Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Tanja Sgraja, Lauris E. Kemp, Nicola Ramsden and William N. Hunter*

Division of Biological Chemistry and Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland

Correspondence e-mail: w.n.hunter@dundee.ac.uk

Received 26 April 2005
Accepted 13 June 2005
Online 30 June 2005

PDB Reference: 2*C*-methyl-D-erythritol-2,4-cyclodiphosphate synthase, 1yqn, r1yqnsf.

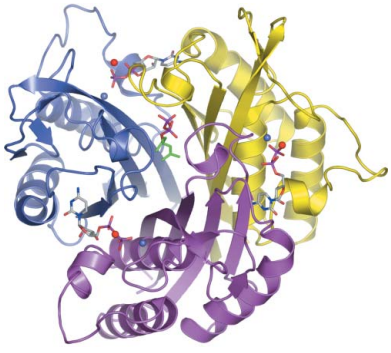
A double mutation of *Escherichia coli* 2*C*-methyl-D-erythritol-2,4-cyclodiphosphate synthase disrupts six hydrogen bonds with, yet fails to prevent binding of, an isoprenoid diphosphate

The essential enzyme 2*C*-methyl-D-erythritol-2,4-cyclodiphosphate (MECP) synthase, found in most eubacteria and the apicomplexan parasites, participates in isoprenoid-precursor biosynthesis and is a validated target for the development of broad-spectrum antimicrobial drugs. The structure and mechanism of the enzyme have been elucidated and the recent exciting finding that the enzyme actually binds diphosphate-containing isoprenoids at the interface formed by the three subunits that constitute the active protein suggests the possibility of feedback regulation of MECP synthase. To investigate such a possibility, a form of the enzyme was sought that did not bind these ligands but which would retain the quaternary structure necessary to create the active site. Two amino acids, Arg142 and Glu144, in *Escherichia coli* MECP synthase were identified as contributing to ligand binding. Glu144 interacts directly with Arg142 and positions the basic residue to form two hydrogen bonds with the terminal phosphate group of the isoprenoid diphosphate ligand. This association occurs at the trimer interface and three of these arginines interact with the ligand phosphate group. A dual mutation was designed (Arg142 to methionine and Glu144 to leucine) to disrupt the electrostatic attractions between the enzyme and the phosphate group to investigate whether an enzyme without isoprenoid diphosphate could be obtained. A low-resolution crystal structure of the mutated MECP synthase Met142/Leu144 revealed that geranyl diphosphate was retained despite the removal of six hydrogen bonds normally formed with the enzyme. This indicates that these two hydrophilic residues on the surface of the enzyme are not major determinants of isoprenoid binding at the trimer interface but rather that hydrophobic interactions between the hydrocarbon tail and the core of the enzyme trimer dominate ligand binding.

1. Introduction

The enzyme 2*C*-methyl-D-erythritol-2,4-cyclodiphosphate (MECP) synthase catalyzes the fifth stage in the biosynthesis of isoprenoid precursors by the 1-deoxy-D-xylulose-5-phosphate (DOXP) pathway. This biosynthetic route, sometimes also referred to as the 2*C*-methyl-D-erythritol-4-phosphate (MEP) pathway (Eisenreich *et al.*, 1998; Lichtenthaler, 1999), is present in many bacteria, the plastids of plants and in apicomplexan parasites. The genes encoding enzymes of the DOXP pathway have been shown to be essential in bacteria. Since mammals use a different biosynthetic route dependent on mevalonate, inhibitors of the DOXP enzymes might assist the development of novel broad-spectrum antimicrobial drugs. The biomedical implications of understanding the DOXP enzymes have in part been responsible for the interest in this area of research.

