



# Crystal Structure of Chorismate Synthase: A Novel FMN-binding Protein Fold and Functional Insights

Hyung Jun Ahn, Hye-Jin Yoon, Byung Il Lee and Se Won Suh\*

Department of Chemistry  
College of Natural Sciences  
Seoul National University  
Seoul 151-0742, South Korea

Chorismate synthase catalyzes the conversion of 5-enolpyruvylshikimate 3-phosphate to chorismate in the shikimate pathway, which represents an attractive target for discovering antimicrobial agents and herbicides. Chorismate serves as a common precursor for the synthesis of aromatic amino acids and many aromatic compounds in microorganisms and plants. Chorismate synthase requires reduced FMN as a cofactor but the catalyzed reaction involves no net redox change. Here, we have determined the crystal structure of chorismate synthase from *Helicobacter pylori* in both FMN-bound and FMN-free forms. It is a tetrameric enzyme, with each monomer possessing a novel “ $\beta$ - $\alpha$ - $\beta$  sandwich fold”. Highly conserved regions, including several flexible loops, cluster together around the bound FMN to form the active site. The unique FMN-binding site is formed largely by a single subunit, with a small contribution from a neighboring subunit. The isoalloxazine ring of the bound FMN is significantly non-planar. Our structure illuminates the essential functional roles played by the cofactor.

© 2004 Elsevier Ltd. All rights reserved.

**Keywords:** *aroC*; chorismate synthase; FMN-binding protein; *Helicobacter pylori*; shikimate pathway

\*Corresponding author

## Introduction

The shikimate pathway, which links metabolism of carbohydrates to biosynthesis of aromatic compounds, is essential to bacteria, fungi, and plants<sup>1</sup> as well as apicomplexan parasites.<sup>2</sup> This seven-step metabolic route leads from phosphoenolpyruvate and erythrose 4-phosphate to chorismate, a common precursor for the synthesis of aromatic amino acids and many aromatic compounds including folic acid, ubiquinone, and vitamins E and K. This pathway is absent in metazoans, thus making it an attractive target for the development of antimicrobial agents,<sup>3</sup> herbicides,<sup>4</sup> and antiparasitic agents. 5-Enolpyruvylshikimate 3-phosphate (EPSP) synthase, which catalyzes the sixth step in the pathway, has already been successfully targeted, with the development of glyphosate, one of

the world's best-selling herbicides.<sup>5</sup> Until now, crystal structures of six shikimate pathway enzymes have been elucidated except for chorismate synthase: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase,<sup>6</sup> 3-dehydroquinate synthase,<sup>7</sup> type I and II dehydroquinase,<sup>8</sup> shikimate dehydrogenase,<sup>9,10</sup> type I and II shikimate kinase,<sup>11,12</sup> and EPSP synthase.<sup>13,14</sup> These enzymes in series catalyze the first six steps of the shikimate pathway.

Chorismate synthase (EC 4.2.3.5, 5-O-(1-carboxyvinyl)-3-phosphoshikimate phosphate-lyase) catalyzes the conversion of EPSP to chorismate in the last step of the shikimate pathway. The enzyme has an absolute requirement for reduced flavin as a cofactor, although the catalyzed reaction involves no net redox change. The reaction involves an unusual *anti*-1,4-elimination of the C-3 phosphate and the C-6 *pro R* hydrogen. The cleavage of a non-activated C–H bond is another unusual aspect of this reaction. A number of non-concerted mechanisms have been proposed to account for the cofactor requirement and stereochemistry of the reaction. Evidence for a non-concerted mechanism includes a tritium kinetic isotope effect at C-3,<sup>15</sup> the slow conversion of (6 *S*)-6-fluoro-EPSP to

Abbreviations used: DHCCP, *cis*-3,4-dihydroxycyclohexene-1-carboxylic acid 3-phosphate; EPSP, 5-enolpyruvylshikimate 3-phosphate; MAD, multiwavelength anomalous diffraction; r.m.s., root-mean-square; SeMet, selenomethionine.

E-mail address of the corresponding author: sewonsuh@snu.ac.kr

6-fluoro-chorismate,<sup>16</sup> transient kinetics studies,<sup>17</sup> and a secondary  $\beta$  deuterium kinetic isotope effect at C-4.<sup>18</sup> It has been suggested that C(6)–H cleavage most likely follows C(3)–O cleavage in a non-concerted radical or E1 reaction mechanism and C(6)–H cleavage is irreversible.<sup>18</sup> It has also been suggested that the enzyme must tightly control the deprotonation at *pro R* C-6 with an appropriately positioned active site base.<sup>18</sup> Studies with substrate and cofactor analogs provided strong evidence for a radical mechanism.<sup>19</sup> Furthermore, a possible regulation of the enzyme activity by means of the availability of reduced FMN remains poorly understood.

Chorismate synthases from *Neurospora crassa* and *Saccharomyces cerevisiae* possess an additional NADPH:FMN oxidoreductase activity<sup>20</sup> and thus are referred to as bifunctional enzymes. The active site responsible for this additional activity appears to overlap with that of chorismate synthase activity, as evidenced by a common binding site for both EPSP and NADPH.<sup>21</sup> In contrast, chorismate synthases from many other organisms, including *Escherichia coli*, *Thermatoga maritima*, *Plasmodium falciparum*, and *Staphylococcus aureus*, do not show such activity and are termed monofunctional.<sup>22</sup> It is not feasible to predict mono/bi-functionality based on the size of the protein alone. Chorismate synthase represents a distinct class of a conserved enzyme family, since it lacks a primary sequence similarity to other classes of enzymes. Three signature motifs have been detected in the chorismate synthase sequences: (i) G-[DES]-S-H-[GC]-X<sub>2</sub>-[LIVM]-[GTIV]-X-[LIVT]-[LIV]-[DEST]-G-X-[PV] (sig1), (ii) [GE]-X<sub>2</sub>-S-[AG]-R-X-[ST]-X<sub>3</sub>-[VT]-X<sub>2</sub>-[GA]-[STAVY]-[LIVMF] (sig2), and (iii) R-[SH]-D-[PSV]-[CSAV]-X<sub>4</sub>-[SGAIV]-X-[IVGSTAP]-[LIVM]-X-E-[STAH]-[LIVMA] (sig3) (PROSITE database of protein families and domains<sup>†</sup>), where X denotes any amino acid, and the amino acids in brackets depict the option at a given position. At present no structural information is available for chorismate synthase.

In order to provide the missing structural information and to better understand the functional role of the cofactor, we have determined the crystal structure of chorismate synthase from *Helicobacter pylori* (encoded by the *aroC* gene), a 365 residue protein ( $M_r = 40,100$ ), in both FMN-bound and FMN-free forms. *H. pylori* chorismate synthase shows a sequence identity of 45% against the *E. coli* enzyme, one of the best-characterized chorismate synthases. This first structure of chorismate synthase reveals a novel FMN-binding protein fold with a unique FMN-binding site. Highly conserved regions including several flexible loops cluster together to form the active site with a deep cofactor-binding pocket. The isoalloxazine moiety of the bound cofactor is significantly non-planar. We propose a model for the chorismate synthase–

FMN–EPSP ternary complex. Our structure provides insights into the essential role of the cofactor and sheds light on the functional roles of highly conserved residues in the active site.

## Results and Discussion

### Structure determination and model quality

The crystal structure of *H. pylori* chorismate synthase in the apo form was determined at 3.0 Å resolution using the multiwavelength anomalous diffraction (MAD) data collected from a crystal of the selenomethionine (SeMet)-substituted enzyme (see Materials and Methods). Subsequently, the structure of the enzyme in complex with FMN was refined to 1.95 Å resolution (Table 1). The refined model of the binary complex with FMN includes a tetramer that accounts for all 365 residues of chorismate synthase, four FMN molecules bound to each of the four monomers, and 455 water molecules in the asymmetric unit of the crystal. The eight extra residues (LEHHHHHH) at the C terminus introduced into the recombinant enzyme to facilitate purification were not visible in the electron density map. The residues 87–96 were modeled as alanine, because their electron density was very poor due to high flexibility. The four independent monomers in the asymmetric unit have highly similar structures; their pair-wise comparisons give root-mean-square (r.m.s.) deviations ranging between 0.67 Å and 0.90 Å for 365 C $\alpha$  atoms. However, two of them had considerably lower average *B*-factors (36.9 Å<sup>2</sup> and 36.8 Å<sup>2</sup> for A and D subunits, respectively) than the other two (49.3 Å<sup>2</sup> and 53.6 Å<sup>2</sup> for B and C subunits, respectively). The FMN molecules bound largely to A, B, C and D subunits have average *B*-factors of 52.4, 64.5, 67.9, and 56.7 Å<sup>2</sup>, respectively. As explained below, each FMN-binding site is formed mostly within a single subunit together with a small contribution from a neighboring subunit. In all four subunits, four loop regions (residues 13–21, 46–64, 83–98, and 114–128) have very high *B*-factors (>50 Å<sup>2</sup>). Two additional loop regions (residues 260–275 and 301–324) in subunits B and C have also very high *B*-factors. These loop regions contain many of the highly conserved residues and will be referred to as F1 through F6, respectively. Mobility of flexible F5 and F6 loop regions in subunits A and D is reduced by intermolecular contacts in the crystal. We chose subunits A and D for discussions below on the cofactor binding, because they are better defined by the electron density than subunits B and C. Furthermore, we have determined the structure in an FMN-free form at 2.25 Å resolution (Table 1). The binding of FMN causes little overall structural change (an r.m.s. deviation of 0.24 Å for 365 C $\alpha$  atoms in subunit A), except in three flexible loop regions F2, F3, and F4, which show maximum deviations in the range between 0.8 Å and 1.2 Å.

