Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 5690–5694, August 1987
Biochemistry

Three-dimensional structure of the bifunctional enzyme *N*-(5′-phosphoribosyl)anthranilate isomerase–indole-3-glycerol-phosphate synthase from *Escherichia coli*

(protein structure/x-ray crystallography/protein evolution/gene fusion/tryptophan biosynthesis)

J. P. PRIESTLE, M. G. GRÜTTER*, J. L. WHITE†, M. G. VINCENT‡, M. KANIA, E. WILSON§, T. S. JARDETZKY¶, K. KIRSCHNER, AND J. N. JANSONIUS‖

Departments of Structural Biology and Biophysical Chemistry, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Communicated by Charles Yanofsky, May 12, 1987 (received for review February 6, 1987)

ABSTRACT ***N*-(5′-Phosphoribosyl)anthranilate isomerase–indole-3-glycerol-phosphate synthase from *Escherichia coli* is a monomeric bifunctional enzyme of M_r 49,500 that catalyzes two sequential reactions in the biosynthesis of tryptophan. The three-dimensional structure of the enzyme has been determined at 2.8-Å resolution by x-ray crystallography. The two catalytic activities reside on distinct functional domains of similar folding, that of an eightfold parallel β-barrel with α-helices on the outside connecting the β-strands. Both active sites were located with an iodinated substrate analogue and found to be in depressions on the surface of the domains created by the outward-curving loops between the carboxyl termini of the β-sheet strands and the subsequent α-helices. They do not face each other, making "channeling" of the substrate between active sites virtually impossible. Despite the structural similarity of the two domains, no significant sequence homology was found when topologically equivalent residues were compared.**

N-(5′-Phosphoribosyl)anthranilate isomerase–indole-3-glycerol-phosphate synthase (PRAI–IGPS) from *Escherichia coli* is a monomeric bifunctional enzyme of M_r 49,500 (1). It catalyzes two sequential reactions in the biosynthesis of tryptophan (2, 3):

$$\text{PRA} \xrightarrow[\text{PRAI}]{} \text{CdRP}$$
$$\text{CdRP} \xrightarrow[\text{IGPS}]{} \text{IGP} + \text{CO}_2 + \text{H}_2\text{O}$$

in which PRA represents *N*-(5′-phosphoribosyl)anthranilate, CdRP is 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate, and IGP is indole-3-glycerolphosphate. These reactions are catalyzed by two separate enzymes in *Bacillus subtilis* and *Pseudomonas putida* (4–7) and by a multienzyme complex in *Brevibacterium flavum* (8). The *E. coli* enzyme is encoded by the *trpCF* gene of the tryptophan operon. The amino acid sequence, as inferred from the gene sequence, consists of 452 amino acid residues (9–11). The study of mutants (12) and limited proteolysis (13) indicate that the two catalytic activities reside on different domains, the N-terminal domain carrying the IGPS activity.

Crystals suitable for x-ray studies have been grown from ammonium sulfate by repeated seeding (14, 15). The three-dimensional structure of the complete bifunctional enzyme has been determined at 2.8-Å resolution by the multiple isomorphous replacement (MIR) method.

MATERIALS AND METHODS

Crystals of PRAI–IGPS were obtained as described (14, 15). They belong to space group $P4_1$ with unit cell parameters $a = b = 104.7$ Å, $c = 68.0$ Å, and one molecule per asymmetric unit (67% solvent content). Screening of heavy atom derivatives was done by soaking crystals in a buffered solution of the appropriate heavy atom compound in 3 M ammonium sulfate (Table 1). For potentially useful derivatives, a low-resolution diffractometer data set was collected, while oscillation data sets to 2.5-Å resolution were collected for the three best derivatives. One derivative was prepared by diffusing an iodinated form of the substrate analogue *N*-(5′-phosphoribit-1-yl)anthranilate (rCdRP) into the crystals. 5-Iodo-rCdRP was prepared as described for rCdRP (16) except that 5-iodoanthranilate (Fluka) was used in place of anthranilate as starting material and was dissolved in dioxane instead of in 50% aqueous ethanol.

The crystals were mounted in thin-walled glass capillaries. Diffractometer data were collected as described elsewhere (17) with the following modifications: after crystal alignment, a single test reflection was used to determine the direction of the *c*-axis of the crystal; several crystals were required per data set due to strong radiation sensitivity; a maximum intensity loss of about 30% was accepted; and intensities were measured by ω-scan over 0.8–1.5°, depending on mosaic spread.