There are seven enzymes in the DOXP pathway, which catalyze the formation of the universal isoprenoid precursor isopentenyl diphosphate (IPP) and the isomer dimethylallyl diphosphate (DMAPP). The first step is the condensation of glyceraldehyde 3-phosphate and pyruvate by DOXP synthase (Lois *et al.*, 1998; Sprenger *et al.*, 1997), followed by conversion of DOXP to MEP by a reductoisomerase (Reuter *et al.*, 2002; Steinbacher *et al.*, 2003). Next, a cytidyltransferase produces 4-diphosphocytidyl-2*C*-methyl-D-erythritol (CDP-ME; Richard *et al.*, 2002). In an ATP-dependent reaction, CDP-ME kinase then phosphorylates the erythritol moiety (Miallau



et al., 2003; Wada *et al.*, 2003). MECP synthase then catalyzes the formation of the unusual cyclic diphosphate MECP with release of CMP (Fig. 1; Kemp *et al.*, 2002). MECP is finally converted to IPP in two further steps catalyzed by IspG and IspH proteins (Hecht *et al.*, 2001; Rohdich *et al.*, 2002, 2004). Some of the IPP produced is isomerized to DMAPP by an IPP isomerase (Eisenreich *et al.*, 2004), thereby providing the reagents necessary for chain elongation leading to more complex isoprenoids. The sequential additions of IPP to the allylic precursor are catalyzed by isoprenyl diphosphate synthases to give geranyl diphosphate (GPP), farnesyl diphosphate (FPP) and longer chain products (Kellogg & Poulter, 1997; Leyes *et al.*, 1999). These compounds then serve as precursors for the biosynthesis of more complex natural products including sterols, dolichols, triterpenes, ubiquinones and plastoquinones (Sacchettini & Poulter, 1997).

Recent studies have revealed that isoprenoids bind to MECP synthase, raising the possibility of a feedback mechanism to regulate this important biosynthetic pathway (Kemp *et al.*, 2005). Electrospray mass spectrometry indicated that recombinant *Escherichia coli* MECP synthase binds a mixture of IPP, GPP and FPP in an approximate ratio of 1:4:2 and it is assumed that these ligands are acquired as the protein is produced and are retained during purification. Crystal structures of MECP synthase in complex with GPP or FPP have been deposited in the PDB (Gabrielsen *et al.*, 2004; Kemp *et al.*, 2005; Ni *et al.*, 2004) and reveal the isoprenoid diphosphate to be placed at an interface created by the three subunits. The terminal phosphate group is positioned by six hydrogen bonds with three Arg142 residues, each held in place by Glu144 on the surface of the trimeric prism. The second phosphate group participates in three hydrogen bonds with amide groups of three Phe139 residues, whilst the tail of the isoprenoid ligands extends down into a hydrophobic cavity (Figs. 2 and 3*a*).

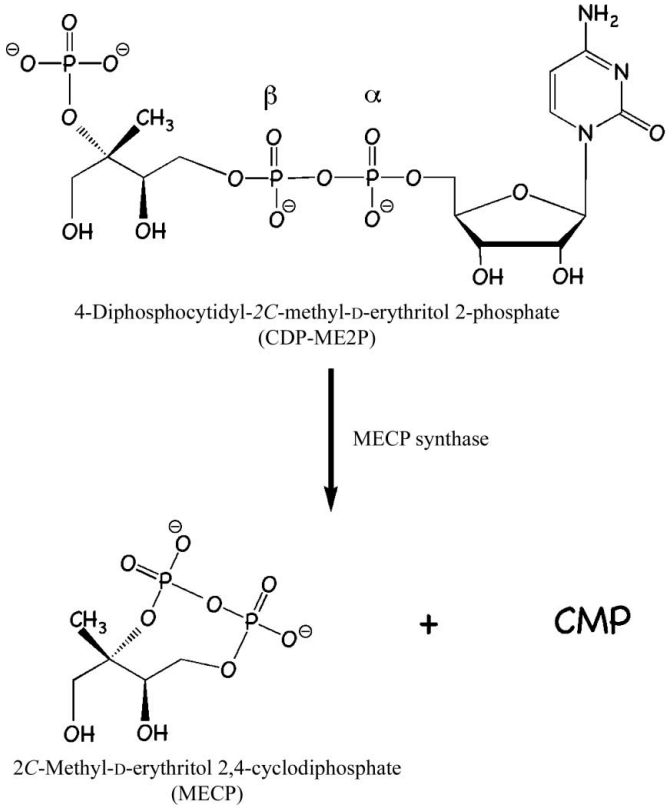


Figure 1
The reaction catalyzed by MECP synthase.

