Role of the N-Terminal Extension of the $(\beta\alpha)_8$-Barrel Enzyme Indole-3-glycerol Phosphate Synthase for Its Fold, Stability, and Catalytic Activity[†,‡]

Birgit Schneider,[§,||,⊥] Thorsten Knöchel,[#,△] Beatrice Darimont,[○,◇] Michael Hennig,[#,□] Susanne Dietrich,[§] Karin Babinger,[§] Kasper Kirschner,[○] and Reinhard Sterner*[,§,||]

Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany, Institut für Biochemie, Universität zu Köln, Otto-Fischer-Strasse 12-14, D-50674 Köln, Germany, and Abteilung für Strukturbiologie und Abteilung für Biophysikalische Chemie, Universität Basel, Biozentrum, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Received August 17, 2005; Revised Manuscript Received October 21, 2005

ABSTRACT: Indole-3-glycerol phosphate synthase (IGPS) catalyzes the fifth step in the biosynthesis of tryptophan. It belongs to the large and versatile family of $(\beta\alpha)_8$-barrel enzymes but has an unusual N-terminal extension of about 40 residues. Limited proteolysis with trypsin of IGPS from both *Sulfolobus solfataricus* (sIGPS) and *Thermotoga maritima* (tIGPS) removes about 25 N-terminal residues and one of the two extra helices contained therein. To assess the role of the extension, the N-terminally truncated variants sIGPSΔ(1−26) and tIGPSΔ(1−25) were produced recombinantly in *Escherichia coli*, purified, and characterized in comparison to the wild-type enzymes. Both sIGPSΔ(1−26) and tIGPSΔ(1−25) have unchanged oligomerization states and turnover numbers. In contrast, their Michaelis constants for the substrate 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate are increased, and their resistance toward unfolding induced by heat and guanidinium chloride is decreased. sIGPSΔ(1−26) was crystallized, and its X-ray structure was solved at 2.8 Å resolution. The comparison with the known structure of sIGPS reveals small differences that account for its reduced substrate affinity and protein stability. The structure of the core of sIGPSΔ(1−26) is, however, unchanged compared to sIGPS, explaining its retained catalytic activity and consistent with the idea that it evolved from the same ancestor as the phosphoribosyl anthranilate isomerase and the α-subunit of tryptophan synthase. These $(\beta\alpha)_8$-barrel enzymes catalyze the reactions preceding and following IGPS in tryptophan biosynthesis but lack an N-terminal extension.

The $(\beta\alpha)_8$- or TIM-barrel structure seems to be nature's favorite scaffold to constitute enzymic active sites (*1*−*3*). The canonical $(\beta\alpha)_8$ barrel is composed of eight $\beta\alpha$ units, in which a β-strand is connected to an α-helix by a $\beta\alpha$ loop. The individual $\beta\alpha$ units are linked by $\alpha\beta$ loops. In all known $(\beta\alpha)_8$-barrel enzymes, active site residues are located at the C-termini of some of the parallel β-strands and also in the $\beta\alpha$ loops. This similarity suggests that a large fraction of $(\beta\alpha)_8$-barrel enzymes may have arisen from an ancestral precursor by divergent evolution, despite the wide range of chemically diverse reactions catalyzed by them (*4*−*6*). Three consecutively operating enzymes of the pathway of tryptophan biosynthesis, namely, phosphoribosyl anthranilate isomerase (PRAI),[1] indole-3-glycerol phosphate synthase (IGPS), and the α-subunit of tryptophan synthase (TSα), possess the $(\beta\alpha)_8$-barrel fold. They have a common phosphate binding motif in $\beta\alpha$ loops 7 and 8 (*7*). The structural and functional relationships between these enzymes, as well as the similarity of their substrates, suggest that they have evolved from a common ancestor.

IGPS, which catalyzes the conversion of 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP) to indole-3-glycerol phosphate (IGP) (Figure 1), carries an N-terminal extension of the first β-strand (Figure 2). This extension is absent in PRAI and TSα, which seems to contradict the hypothesis of divergent evolution. This N-terminal extension is about 40 residues long and includes both helix α0 that lines the active site cleft and helix α00 at the opposite face of the barrel. Earlier deletion studies (*8*) found that the

[†] This work was supported by the Deutsche Forschungsgemeinschaft (Grants STE 891/4-1 and STE 891/4-2 to R.S.) and the Swiss National Science Foundation (Grant 31-32369.91 to K.K.).

[‡] The coordinates of sIGPSΔ(1−26) were deposited in the Protein Data Bank (PDB entry code 2c3z).

* Corresponding author. Phone: +49-941 943 3015. Fax: +49-941 943 2813. E-mail: Reinhard.Sterner@biologie.uni-regensburg.de.

[§] Institut für Biophysik und physikalische Biochemie, Universität Regensburg.

[||] Institut für Biochemie, Universität zu Köln.

[⊥] Current address: Institut für Biophysikalische Chemie, Universität Frankfurt, Marie-Curie-Strasse 9, D-60439 Frankfurt, Germany.