Oscillation data were collected and processed as described elsewhere (18) with the following modifications: a cooling device kept the crystal at about 4°C; data were collected over 90°; two sheets of film per film pack were used; crystal orientation was checked by still photographs at the beginning and end of data collection; and the films were digitized, using a 3.0 OD scale.

The oscillation and diffractometer data sets (Table 1) were not merged at this point. Heavy atom sites were accepted only if found with both sets of data. They were located by difference Patterson and difference Fourier techniques and their positions and occupancies were improved by phase refinement (19). Phases were calculated (20) separately for the diffractometer and oscillation data and combined (Table

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: PRAI, *N*-(5′-phosphoribosyl)anthranilate isomerase; IGPS, indole-3-glycerol-phosphate synthase; PRAI–IGPS, the bifunctional enzyme with both PRAI and IGPS activities; CdRP, 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate; rCdRP, reduced CdRP [*N*-(5′-phosphoribit-1-yl)anthranilate]; MIR, multiple isomorphous replacement.

*Present address: Pharmaceuticals Research Division, CIBA-Geigy Ltd., CH-4002 Basel, Switzerland.
†Present address: Department of Biochemistry, University of Sheffield, Sheffield S10 2TN, England.
‡Present address: Aqua Metro AG, Ringstrasse 75, CH-4106 Therwil, Switzerland.
§Present address: Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94395.
¶Present address: Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138.
‖To whom reprint requests should be addressed.

Table 1. Summary of heavy atom derivatives and data collection

Derivative*	Conc., mM	Soak time, hr	Resolution, Å		No. unique reflections†		f_H/ε‡		R_{merge}§	No. crystals used	
			Dif	Osc	Dif	Osc	Dif	Osc	Osc	Dif	Osc
Native	—	—	3.4	2.5	9944 (92)	22,347 (89)	—	—	11.2	4	13
$KAu(CN)_2$	5.0	2	3.4	2.5	9351 (86)	15,903 (63)	1.67	1.59	12.0	5	8
EMTS	0.5	2	3.8	2.5	6801 (87)	17,070 (68)	1.42	1.25	13.5	2	5
$K_2Pt(CN)_4$	1.5	24	3.6	—	8551 (93)	—	1.15	—	—	3	—
K_2PtCl_4	1.0	3	3.7	—	7644 (90)	—	1.08	—	—	2	—
EMP	0.5	15	3.7	—	7548 (89)	—	1.33	—	—	3	—
Iodo-rCdRP	1.0	15	4.9	2.5	3304 (88)	17,472 (69)	1.14	1.16	13.8	1	4

Dif and Osc refer to data collected on the diffractometer and oscillation camera, respectively.
*EMTS, ethylmercuric thiosalicylate; EMP, ethylmercuric phosphate.
†Number of reflections for the derivatives refers to the number actually used in phase determination. The numbers in parentheses are the percentage of a full data set to the given resolution.
‡ε is the lack-of-closure error = $F_{PH}(obs) - F_{PH}(calc)$, where F_{PH} is the heavy-atom derivative structure factor amplitude.
§$R_{merge} = 100(2\Sigma|(F_1 - F_2)|)/\Sigma(F_1 + F_2)$, where F_1 and F_2 are symmetry-equivalent structure factor amplitudes.

2) by the method of Hendrickson and Lattman (21). The automated protein/solvent boundary determination, solvent flattening, and phase combination procedures of Wang (22) were used to improve the MIR phases.

Electron density maps were computed at 3.6-, 3.2-, 2.8-, and 2.5-Å resolution, using the combined "best" phases and figure-of-merit weighting (20). The 3.6- and 3.2-Å maps displayed the best chain connectivity. Secondary structural elements, especially α-helices, were easily recognizable. Tentative C_α positions were assigned for long stretches of continuous density. An automated search procedure was then used to fit the known sequence to the electron density. The protein structure was built into the electron density on an Evans and Sutherland PS-300 color graphics display system interfaced to a VAX 11/730 computer. The molecular manipulation program FRODO (23), version 6.0, courtesy of Rice University (Houston, TX) was used.

Sim-weighted calculated phases (24) and figure-of-merit-weighted MIR phases (20) were then combined (21). Electron density maps were calculated with these combined phases and figure-of-merit-weighted $(3F_o - 2F_c)$ coefficients, and the structure was rebuilt several times before structure refinement was started. The model was partially refined by using the restrained least-squares refinement programs of Hendrickson and Konnert (25, 26), modified to use fast Fourier transforms (27, 28) and to take intermolecular contacts into account (29). Periodically, the structure was rebuilt.