To investigate the potential contributions that isoprenoid diphosphate binding might make to MECP synthase activity, we sought to prepare a sample with a conserved structure but for which isoprenoid diphosphate binding was disrupted. For this purpose, an *E. coli* MECP synthase double mutant was generated where Arg142 and Glu144 are replaced with methionine and leucine, respectively, and the crystal structure determined to examine the consequences of such mutations.

2. Material and methods

2.1. Mutagenesis and crystallization

The *E. coli* MECP synthase double mutant was obtained *via* the polymerase chain reaction (PCR). In the first reaction, the *E. coli* wild-type gene (*ispF*) already inserted into the pET15b vector (Novagen) provided the template and was amplified with forward (GGATCCATGCGAATTGGACACGGTTTTGACGTACATGCC) and reverse (TTCACAGGCAATCCC**TAG**CCC**CAT**TCCGGTAA-TCCGGTAAATCCCAG) primers which introduced mutations into the gene such that the encoded protein carried methionine and leucine residues at positions 142 and 144 instead of arginine and glutamic acid. The triplets that produce the changes are shown in bold. This PCR product only covered part of the gene; therefore, a second amplification with the extended reverse primer (GGA-TCCTCATTTTGTTGCCTTAATGAGTAGCGCCACCGCTTCAC-AGGCAATCCC) was used to obtain the full-length gene. The mutated gene was cloned into the pCR blunt II TOPO vector (Invitrogen), isolated by digestion with the restriction enzymes *Bam*HI and *Xho*I and cloned into the pET15b expression vector. DNA sequencing confirmed the incorporation of the mutations. This plasmid was used to produce the mutated MECP synthase by established protocols (Kemp *et al.*, 2002).

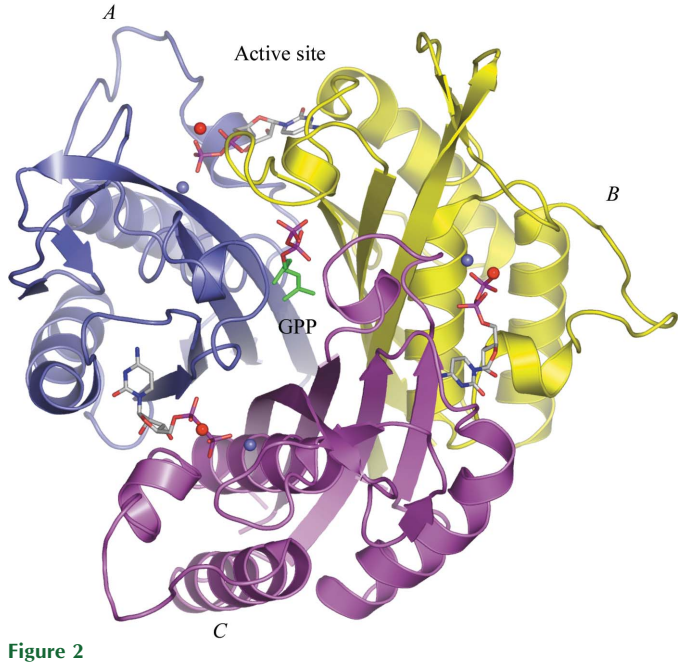


Figure 2
Ribbon representation of the functional MECP synthase trimer with the single subunits colored yellow, magenta and blue. The CDP molecules in stick representation (C positions are white, O red, N blue) bind in the active site formed at the interface of two subunits and are coordinated by Zn^{2+} (blue sphere) and Mn^{2+} ions (red sphere). The GPP molecule is also depicted in stick representation in the core of the homotrimer and colored C green, O red, P purple. Figs. 2, 3 and 4 were prepared with *PyMOL* (DeLano, 2002).

Table 1
Data-collection, refinement and model geometry statistics.

Values in parentheses are for the highest resolution shell.

