Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 5046–5050, June 1991
Biochemistry

Structure and topological symmetry of the glyphosate target 5-*enol*-pyruvylshikimate-3-phosphate synthase: A distinctive protein fold

(x-ray crystallography/aromatic amino acid biosynthesis/herbicide/helical macrodipole)

WILLIAM C. STALLINGS*, SHERIN S. ABDEL-MEGUID*, LOUIS W. LIM*, HUEY-SHENG SHIEH*, HENRY E. DAYRINGER*, NANCY K. LEIMGRUBER*, RODERICK A. STEGEMAN*, KAREN S. ANDERSON†, JAMES A. SIKORSKI†, STEPHEN R. PADGETTE†, AND GANESH M. KISHORE†

*Monsanto Corporate Research and †Technology Division of Monsanto Agricultural Company, Monsanto Company, 700 Chesterfield Village Parkway, St. Louis, MO 63198

Communicated by J. E. Varner, January 28, 1991

ABSTRACT **5-*enol*-Pyruvylshikimate-3-phosphate synthase (EPSP synthase; phospho*enol*pyruvate:3-phosphoshikimate 1-carboxyvinyltransferase, EC 2.5.1.19) is an enzyme on the pathway toward the synthesis of aromatic amino acids in plants, fungi, and bacteria and is the target of the broad-spectrum herbicide glyphosate. The three-dimensional structure of the enzyme from *Escherichia coli* has been determined by crystallographic techniques. The polypeptide backbone chain was traced by examination of an electron density map calculated at 3-Å resolution. The two-domain structure has a distinctive fold and appears to be formed by 6-fold replication of a protein folding unit comprising two parallel helices and a four-stranded sheet. Each domain is formed from three of these units, which are related by an approximate threefold symmetry axis; in each domain three of the helices are completely buried by a surface formed from the three β-sheets and solvent-accessible faces of the other three helices. The domains are related by an approximate dyad, but in the present crystals the molecule does not display pseudo-symmetry related to the symmetry of point group *32* because its approximate threefold axes are almost normal. A possible relation between the three-dimensional structure of the protein and the linear sequence of its gene will be described. The topological threefold symmetry and orientation of each of the two observed globular domains may direct the binding of substrates and inhibitors by a helix macrodipole effect and implies that the active site is located near the interdomain crossover segments. The structure also suggests a rationale for the glyphosate tolerance conferred by sequence alterations.**

Herbicides have become an integral part of modern agriculture providing cost-effective and reliable weed control during crop production. The herbicide Round-up (Monsanto) has outstanding toxicological, environmental, and herbicidal properties (1, 2). Glyphosate, its active ingredient, inhibits 5-*enol*-pyruvylshikimate-3-phosphate synthase‡ (EPSP synthase; phospho*enol*pyruvate:3-phosphoshikimate 1-carboxyvinyltransferase, EC 2.5.1.19), an enzyme leading to the biosynthesis of aromatic compounds in plants and microbes. EPSP synthase catalyzes an unusual, reversible condensation of shikimate 3-phosphate (S3P) and phospho*enol*pyruvate (PEP) with transfer of the carboxyvinyl moiety from PEP to S3P. Glyphosate is a competitive inhibitor with respect to PEP and an uncompetitive inhibitor with respect to S3P (3, 4).

Introduction of glyphosate tolerance into crop plants is of significant academic and commercial interest (5). Recent developments in plant genetic engineering permit the introduction and expression of genes in a wide range of crop plants (6). Expression of glyphosate-tolerant EPSP synthase variants in plants has been demonstrated to confer tolerance to the herbicide (7). The P101S EPSP synthase from *Salmonella typhimurium* shows a 3-fold increase in K_i for glyphosate (8, 9), but most glyphosate-tolerant EPSP synthases have a significant alteration in K_m for PEP; for example, the K_i for glyphosate of the G96A *Escherichia coli* (SM-1) variant is increased 8000-fold relative to the wild-type enzyme, and this is accompanied by 13-fold increase in K_m for PEP (5). An EPSP synthase variant ideal for genetically engineering glyphosate tolerance in crop plants would have no alteration in k_{cat} and K_m for substrates and >10-fold enhancement in the K_i for glyphosate.