When the model had improved sufficiently, as indicated by the crystallographic R factor, only calculated phases were used in the computation of electron density maps. "Omit maps" (30) were calculated to remove potential bias from the model phases. The current R factor at 2.8 Å is 38.3%.

RESULTS

Fig. 1 depicts the overall folding of PRAI–IGPS. The enzyme is organized into two domains with the IGPS domain consisting of residues 1–255 and the PRAI domain consisting of residues 256–452. Both domains have the same folding, that of a central eightfold parallel β-barrel with α-helices on the outside connecting the β-strands. This beautifully symmetric folding pattern was seen in triose-phosphate isomerase in 1981 (31) and is often referred to as the "TIM-barrel." Several other enzymes have since been shown to have this folding, either as the complete enzyme as in triose-phosphate isomerase and 2-keto-3-deoxyphosphogluconate aldolase (32) or as one domain in a multidomain structure as in pyruvate kinase (33), Taka-amylase A (34), xylose isomerase (35), glycolate oxidase (36), ribulose-1,5-bisphosphate carboxylase/oxygenase (37), and muconate lactonizing enzyme (38). The IGPS domain contains an additional six-turn N-terminal α-helix followed by a loop of 24 residues.

Table 2. Mean figures-of-merit of phases as a function of resolution

Data	Figure-of-merit			
	∞–4.0 Å	4.0–3.2 Å	3.2–2.8 Å	2.8–2.5 Å
Diffractometer	0.716	0.472	—	—
Oscillation	0.621	0.527	0.475	0.381
Combined	0.789	0.542	0.475	0.381
Solvent flattened	0.880	0.768	0.659	0.638

There are few noncovalent interdomain interactions. The last α-helix of the IGPS domain interacts with a short loop between the first α-helix and the second β-strand of the PRAI domain. The carboxyl-terminal residue of the PRAI domain fits into a pocket in the IGPS domain. In the present model the interactions consist of three hydrogen bonds and some 15 van der Waals contacts (half of which involve the carboxyl-terminal residue, Tyr-452).

Horowitz (39, 40) proposed that biosynthetic pathways evolved by gene duplication followed by independent evolution of one of the gene copies. If this were true, one should find sequence and structural similarities between consecutive enzymes along a biosynthetic pathway. The enzymes catalyzing two sequential steps in the biosynthesis of methionine in *E. coli* have been shown to possess high sequence homology (41), supporting this proposition. Since the overall folding of the domains of PRAI–IGPS is so similar, they may constitute another example. In this case, the process would have been followed by gene fusion, as has been suggested for bifunctional enzymes in general (42).

C_α atoms of the two separate domains were aligned visually on the graphics system. A rigid-body least-squares fit of the spatially corresponding C_α pairs gave an overall root-mean-square deviation of 2.17 Å between the two domains. Alignment of the spatially corresponding residues (Fig. 2) resulted in 19 identities (10.3%). Jumbling the sequences 1000 times gave an average of 12.2 (6.6%) identities with a standard deviation of 3.3; hence, the 19 identities found are higher than random, but not significantly so (i.e., $< 3\sigma$ from the mean). Likewise, no evidence for relatedness has been found in the nucleotide sequences (10).

The inhibitor rCdRP binds to both active sites of PRAI–IGPS in solution (16, 43). However, difference maps made by using the original MIR phases failed to reveal its position. 5-Iodo-rCdRP, on the other hand, gave very clear iodine peaks in the difference maps, located in the pockets formed by the loops at the carboxyl ends of the two β-barrels (Fig. 1).