[#] Abteilung für Strukturbiologie, Universität Basel.

[△] Current address: Protein Crystallography (ZD-A/ZFA), Merck KGaA, Frankfurter Strasse 250, D-64293 Darmstadt, Germany.

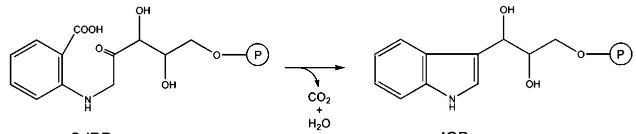
[○] Abteilung für Biophysikalische Chemie, Universität Basel.

[◇] Current address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403-1229.

[□] Current address: Pharma Research Discovery, 65/319, F. Hoffmann-La Roche Ltd., CH-4070 Basel, Switzerland.

[1] Abbreviations: PRAI, phosphoribosyl anthranilate isomerase; IGPS, indole-3-glycerol phosphate synthase; sIGPS, IGPS from *Sulfolobus solfataricus*; tIGPS, IGPS from *Thermotoga maritima*; sIGPSΔ(1−26) and tIGPSΔ(1−25), N-terminally truncated sIGPS and tIGPS; *strpC*, *ttrpC*, *strpC*Δ(1−26), and *ttrpC* Δ(1−25), genes encoding sIGPS, tIGPS, sIGPSΔ(1−26), and tIGPSΔ(1−25); TSα, α-subunit of tryptophan synthase; CdRP, 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate; rCdRP, reduced CdRP.

10.1021/bi051640n CCC: $30.25
Published on Web 11/23/2005

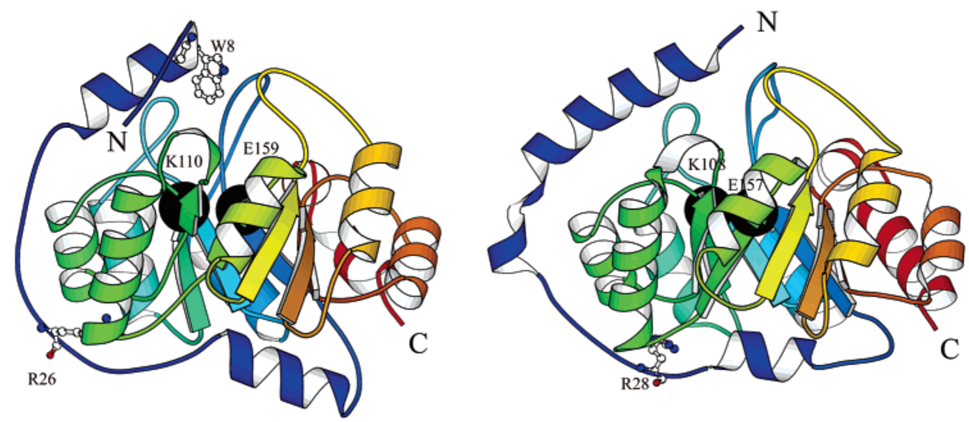


CdRP IGP

FIGURE 1: Reaction catalyzed by IGPS. The substrate 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP) is converted into the products indole-3-glycerol phosphate (IGP), carbon dioxide, and water.

N-terminal portion of helix α0 is not essential for catalysis but rather for effective substrate binding, suggesting that the catalytic function of IGPS is sustained even by the canonical $(\beta\alpha)_8$ barrel.

To elucidate the role of the N-terminal extension of IGPS further, we constructed more extensively truncated variants of IGPS from *Sulfolobus solfataricus* [sIGPS (*9*)] and *Thermotoga maritima* [tIGPS (*10*)] and investigated their catalytic and structural properties as well as their thermostabilities. The K_M values for the substrate 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP) of the recombinantly truncated variants sIGPSΔ(1−26) and tIGPSΔ(1−25) are drastically increased compared to those of the wild-type IGPSs, and their stabilities toward heat inactivation and chemical unfolding are decreased. However, the catalytic turnover numbers are unchanged. The X-ray structure of sIGPSΔ(1−26) has been solved at 2.8 Å resolution by molecular replacement. A comparison with native sIGPS (*9*) shows that the entire helix α0 can be removed without significant effects on the structural core of the enzyme. These results therefore support a common evolutionary origin of the three $(\beta\alpha)_8$-barrel enzymes of tryptophan biosynthesis.