EPSP synthase has been extensively investigated by using chemical and biophysical approaches (10–16); correlation of these data with the three-dimensional structure of the enzyme is essential for a detailed description of the reaction mechanism and its inhibition by glyphosate. Understanding how substrates and the inhibitor, glyphosate, bind to EPSP synthase would facilitate both the design of glyphosate-tolerant variants and the targeted synthesis of new herbicides. As a first step toward understanding the interaction of glyphosate with EPSP synthase, we have crystallized the enzyme from *E. coli* and determined its structure at 3-Å resolution.§

METHODS

E. coli EPSP synthase was isolated from a cloned *E. coli* strain that overproduces the protein, and the enzyme was purified as described (18, 19). Crystals were grown by the method of vapor diffusion in sitting or hanging drops (20) by using 35–45% saturated ammonium sulfate buffered at pH 7.5–8.5 with 100 mM Tris as the precipitant. The protein concentration was adjusted to 15 mg/ml in 29 mM 2-mercaptoethanol buffered at pH 7.5 with 10 mM Tris. The enzyme crystallizes in the orthorhombic space group $P2_12_12_1$ with $a = 92.1$, $b = 83.2$, and $c = 71.7$ Å and with one molecule of the monomeric protein in the crystallographic asymmetric unit. X-ray diffraction data were measured from native and derivative crystals with a dual chamber multiwire detector system (21). The structure was determined by the multiple-isomorphous replacement method (22) using five mercurial derivatives (22, ¶) that all bind at three sites with varying

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EPSP synthase, 3-*enol*-pyruvylshikimate-5-phosphate synthase; S3P, shikimate 3-phosphate; PEP, phospho*enol*pyruvate.

‡Official enzyme nomenclature renumbers the Shikimate ring in the reverse direction; hence the systematic name exchanges the substituents at positions 3 and 5.

§The atomic coordinates have been deposited in the Protein Data Bank (17), Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (reference code 1EPS).

¶The five heavy atom derivatives were prepared by soaking native EPSP synthase crystals in 1 mM ethyl mercurithiosalicylate, 1 mM methyl mercury (II) chloride, 1 mM mercury (II) acetate, saturated mercuriphenyl glyoxal, and saturated dimercuriacetate (23). The latter was a gift from Robert Fletterick of the University of California, San Francisco.

occupancies. The resulting electron density map was further improved by using the density modification procedure described by Wang (24) with 0.84 as the figure-of-merit of the solvent-leveled map.

RESULTS AND DISCUSSION

The chain tracing of a single molecule of *E. coli* EPSP synthase is illustrated in Fig. 1. The structure consists of two distinct globular domains, which are roughly hemispheric, each with a radius of about 25 Å. Both the amino and the carboxyl terminus of the chain are located in the lower domain with two polypeptide crossover segments clearly defined in the 3-Å electron density map. Distributed about the axis of each hemisphere are three buried parallel helices of three to four turns each; the domain surfaces comprise three β-sheets, most with four strands, and three parallel helices that have solvent accessible faces. These constitute the principal elements of secondary structure in the EPSP synthase molecule. The backbone folds of the two domains are therefore remarkably similar. Superposition of α-carbon atoms from one domain on those of the other yields a 2.1-Å rms deviation when 131 atoms (of 208 atoms in each domain) are considered equivalent and used in the least-squares calculation of the transformation matrix (26). These results initially suggested that EPSP synthase is the product of gene duplication. However, the structure displays higher than twofold topological symmetry.

As shown in Fig. 2, in each domain an approximate threefold axis of symmetry relates the principal elements of secondary structure and generates three subdomains or protein folding units. The topology of each of these folding units is illustrated in Fig. 3. The four-stranded β-sheet structures contain both parallel and antiparallel strands; the helices are parallel with equivalent chain polarities. Among the six folding units in the structure, three are individually formed from continuous segments of the polypeptide chain; the connectivity within each of these units is shown schematically in Fig. 3 *Left* and three-dimensionally in Fig. 3 *Right*. As illustrated in Fig. 4, these three folding units are subdomains 4, 5, and 6. The other three units are not formed from continuous polypeptide chain. Folding unit 1 comprises both the amino and carboxyl terminus. This is the only example of an incomplete folding unit that we observe in the EPSP synthase structure; the strand corresponding to strand **e** in Fig. 3 *Left* is missing. Folding units 2 and 3, although complete, are formed from discontinuous polypeptide chain segments (Fig. 4).

