Hexameric Assembly of the Bifunctional Methyerythritol 2,4-Cyclodiphosphate Synthase and Protein-Protein Associations in the Deoxy-xylulose-dependent Pathway of Isoprenoid Precursor Biosynthesis*

Received for publication, August 4, 2004, and in revised form, October 1, 2004
Published, JBC Papers in Press, October 2, 2004, DOI 10.1074/jbc.M408895200

Mads Gabrielsen‡, Charles S. Bond‡, Irene Hallyburton‡, Stefan Hecht§, Adelbert Bacher§, Wolfgang Eisenreich§, Felix Rohdich§, and William N. Hunter‡¶
From the ‡Division of Biological Chemistry and Molecular Microbiology, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, United Kingdom and §Institute für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

The bifunctional methyerythritol 4-phosphate cytidylyltransferase methyerythritol 2,4-cyclodiphosphate synthase (IspDF) is unusual in that it catalyzes non-consecutive reactions in the 1-deoxy-D-xylulose 5-phosphate (DOXP) pathway of isoprenoid precursor biosynthesis. The crystal structure of IspDF from the bacterial pathogen Campylobacter jejuni reveals an elongated hexamer with D3 symmetry compatible with the dimeric 2C-methyl-1-erythritol-4-phosphate cytidylyltransferase and trimeric 2C-methyl-3-erythritol-2,4-cyclodiphosphate synthase monofunctional enzymes. Complex formation of IspDF with 4-diphosphocytidyl-2C-methyl-3-erythritol kinase (IspE), the intervening enzyme activity in the pathway, has been observed in solution for the enzymes from C. jejuni and Agrobacterium tumefaciens. The monofunctional enzymes (2C-methyl-1-erythritol-4-phosphate cytidylyltransferase, IspE, and 2C-methyl-3-erythritol-2,4-cyclodiphosphate synthase) involved in the DOXP biosynthetic pathway of Escherichia coli also show physical associations. We propose that complex formation of the three enzymes at the core of the DOXP pathway can produce an assembly localizing 18 catalytic centers for the early stages of isoprenoid biosynthesis.

Life depends on isoprenoids such as sterols, dolichols, triterpenes, ubiquinone, and plastoquinone and components of macromolecules such as the prenyl groups of prenylated proteins and isopentenylated tRNAs (1, 2). This large family of natural molecules includes electron transport processes in respiration and photosynthesis, hormone-based signaling, the regulation of transcription, and posttranslational processes that control lipid biosynthesis, meiosis, apoptosis, protein cleavage, and degradation. In addition, isoprenoids fulfill an important role as a structural component of cell and organelle membranes.

Nature utilizes two distinct biosynthetic routes to produce isopentenyl diphosphate (IPP)1 and dimethylallyl diphosphate (DMAPP), the universal five-carbon precursors for isoprenoids. In mammals, the cytosol of plants, fungi, and some Gram-positive bacteria, the precursor biosynthesis is carried out by the mevalonate pathway (2–7). This begins with the conversion of acetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA followed by reduction, phosphorylation, and decarboxylation to generate IPP, some of which is isomerized to DMAPP. In chloroplasts, algae, cyanobacteria, eubacteria, and apicomplexa, a non-mevalonate-dependent route, the 1-deoxy-D-xylulose 5-phosphate (DOXP) or 2C-methyl-3-erythritol-4-phosphate pathway, produces IPP and DMAPP (2, 7–12).

The non-mevalonate pathway starts with the condensation of pyruvate and glyceraldehyde-3-phosphate to produce DOXP, which is converted to 2C-methyl-3-erythritol-4-phosphate in reactions catalyzed by DOXP synthase (13, 14) and DOXP reductoisomerase (IspC) (15–17), respectively. 2C-Methyl-3-erythritol-4-phosphate reacts with CTP to produce 4-diphosphocytidyl-2C-methyl-3-erythritol and pyrophosphate in a reaction catalyzed by 2C-methyl-3-erythritol-4-phosphate cytidylyltransferase (IsPD) (Fig. 1) (18–20). The ATP-dependent 4-diphosphocytidyl-2C-methyl-3-erythritol kinase (IspE) (21) phosphorylates 4-diphosphocytidyl-2C-methyl-3-erythritol, producing 4-diphosphocytidyl-2C-methyl-3-erythritol-2-phosphate. Next, the 4-diphosphocytidyl-2C-methyl-3-erythritol-2-phosphate is converted to 2C-methyl-3-erythritol-2,4-cyclodiphosphate and CMP by 2C-methyl-3-erythritol-2,4-cyclodiphosphate synthase (IspF) (Fig. 1) (22–24). The cyclic diphosphate product, 2C-methyl-3-erythritol-2,4-cyclodiphosphate, is reduced to 1-hydroxy-2-methyl-2(1E)-butenyl 4-diphosphate by a reductase encoded by the ispG gene (25–27), and then the ispH gene product converts the butenyl diphosphate to IPP and DMAPP (26, 28).