sIGPS tIGPS

```
                        α0                     α00             β1                     α1        β2
                   HHHHHHHHHHHH               HHHHHH         SSSSSS               HHHHHHHHHHH  SSSSS          HH
                         10        20        30        40           50        60        70        80        90
tIGPS           MRRLWEIVEAKKKDILEIDGENLIVQRRNHRFLEVLSGKE---RVKIIAEFKKASPSAGDINADASLEDFIRMYDELADAISILTEKHYFKGDPAF
                   |   |                 ||      |            |||| |  |||                    |   §||§|  |§ §
sIGPS      MPRYLKGWLKDVVQLSL-RRPSFRAS---RQRPIISLNERILEFNKRNITAIIAEYKRKSPSG-LDV-ERDPIEYSKFMERYAVGLSILTEEKYFNGSYET
                    10         20           30        40        50         60         70        80        90
                 HHHHHHHHHHH                   HHHHHHHHHH     SSSSS               HHHHHHHH     SSSSS         HHH
                      α0                           α00          β1                   α1         β2

              α2        β3          α3         β4           α4          β5        α5          β6             α6
            HHHHH      SSSS      HHHHHHHH    SSSSS     HHHHHHHHHHHHH  SSSSSS   HHHHHHHHHH    SSSSS           HHH
                  100       110       120       130       140       150       160       170       180       190
tIGPS       VRAARNLTCRPILAKDFYIDTVQVKLASSVGADAILIIARILTAEQIKEIYEAAEELGMDSLVEVHSREDLEKVFSVIRPKIIGINTRDLDTFEIKKNVL
             |        §|| §§§     |   |   §§|  § §  |||        | |   ||  § §     ||      |    |§|§ §|§ | || |
sIGPS       LRKIASSVSIPILMKDFIVKESQIDDAYNLGADTVLLIVKILTERELESLLEYARSYGMEPLIEINDENDLDIALR-IGARFIGINSRDLETLEINKENQ
                100       110       120       130       140       150       160       170        180       190
            HHHHHHH    SSSSS    HHHHHHHHHH   SSSS      HHHHHHHHHHHHH    SSSS   HHHHHHHHH     SSSSS          HHHH
              α2        β3          α3        β4             α4           β5      α5           β6            α6

                         β7         α7         β8    α8´        α8
            HH         SSSSS     HHHHHHH      SSSS HHHHH   HHHHHHHHHHH
                  200       210       220        230       240       250
tIGPS       WELLPLVPDDTVVVAESGIKDPRELKDLRG-KVNAVLVGTSIMKAENPRRFLEEMRAWSE
              |    |   | ||§§||    |   ||   ||| § § | §    |     |
sIGPS       RKLISMIPSNVVKVAESGISERNEIEELRKLGVNAFLIGSSLMRN--PE-KIKEFIL
                 200       210       220       230         240
            HHHH       SSSSS     HHHHHHHHH    SSSS HHHHH    HHHHHHH
                        β7         α7          β8    α8´       α8
```

FIGURE 2: (A) Ribbon representations of the sIGPS (PDB entry code 1IGS) and tIGPS (PDB entry code 1I4N) structures. The drawing gives the N-terminal extension in dark blue and the subsequent eight canonical $\beta\alpha$ modules in rainbow colors. Limited tryptic proteolysis of sIGPS and tIGPS occurs at the C-termini of R26 and R28, respectively, which are shown in ball-and-stick representation. W8, which is located on helix α0 of sIGPS and contributes to the hydrophobic pocket at the active site, is also shown in ball-and-stick representation. The catalytically essential lysine (K110 and K108, respectively) and glutamate (E159 and E157, respectively) residues are represented as black spheres. (B) Amino acid sequence alignment between sIGPS and tIGPS, performed with CLUSTAL W (*26*; http://www.ebi.ac.uk/clustalw). Residues belonging to α-helices and β-strands are marked with H and S, respectively. Identical residues between sIGPS and tIGPS are indicated by (|), and invariant residues for a set of 19 IGPS sequences are indicated by (§). The same residues are marked as in (A).

MATERIALS AND METHODS

DNA Manipulation and Sequence Analysis. Preparation of DNA, amplification, extraction with restriction endonucleases, and sequencing were performed as described (*11*).

Cloning of strpC, strpCΔ(1−26), ttrpC, and ttrpCΔ(1−25). The s*trp*C gene encoding sIGPS was amplified by PCR, using s*trp*C-pDS56/RBSII/SphI (*9*) as the template and the oligonucleotides 5′-CTA **GCT AGC** ATG CCA CGT TAT-3′ and 5′-CCG **CTC GAG** TAG TAT AAA TTC-3′ (restriction sites in bold) as 5′- and 3′-primers. The t*trp*C gene encoding tIGPS was also amplified by PCR, using t*trp*C-pET21a(+) (*12*) as the template and the oligonucleotides 5′-GGG AAT TC**C ATA TG**A GAA GAC TC-3′ and 5′-CCG CTC **GAG CTC** TGA CCA TGC-3′ (restriction sites in bold) as 5′- and 3′-primers. Both genes were cloned into pET21a(+), using the restriction endonucleases *Nhe*I and *Xho*I for s*trp*C and *Nde*I and *Xho*I for t*trp*C. The 3′-primers for the amplification of s*trp*C and for the amplification of t*trp*C were designed such that the internal stop codons were deleted. As a consequence, six plasmid-encoded histidine codons were added to the 3′-end of the genes, which allowed the purification of the protein products by metal chelate affinity chromatography. The s*trp*CΔ(1−26) and t*trp*CΔ(1−25) genes encoding sIGPSΔ(1−26) and tIGPSΔ(1−25) were cloned into pET21a(+) in an analogous way as described for the wild-type genes, using the same templates, the same restriction endonucleases, and the same 3′-primers. However, the 5′-primers 5′-CTA **GCT AGC** ATG CCA CGT TAT-3′ and 5′-GGG AAT TC**C ATA TG**A GAA GAC TC-3′ (restriction sites in bold) were designed such that the first 78 bp and 75 bp were deleted in s*trp*C and t*trp*C, respectively.