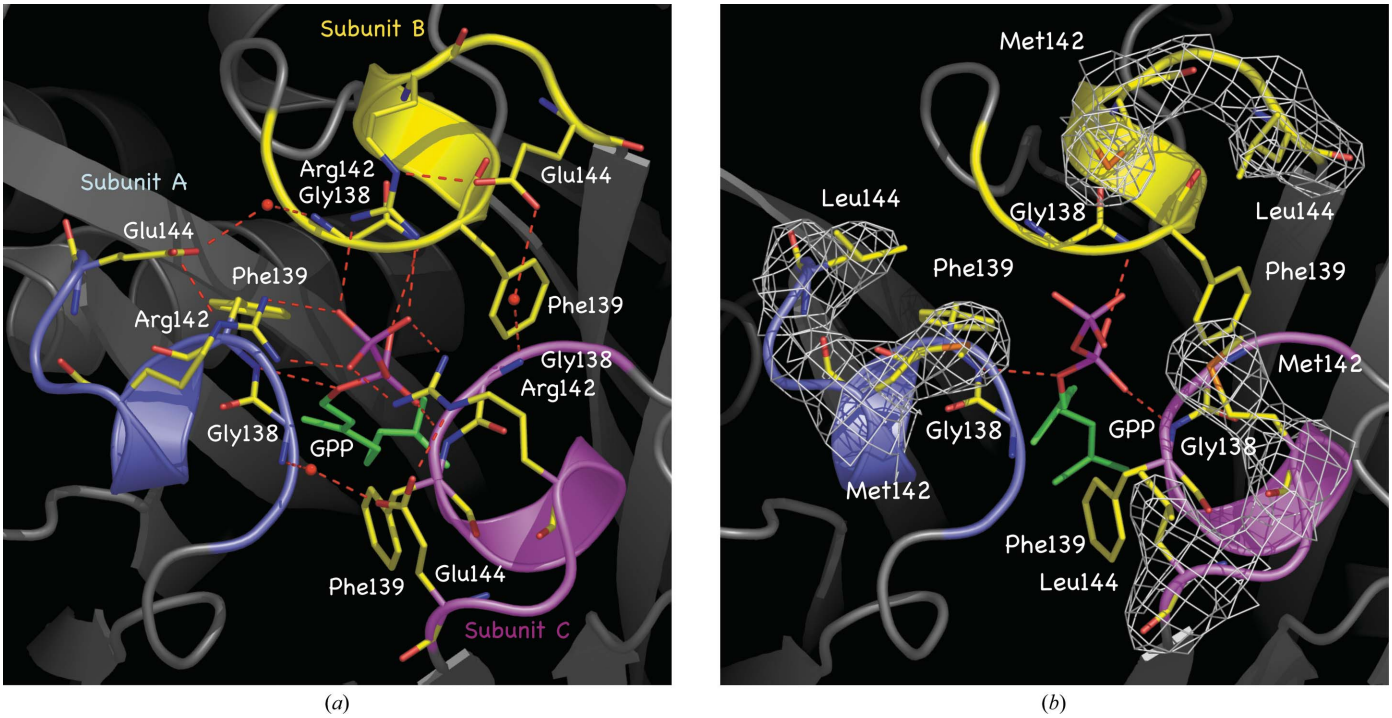
Space group	$I2_13$
Unit-cell parameter (Å)	143.8
Resolution (Å)	31.2–3.1
Wavelength (Å)	1.5418
No. of unique reflections	9138
Completeness (%)	100 (100)
Redundancy	7.3 (7.3)
$I/\sigma(I)$	22.6 (4.5)
R_{sym}† (%)	6.7 (42.3)
R_{work} (%)	18.8
R_{free} (%)	20.9
No. of protein atoms	1182
No. of water molecules	27
R.m.s. deviations from ideal values	
Bond lengths (Å)	0.007
Angles (°)	1.21
Mean B factors (Å^2)	
Protein atoms	73.6
Solvent atoms	67.0
Zn^{2+}	73.9
Mn^{2+}	80.3
CDP	72.0
GPP	63.6
Ramachandran analysis	
Residues in favoured regions (%)	87.6
Residues in additionally allowed regions (%)	12.4

† $R_{sym} = \sum |I - \langle I \rangle| / \sum I$, where I represents the measured intensity, $\langle I \rangle$ the averaged intensity and the summation is over all symmetry-equivalent reflections.