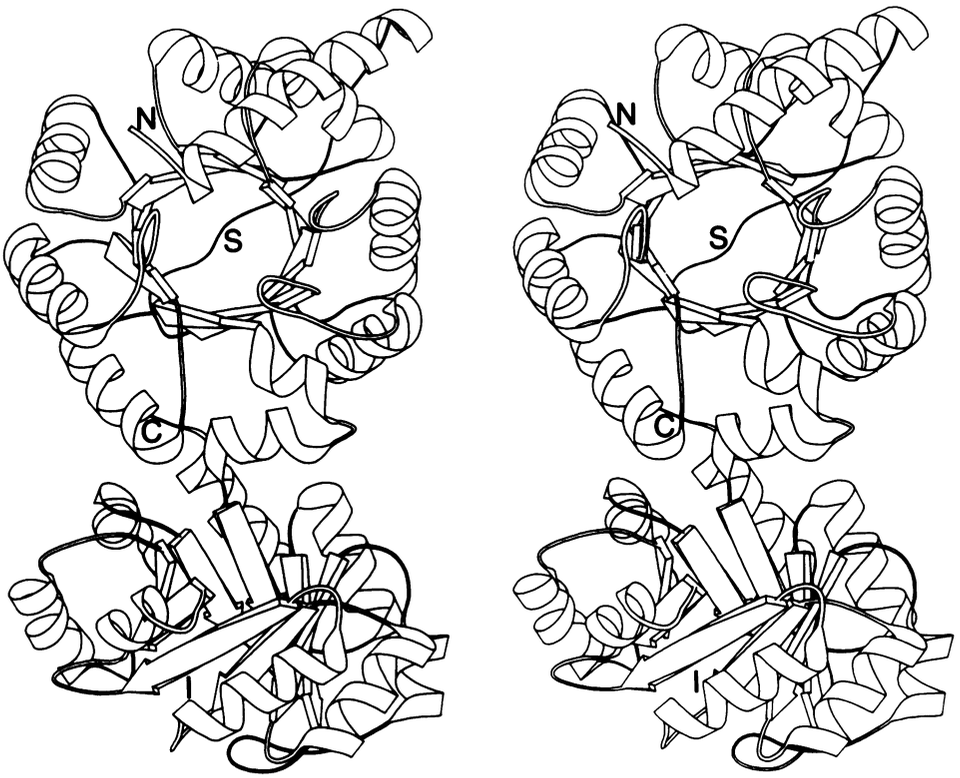

FIG. 1. Stereo cartoon drawing (program written by J.P.P.) of PRAI–IGPS. α-Helices are represented by helical ribbons and β-sheet strands by twisted arrows. The locations of the active sites of the IGPS and PRAI domains are indicated by S and I, respectively. N and C indicate the amino and carboxyl termini.

The active site in the IGPS domain is formed by the carboxyl-terminal residues from the eight strands of the β-barrel and the first two turns of the long N-terminal α-helix. Especially important seem to be the loops between the β-strands 1, 3, and 6 and the α-helices following these. The active site of the PRAI domain is also at the carboxyl terminus of the β-barrel. All β-strands, except the fifth, participate in the formation of the active-site pocket. The loops between β-strands 2, 6, and 8 and the α-helices following these also contribute.

```
       50         60         70         80         90
      <-ß1-->              <-α1-->        <-ß2->           <--
IPGS =TAFILECKKASPSKGVIRDDFDPARIAAIYKHYASAISVLTDEKYFQGSFNF
PRAI =GENKVCGL----------TRGQDAKAAYDAGAIYGGLIFVATSPRCVNVEQA
       <ß1->               <--α1--->     <-ß2-->          <--
    255  260                 270        280        290

100      110        120        130        140         150
-α2--->      <ß3->         <-α3-->       <-ß4->       <---α4--
LPIISQIAPQPILCKDFIIDPYQIYLARYYQA-DACLLMLSVLDDDQ---YRQLAAVA
QEVMA-AAP--LQYVGVFR--NHDIADVVDKAKVLSLVAVQLHGNEEQLYIDTLREAL
α2->         <--ß3-->      <---α3---->   <ß4->        <--α4-->
 300           310          320        330        340

    160          170        180        190        200
->  <-ß5-->      <---α5--->    <ß6>           <---α6---->
HSLEMGVLTEVS---NEEEQERAIALGAKVVGINNRDLRDLSIDLNRTRELAPKLGHN
-PAHVAIWKALSVGETLPAREFQHVD-KYVLDNGQG----GSGQRFDWSLL-NGQSLG
    <--ß5-->      <--α5--->   <-ß6->        <--α6--->
350       360        370        380          390         400

210       220        230        240          250  255
<-ß7->       <---α7--->    <ß8><-α8'->       <--α8-->
VTVISESGINTYAQVRELSHF-ANGFLIGSALMAH--DDLHAAVRRVLLG=
-NVLLAGGLGAD-NCVEAAQTGCAGLDFNSAVESQPGIKDARLLASVFQTLRAY.
 <-ß7->        <-α7->    <ß8>              <---α8---->
         410        420        430        440        450
```

FIG. 2. Comparison of amino acid residues occupying equivalent spatial positions in the IGPS and PRAI domains. The standard one-letter residue symbols are used, and the sequence numbers correspond to the residues under the last digit (except Pro-350). The first 47 residues of the IGPS domain are not shown. Secondary structural elements are shown above the IGPS sequence and below the PRAI sequence. Identities are indicated by a line between residues of the two sequences. A – indicates a deletion relative to the other domain; a . indicates the carboxyl terminus; and an = indicates that the polypeptide chain extends further.

