T., Hernández-Acosta, P., Steinbacher, S. & Culiáñez-Macià, F. A. (2001). *J. Biol. Chem.* **276**, 19190–19196.

Kupke, T., Kempter, C., Gnau, V., Jung, G. & Götz, F. (1994). *J. Biol. Chem.* **269**, 5653–5659.

Kupke, T., Kempter, C., Jung, G. & Götz, F. (1995). *J. Biol. Chem.* **270**, 11282–11289.

Kupke, T., Stevanovic, S., Sahl, H.-G. & Götz, F. (1992). *J. Bacteriol.* **174**, 5354–5361.

Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M. & Steinbacher, S. (2000*a*). *J. Biol. Chem.* **275**, 31838–31846.

Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M. & Steinbacher, S. (2000*b*). *J. Biol. Chem.* **276**, 31838–31846.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Majer, F., Schmid, D. G., Altena, K., Bierbaum, G. & Kupke, T. (2002). *J. Bacteriol.* **184**, 1234–1243.

Minami, Y., Yoshida, K.-I., Azuma, R., Urakawa, A., Kawauchi, T., Otani, T., Komiyama, K. & Omura, S. (1994). *Tetrahedron Lett.* **35**, 8001–8004.

Mota-Meira, M., Lacroix, C., LaPointe, G. & Lavoie, M. C. (1997). *FEBS Lett.* **410**, 275–279.

Muller, Y. A., Schumacher, G., Rudolph, R. & Schulz, G. E. (1994). *J. Mol. Biol.* **237**, 315–335.

Nagano, N., Orengo, C. A. & Thornton, J. M. (2002). *J. Mol. Biol.* **321**, 741–765.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Qi, F., Chen, P. & Caufield, P. W. (1999). *Appl. Environ. Microbiol.* **65**, 3880–3887.

Qi, F., Chen, P. & Caufield, P. W. (2000). *Appl. Environ. Microbiol.* **66**, 3221–3229.

Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241–256.

Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). *Nature (London)*, **250**, 194–199.

Schmid, D. G., Majer, F., Kupke, T. & Jung, G. (2002). *Rapid Commun. Mass Spectrom.* **16**, 1779–1784.

Schnell, N., Entian, K.-D., Schneider, U., Götz, F., Zähner, H., Kellner, R. & Jung, G. (1988). *Nature (London)*, **333**, 276–278.

Schulz, G. E. & Schirmer, R. H. (1974). *Nature (London)*, **250**, 142–144.

Smith, L., Novak, J., Rocca, J., McClung, S., Hillman, J. D. & Edison, A. S. (2000). *Eur. J. Biochem.* **267**, 6810–6816.

Spitzer, E. D., Jimenez-Billine, H. E. & Weiss, B. (1988). *J. Bacteriol.* **170**, 872–876.

Spitzer, E. D. & Weiss, B. (1985). *J. Bacteriol.* **164**, 994–1003.

Steinbacher, S., Hernández-Acosta, P., Bieseler, B., Blaesse, M., Huber, R., Culiáñez-Macià, F. A. & Kupke, T. (2003). *J. Mol. Biol.* **327**, 193–202.

Strauss, E. & Begley, T. P. (2001). *J. Am. Chem. Soc.* **123**, 6449–6450.

Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W. & Begley, T. P. (2001). *J. Biol. Chem.* **276**, 13513–13516.

Turk, D. (1992). PhD thesis. TU München, München, Germany.