<sup>†</sup> <http://us.expasy.org/prosite/>

**Table 1.** Data collection and refinement statistics

A. SeMet MAD data collection			
Unit cell parameters:			
$a = b, c$ (Å)	146.42, 132.06		
$\alpha = \beta = \gamma$ (deg.)	90		
Space group	I4		
Data set	Remote	Peak	Inflection
Wavelength (Å)	0.9500	0.9795	0.9798
Resolution range (Å)	20–3.0	20–3.0	20–3.0
Total/unique reflections	149,993/27,784	150,116/27,785	150,146/27,809
Completeness (%) <sup>a</sup>	99.6 (99.6) <sup>a</sup>	99.6 (99.6)	99.6 (99.6)
$R_{\text{sym}}$ (%) <sup>b</sup>	13.0 (45.8) <sup>b</sup>	11.0 (37.7)	10.9 (37.9)
$\langle I/\sigma \rangle$	5.3 (1.6)	6.0 (1.8)	4.3 (1.9)
Mean figure of merit <sup>c</sup>	0.49/0.65 (before/after RESOLVE)		
B. Refinement			
Data set	FMN-bound	FMN-free	
Wavelength (Å)	1.0000	1.0000	
Unit cell parameters			
$a = b, c$ (Å)	146.69, 132.31	146.48, 132.06	
Resolution range (Å), data completeness (%)	30–1.95, 99.6	30–2.25, 98.5	
No. of reflections (working/free set) <sup>d</sup>	90,603/10,100	58,346/6,567	
Average B-factors (Å <sup>2</sup> )			
For 4 × 365 amino acid residues of a tetramer	44.1	50.8	
For four FMN molecules	60.3	–	
For water molecules (no. water molecules)	32.8 (455)	36.9 (492)	
r.m.s. deviations, bond lengths (Å)/angles (deg.) <sup>e</sup>	0.0057/1.14	0.0060/1.20	
$R_{\text{work}}/R_{\text{free}}$ <sup>d</sup> (%)	21.5/24.4	19.8/24.5	

<sup>a</sup> Completeness for  $I/\sigma(I) > 1.0$ . Numbers in parentheses reflect statistics for the highest-resolution shell (3.10–3.00 Å).

<sup>b</sup>  $R_{\text{sym}} = \sum_i \sum_j |I(h_i) - \langle I(h) \rangle| / \sum_i \sum_j I(h_i)$ , where  $I(h)$  is the intensity of reflection  $h$ ,  $\sum_i$  is the sum over all reflections, and  $\sum_j$  is the sum over  $i$  measurements of reflection  $h$ . Numbers in parentheses reflect statistics for the last shell (3.10–3.00 Å).

<sup>c</sup> Mean figure of merit =  $\langle |\sum P(\alpha) e^{i\alpha} / \sum P(\alpha)| \rangle$  where  $\alpha$  is the phase and  $P(\alpha)$  is the phase probability distribution.

<sup>d</sup>  $R = \sum ||F_o| - |F_c|| / \sum |F_o|$ , where  $R_{\text{free}}$  is calculated for a randomly chosen 10% of reflections, which were not used for structure refinement, and  $R_{\text{work}}$  is calculated for the remaining reflections. The percentage of reflections set aside for  $R_{\text{free}}$  is the same for all structures.

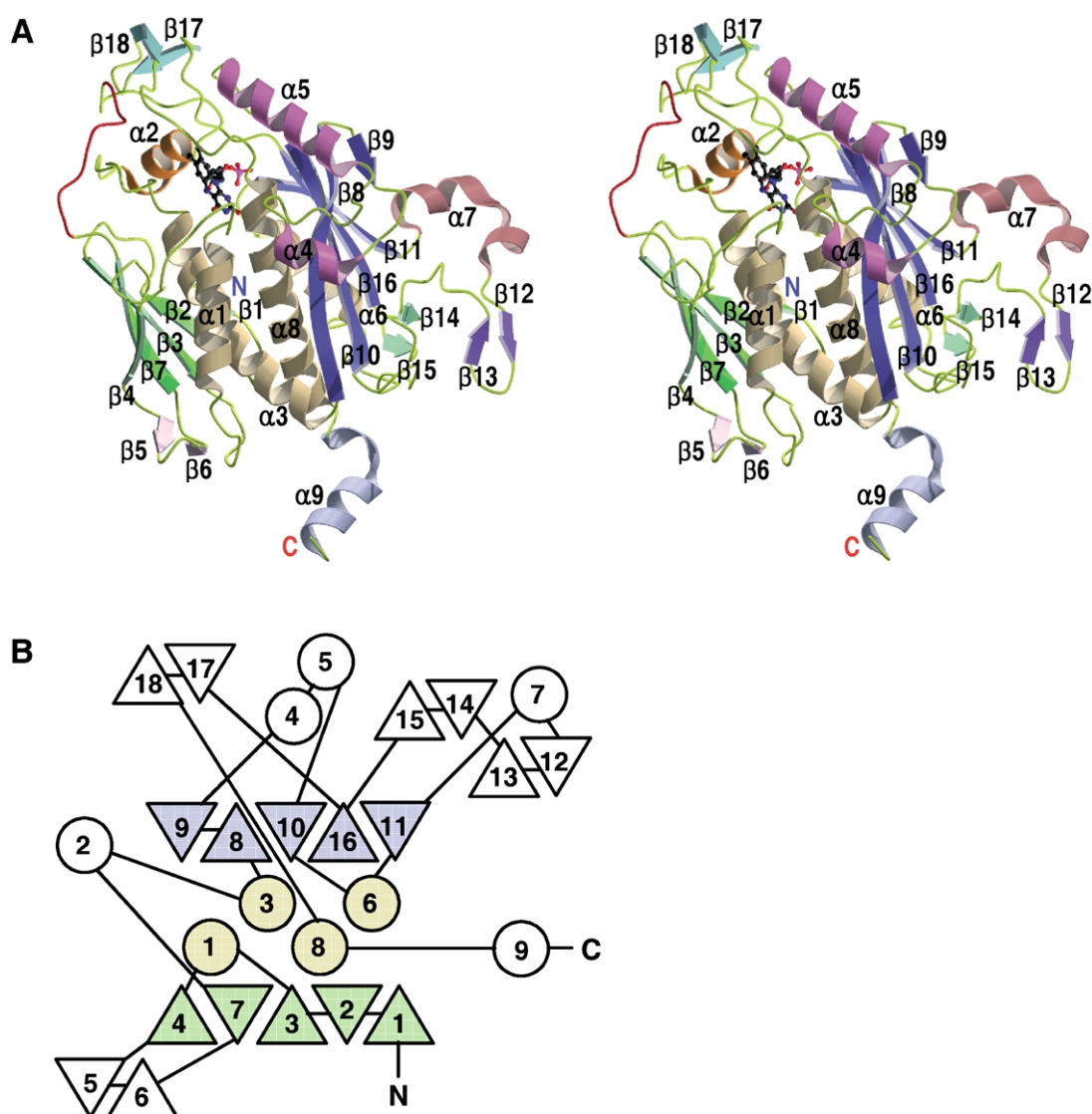
<sup>e</sup> Deviations from ideal bond lengths and angles.

## Chorismate synthase monomer has a novel fold

A structural similarity search using the DALI server<sup>23</sup> indicated that the folding topology of *H. pylori* chorismate synthase monomer has not been previously observed. The highest Z-score was 3.4 for 3-hydroxy-3-methylglutaryl-coenzyme A reductase from *Pseudomonas mevalonii* (PDB ID 1QAX) and *E. coli* aspartyl-tRNA synthetase (PDB ID 1C0A), both of which possess no meaningful structural or functional resemblance to chorismate synthase. The r.m.s. deviations were 5.0 Å and 4.0 Å for superposition of 116 and 88 C $\alpha$  atoms of the equivalenced residues in two proteins, respectively, with corresponding sequence identities between the equivalenced residues of 12 and 17%. A systematic study of the SCOP database<sup>24</sup> revealed that FMN-binding proteins exhibit one of four folds: the ( $\beta/\alpha$ )<sub>8</sub>-barrel fold, the flavodoxin-like fold, the pyridoxine 5'-phosphate oxidase-like  $\beta$ -barrel fold, or the  $\alpha + \beta$  NADH oxidase/flavin reductase fold.<sup>25</sup> The fold of chorismate synthase revealed in this study is novel and thus our chorismate synthase structure represents a new fold for the FMN-binding protein family.