DISCUSSION

PRAI–IGPS is composed of two distinct domains corresponding to the two enzymic functions. Both domains consist of a parallel β-barrel core of eight strands with α-helices surrounding it, as earlier observed for triose-phosphate isomerase ("TIM-barrel"). The IGPS domain contains at the amino terminus an additional 47 residues forming a six-turn α-helix followed by 24 residues without secondary structure. The two active sites are found in the pockets formed by the carboxyl ends of the β-barrels in each domain, as in all other crystallographically studied enzymes with the triose-phosphate isomerase folding (31–33, 35–38, 44). In PRAI–IGPS the two active sites face away from each other, which seems to rule out direct transfer of CdRP, the product of the isomerase reaction, between the two active sites. This finding confirms kinetic evidence for accumulation of CdRP in the bulk solvent (3). It also renders unlikely any transmission of conformational changes between the active sites, for which, indeed, no indication has been found in solution (43).

Noncovalent interactions between the two domains are mainly mediated by the carboxyl-terminal residue Tyr-452 and the amino-terminal domain. This explains why removal of carboxyl-terminal residues by either mutation (12) or carboxypeptidase A (T. Niermann and K.K., unpublished work) leads to loss of both PRAI and IGPS activities. Tyr-452 may act as an "anchor" preventing rotation of the two domains relative to each other about the covalent polypeptide linkage and supporting additional weak interdomain interactions. Isolated PRAI and IGPS domains prepared by manipulation of the *trpCF* gene of *E. coli* fold autonomously *in vivo*. However, the IGPS and PRAI activities are reduced to about

10% of their original values, and these artificial subunits do not associate (45).

To date, the gene sequences of both PRAI and IGPS from several organisms other than *E. coli* have been elucidated: *Neurospora crassa trpGCF* (46), *Saccharomyces cerevisiae trpGC* (47, 48) and *trpF* (49), *Bacillus subtilis trpC* and *trpF* (7), *Acinetobacter calcoaceticus trpC* and *trpF* (ref. 50; J. Kaplan and B. Nichols, personal communication), *Salmonella typhimurium trpCF* (11), *Aspergillus nidulans trpGCF* (51), *Lactobacillus casei trpC* and *trpF* (F. Imamoto and Y. Natori, personal communication), *Brevibacterium lactofermentum trpCF* (52), and *Penicillium chrysogenum trpGCF* (53). All sequences were first aligned pairwise by the maximum match method of Needleman and Wunsch (54). The amino acid similarity matrix of Bacon and Anderson (55) was used, but modified such that the score for identical amino acids was twice, rather than equivalent to, that of the two most similar residues. The multiple sequence alignment was done by eye, based on the pairwise alignments (Fig. 3).

Examination of Fig. 3 in the light of the tertiary structure of the *E. coli* enzyme reveals several interesting features. The single residue Gly-255 bridges the last α-helix of the IGPS domain and the first strand of the β-barrel of the PRAI domain, and it is completely buried. Its position corresponds to one residue beyond the carboxyl terminus of the shortest monofunctional IGPS (*S. cerevisiae*) and one residue within the amino terminus of the shortest monofunctional PRAI (*L. casei*). This fusion (also observed in *S. typhimurium* and *B. lactofermentum*) therefore seems to be the shortest possible. In the trifunctional enzymes from *P. chrysogenum*, *A. nidulans*, and *N. crassa*, the connector between the IGPS and the PRAI domains has an insertion of 13 residues relative to the *E. coli* enzyme. When an extra 28 residues are inserted at the connection in the *E. coli* enzyme, full activities for both