The backbone chain fold of EPSP synthase is therefore not that of a simple symmetric molecule formed by sequential repetition of a folding motif. Were this the case, it would be elegant to imagine that the structure had evolved via 6-fold replication of a primordial gene for the folding unit. As explained below, however, close examination of the order in which the secondary structural elements of the folding units occur along the sequence suggests that such a mechanism is still possible. In Fig. 3 *Left*, the strands and helices of the folding unit are assigned letters **a** through **f** in the order in which they appear in each of the three examples connected by continuous polypeptide chain. In Fig. 4 *Right*, the elements of secondary structure are expressed in the order in which they appear along the chain of the folded protein. Although the elements of secondary structure of folding units 2 and 3 are from discontinuous segments of the chain (Fig. 4 *Left*), in the coding segment of the gene of *E. coli* EPSP synthase they are joined by the two pentapeptide interdomain crossover segments (Fig. 4 *Right*). At the first crossover, the union of the amino terminus of unit 2 and the carboxyl terminus of unit 3 yields continuous expression of secondary structural elements **a** through **f**. Likewise, crossback from folding unit 3 to unit 2 is also accomplished with related continuity.

Consistent with the blocking shown in Fig. 4 *Right*, the discontinuous DNA coding sequences for folding units 2 and 3 may have evolved from ancestral genes for originally independent units that later incorporated the messages for the domain crossovers, even though the elements of secondary structure in these units are now formed from discontinuous polypeptide chain segments. Thus, the gene sequence for the enzyme should exhibit six consecutive homologous regions. We are examining the protein and gene sequences of plant (27, 28), fungal (29, 30), and bacterial (9, 31, 32) EPSP synthases in efforts to detect linear evidence for the three-dimensionally observed pseudo-symmetry. The searches will intensify when side-chain positions are more firmly established and coordinates for the fully refined structure are suitable for deposition in the Protein Data Bank (17).

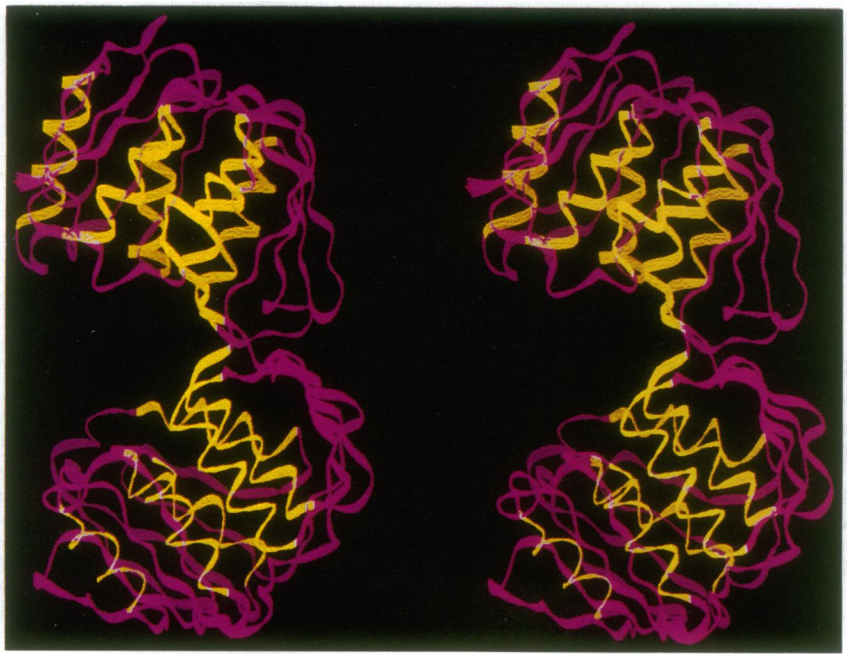


FIG. 1. Stereo ribbon diagram of the *E. coli* EPSP synthase polypeptide fold as seen in the 3-Å electron density map. The molecule folds into two globular hemispheric domains, each with a radius of about 25 Å. The domains are linked by two crossover chain segments with both the amino and carboxyl termini of the protein in the lower domain. The clarity of the 3-Å electron density map in these regions and observation of the topological symmetry and approximate equivalence of the two domains (see text) indicate that the linkages between principal secondary structural elements and the overall fold of the molecule are reported correctly. The two flat surfaces of the hemispheres, which in projection form a "V," are almost normal and accommodate the amino termini of the six helices in each domain. Helical macrodipolar effects (25) thereby can create a potential well that would guide anion ligands to an active site near the intersection at the domain crossovers.