The protein was incubated with 2 m*M* manganese chloride and 2 m*M* cytidine diphosphate (CDP) for 2 h prior to crystallization by the hanging-drop vapor-diffusion method. Drops were assembled from 3 µl protein solution (in 50 m*M* Tris–HCl pH 7.7 and 50 m*M* NaCl) with a concentration of 5.5 mg ml^{-1} and 3 µl reservoir solution comprised of 0.1 *M* sodium acetate pH 5.0, 0.1 *M* ammonium sulfate and 18% polyethyleneglycol monomethylether molecular weight 2000. Cubic crystals with maximum dimensions of 0.4 mm were grown overnight at 289 K.

2.2. Data collection and refinement

A crystal was soaked in mother liquor containing 20% glycerol as cryoprotectant and frozen in a nitrogen-gas cryostream at 113 K. A data set was collected using a Rigaku RU-200 rotating-anode generator (Cu $K\alpha$) operating at 40 kV and 80 mA and an R-AXIS IV image-plate system. Data were processed with *MOSFLM* (Powell, 1999) and scaled in *SCALA* (Collaborative Computational Project, Number 4, 1994) to a resolution of 3.1 Å. The wild-type MECP synthase with PDB code 1gx1 provided the model for molecular replacement using *MOLREP* (Vagin & Teplyakov, 2000). Although the biological assembly of the MECP synthase is a trimer, the asymmetric unit contains only a monomer. The initial correlation coefficient and R factor were 75.1 and 36.3%, respectively. The electron-density map was examined with *O* (Jones *et al.*, 1991). A CDP molecule, Mn^{2+} and Zn^{2+} ions were placed into the active site and the model was subjected to refinement in *REFMAC* (Murshudov *et al.*, 1997) after setting aside 5% of the data for the calculation of R_{free}. At this stage, the electron-density and difference-density maps clearly indicated that residues Arg142 and Glu144 should be altered to methionine and leucine (Fig. 3*b*) and that an isoprenoid diphosphate molecule was present at the interface of the three subunits. This could be modelled as GPP and was assigned an occupancy of 0.33 as the molecule is situated on the crystallographic threefold axis. The careful placement of 27 solvent molecules concluded the analysis. Relevant statistics are presented in Table 1.



(*a*) (*b*)

Figure 3
The entrance of the isoprenoid diphosphate-binding site in (*a*) the wild-type and (*b*) the mutated enzyme. GPP molecules are shown as in Fig. 2. Selected amino-acid residues are depicted in stick mode and colored C yellow, N blue, O red. Hydrogen-bonding interactions between the MECP synthase and GPP are displayed as red dashed lines and water molecules as red spheres. In (*b*) the white chicken wire represents an omit $(2F_o - F_c)$ electron-density map corresponding to the mutated residues and contoured at the 1.4σ level. F_o and F_c represent observed and calculated structure factors, respectively.

3. Results and discussion

3.1. Overall structure and the active site

MECP synthase is an α/β-protein with a subunit fold consisting of a four-stranded β-sheet with three α-helices on one side. In addition, one β-hairpin and a short β-strand following $\beta1$ as well as two 3_{10}-helices are attached to the overall structure. The functional unit of the enzymes is a trimer. The β-sheets of three subunits facing each other form a triangular prism in the core and are surrounded by helical structural elements (Fig. 2). The biological assembly of the MECP synthase carries three active sites formed at the interface of adjacent subunits.