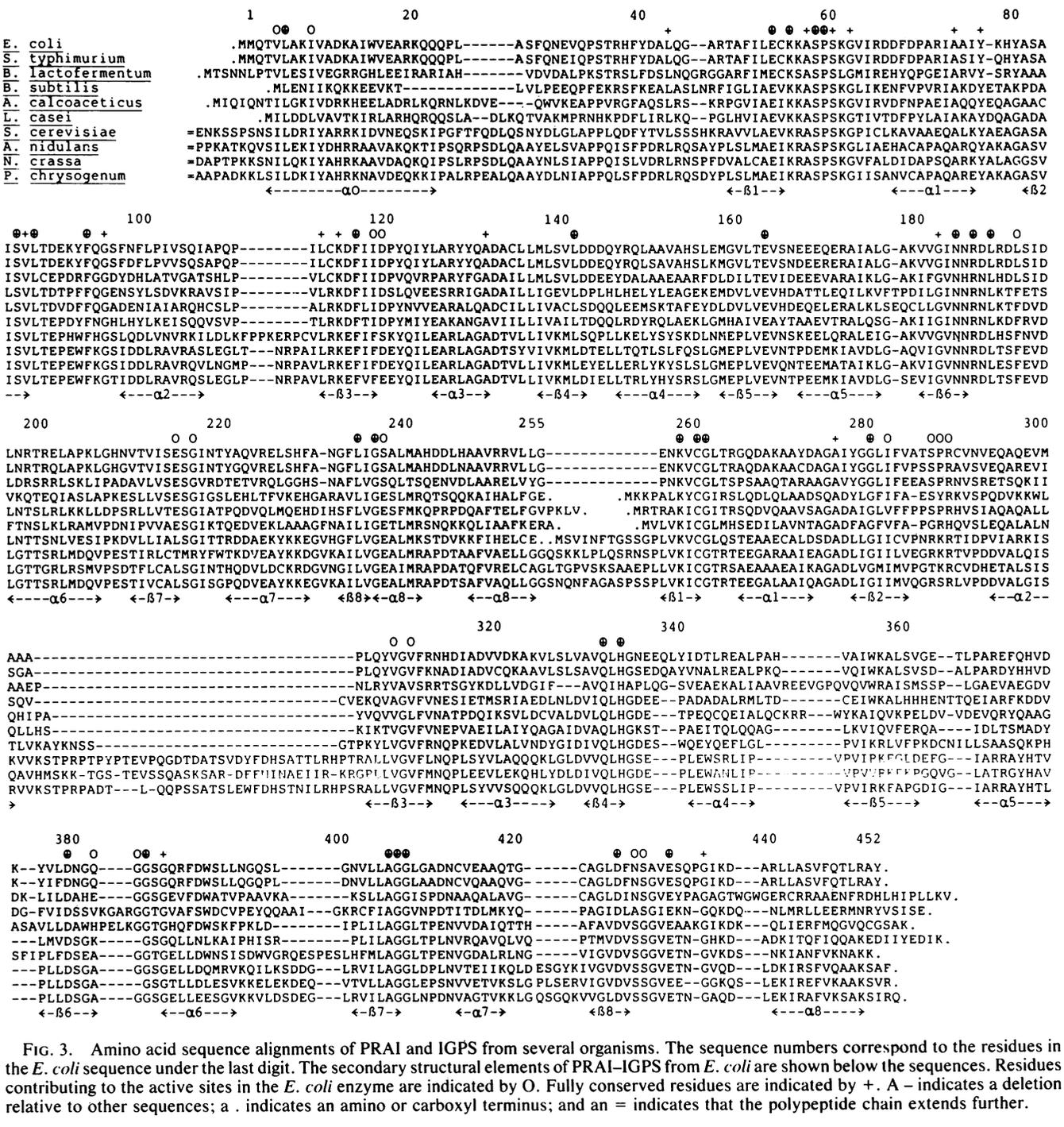

FIG. 3. Amino acid sequence alignments of PRAI and IGPS from several organisms. The sequence numbers correspond to the residues in the *E. coli* sequence under the last digit. The secondary structural elements of PRAI–IGPS from *E. coli* are shown below the sequences. Residues contributing to the active sites in the *E. coli* enzyme are indicated by O. Fully conserved residues are indicated by +. A – indicates a deletion relative to other sequences; a . indicates an amino or carboxyl terminus; and an = indicates that the polypeptide chain extends further.

domains are maintained (56). This observation implies that the length of the connecting segment can vary without interfering with the noncovalent interactions that stabilize the enzymically active states of the domains.

Insertions with respect to the *E. coli* enzyme occur far more frequently in the loops at the amino ends of the two β-barrels, opposite the active sites, than at the carboxyl ends. The particularly large insertion between α2 and β3 of the PRAI domain of the trifunctional enzymes of *P. chrysogenum*, *A. nidulans*, and *N. crassa* might create either additional stabilizing interactions between the PRAI and IGPS domains or possible sites of interaction with the third functional domain, glutamine amidotransferase, or with the second subunit, anthranilate synthase.

Regions of high sequence homology can be associated with catalytic importance. For example, 16 of the 24 amino acid residues forming the IGPS active site (66.7%) are completely invariant, compared to 12 out of 231 (5.2%) for the rest of the domain. For the PRAI domain, 12 of the 23 residues forming the active site (52.2%) are invariant, compared to 3 out of 174 (1.7%) for the rest of the domain.