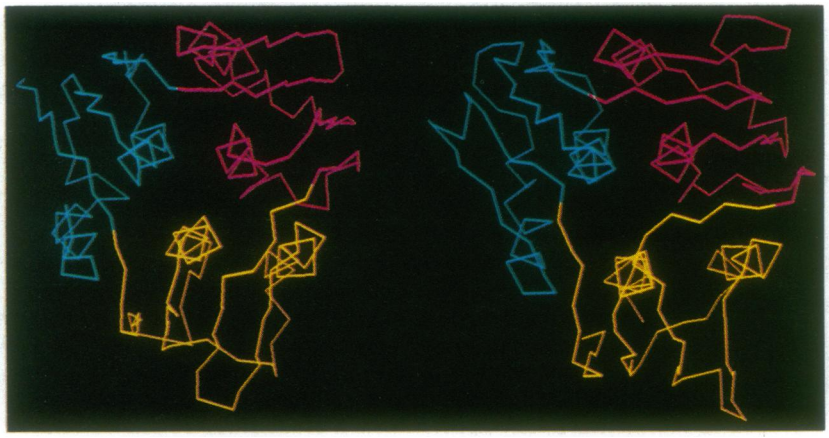


FIG. 2. Views down the axes of the two hemispheric domains illustrate the general equivalence of their secondary structural features. In both domains, three buried parallel helices are surrounded by a surface that includes three helices with solvent-accessible faces and three β-sheets, most with four strands. In each domain an approximate threefold symmetry axis, parallel to the helix axes, appears to relate the principal elements of secondary structure.

It is also interesting to consider the relationship between the folded backbone structure and the molecular event that it catalyzes. The reaction is selectively inhibited by glyphosate (4) with a K_i of 0.16 μM (12), and the mechanism has been shown recently to involve a stable tetrahedral intermediate (33–36) as shown in Scheme 1.

In the EPSP synthase reaction, reactants, products, the tetrahedral intermediate, and the inhibitor glyphosate are all multiply charged anions. It is reasonable to speculate that the binding of anion ligands at the active site may be electrostatically facilitated by a helical macrodipole effect (25), which results from the spatial orientation of the two domains. The flat surfaces of the hemispheric domains are almost normal, and the amino termini of the 12 helices in the molecule are all positioned near these two planes (Fig. 1). Positive charge should accumulate on the two surfaces with dipolar contributions from each helix. The generated field would have a gradient, created by the perpendicular orientation of the two planes, with maximal positive charge density at their intersection near the domain crossovers. We postulate that an anion in the vicinity of an EPSP synthase molecule should be attracted and guided to an active site near the crossover region. Model building and side-chain placement, as well as the structure elucidation of enzyme-ligand complexes, will provide ancillary data to confirm or to disprove this hypothesis, but our observations also suggest experiments that investigate the ionic-strength dependence of the rate constants in the well-characterized steps (13) of the reaction mechanism.

Side-chain placement has been difficult in the 3-Å map and will require data at higher resolution; the present crystals show order to about 2.1-Å resolution. At the present stage of the analysis it is not possible to offer a structural rationale for the glyphosate tolerance conferred by the substitution of alanine for glycine at position 96; nevertheless, it has been possible to examine the region of proline-101, where the mutation to serine in the *S. typhimurium* enzyme confers a measure of resistance to glyphosate inhibition (9). We have engineered by site-directed mutagenesis the corresponding substitution in *E. coli* K-12 EPSP synthase. The serine-101-containing *E. coli* enzyme is also glyphosate tolerant, exhibiting an IC_{50} value of 5 mM, compared with 20 μM for the wild-type protein. The *E. coli* and the *S. typhimurium* enzymes show 88% identity in their amino acid sequences (28), with the differences well scattered within the molecules. Tentative placement of the proline ring about one turn from the amino terminus of a buried helix near the domain crossovers is consistent with our speculations on the active-site