Heterologous Gene Expression and Purification of the Protein Products. Expression was conducted in *Escherichia coli* BL21-CodonPlus(DE3) RIPL cells (Stratagene), which contain the gene for the phage T7 polymerase on their chromosome and plasmids for tRNA genes that are rarely used in *E. coli*. An overnight culture of freshly transformed cells containing LB medium with ampicillin (0.15 mg/mL), chloramphenicol (0.03 mg/mL), and streptomycin (0.1 mg/mL) was used to inoculate 2 L of the same medium, which was stirred at 37 °C until the OD_{600} reached a value between 0.5 and 0.6. Expression was induced by adding 1 mM IPTG and continued overnight. Cells were harvested by centrifugation (Sorvall RL 5B plus, GS3 rotor, 4000 rpm, 20 min, 4 °C), resuspended in 30 mL of 100 mM potassium phosphate buffer, pH 7.5, lysed by sonification (Branson sonifier W-250 D, 2 × 2 min, amplitude 50%), and centrifuged again (Sorvall RC 5 plus, rotor 22500, 12000 rpm, 20 min, 4 °C). The analysis by SDS−PAGE showed that sIGPS was present in equal amounts in the soluble and the insoluble fractions of the cell extract. In contrast, both sIGPSΔ(1−26) and tIGPSΔ(1−25) were found almost exclusively in the soluble fraction. For unknown reasons, tIGPS was found mainly in the soluble fraction of cultures that were grown without IPTG but was absent from both the soluble and insoluble fraction of cultures containing 1 mM IPTG. All proteins were purified from the soluble fraction of the cell extract, which was first dialyzed against 10 mM potassium phosphate, pH 7.5. To fragment nucleic acids, 250 units of benzonase and 2 mM $MgCl_2$ were added, followed by an incubation for 1 h at 37 °C. In the case of the wild-type enzymes sIGPS and tIGPS the extracts were then incubated at 75 °C for 15 min to precipitate thermolabile host proteins. In the case of the N-terminally truncated variants, heating was performed at 55 and 65 °C, respectively. However, these temperatures were not high enough for an efficient enrichment of either sIGPSΔ(1−26) and tIGPSΔ(1−25) and were therefore omitted in most purifications. The heat-precipitated host proteins were removed by centrifugation (Sorvall RC 5 plus, rotor 22500, 12000 rpm, 20 min, 4 °C), and the supernatant was applied at room temperature onto a Ni^{2+}-chelate affinity column (ÄKTA basic 10 better chromatography system, Amersham Bioscience; POROS MC/20 column, volume 7.8 mL, Applied Biosystems) that had been equilibrated with 100 mM potassium phosphate, pH 7.5, and 300 mM KCl. The bound proteins were eluted with 10 column volumes of a linear gradient of 0−600 mM imidazole dissolved in the same buffer. sIGPS, tIGPS, and their N-terminally truncated variants eluted between 50 and 200 mM imidazole. The fractions containing pure protein (as judged by SDS−PAGE) were pooled, dialyzed against 100 mM potassium phosphate buffer, pH 7.5, concentrated by ultrafiltration, dripped into liquid nitrogen, and stored at −80 °C. The yields were between 10 and 40 mg of protein from 2 L of growth medium.

Trypsin Digestion and N-Terminal Sequencing. sIGPS and tIGPS were incubated with trypsin at a molar ratio of 10:1 in 100 mM potassium phosphate, pH 7.5 at 25 °C. After different time intervals, aliquots were taken, and the reaction was stopped by adding an equal volume of 2 × SDS−PAGE buffer, followed by SDS−PAGE. The proteins were Western-blotted (semidry transfer unit model TE70, Hoefer Pharmacia Biotech) onto a polyvinylidene fluoride membrane, stained by Coomassie blue, cut from the dried membrane, and N-terminally sequenced by automated Edman degradation (Applied Biosystems Procise).