Enzymes of the DOXP pathway present attractive targets for the development of broad spectrum antimicrobial drugs targeting some of the world’s most serious diseases, including tuberculosis, malaria, and a range of sexually transmitted infections (6, 7, 29). These enzymes are absent from humans, provide the

* This work was supported by a Wellcome Trust Senior Research Fellowship (to W. N. H.), a Biotechnology and Biological Sciences Research Council Sir David Phillips Research Fellowship (to C. S. B.), the Deutsche Forschungsgemeinschaft, Fonds der Chemischen Industrie, and Hans-Fischer Gesellschaft (A. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The abbreviations used are: IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; DOXP, 1-deoxy-D-xylulose 5-phosphate; GPP, geranyl diphosphate; IspDF, methylerythritol 2,4-cyclodiphosphate synthase; IspD, 2C-methyl-3-erythritol-4-phosphate cytidylyltransferase; IspF, 2C-methyl-3-erythritol-2,4-cyclodiphosphate synthase; IspE, 4-diphosphocytidyl-2C-methyl-3-erythritol kinase.

2 1 The abbreviations used are: IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; DOXP, 1-deoxy-D-xylulose 5-phosphate; GPP, geranyl diphosphate; IspDF, methylerythritol 2,4-cyclodiphosphate synthase; IspD, 2C-methyl-3-erythritol-4-phosphate cytidylyltransferase; IspF, 2C-methyl-3-erythritol-2,4-cyclodiphosphate synthase; IspE, 4-diphosphocytidyl-2C-methyl-3-erythritol kinase.

This paper is available on line at http://www.jbc.org
pathogen's only source of IPP and DMAPP, and have been validated (i.e. proven essential for survival) as therapeutic targets by genetic approaches (29). Such genetic validation is proven essential for survival as therapeutic validated (29).}

To investigate these aspects of structure and organization in the non-mevalonate biosynthetic pathway, we have determined the crystal structure of IspDF from the highly prevalent foodborne pathogen Campylobacter jejuni (CjIspDF). In particular, with the use of analytical ultracentrifugation, we studied the possible association of CjIspDF with CjIspE, the possible association of Agrobacterium tumefaciens IspDF (AtIspDF) with AtIspE, and the possible association of monofunctional Escherichia coli enzymes (EcIspD, EcIspE, and EcIspF) with each other.

**EXPERIMENTAL PROCEDURES**

Sample Preparation—Recombinant CjIspDF, EcIspD, EcIspE, and EcIspF were prepared by established methods, and purity was checked by SDS-PAGE and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (33, 38, 39). The genes (ispE from C. jejuni and ispDF and ispE from A. tumefaciens) were obtained by PCR of genomic DNA using the oligonucleotide primers listed in Table I. The forward primers contained an NdeI restriction site, and the reverse primers contained a BamHI restriction site. The genes were cloned into pCR-Blunt II TOPO vector (Invitrogen) and then into pET15B expression vector (Novagen). Protein expression and purification again followed published protocols.