Each monomer of chorismate synthase consists of nine  $\alpha$ -helices and 18  $\beta$ -strands, which fold into a large single domain of approximate dimensions 70 Å × 55 Å × 42 Å with several protruding parts (Figure 1A). The core of each monomer may be

termed a three-layered, “ $\beta$ - $\alpha$ - $\beta$  sandwich fold”. This novel fold is composed of two antiparallel five-stranded  $\beta$ -sheets ( $\uparrow\beta 1 \downarrow\beta 2 \uparrow\beta 3 \downarrow\beta 7 \uparrow\beta 4$  and  $\uparrow\beta 9 \downarrow\beta 8 \uparrow\beta 10 \downarrow\beta 16 \uparrow\beta 11$ ) and four  $\alpha$ -helices ( $\uparrow\alpha 1 \downarrow\alpha 3 \downarrow\alpha 8 \uparrow\alpha 6$ ) that are sandwiched between these two  $\beta$ -sheets (Figure 1(B)). We refer to these two  $\beta$ -sheets as N-terminal and C-terminal  $\beta$ -sheets, respectively. The monomer structure has four  $\beta$ -hairpins ( $\beta 5$ - $\beta 6$ ,  $\beta 12$ - $\beta 13$ ,  $\beta 14$ - $\beta 15$ , and  $\beta 17$ - $\beta 18$ ). The  $\beta 17$ - $\beta 18$  hairpin has a three-residue turn, whereas the other three hairpins have two-residue turns. The turn in the  $\beta 14$ - $\beta 15$  hairpin is type II with a glycine residue at the second position of the turn,<sup>26</sup> whereas the turns in other hairpins ( $\beta 5$ - $\beta 6$  and  $\beta 12$ - $\beta 13$ ) are type I. Between two adjacent antiparallel  $\beta$ -strands  $\beta 4$  and  $\beta 7$ , the  $\beta 5$ - $\beta 6$  hairpin is inserted. The protruding  $\alpha$ -helix  $\alpha 7$  and two  $\beta$ -hairpins ( $\beta 12$ - $\beta 13$  and  $\beta 14$ - $\beta 15$ ) make a winding excursion between two adjacent  $\beta$ -strands  $\beta 11$  and  $\beta 16$ . The last  $\beta$ -strand  $\beta 16$  of the C-terminal  $\beta$ -sheet and  $\alpha$ -helix  $\alpha 8$  are connected by a long winding loop including the  $\beta 17$ - $\beta 18$  hairpin. Between the  $\beta$ -strand  $\beta 7$  of the N-terminal  $\beta$ -sheet and  $\alpha$ -helix  $\alpha 3$ , there is a long winding loop and  $\alpha$ -helix  $\alpha 2$ . Two  $\alpha$ -helices,  $\alpha 4$  and  $\alpha 5$ , make a winding excursion between  $\beta$ -strands  $\beta 9$  and  $\beta 10$ , and they partly cover one edge of the C-terminal  $\beta$ -sheet. The C-terminal  $\alpha$ -helix  $\alpha 9$  protrudes from the main body of the monomer (Figure 1A). The unusual  $\beta$ - $\alpha$ - $\beta$



**Figure 1.** Stereo view of a chorismate synthase monomer. A, Ribbon diagram. Ten residues (87–96) built of alanine residues are shown as a red coil. FMN is shown in a ball-and-stick model. Carbon, oxygen, nitrogen, and phosphorus atoms are colored in black, red, blue, and purple, respectively. B, Topology diagram of a chorismate synthase monomer. The secondary structure elements that belong to the three-layered,  $\beta$ - $\alpha$ - $\beta$  sandwich fold are colored green, yellow, and blue, respectively. Triangles and circles denote  $\beta$ -strands and  $\alpha$ -helices, respectively.

sandwich fold provides a framework for building the unique cofactor-binding site, which also requires dimerization (further discussed below).

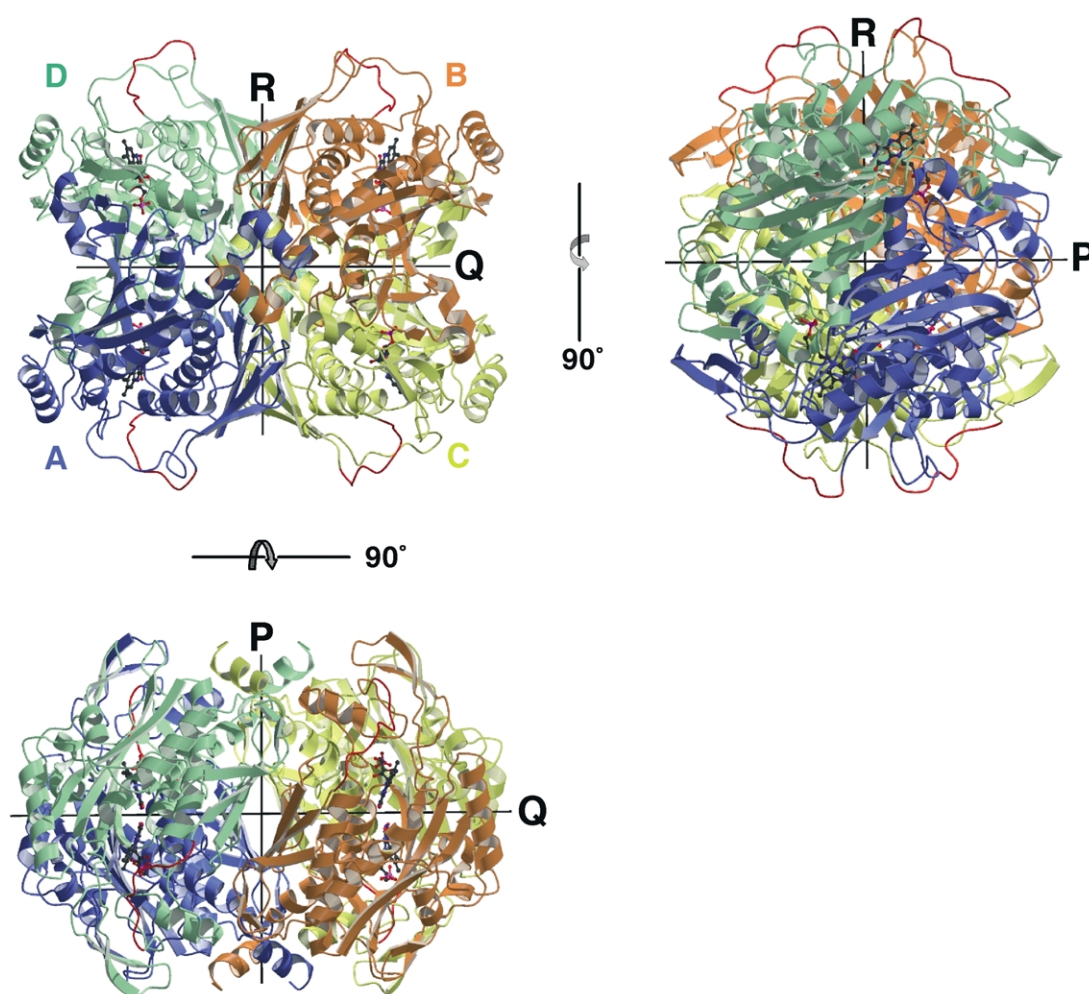
### Tetrameric structure

Dynamic light-scattering analysis indicated that *H. pylori* chorismate synthase exists as a tetramer in solution.<sup>27</sup> Consistent with this observation, four independent monomers in the crystallographic asymmetric unit forms a tight tetramer of 222 point group symmetry with approximate dimensions 80 Å × 75 Å × 65 Å (Figure 2). Three mutually perpendicular molecular dyads are here referred to as *P*, *Q*, and *R* axes. The approximate size and shape of the tetramer as deduced from

the small-angle X-ray scattering experiments<sup>28</sup> are in reasonable agreement with our structure.

Among the inter-subunit interfaces between monomers within the tetramer, that between the *Q*-axis related subunit pair A and D (or the pair B and C) has the most extensive interactions. Dimerization across the *Q*-axis plays a functionally critical role, because it is required to form a complete FMN-binding pocket. This interface involves residues located on  $\alpha$ 2 (**Lys112**),  $\alpha$ 3 (**Glu128**, **Arg132**), the  $\alpha$ 5- $\beta$ 10 loop (**Asp198**), the  $\beta$ 10- $\alpha$ 6 loop (Gln223, Leu225, Tyr226),  $\alpha$ 6 (Lys232, Glu235),  $\alpha$ 7 (Ser255, Ser260), the  $\beta$ 12- $\beta$ 13 hairpin (Gln268),  $\beta$ 16 (His294), the  $\beta$ 17- $\beta$ 18 hairpin (Gln305, Glu314), and  $\alpha$ 8 (Asn350). Bold-faced residues at this interface are highly conserved among six chorismate synthases, whose sequences are aligned in Figure 3. The  $\alpha$ 5- $\beta$ 10 loop is not directly





**Figure 2.** Ribbon diagram of a tetramer. Four subunits (A–D) are colored in blue, orange, yellow, and green, respectively. FMN is drawn and colored in the same manner as in [Figure 1A](#). *P*, *Q*, and *R* denote the 222 point group symmetry axes.