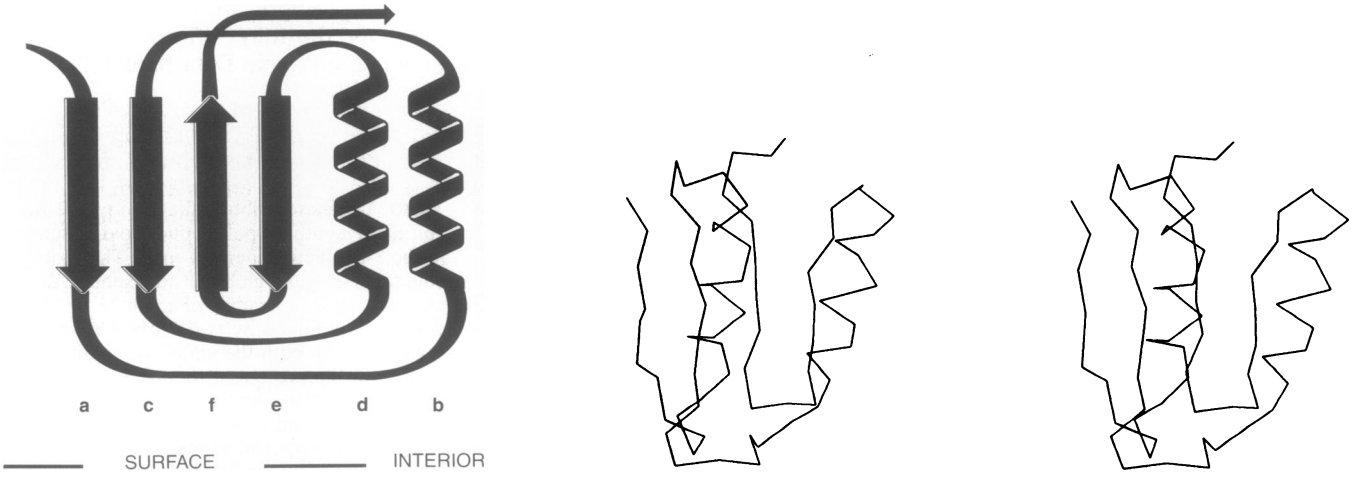


FIG. 3. (*Left*) Topology of the principal elements of secondary structure in the folding units that are related by the pseudo threefold axes is illustrated. The folding unit includes a β-sheet and two parallel helices with equivalent chain polarities. One of the helices is completely buried in the interior of the EPSP synthase molecule; the other is only partially buried and includes a solvent-exposed face. The β-sheet is also located on the surface of the molecule; the first two strands of the four-stranded sheet are parallel, with the last three strands antiparallel. The structure folds by forming six of these folding units or subdomains, and only one of the units in the structure is incomplete. The chain connectivity of the three folding units formed by continuous segments of the polypeptide chain is also shown. The principal elements of secondary structure are assigned letters **a** through **f** according to the sequence in which they occur along the chain. (*Right*) Stereo diagram of the α-carbon backbone of one of the three folding units formed by continuous polypeptide chain; this view is from the solvent into the center of the upper domain in Fig. 1. The overall topology (37) is +1X, +2X, −1 with one of the parallel helices in each of the two crossover connections. With about 70 residues per folding unit, 6-fold repetition of this motif accounts for almost all 427 amino acid residues of the *E. coli* protein.

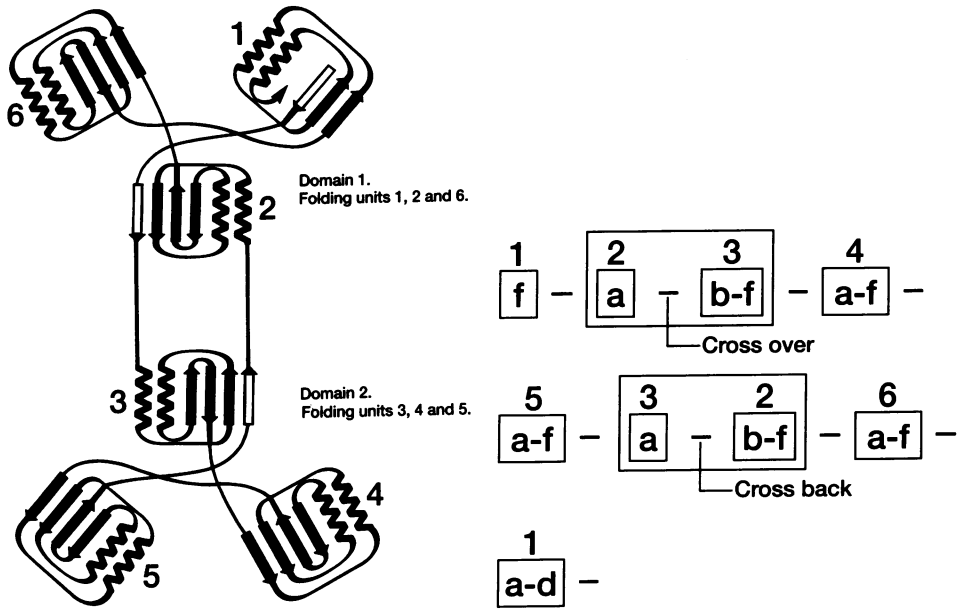
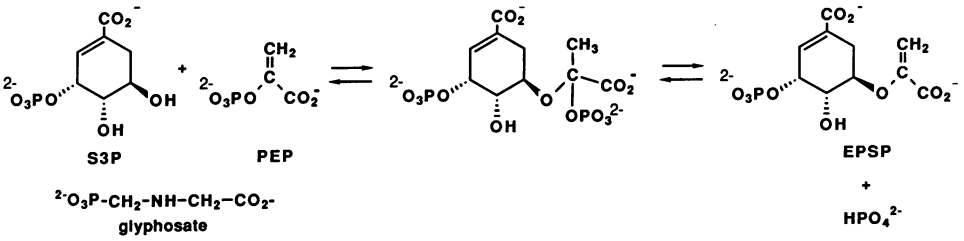