Crystallographic and Diffractive Data Collection of CjIspDF—Purified enzyme was dialyzed against 100 mM Tris-HCl (pH 7.6) and 50 mM NaCl and then concentrated to ~10 mg ml⁻¹. Hexagonal rods were obtained at 4 °C in 3 μl of hanging drops containing a 2:1 mixture of 10 mg ml⁻¹ protein solution (100 mM Tris, pH 7.6, 50 mM NaCl, 10 mM MgCl₂, and 2 mM CMP and CTP) with reservoir (35% (v/v) ethylene glycol). The crystals have space group P6₁22 with one monomer per asymmetric unit and a solvent content of ~62% (structure I). Isomorphous crystals, subsequently shown to contain Zn²⁺ in the IspF active site (structure II), were obtained from a separate protein preparation using 2 μl of hanging drops containing a 1:1 mixture of 12 mg ml⁻¹ protein solution (100 mM Tris-HCl, pH 7.6, 50 mM NaCl, 10 mM MgCl₂, and 7 mM CMP and CTP) with reservoir (35% (v/v) ethylene glycol). Crystals were harvested directly from the drops and then cooled in a nitrogen gas stream at ~170 °C for data collection. Diffraction data were measured with Q4 charge-coupled device detectors (Area Detector Systems Corp.) at the European Synchrotron Radiation Facility, processed, and scaled using the HKL program O (45), with inclusion of solvent positions, CMP, GPP, and an ethylene glycol molecule. The R-free and R-work were monitored as R-free and R-work were monitored as 23% and 38%. This model was refined using REFMAC (44) and the graphics program O (45), with inclusion of solvent positions, CMP, GPP, and an ethylene glycol molecule. The R-free and R-work were monitored as guides for refinement, and stereochemistry was assessed with PROCHECK (46). Structure II was derived from structure I by rigid body refinement and then completed using similar refinement protocols. Crystallographic statistics are presented in Table I. Surface areas were calculated using GRASP (47); molecular figures were generated.
Structure of IspDF and an Association with IspE

Table I

| Oligonucleotide primers | Forward | Reverse |
|------------------------|---------|---------|
| C. jejuni ispE         | CATATGGAGCTTACGCTAAGC | GGAATCCTGACTACGATACACTAAAC |
| A. tumefaciens ispDF   | CATATGCGAATAATGACGATGAAATTC | GGAATCCTGACTACGATACACTAAAC |
| A. tumefaciens ispE    | CATATGGCCCTCAGCAAGTGGTTCCC | GGAATCCTGACTACGATACACTAAAC |

Table II

| Crystallographic statistics | Structure | I | II |
|-----------------------------|-----------|---|---|
| Unit cell dimensions a= b = c (Å) | 108.8, 161.3 | 107.8, 161.0 |
| Resolution (Å)/wavelength (Å) | 2.3/0.9330 | 3.1/0.9761 |
| No. of measurements/unique reflections | 489,834/25,216 | 367,588/11,705 |
| Redundancy/completeness (%) | 19.4/98.4 (87.2) | 31.4/100 (100) |
| LoG(I) | 41.5 (5.3) | 41.7 (8.4) |
| R-merge (%) | 6.8 (36.2) | 8.8 (52.5) |
| Protein residues/atoms | 3690/2922 | 3680/908 |
| Water molecules | 276 | 86 |
| CMP/Zn²⁺/Mg²⁺/GPP/ethylene glycol | 2/1/1/1 | 2/1/1/1/1 |
| R-work (%)/no. of reflections | 17.6/23,826 | 21.9/10,027 |
| R-free (%)/no. of reflections | 23.0/1,278 | 29.0/508 |
| Wilson B (Å²) | 43.8 | 81.4 |
| Average isotropic thermal parameters (Å²) | 42.3/45.8 | 43.7/45.3 |
| Main chain/side chain atoms | 54.0 | 35.4 |
| Water molecules | 35.6/35.2 | 44.1/31.7 |
| r.m.s.d. for bond lengths (Å)/bond angles (°) | 0.011/1.25 | 0.011/1.42 |
| Ramachandran analysis | 91.1 | 86.0 |
| % Favored regions | 83.3 | 13.1 |
| % Additional allowed regions | 0.3 (1 residue) | 0.9 (3 residues) |
| % Generously allowed regions | | |
| % Disallowed regions | | |

* r.m.s.d., root mean square deviation.

using MOLSCRIPT (48) and Raster3D (49).

Size Exclusion Chromatography and Analytical Ultracentrifugation—Size exclusion chromatography was performed with samples applied to a 24-ml (1 × 30-cm) Superdex 200 HR 10/30 size exclusion column equilibrated with buffer (50 mM HEPES, 0.3 mM NaCl, 0.01% (w/v) NaN₃, pH 7.5) and eluted at a flow rate of 0.5 ml min⁻¹. The analytical standards were carbonic anhydrase (29 kDa), BSA (66 kDa), (w/v) NaN₃, pH 7.5) and eluted at a flow rate of 0.5 ml min⁻¹. The analytical standards were carbonic anhydrase (29 kDa), BSA (66 kDa), thyroglobulin (669 kDa). Blue dextran was used to measure the void volume (V₀). The elution volume (Vₑ) of each standard was measured and corrected by the subtraction of half the sample volume. The retention coefficient (R) for each standard was calculated as R = Vₑ/V₀. A plot of log 10 molecular mass of standards against R gave a straight line used to calculate molecular mass values (data not shown).