Analytical Methods, Steady-State Enzyme Kinetics, and Ligand Titration Studies. SDS−PAGE, analytical gel filtration, fluorescence emission and circular dichroism spectroscopy, and estimation of protein concentrations were performed as described (*13*). The catalytic activity of IGPS variants using the enzymatically produced substrate CdRP was monitored fluorometrically as described (*14*, *15*). The values for the steady-state enzyme kinetic parameters k_{cat} and K_M^{CdRP} were determined from saturation curves, which were constructed from initial velocity measurements recorded at 25 °C in 50 mM EPPS buffer, pH 7.5, 4 mM EDTA, and 1 mM DTT in the presence of various concentrations of CdRP. Whereas sIGPS, tIGPS, and tIGPSΔ(1−25) showed hyperbolic saturation curves, the saturation curve of sIGPSΔ(1−26) was linear over the entire tested concentration range (0−312 μM) of CdRP. As a consequence, only the catalytic efficiency paramter k_{cat}/K_M^{CdRP} could be determined for this variant but not the separate k_{cat} and K_M^{CdRP} values. The binding of the substrate analogue reduced CdRP (rCdRP) to tIGPS and tIGPSΔ(1−25) at 25 °C in 50 mM EPPS buffer, pH 7.5, 4 mM EDTA, and 1 mM DTT was followed by fluorescence energy transfer and analyzed as described (*16*).

Heat Inactivation. The kinetics of irreversible heat inactivation were measured as described (*9*). The proteins were incubated in 50 mM EPPS buffer, pH 7.0, containing 4 mM EDTA at 75 °C [sIGPS and sIGPSΔ(1−26)] or 80 °C [tIGPS

Table 1: Data Collection and Refinement Statistics of sIGPSΔ(1−26)[a]

Data Collection	
unique reflections	5474
max resolution (Å)	2.7
redundancy	2.8
completeness (%)	96.4
$\langle I/\sigma \rangle$	5.6
reflections $>3\sigma$ (%)	71.5
R_{sym}	0.102
Refinement	
resolution range (Å)	10.0−2.8
mean *B* value of protein atoms (Å^2)	19.8
mean *B* value of ligand atoms (Å^2)	23.4
mean *B* value of water atoms (Å^2)	31.7
rms distances (Å)	0.011
rms bond angles (deg)	1.874
R-factor (all data)	0.18

[a] $R_{sym} = \sum_{hkl}\sum_i||F_i(hkl)| - |\overline{F(hkl)}||/\sum_{hkl}\sum_i|F_i(hkl)|$. $R = \sum_{hkl}||F_o| - k|F_c||/\sum_{hkl}|F_o|$.

and tIGPSΔ(1−25)]. Aliquots were taken after different times, chilled on ice, and centrifuged (Eppendorf 5415R, 13000 rpm, 3 min, 4 °C). The residual activity of the supernatant was measured at 50 °C in 50 mM EPPS buffer, pH 7.0, 4 mM EDTA, and 2 mM DTT. The half-lives of inactivation were calculated by fitting the decrease of the initial velocities to a single exponential function (program SigmaPlot 5.0).

Chemical Unfolding. sIGPS, sIGPSΔ(1−26), tIGPS, and tIGPSΔ(1−25) were incubated with various concentrations of guanidinium chloride dissolved in 50 mM potassium phosphate, pH 7.5 at 25 °C. Protein unfolding was followed by the decrease of the circular dichroism signal at 220 nm [not for tIGPSΔ(1−25), which generated ambiguous results] and of the fluorescence emission at 320 nm after excitation with 280 nm [not for sIGPSΔ(1−26), which contains no tryptophan residues]. The signals were followed after different time intervals until no further change was observed. Moreover, the proteins were completely unfolded in 6 M guanidinium chloride, and refolding was induced by removing the chaotropic reagent by means of dilution into 50 mM potassium phosphate, pH 7.5 at 25 °C.

Crystallization of sIGPSΔ(1−26) and Data Collection. Crystallization was performed by the hanging drop vapor diffusion method at 4 °C. Crystals were grown by mixing equal volumes of a reservoir solution containing 30% PEG 8000 and 0.2 M ammonium sulfate as precipitants in 0.1 M MES buffer at pH 6.5, with protein concentrated to 40 mg/mL. X-ray diffraction data from a single crystal of approximate size 0.1 × 0.1 × 0.3 mm^3 were collected on a MAR-Research image plate area detector mounted on a rotating anode generator. Data evaluation was performed with MOSFLM and the CCP4 suite of programs (*17*). The space group of the measured crystal is $P2_12_12_1$ with cell dimensions $a = 39.2$ Å, $b = 61.4$ Å, and $c = 80.4$ Å. The crystal, containing one molecule per asymmetric unit ($V_M = 1.92$), had a solvent content of 35.9%. Statistics on the data collection are summarized in Table 1.

Structure Solution and Refinement. The structure of sIGPSΔ(1−26) was solved by molecular replacement [X-PLOR (*18*)] using the structure of sIGPS that was truncated by the first 26 amino acids as the search model (*9*) (PDB entry code 1IGS). The final *R*-factor of the refined model including 130 water molecules and one sulfate ion is 18.0% including all data in the range from 10.0 to 2.8 Å. For a summary of refinement statistics see Table 1. Of all residues 85.6% are found in the most favored region of the Ramachandran diagram. All other residues except S211 are located in additionally allowed areas. As in wild-type sIGPS (*9*, *19*) S211 adopts an unfavorable conformation with $\varphi = 49.7°$ and $\psi = 152.7°$.