FIG. 4. Schematic illustration of the chain tracing of *E. coli* EPSP synthase (*Left*) and the order of the principal secondary structural elements as they occur in the EPSP synthase gene (*Right*). Folding unit 1 comprises both the amino and carboxy termini. In *Left*, open arrows distinguish strands that are not directly connected to other elements of secondary structure in the folding units (1, 2, and 3) in which they occur. With reference to Fig. 3 *Left* and beginning at the amino terminus, formation of strand **f** of folding unit 1 is followed by formation of strand **a** of the unit 2 with subsequent crossover to the other domain. Elements **b** through **f** of folding unit 3 then form, followed by the formation of units 4 and 5 from the continuous chain. The remainder of folding unit 3, strand **a**, forms next, with cross-back to the initial domain and formation of the remainder, elements **b** through **f**, of unit 2. This is followed by formation of unit 6 from the continuous chain. The polypeptide terminates by completing most of folding unit 1, elements **a** through **d**. In *Right* the principal elements of secondary structure are listed as they occur sequentially in the chain tracing (Fig. 1) and therefore as they occur along the coding segment of the gene of *E. coli* EPSP synthase. Numbers above the boxes indicate the folding unit of the gene product, and letters within the boxes indicate the elements of secondary structure identified in Fig. 3 *Left*. Folding units 2 and 3 are not formed from continuous polypeptide chain segments. The amino terminus of folding unit 2 is joined directly to the carboxyl terminus of folding unit 3 by a domain crossover segment, and their union, indicated by blocking (the two larger boxes), results in the expression (**a** through **f**) of a set of the elements of secondary structure of a complete folding unit. Similarly, at the other domain crossover, strand **a** of folding unit 3 is joined directly to structural elements **b** through **f** in folding unit 2 by an interdomain polypeptide linkage. Coding for elements **a** through **d** of folding unit 1 should appear at the 3′ end of the gene, while strand **f** is apparently encoded at the 5′ end; code for strand **e** of folding unit 1, not present in the structure, should therefore not be present in the *E. coli* gene sequence.

locale. Since it is unlikely that a proline residue could be directly involved in substrate or inhibitor binding, we have focused our attention on nearby residues and suggest that the mutation straightens a kink in the helix near its amino terminus. Therefore, it is possible that, upon mutation, the side chains of residues neighboring proline-101, such as the guanidinium group of arginine-100, would experience reorientation in the active site, thereby influencing ligand binding.

In summary, chain tracing establishes *E. coli* EPSP synthase as a structure with two globular domains. Taxonomic description (37, 38) of the symmetric domains is problematic: the interacting helices are parallel and are completely or partly buried, while extensive areas of the domain surfaces are covered by β-sheets. It is tempting to propose a new term such as "inside-out α/β-barrel," but this may be inappropriate. A more vivid descriptor is the "mushroom button," wherein we suggest that the buried helices form the chopped stems and that the surface helices and sheet strands form the caps of a molecular fungus. The topological symmetry of the molecule may reflect 6-fold gene replication, and the approach of anion ligands can be directed to an active site near the domain junctures by macrodipolar effects. Tolerance to glyphosate inhibition offered by the mutation at position 101 probably reflects spatial rearrangements of active site residues as they adopt conformations that straighten a proline-induced helix kink in the wild-type protein. Given that sequence homologies among the bacterial, fungal, and plant EPSP synthases are substantial (28), we anticipate that the three-dimensional structures of the proteins from other sources will resemble the fold of the enzyme from *E. coli*. The use of protein crystal structures as templates in the design of candidate drug molecules has become standard. By contrast, related applications in agricultural science are limited (39–43). EPSP synthase represents a unique opportunity in molecular design; its structure will be used both as a template in the biorational synthesis of new inhibitors and in the design of catalytically competent variants useful for the engineering of crop selectivity to an environmentally acceptable herbicide.