Sedimentation velocity experiments were performed with a Beckman XL-1 analytical ultracentrifuge. An-50Ti rotor (28,000 rpm at 20 °C), using absorption optics at a wavelength of 280 nm. Three different sample concentrations (0.25, 0.5, and 1 mg ml⁻¹) in 100 mM Tris-HCl buffer, pH 7.6, 50 mM NaCl, with the addition of 10 mM dithiothreitol in the presence of E. coli and IspE, were used in the study, and all data were analyzed with SEDFIT (50). The IspDF samples were analyzed alone and then in combination with the IspE derived from the same organism. All combinations of the two and then three protein mixtures were analyzed for the E. coli enzymes.

RESULTS

Subunit, Domain, and Quaternary Structure of CjIspDF—Two structures (I and II) have been determined at resolutions of 2.3 and 3.1 Å, respectively. The structure of CjIspDF presents one subunit of 371 amino acids (~41.7 kDa) per asymmetric unit and comprises two distinct domains (Fig. 2n). The N-terminal domain (residues 1–209) corresponds to the cytidyltransferase IspD, and the C-terminal domain (residues 210–371) corresponds to the synthase IspF. The N-terminal CjIspD segment forms a Rossmann fold-like α/β domain, into which is inserted an extended “β-arm” (β8 and β9). The core of the domain consists of a seven-stranded twisted β sheet (β5, β4, β1, β6, β10, β7, and β11), where all but β10 are parallel. This domain shares 25% sequence identity with EcIspD (Fig. 2b), and an overlay of the two structures (Protein Data Bank code 1152) gives a root mean square deviation of 1.75 Å for 191 Ca atoms. The C-terminal CjIspF domain of 163 residues displays an α/β fold constructed with a four-stranded β-sheet of β12, β18, β16, and β17, with β18 antiparallel to the others, on one side and a four-helix bundle of α8 to α12 on the other. This domain shares a sequence identity of ~48% with EcIspF (Fig. 2b), and this pair is more similar than the CjIspD and EcIspD domains, reflected in a root mean square deviation of 0.83 Å for superposition of 149 Ca atoms of CjIspF and EcIspF (Protein Data Bank 1GX1). The major difference between these two IspF structures occurs in the loop between α9 and α10 (CjIspD numbering) on one side of the active site, where deviations of nearly 8 Å occur between several equivalent Ca pairs (data not shown).

Eight segments of polypeptide form the CjIspD-IspF domain-interface over an area of ~850 Å². These segments include the loop leading into θ1, the β-strands β2 and β3, and the helix α7 on the N-terminal domain, which interact with the link section between α7 and β12; the C-terminal sections of α8 and α10; and the C terminus itself. The interface is mainly hydrophobic in character and involves the residues Val21, Phe25, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368. In addition, the aliphatic component of polar side chains Lys20, Lys22, and Glu295 interacts with the aromatic ring of Tyr365. Only three direct hydrogen bonding interactions link the domain, Lys22, Leu26, Tyr59, Leu198, Leu203, Pro206, Ile296, Phe366, and Trp368.
The hexamer, for which subunits are assigned labels A to F, has a total mass of about 250.2 kDa and longest dimensions of \( \sim 100 \times 130 \) Å in the equatorial and axial directions, respectively. Two trimers (subunits A, B, and C and then subunits D, E, and F) of IspF domains form the apices, and three dimers of IspD domains form the middle of the assembly (Fig. 3). The IspD dimers involve subunit pairs A with D, B with E, and C with F. A distance of about 35 Å separates the IspD active sites within a dimer, the closest distance between two IspD active sites on separate dimers is 56 Å, and the closest distance between the IspD active site and IspF active site on the same monomer is \( \sim 38 \) Å. The closest distance between two IspF active sites is \( \sim 30 \) Å.