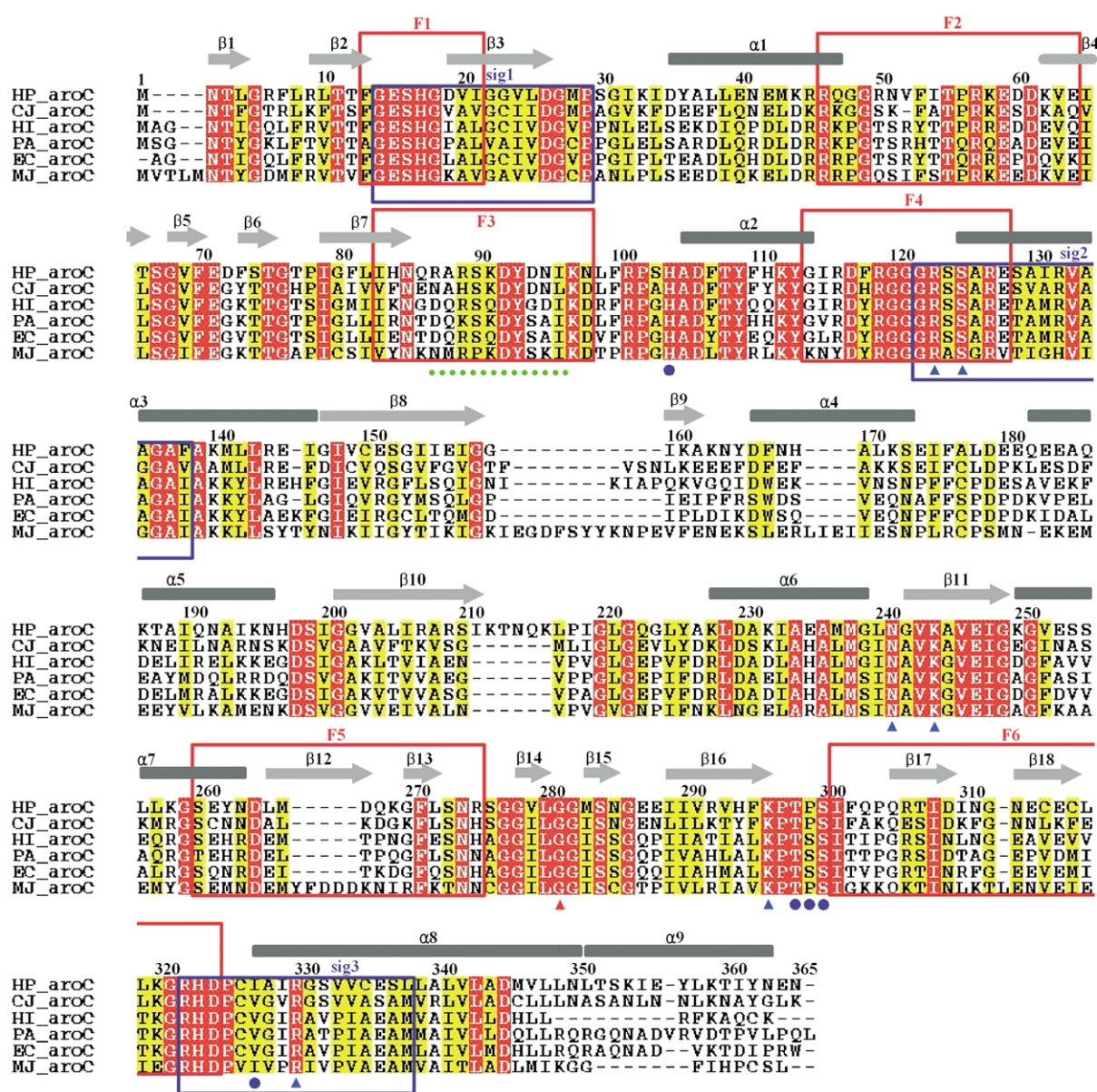
involved in the formation of the active site but is highly conserved, possibly because it indirectly influences the geometry of the active site and is also involved in this functionally critical dimerization. This inter-subunit interface buries the solvent-accessible surface area of each monomer ( $19,300 \text{ \AA}^2$ ) by 20.7% and there are 39 inter-subunit hydrogen bonds. Two C-terminal  $\beta$ -sheets from subunits A and D ( $\uparrow 9 \downarrow 8 \uparrow 10 \downarrow 16 \uparrow 11$  and  $\downarrow 11' \uparrow 16' \downarrow 10' \uparrow 8' \downarrow 9'$ ) are related by the *Q*-axis and form a ten-stranded antiparallel  $\beta$ -sheet ( $\uparrow 9 \downarrow 8 \uparrow 10 \downarrow 16 \uparrow 11 \downarrow 11' \uparrow 16' \downarrow 10' \uparrow 8' \downarrow 9'$ ), which span the whole tetramer ([Figure 2](#)). Two  $\beta$ -hairpin loops  $\beta_{12}$ - $\beta_{13}$  and  $\beta_{17}$ - $\beta_{18}$ , each from monomers A and D, come close together in the tetramer but they do not form a contiguous  $\beta$ -sheet.

Between the *R*-axis-related subunit pair A and C (or the pair B and D), mostly residues from three N-terminal strands  $\beta_1$  (Met1, **Asn2**, Thr3),  $\beta_2$  (Arg9, **Phe13**), and  $\beta_3$  (**Asp26**) participate in the interactions, which include 11 inter-subunit hydrogen bonds and one salt-bridge between the amino terminus and the side-chain of Asp26. The bold-faced residues are highly conserved ([Figure](#)

3). The two  $\beta$ -strands  $\beta_1$  and  $\beta_{1'}$  from subunits A and C, respectively, are nearly orthogonal to each other. This subunit interface buries  $2820 \text{ \AA}^2$  of the solvent-accessible surface area, which corresponds to 7.4% of the monomer solvent-accessible surface area. The inter-subunit interface across the *P*-axis involves residues located on the  $\beta_1$ - $\beta_2$  loop (Phe7),  $\alpha_3$  (Glu145),  $\alpha_8$  (Asn350), and  $\alpha_9$  (Thr352, Lys359, Tyr362), and the inter-subunit interactions include 16 hydrogen bonds. There are few conserved residues at this interface. Two protruding C-terminal helices  $\alpha_9$  and  $\alpha_{9'}$  from subunits A and B, respectively, are intertwined around each other, fitting into the crevice of the neighboring monomer between the C-terminal helix and the main body of the monomer. This subunit interface buries  $2770 \text{ \AA}^2$  of the solvent-accessible surface area, which corresponds to 7.2% of the monomer solvent-accessible surface area.