The incorporation of two mutations has little influence on both the overall structure and the detail in the active site. An overlay of 156 C^{α} atoms of one subunit of PDB code 1gx1 with the subunit of the mutant protein gives a root-mean-square deviation of 0.5 Å. The mutated MECP synthase was co-crystallized with CDP in the presence of $MnCl_2$ and these ligands, together with the endogenous Zn^{2+}, are clearly defined in the active site. The binding mode of the ligands is highly similar to the conformation in the structures described previously (Kemp *et al.*, 2002, 2005). In brief, the cytidine moiety of CDP is bound in a hydrophobic pocket situated at the C-terminus of the parallel β-strands of one subunit and is held in place by hydrogen-bonding interactions with the main-chain atoms of Ala100, Lys104, Met105 and Leu106. The ribose hydroxyl groups interact with the side chain of Asp56 and amide of Gly58 provided by the partner subunit (not shown). The Mn^{2+} ion coordinates the α- and β-phosphate groups of CDP, the Glu135 carboxylate and three water molecules. A Zn^{2+} ion binds with tetrahedral geometry to the Asp8, His10 and His42 side chains and the β-phosphate of CDP. The amino-acid residues 63–71 adjacent to the Zn^{2+}-binding site are poorly ordered and exhibit high thermal parameters in excess of 100 $Å^2$. Crystal structures in complex with CDP-ME or the cyclic product MECP showed that the loop region is well defined only in the presence of the erythritol moiety (Richard *et al.*, 2002; Steinbacher *et al.*, 2002).

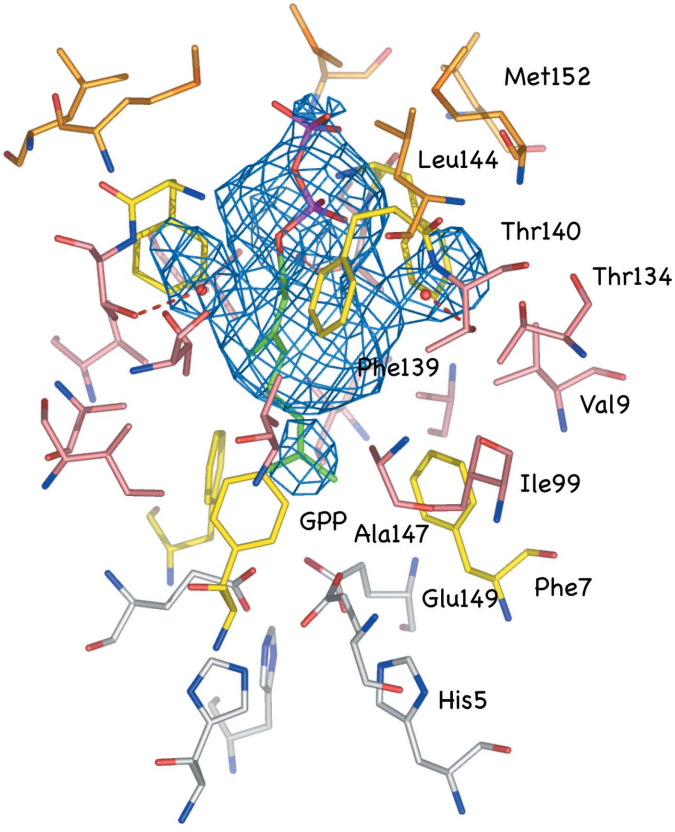


Figure 4
Stick representation of the GPP-binding site in the Met142/Leu144 mutant MECP synthase structure. An omit ($F_o - F_c$) difference density map (blue chicken wire) covering GPP and the adjacent water molecules, which are shown as red dots, is contoured at the 1.5σ level. The GPP-binding cavity is lined by Met142 and Leu144 (side-chain C positions colored orange), six phenylalanines (C yellow), aliphatic residues (C pink) and His5 and Glu149, which seal the pit bottom (C white).

3.2. The isoprenoid diphosphate-binding site

At the 'top' end (designated by the C-terminus of strands $\beta1$, $\beta4$ and $\beta5$) of the homotrimer structure, there is an opening into a hydrophobic pit approximately 15 Å in depth and it is here that isoprenoid diphosphates bind (Fig. 4). The pit bottom is sealed by the side chains of three interacting His5 and three Glu149 residues and the depth limits the length of the ligands that can bind. Several crystal structures in complex with GPP or FPP in a variety of different space groups have been reported (Gabrielsen *et al.*, 2004; Kemp *et al.*, 2002; Ni *et al.*, 2004). In all structures, the side chains of three arginine residues donate hydrogen bonds, six in total, to the terminal phosphate and position the isoprenoids at the entrance of the hydrophobic pit in the wild-type enzyme (Fig. 3*a*). Each arginine is fixed by association with a glutamic acid in turn held in place by a solvent-mediated contact to a main-chain amide.