RESULTS AND DISCUSSION

Preparation of sIGPSΔ(1−26) and tIGPSΔ(1−25). Because both sIGPS and tIGPS carry about 10 Arg and Lys residues in the N-terminal sequences preceding strand β1 (Figure 2), we tested their partial or complete removal by trypsin at 25 °C. The progress of the reaction was monitored by SDS−PAGE, which identified one major digestion product for each enzyme. N-Terminal sequencing revealed that trypsin cleaved sIGPS about as efficiently at the C-termini of either R23 or R26 and tIGPS exclusively at the C-terminus of R28. The sequence alignment (Figure 2B) shows that R26 in sIGPS corresponds to V25 in tIGPS, which are located in the loop connecting helices α0 and α00. The canonical $(\beta\alpha)_8$ barrels are apparently resistant to tryptic cleavage under these conditions. To obtain large amounts of sIGPS, sIGPSΔ(1−26), tIGPS, and tIGPSΔ(1−25), the corresponding genes were cloned by PCR into pET21a(+) and were heterologously expressed in *E. coli*. The recombinant proteins were enriched from the soluble cell extract. As judged by SDS−PAGE, the final preparations were more than 95% pure.

Oligomerization States. Both sIGPS (*9*) and tIGPS (*10*) are intrinsically monomeric proteins. tIGPS can dimerize, however, by oxidative disulfide formation between its unique, exposed cysteine (C102) residue (*20*). Analytical gel filtration was performed to check whether the N-terminal deletion of about 25 residues affects the oligomerization states of both proteins. sIGPS and sIGPSΔ(1−26) eluted as single symmetric peaks corresponding to molecular masses of 23.5 and 26.3 kDa (data not shown), which are somewhat lower and identical, respectively, than the calculated monomer values of 28.6 and 26.4 kDa. In contrast, both tIGPS and tIGPSΔ(1−25) eluted as two peaks. The trailing peaks correspond to molecular masses of 24.2 and 28.2 kDa, similar to the calculated monomer values of 28.7 and 26.6 kDa. The leading peaks correspond to dimer masses of 39.3 and 47.1 kDa. Obviously, the tendency of tIGPS to form covalently linked dimers seems to be retained in tIGPSΔ(1−25). Thus, we conclude that the N-terminal truncations do not change the association states of sIGPS or tIGPS.

Catalytic Activities. To better understand the functional role of the N-terminal extensions, steady-state enzyme kinetic studies were undertaken to compare the wild-type enzymes with their truncated variants. Table 2 summarizes the turnover numbers (k_{cat}) and the Michaelis constants (K_M^{CdRP}) as well as the catalytic efficiency parameters (k_{cat}/K_M^{CdRP}) at 25 °C that were deduced from substrate saturation curves. Since the catalytic rate of sIGPSΔ(1−26) is a linear function of CdRP up to the highest assayed concentration of 312 μM, its K_M^{CdRP} value must be strongly elevated compared to that of sIGPS. From the k_{cat}/K_M^{CdRP} value, which was deduced

Table 2: Steady-State Enzyme Kinetic Constants of sIGPS, tIGPS, and Their Truncated Variants at 25 °C[a]

protein	k_{cat} (s^{-1})	K_M^{CdRP} (μM)	k_{cat}/K_M^{CdRP} ($s^{-1}\,\mu M^{-1}$)	R[d]
sIGPS	0.025 ± 0.007	0.09 ± 0.05	0.28 ± 0.2	1.0
sIGPSΔ(1−9)[b]	0.041	9.9	0.0042	5.8×10^{-3}
sIGPSΔ(1−26)[c]	>0.028	>312	9.2×10^{-5}	3.3×10^{-4}
tIGPS	0.095 ± 0.03	0.025 ± 0.02	3.8 ± 1.0	1.0
tIGPSΔ(1−25)	0.06 ± 0.02	0.69 ± 0.03	0.087 ± 0.06	2.3×10^{-2}

[a] Conditions: 50 mM EPPS buffer, pH 7.5, 4 mM EDTA, and 1 mM DTT. The mean values and standard deviations from three independent measurements are shown. [b] Values from ref *8*. [c] v_i is a linear function of the substrate concentration up to at least [CdRP] = 312 μM. [d] Ratio $(k_{cat}/K_M)^{var}/(k_{cat}/K_M)^{wt}$.