#### The FMN-binding site of chorismate synthase is unique

In a tetramer of the binary complex, each of four

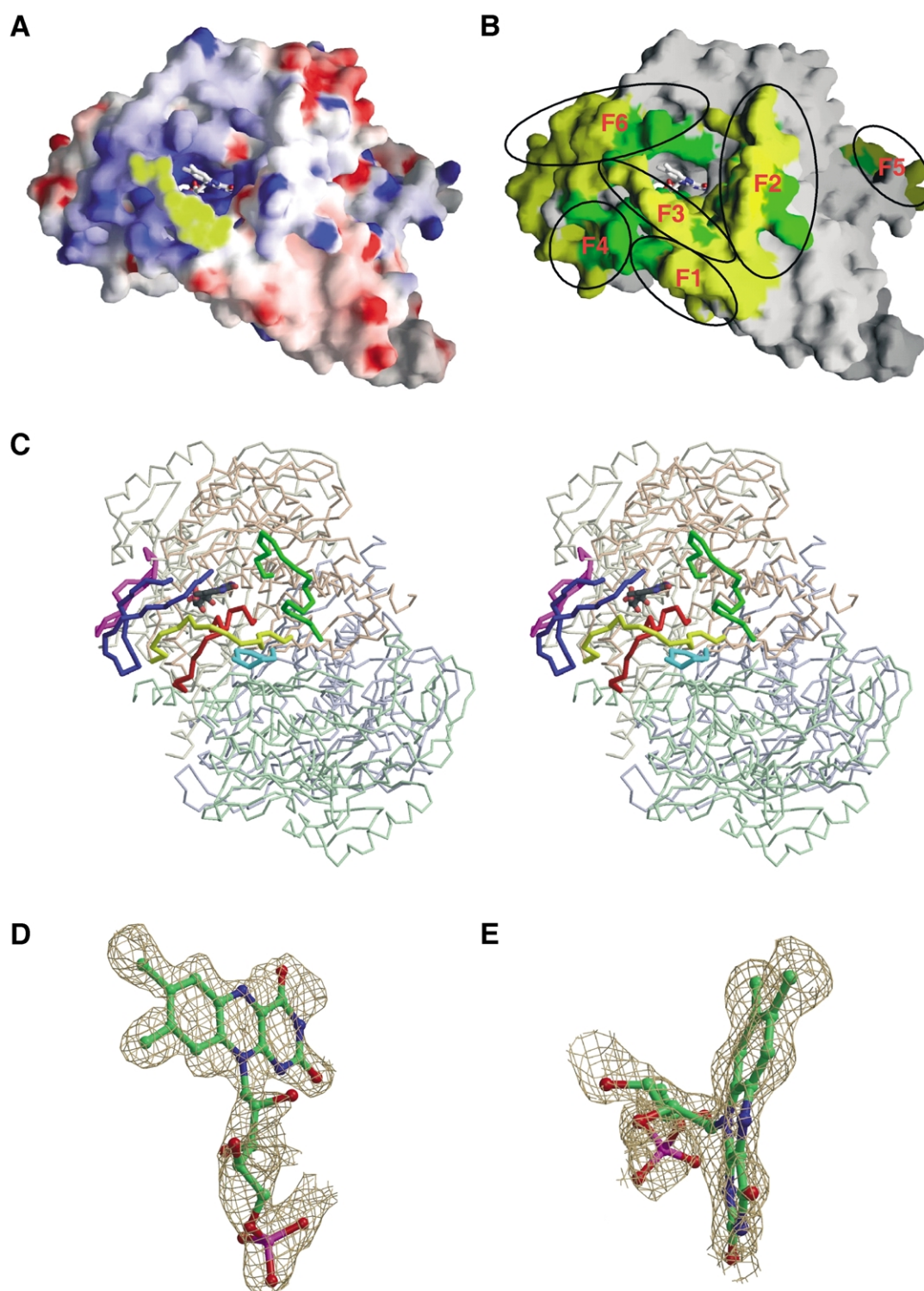


**Figure 3.** Amino acid sequence alignment of six chorismate synthases. Secondary structure elements in *H. pylori* chorismate synthase are colored in light and dark gray. Strictly conserved residues and semi-conserved residues are colored in red and yellow, respectively. Blue boxes represent three signature motifs (sig1, sig2, and sig3). Six red boxes denote the flexible loop regions (F1–F6), as defined in the text. Blue triangles and circles indicate the residues that make hydrophilic and hydrophobic interactions with bound FMN, respectively. A red triangle indicates that the residue Gly281' from an adjacent subunit related by the Q-axis makes a hydrogen bond to FMN. Small green circles indicate ten residues (87–96) modeled as alanine due to poor electron density. The residue numbers are for chorismate synthase from *H. pylori*. HP\_aroC is for chorismate synthase from *H. pylori* (SWISS-PROT entry P56122), CJ\_aroC for *Campylobacter jejuni* (Q9PM41), HI\_aroC for *Haemophilus influenzae* (P43875), PA\_aroC for *Pseudomonas aeruginosa* (Q9I344), EC\_aroC for *E. coli* (P12008), and MJ\_aroC for *Methanococcus jannaschii* (Q58575).

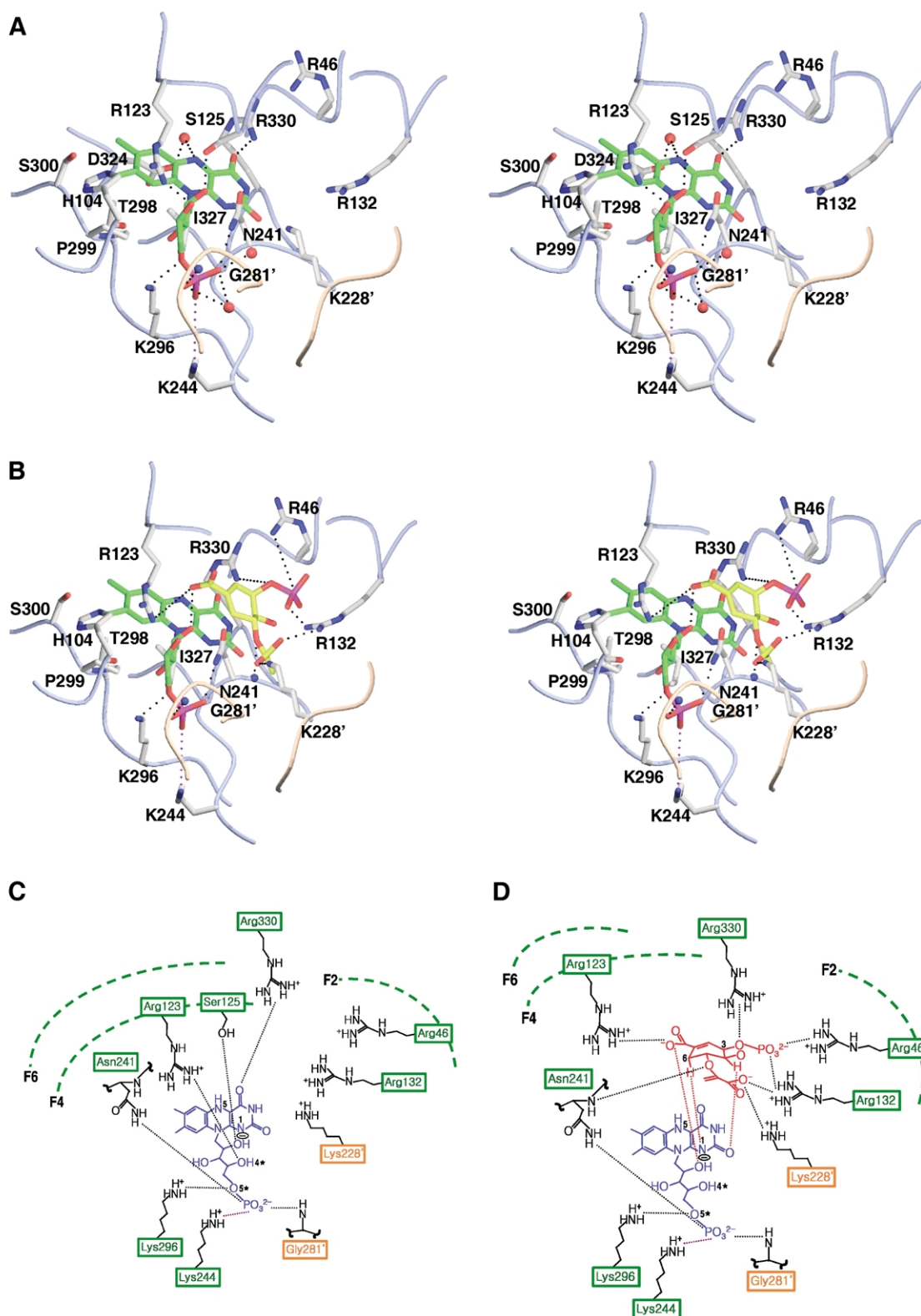
FMN molecules are non-covalently bound to a deep pocket and the bulk of it is buried within the protein, with only an edge of the isoalloxazine ring opposite to the ribityl chain being visible from the surface (Figures 2 and 4A). The electrostatic potential at the molecular surface is highly positive around the FMN binding sites (Figure 4A) and the six flexible loop regions F1–F6 are clustered around or near the bound cofactor

(Figure 4B and C). These flexible regions are rich in strictly or highly conserved residues (Figure 4B). The four FMN-binding sites in the tetramer are well separated from each other. The bound cofactor is most likely in the reduced state, since crystallization of the FMN-bound form was possible only in the presence of a high concentration (40 mM) of dithiothreitol and *E. coli* chorismate synthase binds the reduced FMN 1660-fold more





**Figure 4.** The active site and electron density of FMN. A, Electrostatic potential of the FMN-binding pocket. The positive electrostatic potential at the molecular surface of a monomer is colored in blue and the negative potential in red. Ten residues modeled as alanine are colored in dark yellow. FMN is shown in a stick model and colored in the same manner as in Figure 1A, except carbon atoms in white. B, Molecular surface of a chorismate synthase monomer covering the active site surrounded by six flexible regions. It is in the same orientation as in A. FMN is drawn in the same manner as in A. Strictly conserved residues and highly conserved residues on the molecular surface are colored in green and yellow, respectively. Each circle indicates six flexible regions, respectively. C, Stereo diagram of six flexible regions surrounding the active site in a tetramer. Only one of four active sites is emphasized. Thick traces denoting six flexible regions are colored as follows: F1 in cyan, F2 in green, F3 in yellow, F4 in red, F5 in violet, and F6 in blue. F5 is contributed from an adjacent subunit related by the *Q*-axis. Each monomer is colored in the same manner as in Figure 1A. D, Electron density of bound FMN. The  $2F_o - F_c$  electron density map, contoured at  $1.0\sigma$ , is superimposed on the refined model. E, A view from the edge of the isoalloxazine ring, showing its significant non-planarity.