The double mutation has little effect on the overall structure, as planned, and the hydrophobic substitutions are well defined in the electron density (Fig. 3*b*). The methionine and leucine side chains no longer serve as an anchor for the first phosphate moiety, yet the change has failed to provide a protein that is not loaded with isoprenoid diphosphate. The removal of the six hydrogen bonds that hold the terminal phosphate influence the thermal parameters in that the terminal P atom now has a *B* factor of 72.5 $Å^2$, which decreases to 66.5 $Å^2$ for the second P atom. In wild-type structures (*e.g.* Kemp *et al.*, 2005), the terminal P atom always has a lower *B* factor, by approximately 10 $Å^2$, compared with the second P atom, so the removal of the hydrogen-bonding interactions between the phosphate group and the enzyme has provided a greater degree of flexibility to the terminal phosphate group. Owing to experimental factors involved in the determination of different structures, we confine ourselves to this comment about the trend observed rather than a direct comparison of *B* factors.

The remaining interactions in the hydrophobic pocket of the MECP synthase Met142/Leu144 are highly conserved with those present in the wild-type enzyme. The second phosphate group accepts hydrogen bonds from three main-chain amides of residue Phe139. The failure to prevent isoprenoid diphosphate binding by introducing the double mutations can probably be explained by the strength of the hydrophobic associations that remain in place. The ligand-binding pit is lined by three sets of Phe7, Val9, Ile99, Phe139, Thr134, Thr140 and Ala147 residues provided by each subunit. A water molecule hydrogen bonds to the hydroxyl group of the Thr140 side chain and forms weak contacts to the lipid tail of the ligand. This hydration point was also observed in the *Shewanella oneidensis* enzyme (Ni *et al.*, 2004), in the bifunctional *Campylobacter jejuni* MEP/MECP synthase (Gabrielsen *et al.*, 2004) and in the *E. coli* MECP synthase in the monoclinic space group $P2_1$ (Kemp *et al.*, 2005).

There is a high conservation of sequence for residues surrounding the isoprenoid diphosphate ligands in MECP synthase (Kemp *et al.*, 2005). A notable exception is the enzyme from the thermophilic organism *Thermus thermophilus*, which does not bind an isoprenoid

diphosphate at the trimer core. The corresponding *E. coli* Arg142 and Glu144 residues are replaced by leucine and proline, respectively, in the *T. thermophilus* enzyme (Kishida *et al.*, 2003), similar to the mutations we engineered. The *T. thermophilus* enzyme, in striking contrast to other MECP synthases, carries bulky hydrophilic side chains at the trimer core and has no ligand-binding pocket.

4. Conclusion

Our first attempt to design a mutated *E. coli* MECP synthase, removing six hydrogen-bond interactions with a phosphate and introducing hydrophobic side chains in proximity of this hydrophilic group, rendering the protein incapable of binding isoprenoid diphosphate species at the trimer core, has failed. The loss of direct interactions between enzyme and ligand allows a greater degree of flexibility of the terminal phosphate group and this is reflected in an altered *B*-factor distribution in the ligand compared with the wild-type complex. It would appear, however, that the conservation of van der Waals associations between ligand and the residues that line the hydrophobic ligand-binding pit are the critical determinants of isoprenoid diphosphate binding. A consideration was that the double mutation should leave the enzyme structure and in particular the active site unperturbed and in this we were successful.

In order to occlude isoprenoid diphosphate binding to *E. coli* MECP synthase and to obtain a protein that would allow us to address the issue of what such ligands might contribute to enzyme activity will require mutations that fill the ligand-binding pit yet which retain the quaternary structure of the enzyme. This may not be simple but in the future it could be instructive to base such mutations on the structure of the *T. thermophilus* MECP synthase.