Table 3: Stabilities of sIGPS, tIGPS, and Their N-Terminally Truncated Variants toward Irreversible Heat-Induced Inactivation[a] and Reversible Chemical Unfolding[b]

protein	incubation temp (°C)	half-life (min)	$D_{1/2}$[c] (M GdmCl)
sIGPS	75	123	2.7
sIGPSΔ(1−26)	75	2.9	2.1
tIGPS	80	258	2.0
tIGPSΔ(1−25)	80	6.3	1.9

[a] Incubation buffer: 50 mM EPPS, pH 7.0 at the indicated temperature, and 4 mM EDTA. [b] Incubation buffer: 50 mM potassium phosphate, pH 7.5 at 25 °C, and various concentrations of guanidinium chloride (GdmCl). [c] Concentration of GdmCl that is necessary to unfold 50% of the protein.

from the slope of the graph, it follows that its turnover number must also be increased compared to sIGPS. Similarly, sIGPSΔ(1−9) lacking the N-terminal loop and the first turn of helix α0 has a k_{cat} value comparable to that of sIGPS and a 100-fold increased K_M^{CdRP} value (*8*). By selecting mutants of sIGPS that promote increased turnover at low temperatures, among others, the variant sIGPS-P2QΔ(4−28) was isolated (*21*). It contains a large deletion similar to sIGPSΔ(1−26). The k_{cat} value of tIGPSΔ(1−25) is practically identical to that of tIGPS. However, its K_M^{CdRP} value is increased almost 30-fold. Fluorometric titration of tIGPSΔ(1−25) with the substrate analogue reduced CdRP (rCdRP) yielded a value for the thermodynamic dissociation constant (K_D^{rCdRP}) of 0.46 μM, which is similar to the K_M^{CdRP} value of 0.69 μM (Table 2). Titration of 1 μM tIGPS with rCdRP resulted in a linear curve up to about equal molar concentrations of the ligand. These findings suggest that the K_D^{rCdRP} of tIGPS is at least 1 order of magnitude smaller than the K_D^{rCdRP} of tIGPSΔ(1−25). Taken together, these results show that the entire helix α0 is expendable for catalysis of the IGPS reaction but that its stepwise deletion results in a stepwise and dramatic loss of substrate binding affinity.

Stability Measurements. Thermal stabilities were assessed from the kinetics of irreversible heat inactivation. The proteins were incubated at 75 °C [sIGPS and sIGPSΔ(1−26)] or 80 °C [tIGPS and tIGPSΔ(1−25)]. Samples were drawn after different time intervals, chilled on ice, and tested for their residual activities at 50 °C. The resulting half-lives ($t_{1/2}$) show that the N-terminally truncated variants lose activity about 40-fold faster than the wild-type enzymes (Table 3). The resistance toward unfolding induced by guanidinium chloride (GdmCl) was monitored by circular dichroism and fluorescence spectroscopy to follow the loss in secondary and tertiary structure, respectively. For all proteins, chemical denaturation was reversible. The concentration of GdmCl ($D_{1/2}$) necessary to unfold 50% of the wild-type proteins compared to the truncated variants was somewhat higher [sIGPS versus sIGPSΔ(1−26)] and almost identical [tIGPS versus tIGPSΔ(1−25)] (Table 3). These results suggest that the native extension, including helix α0, considerably increases the kinetic barrier for inactivation of both sIGPS and tIGPS but has only a limited influence on their conformational stabilities. The removal of the N-terminal extension also seems to modify the chemically induced unfolding mechanism of IGPS at equilibrium. At pH 7 only the fully native and unfolded states of sIGPS are populated. At pH 9, however, an equlibrium unfolding intermediate of sIGPS was detected (*22*). In contrast, sIGPSΔ(1−26) seems to unfold via an intermediate even at pH 7 (*23*). It was argued that, at pH 7, the N-terminal extension may selectively stabilize the native state relative to the intermediate by forming salt bridges and hydrogen bonds with the core of the barrel (Table 4). These interactions would be absent in sIGPSΔ(1−26) and weakened at pH 9 in sIGPS, due to the titration of positively charged groups with abnormally low pK_a values (*22*). As a consequence, the relative stabilization of the native state would no longer occur, and the population of the intermediate would increase to a detectable level (*23*). Along the same lines, the unfolding of tIGPSΔ(1−25) at pH 7 takes place in several phases, which suggests that an intermediate is significantly populated. At the same pH value, the unfolding of tIGPS occurs in a

Table 4: Hydrogen Bonds[a] between the 26 N-Terminal Residues and the $(\beta\alpha)_8$-Barrel Core of sIGPS[b]

N-terminal extension	barrel core	distance (Å)
Pro2 N	Glu141 Oε2	2.73
Arg3 N	Ile136 O	2.98
Arg3 Nε	Leu137 O	2.93
Arg3 Nη2	Val134 O	3.14
Lys6 Nζ	Glu185 Oε2	3.56
Trp8 Nε1	Leu184 O	2.76
Asp11 Oδ2	Lys87 Nζ	3.05
Asp11 Oδ2	Tyr88 Oη	2.58
Leu17 O	Lys115 Nζ	2.83
Arg19 O	Lys115 Nζ	2.66
Arg19 Nη1	Thr84 O	3.22
Arg19 Nη1	Asp111 Oδ2	3.06
Arg19 Nη2	Thr84 O	2.95
Phe22 N	Asp121 Oδ1	2.76
Phe22 O	Asn124 Oδ1	3.44
Arg26 Nε	Arg28 O	2.60
Arg26 Nη1	Tyr123 O	2.83
Arg26 Nη1	Ala127 O	2.73
Arg26 Nη2	Pro29 O	3.55
Arg26 Nη2	Gly126 O	3.57
Arg26 Nη2	Asp128 Oδ1	3.23

[a] As judged by the following criteria: distance between donor (D) and acceptor (A) ≤ 3.7 Å; angle C_{adj}−D···A > 90° and angle C_{adj}−A···D > 90°. C_{adj} denote the carbon atoms covalently bound to the respective donor (D) or acceptor (A) atoms. [b] PDB entry code 1IGS (*9*).