Scheme 1

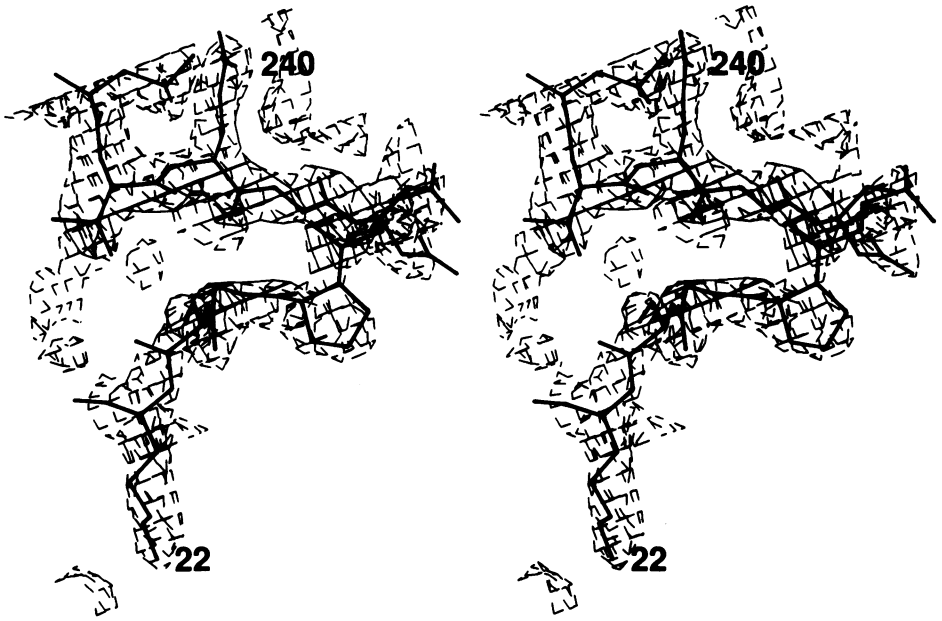


FIG. 5. Double-stranded hinge connecting the two domains.

Note. We have now calculated an isomorphous replacement electron density map at 2.5-Å resolution and used it to build a complete model of the structure with side-chain placement. Refinement of this structure leads to a current crystallographic residual of 23.2%. Analysis of these difference maps confirms the correctness of the chain tracing derived from the electron density map calculated at 3-Å resolution. Our present interpretation of the regions that compose the active site is complicated because it is becoming increasingly clear that EPSP synthase is a hinged molecule that experiences a substantial conformational change during catalysis and on inhibition. A map illustrating the clarity of the double-stranded hinge that connects the two domains is shown in Fig. 5. Phases were calculated from the refined structure, and the Fourier amplitudes are $(|2F_o| - |F_c|)$; side chains from Lys-22 and Glu-240 are labeled.

We are grateful to Robert Fraley and Mel Rueppel for continuous encouragement and support in this project. We also acknowledge useful discussions with Jeff Borgmeyer, Quai Huynh, Richard Leimgruber, and Chuck McWherter, and we thank Anna Stevens for taking precession photographs.