The major contribution to the CjIspD dimer interface is provided by the association of the \( \beta \)-arm of each subunit, with a lesser contribution from side chain interactions between residues on helix \( \alpha 7 \). The active site is formed at the dimer interface by six segments of polypeptide from one subunit and one segment from the partner subunit. The dimer interface covers an area of \( \sim 1670 \) Å\(^2\), 9% of the subunit surface area, and involves 26 intersubunit hydrogen bonds. The surface area is reduced by about 200 Å\(^2\) in comparison with EcIspD mainly due to a smaller loop, reduced by 8 residues, between residues 19 and 20 in CjIspD (Fig. 2b).

IspF forms a compact homotrimer shaped like an extended trigonal prism. The trimer is formed using a single type of subunit-subunit interface with 18 intersubunit hydrogen bonds, mainly involving side chain atoms. The surface area of a CjIspF monomer buried in the homotrimer is 2630 Å\(^2\), about 15% of the total surface area of CjIspDF and significantly larger than the value of 2120 Å\(^2\) observed in EcIspF (35). At the core of the trimer, on the 3-fold axis, is a narrow cavity into which is modeled GPP. This assignment was based on the electron density and previous characterization of ligands that bind EcIspF at the central region of the trimer. These ligands also include phosphate, IPP and/or DMAPP, and farnesyl dipiphosphate. In CjIspDF, the dipiphosphate moiety of GPP is placed at the entrance to the cavity and participates in 10 hydrogen bonds with the side chains of Arg\(^{351}\) and amides of Phe\(^{348}\) from the three subunits (data not shown). The geranyl tail is directed down into the core of the trimer participating in van der Waals interactions with the side chains of Phe\(^{216}\), Met\(^{257}\), and Phe\(^{219}\), which together with Phe\(^{209}\), Ile\(^{210}\), Val\(^{218}\), Leu\(^{346}\), and Leu\(^{358}\) form a hydrophobic lining to the cavity. The presence of GPP likely makes a significant contribution to the stability of the IspF trimer.

There is only one small area of interaction, about 210 Å\(^2\), between the IspD domain of one subunit with the IspF domain of another subunit. This is a single type of interface involving residues on the \( \alpha 2 \) and \( \beta 2 \) of subunit A interacting with the \( \alpha 7 \) segment of subunit C, for example. Two hydrogen bonds are involved here; Lys\(^{86}\) NZ and O donate and accept hydrogen bonds, respectively, with Glu\(^{319}\) OE1 and Phe\(^{354}\) N (data not shown).

The Active Sites of IspD and IspF Are Highly Conserved—Previous work has produced an excellent understanding of substrate recognition and the mechanism of catalysis by EcIspD and EcIspF (20, 34–37, 51). We therefore only comment briefly on the two active sites in CjIspDF.

The active site of CjIspD is formed at the dimer interface by polypeptide segments 9–16, 73–76, 79, 94–99, 191–193, 144–148, and 167–168 from one subunit and segment 127–129 from the other. The recognition and interactions of EcIspD, by direct hydrogen bonds, with substrate and/or product involve 19 residues (Fig. 2b). Ten of these residues only use main chain functional groups, and five are strictly conserved in CjIspDF (Ala\(^{10}\), Ala\(^{11}\), Gly\(^{12}\), Gly\(^{73}\), and Asp\(^{74}\)). A noteworthy difference with respect to the active site is the Gly\(^{18}\) to Ser\(^{14}\) alteration...
and an Arg$^{19}$ to Thr$^{15}$ change (Fig. 2b). In EcIspD, Arg$^{19}$ interacts with a CTP phosphate using main chain amide and side chain groups. Alteration to a threonine does not influence the main chain amide and leaves the hydroxyl in combination with the adjacent Ser$^{14}$ hydroxyl to likely interact with substrate.

Eight of the remaining nine residues that use side chains (Arg$^{16}$, Lys$^{23}$, Ser$^{79}$, Asp$^{96}$, Arg$^{99}$, Thr$^{129}$, Arg$^{139}$, and Lys$^{191}$ in CjIspDF) are strictly conserved. Arg$^{16}$ and Thr$^{129}$ use both main chain and side chain functional groups. Four of these conserved residues, Arg$^{16}$, Lys$^{23}$, Lys$^{191}$, and Arg$^{139}$, bind and polarize the substrate for nucleophilic attack and then serve to stabilize the negatively charged transition state (19). Two residues, Arg$^{76}$ and Thr$^{147}$ in CjIspDF, are important for van der Waals interactions with ligands and are conserved in EcIspD (Fig. 2b).