**Figure 5.** FMN and EPSP-binding modes. A, Stereo diagram of FMN binding in the binary complex. FMN is drawn in a stick model and colored in the same manner as in Figure 1A, except carbon atoms in green. The residues interacting with FMN are also shown in stick models and colored in the same manner as in Figure 1A, except carbon atoms in light gray. Three red balls indicate water molecules. Black dotted lines represent hydrogen bonds. Purple dotted line denotes an ionic interaction between Lys244 and the phosphate group of FMN. The loops colored in light blue are from subunit A and the loop in light orange from the other subunit D, which is related by the Q-axis. Primed residue numbers denote that the residues come from the adjacent subunit. B, Stereo diagram of the FMN-EPSP-binding mode in the proposed model of the ternary complex. EPSP is colored in the same manner as FMN, except carbon atoms in yellow. Black and purple dotted lines are the same as in A. C, A schematic diagram of FMN binding in the



tightly than the oxidized FMN.<sup>29</sup> It is also likely to be the anionic form (FMN<sup>-</sup>) with the deprotonated N1 atom, since the N1 proton of the enzyme-bound reduced flavin has a significantly lower pK<sub>a</sub> value than that of the free reduced FMN.<sup>29</sup> The FMN-binding site in chorismate synthase is unique. The bulk of each FMN-binding site is formed by a monomer, with a small contribution from an adjacent monomer that is related by the molecular Q-axis (Figures 1A and 2). The residues located at the start of  $\alpha$ 2 (**His104**), the  $\alpha$ 2- $\alpha$ 3 loop (**Arg123**), the N terminus of  $\alpha$ 3 (**Ser125**), the N-terminal part of  $\beta$ 11 (**Asn241**, **Lys244**), the C terminus of  $\beta$ 16 (**Lys296**), the  $\beta$ 16- $\beta$ 17 loop (**Thr298**, **Pro299**, **Ser300**), and the N-terminal part of  $\alpha$ 8 (**Ile327**, **Arg330**), as well as the  $\beta$ 14'- $\beta$ 15' loop (**Gly281'**), contribute to the formation of the FMN-binding pocket (Figure 5A and C). The primed residue and secondary structures belong to the adjacent subunit related by the molecular Q-axis. The bold-faced residues are highly conserved (Figure 3).

A notable feature of the bound FMN is a significant deviation from planarity displayed by the isoalloxazine ring (Figure 4D and E). The pyrimidine ring and the dimethylbenzene ring make an angle of approximately 10° and the *re*-face of the isoalloxazine ring is convex. The convex *re*-face appears to provide the substrate EPSP-binding site, as discussed below. Non-planar conformations of the isoalloxazine ring have been previously observed in a few other flavoenzymes such as polyamine oxidase,<sup>30</sup> cholesterol oxidase,<sup>31</sup> and trimethylamine dehydrogenase.<sup>32</sup>

The isoalloxazine ring of FMN faces toward the molecular surface, with its edge opposite to the ribityl chain being exposed to the bulk solvent, whereas the hydrophilic ribityl chain and the phosphate group point away from the molecular surface and are deeply buried in the FMN-binding pocket (Figures 2 and 4A). The hydrophobic dimethylbenzene part of the isoalloxazine ring is positioned in a hydrophobic environment formed by residues **His104**, **Thr298**, **Pro299**, **Ser300**, and **Ile327** (Figure 5A). The bold-faced residues **His104**, **Thr298**, and **Ser300** are strictly conserved, while **Pro299** and **Ile327** are semi-conserved among six chorismate synthases, whose sequences are aligned in Figure 3. The O4 atom of the pyrimidine moiety forms a hydrogen bond with the NH2 atom of **Arg330**. The isoalloxazine N5 atom is not hydrogen-bonded to any protein atoms; it is hydrogen-bonded to a water molecule, which is in turn hydrogen-bonded to the side-chain of conserved **Asp324** (Figure 5A). In most flavin-depen-

dent enzymes, this N5 atom takes part in redox reactions and is within hydrogen-bonding distances from a hydrogen-bond donor, typically a backbone or a side-chain nitrogen atom.<sup>33</sup> The ribityl hydroxyl groups OH2\*, OH4\*, and OH5\* make hydrogen bonds to the side-chains of highly conserved residues **Arg123**, **Ser125**, and **Lys296**, respectively (Figure 5A). For the phosphate group, hydrogen bonds exist between the side-chain of conserved **Asn241** and the O2P oxygen atom, and between the main-chain nitrogen atom of conserved **Gly281'** from an adjacent subunit and O3P oxygen. Two buried water molecules also form hydrogen bonds to the phosphate group. The conserved residue **Lys244** makes a weak salt-bridge with the phosphate group. As far as we are aware, the observed FMN-binding mode is distinct in that both the ribityl chain and the phosphate group are well buried within the protein.

In general, the proximity of a positive charge to the isoalloxazine ring is thought to increase the redox potential of FMN, whereas a negative charge or a hydrophobic environment is expected to lower it.<sup>34</sup> In chorismate synthase, a hydrophobic environment surrounds the dimethylbenzene moiety of the isoalloxazine ring and three of the four positively charged residues (**Arg123**, **Lys244**, and **Lys296**) in the FMN-binding pocket interact with the ribityl chain and the phosphate group, but not with the isoalloxazine ring. The side-chain of remaining **Arg330** is in close contact with the N3-C4=O4 locus of the pyrimidine moiety. A positive charge at this location can possibly stabilize negatively charged N1 in the reduced flavin, thus keeping the electron until the substrate EPSP is bound. The presence of a positively charged residue near the N3-C4=O4 locus has been known to serve two purposes: to favor flavin binding and to regulate redox properties of the cofactor.<sup>33</sup> Therefore, **Arg330** appears to play a key role in binding and activation of reduced FMN in chorismate synthase.

### A model for the substrate binding

In order to shed light on the functional roles of highly conserved residues in the active site, we have attempted to model the substrate binding, which was assisted by the following information. First, it is most probable that EPSP is in a half-chair conformation such that the bonds to be broken are pseudo-axial.<sup>18</sup> Second, a theoretical study of the chorismate synthase reaction provided some information about the geometry of FMN-EPSP complex.<sup>35</sup> The energy minimum of the

binary complex. FMN is drawn in blue lines and three broken lines indicate three flexible loop regions F2, F4, and F6. Residues boxed in green color belong to subunit A and those in orange color to the other subunit D. Black and purple dotted lines are the same as in A. The reduced FMN is drawn in the mono-anionic form. D, A schematic diagram of the FMN-EPSP-binding mode in the proposed ternary complex model. EPSP is drawn in red lines. Black and purple dotted lines are the same as in A. Red dotted lines indicate interactions between FMN and EPSP that have been suggested by density functional calculations.<sup>35</sup>

supermolecule formed between FMNH<sup>-</sup> and *cis*-3,4-dihydroxycyclohexene-1-carboxylic acid-3-phosphate (DHCCP), which lacks the C(5)-enol-pyruvate side-chain of EPSP, indicated that the carbocyclic ring of DHCCP is located on the *re*-face of FMN. This study also indicated that the C-6 hydrogen is unusually acidic and may require only a weak base for deprotonation. Thirdly, when a ternary complex is formed, the environment of the bound FMN becomes more hydrophobic.<sup>28</sup> Fourthly, we expect the substrate EPSP to interact with a number of positively charged residues, since it is a highly negatively charged molecule. We could place EPSP on the convex *re*-face of FMN (Figure 5B) without a major structural alteration of the active site except a small movement of the flexible loop F4 around Ser125. Accommodation of EPSP on the *re*-face of FMN is made more favorable by bending of the isoalloxazine ring. This ternary complex model does not rule out a possibility of more extensive conformational changes in the enzyme, which may involve changes at the inter-subunit interface and in some of the other flexible loops.

In the proposed model for the chorismate synthase-FMN-EPSP ternary complex, EPSP makes the following interactions with FMN (Figure 5(B) and (D)), which have been suggested by the density functional calculations on a complex between an EPSP analog and reduced FMN.<sup>35</sup> The C(1)-carboxylate group of EPSP is hydrogen-bonded with O2<sup>\*</sup> hydroxyl in the ribityl chain. The breaking C(6)-H bond of EPSP and N1 of deprotonated flavin make a bonding interaction. The C-4 hydrogen atom of EPSP interacts with the C2 carbonyl group of FMN. EPSP might make additional interactions with protein residues, most of which are highly conserved arginine and lysine residues. The C(1)-carboxylate group of EPSP forms a salt-bridge with the positively charged side-chain of Arg123. The negative charges on the C(3)-O-PO<sub>3</sub><sup>-</sup> locus of EPSP are compensated for by the positive charges on Arg46, Arg132, and Arg330. This environment, together with the formation of hydrogen bonds between the C(3)-phosphate group of EPSP and these arginine residues, would be expected to promote phosphorolysis at C-3. The negatively charged carboxylate group of the carboxyvinyl moiety in EPSP could interact with the positively charged residues Arg132 and Lys228', with Lys228' coming from the adjacent Q-axis related subunit. The C-5 oxygen of EPSP is likely to be hydrogen-bonded to the backbone nitrogen atom of Asn241. Compared to the binary complex structure, Ser125 on the flexible loop F4 is expected to shift its position in the ternary complex model such that it would no longer interact with FMN, and the side-chains of Arg123 and Arg330 would interact with EPSP instead of FMN (Figure 5). Interestingly, nine of the ten residues that interact or are proposed to interact with the cofactor and the substrate are also conserved in a

bifunctional chorismate synthase from *N. crassa*. The corresponding residues are Arg46, Arg125, Ser127, Arg134, Lys246', Pro259, Lys262, Gly323', Lys338, and Arg373, where Pro259 in the *N. crassa* enzyme replaces Asn241 of *H. pylori* chorismate synthase.