This work was supported by a fellowship of the German Academic Exchange Service (DAAD), the Wellcome Trust, the Biotechnology and Biological Sciences Research Council (UK) and Inpharmatica. We acknowledge support from the European Synchrotron Radiation Facility, Grenoble, France for the study of MECP synthase.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *PyMOL.* http://www.pymol.org.
Eisenreich, W., Bacher, A., Arigoni, D. & Rohdich, F. (2004). *Cell Mol. Life Sci.* **61**, 1401–1426.
Eisenreich, W., Schwarz, M., Cartayrade, A., Arigoni, D., Zenk, M. H. & Bacher, A. (1998). *Chem. Biol.* **5**, R221–R233.
Gabrielsen, M., Bond, C. S., Hallyburton, I., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F. & Hunter, W. N. (2004). *J. Biol. Chem.* **279**, 52753–52761.
Hecht, S., Eisenreich, W., Adam, P., Amslinger, S., Kis, K., Bacher, A., Arigoni, D. & Rohdich, F. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 14837–14842.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kellogg, B. A. & Poulter, C. D. (1997). *Curr. Opin. Chem. Biol.* **1**, 570–578.
Kemp, L. E., Alphey, M. S., Bond, C. S., Ferguson, M. A., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F. & Hunter, W. N. (2005). *Acta Cryst.* D**61**, 45–52.
Kemp, L. E., Bond, C. S. & Hunter, W. N. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 6591–6596.
Kishida, H., Wada, T., Unzai, S., Kuzuyama, T., Takagi, M., Terada, T., Shirouzu, M., Yokoyama, S., Tame, J. R. & Park, S. Y. (2003). *Acta Cryst.* D**59**, 23–31.
Leyes, A. E., Baker, J. A. & Poulter, C. D. (1999). *Org. Lett.* **1**, 1071–1073.
Lichtenthaler, H. K. (1999). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 47–65.
Lois, L. M., Campos, N., Putra, S. R., Danielsen, K., Rohmer, M. & Boronat, A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2105–2110.
Miallau, L., Alphey, M. S., Kemp, L. E., Leonard, G. A., McSweeney, S. M., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F. & Hunter, W. N. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 9173–9178.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Ni, S., Robinson, H., Marsing, G. C., Bussiere, D. E. & Kennedy, M. A. (2004). *Acta Cryst.* D**60**, 1949–1957.
Powell, H. R. (1999). *Acta Cryst.* D**55**, 1690–1695.
Reuter, K., Sanderbrand, S., Jomaa, H., Wiesner, J., Steinbrecher, I., Beck, E., Hintz, M., Klebe, G. & Stubbs, M. T. (2002). *J. Biol. Chem.* **277**, 5378–5384.
Richard, S. B., Ferrer, J. L., Bowman, M. E., Lillo, A. M., Tetzlaff, C. N., Cane, D. E. & Noel, J. P. (2002). *J. Biol. Chem.* **277**, 8667–8672.
Rohdich, F., Bacher, A. & Eisenreich, W. (2004). *Biorg. Chem.* **32**, 292–308.
Rohdich, F., Hecht, S., Gärtner, K., Adam, P., Krieger, C., Amslinger, S., Arigoni, D., Bacher, A. & Eisenreich, W. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 1158–1163.
Sacchettini, J. C. & Poulter, C. D. (1997). *Science*, **277**, 1788–1789.
Sprenger, G. A., Schorken, U., Wiegert, T., Grolle, S., De Graaf, A. A., Taylor, S. V., Begley, T. P., Bringer-Meyer, S. & Sahm, H. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 12857–12862.
Steinbacher, S., Kaiser, J., Eisenreich, W., Huber, R., Bacher, A. & Rohdich, F. (2003). *J. Biol. Chem.* **278**, 18401–18407.
Steinbacher, S., Kaiser, J., Wungsintaweekul, J., Hecht, S., Eisenreich, W., Gerhardt, S., Bacher, A. & Rohdich, F. (2002). *J. Mol. Biol.* **316**, 79–88.
Vagin, A. & Teplyakov, A. (2000). *Acta Cryst.* D**56**, 1622–1624.
Wada, T., Kuzuyama, T., Satoh, S., Kuramitsu, S., Yokoyama, S., Unzai, S., Tame, J. R. & Park, S. Y. (2003). *J. Biol. Chem.* **278**, 30022–30027.