The active site of CjIspF is also formed with contributions from two subunits, in this case, by polypeptide segments 217–219, 241–243, 251–252, and 265–278 from one subunit and segments 309–315 and 340–344 from the partner. The catalytic function of IspF depends on two divalent cations, which orient and polarize the substrate (35–37), and the residues that bind these cations are strictly conserved. Structure II contains a Zn$^{2+}$ ion in the CjIspF active site, confirmed by an x-ray Absorption Near Edge Scan and anomalous dispersion measurements (data not shown). The cation is coordinated by Asp$^{217}$, His$^{219}$, His$^{251}$, and a water molecule. In contrast, structure I does not contain Zn$^{2+}$, and we presume that the cation leached out of the protein during the longer dialysis step (overnight as opposed to 2 h) employed for the batch of enzyme from which that structure was derived. Structure I carries Mg$^{2+}$ in the CjIspF active site coordinated to the phosphate of CMP, side chains of Asp$^{217}$ and Thr$^{241}$, and water molecules. Although nearby, this is not the same divalent cation-binding site used by EcIspF when it coordinates the diphosphate containing CDP or substrate 4-diphosphocytidyl-2C-methyl-D-erythritol (34,
It is likely that Mg$^{2+}$ is brought into the active site with the ligand, be it substrate or, in this case, CMP. For catalysis by the CjIspF domain, we anticipate that similar coordination would occur as in EcIspF, i.e. the ion interacts with two phosphates, the conserved Glu$^{344}$ and water molecules.

Hydrogen bonding interactions with the substrate/products in EcIspF involve 10 amino acids, of which 7 (Ile$^{57}$, Gly$^{58}$, Phe$^{61}$, Ala$^{100}$, Pro$^{103}$, Met$^{105}$, and Leu$^{106}$) use only main chain groups. These residues correspond to Ile$^{266}$, Gly$^{267}$, Tyr$^{270}$, Ala$^{309}$, Pro$^{312}$, Leu$^{314}$, and Lys$^{315}$ in CjIspDF. Despite the use of only main chain groups to interact with the ligands, these residues remain well conserved (Fig. 2b). Five other amino acids provide side chains to hydrogen bond with active site ligands, and these residues, His$^{241}$, Ser$^{244}$, Asp$^{265}$, Lys$^{313}$, and Thr$^{342}$ of CjIspDF, are strictly conserved in EcIspF (Fig. 2b). There is a significant difference in the active site of the two enzymes. In CjIspDF, Lys$^{318}$ NZ forms a hydrogen bond with the CMP O$^{2}$ ribose hydroxyl. In EcIspF, the corresponding position is Ile$^{109}$ placed so the side chain does not contribute to the active site (data not shown).

**Size Exclusion Chromatography—**Experiments were carried out on all enzymes individually and then carried out on the various mixtures. EcIspD showed the expected monomer and dimer species but also showed a tetramer, which disappeared on the addition of dithiothreitol. Previously, we showed that this enzyme forms a dimer-dimer association due to a disulfide

---

**Structure of the IspDF hexamer.** a, stereoview depiction viewed parallel to one of the 2-fold axis. Subunits are colored separately and labeled A to F. Helices are shown as cylinders, $\beta$-strands are shown as arrows, metal ions are shown as spheres, and GPP and CMP are shown as green and yellow sticks. b, view parallel to the molecular 3-fold axis.

---

**Fig. 3.** Structure of the IspDF hexamer. a, stereoview depiction viewed parallel to one of the 2-fold axis. Subunits are colored separately and labeled A to F. Helices are shown as cylinders, $\beta$-strands are shown as arrows, metal ions are shown as spheres, and GPP and CMP are shown as green and yellow sticks. b, view parallel to the molecular 3-fold axis.
linkage (34). Analysis of the CjIspDF trace showed trimer, hexamer, and dodecamer assemblies. A mixture of CjIspDF and CjIspE only showed peaks corresponding to the individual enzymes. In similar fashion, EcIspD and EcIspE did not appear to associate, nor did EcIspD and EcIspF. However, analysis of EcIspE mixed with EcIspF revealed a peak of ~500 kDa not observed in the experiments on individual enzymes (data not shown). This may represent a complex of EcIspF and EcIspE.