### Implications for the reaction mechanism

Several lines of evidence favor a non-concerted radical mechanism for the reaction catalyzed by chorismate synthase.<sup>35</sup> All the proposed radical mechanisms begin with one-electron transfer from FMNH<sup>-</sup> to EPSP, followed by loss of phosphate to form an allylically stabilized radical intermediate.<sup>20</sup> A flavin intermediate, which has an increased absorbance at 400 nm, is formed transiently during the reaction.<sup>36</sup> After the bond-cleavage steps and phosphate release, the reduced FMN decays back to the anionic form.<sup>18</sup> Our structure of chorismate synthase in the FMN-bound form and the proposed model of the ternary complex with the substrate illuminate how the reaction may proceed.

The anionic form of reduced FMN binds first to the apo enzyme and then the substrate EPSP binds to the binary complex to form a productive ternary complex. The order of these early events was indicated by previous spectroscopic and kinetic studies.<sup>28,36</sup> Binding of EPSP as well as possible protein conformational changes involving movements in some of the flexible loop regions result in a more apolar environment for the bound cofactor. A key question regarding the reaction mechanism is how the enzyme substantially lowers the reduction potential of the bound FMN in order to allow the unusual *anti*-1,4-elimination of phosphate to occur without undergoing decarboxylation. The structural changes in the active site as well as close proximity to the negative charges in EPSP may lower the redox potential of the bound FMN so that an electron transfers spontaneously from the cofactor to EPSP.<sup>29</sup> In this process, Arg330 may play a key role as a possible switch that controls the redox property of the enzyme-bound FMN. As a result, the semiquinone FMN<sub>sq</sub> and an allylic radical intermediate of the substrate are formed. This explanation is consistent with the previously proposed "substrate-induced cofactor activation".<sup>36</sup> A cluster of the positively charged Arg46, Arg132, and Arg330 likely contribute to the dissociation of the phosphate group from EPSP and the C(3)-O bond is cleaved. In summary, activation of reduced FMN upon binding of EPSP, an electron transfer from the reduced flavin to EPSP, and cleavage of the phosphate group assisted by the positively charged residues occur sequentially, leading to the formation of flavin semiquinone and the allylic radical intermediate.

After phosphorolysis at C-3, the interaction between Arg330 and C(3)-O oxygen disappears and the negative charge of phosphate group dissociated from EPSP can no longer contribute to

lowering the redox potential of the flavin. Furthermore, the positively charged side-chain of Arg330 may move to interact with the C-4 carbonyl group of FMN<sub>sq</sub> again. These changes in the environment surrounding FMN<sub>sq</sub> raise the redox potential so that the flavin semiquinone accepts an electron from the allylic radical intermediate and is reversed back to the reduced flavin. Concurrently or in the following step, an additional interaction between the breaking C(6)–H bond of EPSP and the deprotonated N1 atom of reduced flavin, which serves as the base, results in the *anti*-1,4-elimination leading to the formation of chorismate. Alternatively, the flavin semiquinone could simply remove C(6)–H hydrogen as a radical from the allylic radical intermediate to form chorismate directly. Both possibilities explain the direct role of reduced flavin in catalysis. After chorismate is released from the active site, the environment surrounding the isoalloxazine ring of the bound cofactor becomes less apolar again and FMN is deprotonated at N1, with Arg330 playing a key role in stabilizing the mono-anionic form of the reduced FMN. The enzyme returns to the initial binary complex state and is now ready to start the next cycle of catalysis.

## Conclusions

We have determined the first crystal structure of chorismate synthase. It reveals several interesting features. First, each monomer is folded into a novel three-layered,  $\beta$ - $\alpha$ - $\beta$  sandwich fold in its core. Second, highly conserved regions that include several flexible loops cluster together to form the active site with a unique FMN-binding pocket. Binding of FMN causes little overall structural changes except in three flexible loop regions (F2–F4). Thirdly, most part of the cofactor is bound deeply within the protein, and conserved residues Arg123, Ser125, Asn241, Lys244, Gly281', Lys296, and Arg330 interact with the cofactor. Arg330 seems to play an important role in controlling the redox property of the bound FMN. Fourthly, we have proposed a model for EPSP binding to the *re*-face of the isoalloxazine ring of FMN. Highly conserved arginine and lysine residues Arg46, Arg123, Arg132, and Lys228 appear to play important roles in EPSP recognition.

## Materials and Methods

### Crystallization and data collection

SeMet-substituted chorismate synthase from *H. pylori* with a C-terminal purification tag (LEHHHHHH) was expressed in *E. coli* B834(DE3) cells, using the M9 cell culture medium containing SeMet and 14 other amino acids. During purification, 10 mM dithiothreitol was added. Crystals of the SeMet-substituted enzyme were grown in an FMN-free form under the same conditions as the native protein,<sup>27</sup> except for the presence of 5 mM

dithiothreitol. The SeMet-substituted crystals belong to the space group *I4*, with unit cell dimensions of  $a = b = 146.42$  Å and  $c = 132.06$  Å. A set of three-wavelength MAD data was collected at 100 K from a flash-frozen crystal of the SeMet-substituted protein on an Area Detector Systems Corporation Quantum 4R CCD detector at the experimental station BL-18B of Photon Factory, Japan. Raw data were processed and scaled with the programs MOSFLM and SCALA.<sup>37</sup> The SeMet crystals were partially merohedrally twinned<sup>27</sup> but the structure could be solved readily without detwinning the data.

Crystals of the native protein in a binary complex with FMN were grown in the presence of FMN in a fivefold molar excess over the protein monomer. The reservoir condition was the same as the native crystals in the apo-form,<sup>27</sup> except for the presence of 40 mM dithiothreitol. Data from a crystal of the binary complex were collected at 100 K using synchrotron radiation from beamline 6B at Pohang Light Source, Korea. The binary complex crystals belong to the same space group *I4*, with unit cell dimensions of  $a = b = 146.69$  Å and  $c = 132.31$  Å. Crystals of the native protein in an FMN-free form were grown under conditions as described.<sup>27</sup> These crystals have unit cell dimensions of  $a = b = 146.48$  Å and  $c = 132.06$  Å.

### Structure solution and refinement

Of the 36 expected selenium sites, 35 were located with the program SOLVE,<sup>38</sup> except the selenium site at the N terminus of the protein. The phases were then improved by density modification using RESOLVE.<sup>39</sup> The mean figure of merit was 0.49/0.65 before/after RESOLVE in the resolution range of 20–3.0 Å. The resulting electron density map was readily interpretable and allowed model building of an initial model that accounted for ~80% of the backbone of the polypeptide chain. The model was refined with the program CNS<sup>40</sup> and manual model building was done using the program O.<sup>41</sup> Subsequently, the structure of the binary complex in the FMN-bound form, as well as the FMN-free form, was refined using the model of the SeMet-substituted enzyme. FMN and water molecules were assigned based on the  $F_o - F_c$  maps calculated with model phases. In the  $2F_o - F_c$  map, it was evident that the flavin ring of FMN was slightly bent. Therefore, the planarity restraint was applied to each of the dimethylbenzene and pyrimidine rings only, but not to the central ring of the isoalloxazine moiety. Table 1 summarizes data collection and refinement statistics.

### Sequence alignment and graphics

A BLAST search against SWISS-PROT database† with the amino acid sequence of *H. pylori* chorismate synthase picked up more than 70 other members of chorismate synthase from various sources. They were aligned to decide the conserved residues. Six selected sequences are aligned in Figure 3, which was generated using the programs AMPS<sup>42</sup> and ALSCRIPT.<sup>43</sup> The structure Figures were drawn with the programs MOLSCRIPT,<sup>44</sup> BOBSCRIPT,<sup>45</sup> and Raster3D.<sup>46</sup> Figures of the molecular surface were prepared with the program GRASP.<sup>47</sup>

† <http://www.expasy.org>



### Protein data bank accession numbers

The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank under accession codes 1UM0 (binary complex with FMN) and 1UMF (FMN-free form).

### Acknowledgements

We thank Professor N. Sakabe and his staff for assistance during data collection at Photon Factory, beamline BL-18B. We also thank Dr H. S. Lee and his staff during data collection at Pohang Light Source, beamline 6B. This work was supported by grants from the 21st Century Frontier Program for Functional Analysis of Human Genome and from the Korea Ministry of Science and Technology (NRL-2001, grant no. M10318000132). H.J.A., B.I.L., & H.J.Y. are recipients of the BK21 fellowship.