PRAI–IGPS is, to our knowledge, the first native bifunctional enzyme for which the x-ray crystallographic structure has been determined. The three-dimensional structure of the bifunctional proteolytic Klenow fragment of DNA polymerase I from *E. coli* has recently been elucidated (57). The two active sites seem to be perfectly disposed to catalyze the sequential reactions (polymerase, 3′-5′ "editing" exonuclease) on the same (polymeric) substrate, DNA. In contrast, the two active sites of PRAI–IGPS seem to function independently. In the case of PRAI–IGPS, the advantage of gene fusion appears to be the mutually stabilizing interactions between the two functional domains.

We thank Drs. F. Imamoto and B. Nichols for making their sequence data available prior to publication. Dr. G. C. Ford is thanked for computational assistance and T. Teshiba for technical help. This work was supported in part by grants from the Swiss National Science Foundation (3.415-0.78, 3.224-0.82, and 3.098-0.85 to J.N.J. and 3.315-0.82 to K.K.).

1. McQuade, J. F., III, & Creighton, T. E. (1970) *Eur. J. Biochem.* **16,** 199–207.
2. Creighton, T. E. & Yanofsky, C. (1966) *J. Biol. Chem.* **241,** 4616–4624.
3. Creighton, T. E. (1970) *Biochem. J.* **120,** 699–707.
4. Hoch, S. O., Anagnostopoulos, C. & Crawford, I. P. (1969) *Biochem. Biophys. Res. Commun.* **35,** 838–844.
5. Hoch, S. O. & Crawford, I. P. (1973) *J. Bacteriol.* **116,** 685–693.
6. Enatsu, T. & Crawford, I. P. (1968) *J. Bacteriol.* **95,** 107–112.
7. Henner, D. J., Band, L. & Shimotsu, H. (1984) *Gene* **34,** 169–177.
8. Sugimoto, S. & Shiio, I. (1977) *J. Biochem. (Tokyo)* **81,** 823–833.
9. Christie, G. E. & Platt, T. (1980) *J. Mol. Biol.* **142,** 519–530.
10. Yanofsky, C., Platt, T., Crawford, I. P., Nichols, B. P., Christie, G. E., Horowitz, H., van Cleemput, M. & Wu, A. M. (1981) *Nucleic Acids Res.* **9,** 6647–6668.
11. Horowitz, H., van Arsdell, J. & Platt, T. (1983) *J. Mol. Biol.* **169,** 775–797.
12. Yanofsky, C., Horn, V., Bonner, M. & Stasiowski, S. (1971) *Genetics* **69,** 409–433.
13. Kirschner, K., Szadkowski, H., Henschen, A. & Lottspeich, F. (1980) *J. Mol. Biol.* **143,** 395–409.
14. White, J. L., Grütter, M. G., Wilson, E., Thaller, C., Ford, G. C., Smit, J. D. G., Jansonius, J. N. & Kirschner, K. (1982) *FEBS Lett.* **148,** 87–90.
15. Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R. & Jansonius, J. N. (1981) *J. Mol. Biol.* **147,** 465–469.
16. Bisswanger, H., Kirschner, K., Cohn, W., Hager, V. & Hansson, E. (1979) *Biochemistry* **18,** 5946–5953.
17. Eichele, G., Ford, G. C., Glor, M., Jansonius, J. N., Mavrides, C. & Christen, P. (1979) *J. Mol. Biol.* **133,** 161–180.
18. Burnett, R. M., Grütter, M. G. & White, J. L. (1984) *J. Mol. Biol.* **185,** 105–123.
19. Dickerson, R. E., Weinzierl, J. E. & Palmer, R. A. (1968) *Acta Crystallogr. Sect. B* **24,** 997–1003.
20. Blow, D. M. & Crick, F. H. C. (1959) *Acta Crystallogr.* **12,** 794–802.
21. Hendrickson, W. A. & Lattman, E. E. (1970) *Acta Crystallogr. Sect. B* **26,** 136–143.
22. Wang, B.-C. (1985) *Methods Enzymol.* **115,** 90–112.
23. Jones, T. A. (1985) *Methods Enzymol.* **115,** 157–189.
24. Sim, G. A. (1959) *Acta Crystallogr.* **12,** 813–815.