1. Franz, J. E. (1985) in *The Herbicide Glyphosate*, eds. Grossman, E. & Atkinson, D. (Butterworth, London), pp. 3–17.
2. Malik, J., Barry, G. & Kishore, G. (1989) *BioFactors* **2**, 17–25.
3. Boocock, M. R. & Coggins, J. R. (1983) *FEBS Lett.* **154**, 127–133.
4. Steinrücken, H. C. & Amrheim, N. (1980) *Biochem. Biophys. Res. Commun.* **94**, 1207–1212.
5. Padgette, S. R., della-Cioppa, G., Shah, D. M., Fraley, R. T. & Kishore, G. M. (1989) in *Cell Culture and Somatic Cell Genetics of Plants*, eds. Vasil, I. & Schell, J. (Academic, London), Vol. 6, pp. 441–476.
6. Gasser, C. S. & Fraley, R. T. (1989) *Science* **244**, 1293–1299.
7. Comai, L., Facciotti, D., Hiatt, W. R., Thompson, G., Rose, R. E. & Stalker, D. M. (1985) *Nature (London)* **317**, 741–744.
8. Comai, L. (1985) U.S. Patent 4,535,060.
9. Stalker, D. M., Hiatt, W. R. & Comai, L. (1985) *J. Biol. Chem.* **260**, 4724–4728.
10. Huynh, Q. K., Bauer, S. C., Bild, G. S., Kishore, G. M. & Borgmeyer, J. R. (1988) *J. Biol. Chem.* **263**, 11636–11639.
11. Padgette, S. R., Huynh, Q. K., Aykent, S., Sammons, R. D., Sikorski, J. A. & Kishore, G. M. (1988) *J. Biol. Chem.* **263**, 1798–1802.
12. Anderson, K. S., Sikorski, J. A. & Johnson, K. A. (1988) *Biochemistry* **27**, 1604–1610.
13. Anderson, K. S., Sikorski, J. A. & Johnson, K. A. (1988) *Biochemistry* **27**, 7395–7406.
14. Alberg, D. C. & Bartlett, P. A. (1989) *J. Am. Chem. Soc.* **111**, 2337.
15. Grimshaw, C. E., Sogo, S. G., Copley, S. D. & Knowles, J. R. (1984) *J. Am. Chem. Soc.* **106**, 2699–2700.
16. Asano, Y., Lee, J. J., Spreafico, F., Kowal, C. & Floss, H. G. (1985) *J. Am. Chem. Soc.* **107**, 4314–4320.
17. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542.
18. Rogers, S. G., Brand, L. A., Holder, S. B., Sharps, E. S. & Bracken, M. J. (1983) *Appl. Environ. Microbiol.* **1**, 37–43.
19. Castellino, S., Leo, G. C., Sammons, R. D. & Sikorski, J. A. (1989) *Biochemistry* **28**, 3856–3868.
20. McPherson, A. (1976) *Methods Biochem. Anal.* **23**, 249–345.
21. Hamlin, R. (1985) *Methods Enzymol.* **114**, 416–452.
22. Blundell, T. L. & Johnson, L. N. (1976) *Protein Crystallography* (Academic, London).
23. Sprang, S. & Fletterick, R. J. (1979) *J. Mol. Biol.* **131**, 523–551.
24. Wang, B.-C. (1985) *Methods Enzymol.* **115**, 90–112.
25. Hol, W. G. J., van Duijnen, P. T. & Berendsen, H. J. C. (1978) *Nature (London)* **273**, 443–446.
26. Rossmann, M. G. & Argos, P. (1976) *J. Mol. Biol.* **105**, 75–95.
27. Klee, H. J., Muskopf, Y. M. & Gasser, C. S. (1987) *Mol. Gen. Genet.* **210**, 437–442.
28. Gasser, C. S., Winter, J. A., Hironaka, C. M. & Shah, D. M. (1988) *J. Biol. Chem.* **268**, 4280–4289.
29. Duncan, K., Edwards, R. M. & Coggins, J. R. (1987) *Biochem. J.* **246**, 375–386.
30. Charles, I. G., Keyte, J. W., Brammar, W. J., Smith, M. R. & Hawkin, A. R. (1986) *Nucleic Acids Res.* **14**, 2201–2213.
31. Duncan, K., Lewendon, A. & Coggins, J. R. (1984) *FEBS Lett.* **170**, 59–63.
32. Maskell, D. J., Morrissey, P. & Dougan, G. (1988) *J. Bacteriol.* **170**, 2467–2471.
33. Anderson, K. S. & Johnson, K. A. (1990) *J. Biol. Chem.* **265**, 5567–5572.
34. Anderson, K. S., Sikorski, J. A., Benesi, A. J. & Johnson, K. A. (1988) *J. Am. Chem. Soc.* **110**, 6577–6579.
35. Barlow, P. N., Appleyard, R. J., Wilson, B. J. O. & Evans, J. N. (1989) *Biochemistry* **28**, 7985–7991.
36. Anderson, K. S., Sammons, R. D., Leo, G. C., Sikorski, J. A., Benesi, A. J. & Johnson, K. A. (1990) *Biochemistry* **29**, 1460–1465.
37. Richardson, J. S. (1981) *Adv. Protein Chem.* **34**, 167–339.
38. Levitt, M. & Chothia, C. (1976) *Nature (London)* **261**, 552–588.
39. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. (1985) *Nature (London)* **318**, 618–624.
40. Chapman, M. S., Suh, S. W., Curmi, P. M. G., Casio, D., Smith, W. W. & Eisenberg, D. S. (1988) *Science* **241**, 71–74.
41. Andersson, I., Knight, S., Schneider, G., Lindqvist, T., Lundqvist, T., Branden, C. I. & Lorimer, G. H. (1989) *Nature (London)* **337**, 229–234.
42. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W. & Davies, D. R. (1988) *J. Biol. Chem.* **263**, 17857–17871.
43. Priestle, J. P., Grütter, M. G., White, J. L., Vincent, M. G., Kania, M., Wilson, E., Jardetzky, T. S., Kirschner, K. & Jansonius, J. N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5690–5694.