### References

- Herrmann, K. M. & Weaver, L. M. (1999). The shikimate pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 473–503.
- Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, J. R. *et al.* (1998). Evidence for the shikimate pathway in apicomplexan parasites. *Nature*, **393**, 801–805.
- Davies, G. M., Barrett-Bee, K. J., Jude, D. A., Lehan, M., Nichols, W. W., Pinder, P. E. *et al.* (1994). (6S)-6-fluoroshikimic acid, an antibacterial agent acting on the aromatic biosynthetic pathway. *Antimicrob. Agents Chemother.* **38**, 403–406.
- Kishore, G. M. & Shah, D. M. (1988). Amino acid biosynthesis inhibitors as herbicides. *Annu. Rev. Biochem.* **57**, 627–663.
- Steinrücken, H. C. & Armhein, N. (1980). The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase. *Biochem. Biophys. Res. Commun.* **94**, 1207–1212.
- Wagner, T., Shumilin, I. A., Bauerle, R. & Kretsinger, R. H. (2000). Structure of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*: comparison of the  $Mn^{2+}$ -2-phosphoglycolate and the  $Pb^{2+}$ -2-phosphoenolpyruvate complexes and implications for catalysis. *J. Mol. Biol.* **301**, 389–399.
- Carpenter, E. P., Hawkins, A. R., Frost, J. W. & Brown, K. A. (1998). Structure of dehydroquinase synthase reveals an active site capable of multistep catalysis. *Nature*, **394**, 299–302.
- Gourley, D. G., Shrive, A. K., Polikarpov, I., Krell, T., Coggins, J. R., Hawkins, A. R. *et al.* (1999). The two types of 3-dehydroquinase have distinct structures but catalyze the same overall reaction. *Nature Struct. Biol.* **6**, 521–525.
- Benach, J., Lee, I., Edstrom, W., Kuzin, A. P., Chiang, Y., Acton, T. B. *et al.* (2003). The 2.3-Å crystal structure of the shikimate 5-dehydrogenase orthologue YdiB from *Escherichia coli* suggests a novel catalytic environment for an NAD-dependent dehydrogenase. *J. Biol. Chem.* **278**, 19176–19182.
- Michel, G., Roszak, A. W., Sauve, V., Maclean, J., Matte, A., Coggins, J. R. *et al.* (2003). Structures of shikimate dehydrogenase AroE and its paralog YdiB. A common structural framework for different activities. *J. Biol. Chem.* **278**, 19463–19472.
- Romanowski, M. J. & Burley, S. K. (2002). Crystal structure of the *Escherichia coli* shikimate kinase I (AroK) that confers sensitivity to mecillinam. *Proteins: Struct. Funct. Genet.* **47**, 558–562.
- Krell, T., Coggins, J. R. & Laphorn, A. J. (1998). The three-dimensional structure of shikimate kinase. *J. Mol. Biol.* **278**, 983–997.
- Stallings, W. C., Abdel-Meguid, S. S., Lim, L. W., Shieh, H. S., Dayringer, H. E., Leimgruber, N. K. *et al.* (1991). Structure and topological symmetry of the glyphosate target 5-enolpyruvylshikimate-3-phosphate synthase: a distinctive protein fold. *Proc. Natl Acad. Sci. USA*, **88**, 5046–5050.
- Schönbrunn, E., Eschenburg, S., Shuttleworth, W. A., Schloss, J. V., Amrhein, N., Evans, J. N. & Kabsch, W. (2001). Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc. Natl Acad. Sci. USA*, **98**, 1376–1380.
- Balasubramanian, S., Coggins, J. R. & Abell, C. (1995). Observation of a secondary tritium isotope effect in the chorismate synthase reaction. *Biochemistry*, **34**, 341–348.
- Bornemann, S., Ramjee, M. K., Balasubramanian, S., Abell, C., Coggins, J. R., Lowe, D. J. & Thorneley, R. N. F. (1995). *Escherichia coli* chorismate synthase catalyzes the conversion of (6S)-6-fluoro-5-enolpyruvylshikimate-3-phosphate to 6-fluorochorismate. Implications for the enzyme mechanism and the antimicrobial action of (6S)-6-fluoroshikimate. *J. Biol. Chem.* **270**, 22811–22815.
- Bornemann, S., Lowe, D. J. & Thorneley, R. N. F. (1996). The transient kinetics of *Escherichia coli* chorismate synthase: substrate consumption, product formation, phosphate dissociation, and characterization of a flavin intermediate. *Biochemistry*, **35**, 9907–9916.
- Bornemann, S., Theoclitou, M. E., Brune, M., Webb, M. R., Thorneley, R. N. F. & Abell, C. (2000). A secondary  $\beta$  deuterium kinetic isotope effect in the chorismate synthase reaction. *Bioorg. Chem.* **28**, 191–204.
- Osborne, A., Thorneley, R. N. F., Abell, C. & Bornemann, S. (2000). Studies with substrate and cofactor analogues provide evidence for a radical mechanism in the chorismate synthase reaction. *J. Biol. Chem.* **275**, 35825–35830.
- Macheroux, P., Schmid, J., Amrhein, N. & Schaller, A. (1999). A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. *Planta*, **207**, 325–334.
- Kitzing, K., Macheroux, P. & Amrhein, N. (2001). Spectroscopic and kinetic characterization of the bifunctional chorismate synthase from *Neurospora crassa*: evidence for a common binding site for 5-enolpyruvylshikimate 3-phosphate and NADPH. *J. Biol. Chem.* **276**, 42658–42666.
- Bornemann, S. (2002). Flavoenzymes that catalyze reactions with no net redox change. *Nature Prod. Rep.* **19**, 761–772.
- Holm, L. & Sander, C. (1994). Searching protein structure databases has come of age. *Proteins: Struct. Funct. Genet.* **19**, 165–173.
- Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins

- database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540.
25. Albert, A., Martinez-Ripoll, M., Espinosa-Ruiz, A., Yenush, L., Culianez-Macia, F. A. & Serrano, R. (2000). The X-ray structure of the FMN-binding protein AtHal3 provides the structural basis for the activity of a regulatory subunit involved in signal transduction. *Structure*, **8**, 961–969.
  26. Sibanda, B. L. & Thornton, J. M. (1985).  $\beta$ -Hairpin families in globular proteins. *Nature*, **316**, 170–174.
  27. Ahn, H. J., Yang, J. K., Lee, B. I., Yoon, H. J., Kim, H. W. & Suh, S. W. (2003). Crystallization and preliminary X-ray crystallographic studies of chorismate synthase from *Helicobacter pylori*. *Acta Crystallog. sect. D*, **59**, 569–571.
  28. Macheroux, P., Schönbrunn, E., Svergun, D. I., Volkov, V., Koch, M. H., Bornemann, S. & Thorneley, R. N. F. (1998). Evidence for a major structural change in *Escherichia coli* chorismate synthase induced by flavin and substrate binding. *Biochem. J.* **335**, 319–327.
  29. Macheroux, P., Petersen, J., Bornemann, S., Lowe, D. J. & Thorneley, R. N. F. (1996). Binding of the oxidized, reduced, and radical flavin species to chorismate synthase. An investigation by spectrophotometry, fluorimetry, electron paramagnetic resonance, and electron nuclear double resonance spectroscopy. *Biochemistry*, **35**, 1643–1652.
  30. Binda, C., Coda, A., Angelini, R., Federico, R., Ascenzi, P. & Mattevi, A. (1999). A 30 Å long U-shaped catalytic tunnel in the crystal structure of polyamine oxidase. *Structure*, **7**, 265–276.
  31. Yue, Q. K., Kass, I. J., Sampson, N. S. & Vrielink, A. (1999). Crystal structure determination of cholesterol oxidase from *Streptomyces* and structural characterization of key active site mutants. *Biochemistry*, **38**, 4277–4286.
  32. Barber, M. J., Neame, P. J., Lim, L. W., White, S. & Matthews, F. S. (1992). Correlation of X-ray deduced and experimental amino acid sequences of trimethylamine dehydrogenase. *J. Biol. Chem.* **267**, 6611–6619.
  33. Fraaije, M. W. & Mattevi, A. (2000). Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem. Sci.* **25**, 126–132.
  34. Ghisla, S. & Massey, V. (1989). Mechanisms of flavo-protein-catalyzed reactions. *Eur. J. Biochem.* **181**, 1–17.
  35. Dmitrenko, O., Wood, H. B., Bach, R. D. & Ganem, B. (2001). A theoretical study of the chorismate synthase reaction. *Org. Letters* **3**, 4137–4140.
  36. Macheroux, P., Bornemann, S., Ghisla, S. & Thorneley, R. N. F. (1996). Studies with flavin analogs provide evidence that a protonated reduced FMN is the substrate-induced transient intermediate in the reaction of *Escherichia coli* chorismate synthase. *J. Biol. Chem.* **271**, 25850–25858.
  37. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
  38. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
  39. Terwilliger, T. C. (2000). Maximum-likelihood density modification. *Acta Crystallog. sect. D*, **56**, 965–972.
  40. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
  41. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
  42. Barton, G. J. (1990). Protein multiple sequence alignment and flexible pattern matching. *Methods Enzymol.* **183**, 403–428.
  43. Barton, G. J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.
  44. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
  45. Esnouf, R. M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallog. sect. D*, **55**, 938–940.
  46. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photo-realistic molecular graphics. *Methods Enzymol.* **277**, 505–524.
  47. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
  48. Brünger, A. T. (1992). The free R-value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–474.

Edited by I. Wilson

(Received 20 October 2003; received in revised form 18 December 2003; accepted 23 December 2003)

*Note added in proof.* After submission of this manuscript, three crystal structures of chorismate synthase have been reported: chorismate synthase from *Streptococcus pneumoniae* as a ternary complex with oxidized FMN and EPSP (*Structure* (2003). **11**, 1499–1511), chorismate synthase from *Aquifex aeolicus* in an apo form (*Proteins: Struct. Funct. Genet.* (2004). **54**, 166–169), and a bifunctional chorismate synthase from *Saccharomyces cerevisiae* in an apo form (*J. Biol. Chem.* (2004). **279**, 619–625).