25. Hendrickson, W. A. & Konnert, J. H. (1980) in *Computing in Crystallography*, eds. Diamond, R., Ramaseshan, S. & Venkatesan, K. (Indian Academy of Sciences, Bangalore), pp. 13.01–13.23.
26. Hendrickson, W. A. & Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation, and Evolution*, ed. Srinivasan, R. (Pergamon, Oxford), Vol. 1, pp. 43–57.
27. Agarwal, R. C. (1978) *Acta Crystallogr. Sect. A* **34,** 791–809.
28. Jack, A. & Levitt, M. (1978) *Acta Crystallogr. Sect. A* **34,** 931–935.
29. Vincent, M. G. & Priestle, J. P. (1985) *J. Appl. Crystallogr.* **18,** 185–188.
30. Artymiuk, P. J. & Blake, C. C. F. (1981) *J. Mol. Biol.* **152,** 737–762.
31. Alber, T., Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Rivers, P. S. & Wilson, I. A. (1981) *Philos. Trans. R. Soc. London Ser. B* **293,** 159–171.
32. Mavridis, I. M., Hatada, M. H., Tulinsky, A. & Lebioda, L. (1982) *J. Mol. Biol.* **162,** 419–444.
33. Stuart, D. I., Levine, M., Muirhead, H. & Stammers, D. K. (1979) *J. Mol. Biol.* **134,** 109–142.
34. Matsuura, Y., Kusunoki, M., Harada, W., Tanaka, N., Iga, Y., Yasuoka, N., Toda, H., Narita, K. & Kakudo, M. (1980) *J. Biochem. (Tokyo)* **87,** 1555–1558.
35. Carrell, H. L., Rubin, B. H., Hurley, T. J. & Glusker, J. P. (1984) *J. Biol. Chem.* **259,** 3230–3236.
36. Lindqvist, Y. & Brändén, C.-I. (1980) *J. Mol. Biol.* **143,** 201–211.
37. Schneider, G., Lindqvist, Y., Brändén, C.-I. & Lorimer, G. (1986) *EMBO J.* **5,** 3409–3415.
38. Goldman, A., Ollis, D. L. & Steitz, T. A. (1987) *J. Mol. Biol.* **194,** 143–153.
39. Horowitz, N. H. (1945) *Proc. Natl. Acad. Sci. USA* **31,** 153–157.
40. Horowitz, N. H. (1965) in *Evolving Genes and Proteins*, eds. Bryson, V. & Vogel, H. J. (Academic, New York), pp. 15–23.
41. Belfaiza, J., Parsot, C., Martel, A., Bouthier de la Tour, C., Margarita, D., Cohen, G. N. & Saint-Girons, I. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 867–871.
42. Crawford, I. P. (1975) *Bacteriol. Rev.* **39,** 87–120.
43. Cohn, W., Kirschner, K. & Paul, C. (1979) *Biochemistry* **18,** 5953–5959.
44. Levine, M., Muirhead, H., Stammers, D. K. & Stuart, D. I. (1978) *Nature (London)* **271,** 626–630.
45. Pflugfelder, M. (1986) Dissertation (Univ. of Basel, Basel, Switzerland).
46. Schechtman, M. G. & Yanofsky, C. (1983) *J. Mol. Appl. Genet.* **2,** 83–99.
47. Aebi, M., Furter, R., Prantl, F., Niederberger, P. & Hütter, R. (1984) *Curr. Genet.* **8,** 165–172.
48. Zalkin, H., Paluh, J. L., van Cleemput, M., Moye, W. S. & Yanofsky, C. (1984) *J. Biol. Chem.* **259,** 3985–3992.
49. Tschumper, G. & Carbon, J. (1980) *Gene* **10,** 157–166.
50. Kaplan, J. B., Goncharoff, P., Seibold, A. M. & Nichols, B. P. (1984) *Mol. Biol. Evol.* **1,** 456–472.
51. Mullaney, E. J., Hamer, J. E., Roberti, K. A., Yelton, M. M. & Timberlake, W. E. (1985) *Mol. Gen. Genet.* **199,** 37–45.
52. Matsui, K., Sano, K. & Ohtsubo, E. (1986) *Nucleic Acids Res.* **14,** 10113–10114.
53. Peñalva, M. A. & Sánchez, F. (1987) *Nucleic Acids Res.* **15,** 1874.
54. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453.
55. Bacon, D. J. & Anderson, W. F. (1986) *J. Mol. Biol.* **191,** 153–161.
56. Jardetzky, T. S. (1986) Dissertation (Univ. of Basel, Basel, Switzerland).
57. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. (1985) *Nature (London)* **313,** 762–766.