Encouraged by this observation, we investigated further using analytical ultracentrifugation.

**Sedimentation Characteristics**—Initial experiments to investigate the association of IspDF with IspE used recombinant *C. jejuni* enzymes. Aggregation problems with CjIspE were noted, and we therefore generated recombinant AtIspDF and AtIspE, which proved more amenable for analysis and, importantly, provided another system for comparison. The amino acid sequence identity is 32% between the bifunctional enzymes and 25% between the IspE of both species.

The sedimentation velocity runs for CjIspDF and AtIspDF were consistent, and each showed only two peaks, of ~125 and ~250 kDa (Fig. 4a). These peaks correspond to a trimer and a hexamer, and the absence of any other species suggests that IspDF can best be described as a dimer of trimers. The IspD dimers appear less stable than the IspF trimers, which may be explained by the much larger interface area associated with the latter domain, as described earlier. The bifunctional IspDF samples were mixed with the cognate IspE samples, and the results indicated the presence of the trimeric and hexameric IspDF together with species of higher mass (~380 and ~580 kDa) (Fig. 4b). A complex of CjIspDF hexamer plus three IspE dimers would have a mass of about 408 kDa. In the case of AtIspDF + AtIspE (Fig. 4c), an additional high mass species of nearly 1 MDa is observed.

A mixture of EcIspD with EcIspF showed the expected dimer and trimer species for the individual enzymes in a single broad peak, but no higher order complexes were seen (data not shown). In contrast, a mixture of EcIspD with EcIspE displayed a dimer for each enzyme of ~51 and ~62 kDa, respectively, and also displayed a peak of ~130 kDa. The analysis of EcIspD mixed with EcIspF revealed a broad peak corresponding to their dimeric and trimeric states, respectively, as well as a peak at ~250 kDa, which fits with a model of two IspF trimers and three IspE dimers. When the individual enzymes EcIspD, EcIspE, and EcIspF were mixed, a complex of ~430 kDa was observed (Fig. 4d). This likely corresponds to an assembly of three IspD dimers, three IspE dimers, and two IspF trimers, similar to that observed for the complex assembly of IspD and IspE derived from both *C. jejuni* and *A. tumefaciens*. A summary of these results is given in Table III.

Note that in all cases, the individual enzymes were first characterized as controls, and we only observed the higher mass species when enzyme mixtures were analyzed.

**DISCUSSION**

Implications of Protein-Protein Complex Formation for Isoprenoid Precursor Biosynthesis—We observe associations between three enzymes of the DOXP pathway. The clue that set us searching for this is the presence of an unusual bifunctional
enzyme catalyzing nonconsecutive steps in the pathway. Bi-functional enzymes are distinctive and highly conserved products of relatively infrequent gene fusion events that generally link proteins with related yet distinct functions. Some interaction with the intervening kinase IspE therefore seemed plausible. Interestingly, associations were also observed for the monofunctional \textit{E. coli} enzymes. We can now consider likely biological implications of such associations.

There are distinct mechanisms by which metabolic pathways are controlled: firstly, covalent modification, as exemplified by phosphorylation of mammalian HMG-CoA reductase (5), which regulates the mevalonate pathway. Secondarily, the compartmentation of enzymes in organelles or organelles can control pathway flux. For example, the mevalonate pathway occurs in the cytosol and mitochondria of plants, whereas the DOXP pathway is compartmentalized into chloroplasts. Apicomplexan parasites likewise carry the DOXP pathway enzymes in their apicoplast (52). Thirdly, the binding of effector molecules can provide feedback control or link pathways together. There is also control by repression or activation of gene expression. In Gram-positive bacteria that utilize the mevalonate pathway, the relevant genes are organized into operons (7) and are likely regulated at the level of transcription.

The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely controlled: firstly, covalent modification, as exemplified by phosphorylation of mammalian HMG-CoA reductase (5), which regulates the mevalonate pathway. Secondarily, the compartmentation of enzymes in organelles or organelles can control pathway flux. For example, the mevalonate pathway occurs in the cytosol and mitochondria of plants, whereas the DOXP pathway is compartmentalized into chloroplasts. Apicomplexan parasites likewise carry the DOXP pathway enzymes in their apicoplast (52). Thirdly, the binding of effector molecules can provide feedback control or link pathways together. There is also control by repression or activation of gene expression. In Gram-positive bacteria that utilize the mevalonate pathway, the relevant genes are organized into operons (7) and are likely regulated at the level of transcription.

The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely regulated at the level of transcription. The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely regulated at the level of transcription. The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely regulated at the level of transcription.

The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely regulated at the level of transcription. The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely regulated at the level of transcription. The first two types of control mechanisms are not applicable to the DOXP pathway in eubacteria, and there is no evidence of a transcriptional regulator to coordinate expression of the relevant genes. The recent identification of isoprenoids binding to a transcriptional regulator to coordinate expression of the rel-to the DOXP pathway in eubacteria, and there is no evidence of the relevant genes are organized into operons (7) and are likely regulated at the level of transcription.
ski, H. (2001) J. Mol. Microbiol. Biotechnol. 3, 483–489
30. Lell, B., Ruangwoerayut, R., Wiessner, J., Missinou, M. A., Schindler, A., Baranek, T., Hintz, M., Hutchinson, D., Jomaa, H., and Kremsner, P. G. (2003) Antimicrob. Agents Chemother. 47, 735–738
31. Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 13171–13177
32. von Mering, C., Huynen, M., Jaeger, D., Schmidt, S., Bork, P., and Snel, B. (2003) Nucleic Acids Res. 31, 258–261
33. Gabrielsen, M., Rohdich, F., Eisenreich, W., Graswert, T., Hecht, S., Bacher, A., and Hunter, W. N. (2004) Eur. J. Biochem. 271, 1–8
34. Kemp, L. E., Bond, C. S., and Hunter, W. N. (2003) Acta Crystallogr. Sect. D Biol. Crystallogr. 59, 607–610
35. Kemp, L. E., Bond, C. S., and Hunter, W. N. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 6591–6596
36. Steinbacher, S., Kaiser, J., Wungsintaweekul, J., Hecht, S., Eisenreich, W., Gerhardt, S., Bacher, A., and Rohdich, F. (2002) J. Mol. Biol. 316, 79–88
37. Richard, S. B., Ferrer, J.-L., Bowman, M. E., Lillo, A. M., Tetzlaff, C. N., Cano, D. E., and Noel, J. P. (2002) J. Biol. Chem. 277, 8667–8672
38. Miallau, L., Alphey, M. S., Kemp, L. E., Leonard, G. A., McSweeney, S. M., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F., and Hunter, W. N. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 9173–9178
40. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
41. Collaborative Computational Project Number 4. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
42. Navaza, J. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr. 57, 1367–1372
43. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Nat. Struct. Biol. 6, 458–463
44. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D Biol. Crystallogr. 53, 240–255
45. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
46. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
47. Nicholls, A., Sharp, K., and Honig, B. (1991) Proteins Struct. Funct. Genet. 11, 281–296
48. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
49. Merrit, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505–524
50. Schuck, P. (2000) Biophys. J. 78, 1696–1619
51. Gardner, M. J., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pai, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M. S., Nene, V., Shallow, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaudya, A. B., Martin, D. M., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., Mcfadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M., and Barrell, B. (2002) Nature 419, 498–511
52. Dandekar, T., Snel, B., Huynen, M., and Bork, P. (1998) Trends Biochem. Sci. 23, 324–328
53. Enright, A. J., Iliopoulos, I., Kyprides, N. C., and Ouzounis, C. A. (1999) Nature 402, 86–90
54. Enright, A. J., and Ouzounis, C. A. (2001) Genome Biology http://genomebiology.com/2001/2/9/research/0034
Hexameric Assembly of the Bifunctional Methyerythritol 2,4-Cyclodiphosphate Synthase and Protein-Protein Associations in the Deoxy-xylulose-dependent Pathway of Isoprenoid Precursor Biosynthesis
Mads Gabrielsen, Charles S. Bond, Irene Hallyburton, Stefan Hecht, Adelbert Bacher, Wolfgang Eisenreich, Felix Rohdich and William N. Hunter

J. Biol. Chem. 2004, 279:52753-52761.
doi: 10.1074/jbc.M408895200 originally published online October 2, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M408895200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 55 references, 15 of which can be accessed free at http://www.jbc.org/content/279/50/52753.full.html#ref-list-1