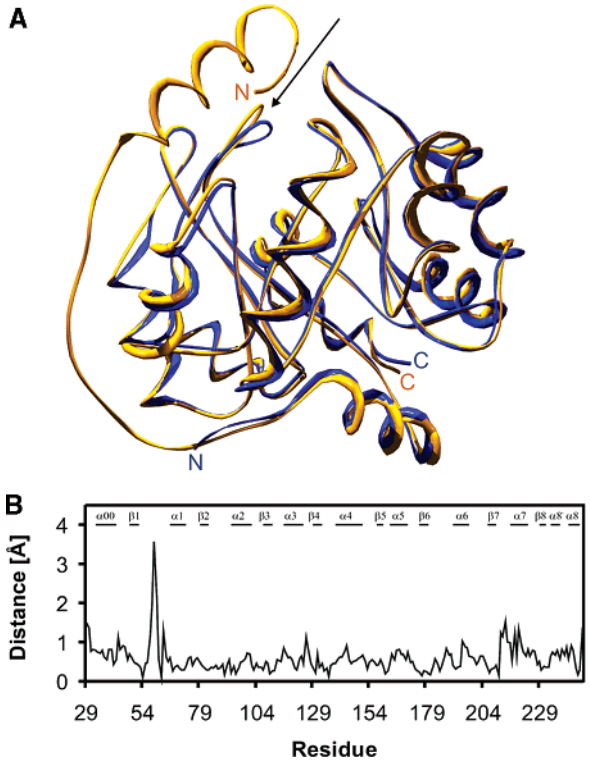


FIGURE 3: (A) Superposition of the Cα traces of sIGPS (yellow; PDB entry code 1IGS) and sIGPSΔ(1−26) (blue; PDB entry code 2c3z) with Swiss PDB viewer. The N- and C-termini of the protein chains are labeled. (B) Distances between equivalent Cα positions of wild-type sIGPS and sIGPSΔ(1−26). The arrow in (A) and the large peak in (B) mark residues 57−60. The deviation between the two structures in this region is not related to the truncation of the N-terminal extension but is due to the high flexibility in the wild-type structure (*19*).

more cooperative manner, in accordance with the presence of lower concentrations of the intermediate (data not shown).

X-ray Structure of sIGPSΔ(1−26). The crystal structure of sIGPSΔ(1−26) was solved to understand better the different performances and stabilities of sIGPS and sIGPSΔ(1−26). It has the same $(\beta\alpha)_8$-barrel fold as the wild-type enzyme except that the entire helix α0 and the first half of loop α0α00 are missing. No electron density was found for the N-terminal residues Q27 and R28. These residues were therefore discarded from the model. For a more detailed comparison, the C_α traces of sIGPS (PDB entry code 1IGS) and sIGPSΔ(1−26) (PDB entry code 2c3z) were superimposed (Figure 3A). Figure 3B depicts the distances between equivalent Cα positions as a function of the residue number (rms deviation 0.68 Å). The largest deviations are found between residues P57 and L60 in loop β1α1. However, this region is highly variable in different crystal forms of wild-type sIGPS as well, depending on pH and ionic strength (*19*). Therefore, the observed deviations seem not to be caused by deleting the N-terminal extension, and the structure of the canonical $(\beta\alpha)_8$-barrel core of sIGPS is unaffected by removal of helix α0.

The IGPS reaction consists of a sequence of condensation, decarboxylation, and dehydration steps and is practically irreversible, due to both the formation of the pyrrole ring of indole and the release of CO_2. X-ray structure (*24*) and mutational analysis (*24*, *25*) imply that the side chains of K110 and E159 of sIGPS act in the reaction as general acid and general base and that residues E51 and K53 assist catalysis by forming specific salt bridges with K110 and the carboxylate moiety of the substrate CdRP. The location and conformations of all of these residues are identical between sIGPS and sIGPSΔ(1−26) within experimental error, which explains why the k_{cat} values of the N-terminally truncated variants are practically identical to those of the wild-type enzymes (Table 2). In the wild-type enzyme, the hydrophobic pocket that accommodates the benzene moiety of the substrate CdRP and the substrate analogue reduced CdRP (rCdRP) (*24*) is formed by residues P57 (loop β1α1), F89 (loop β2α2), and F112 (loop β3α3), the hydrophobic tail of R182, L184, and L187 (loop β6α6), and W8 (on helix α0). Figure 4 shows the hydrophobic pocket of sIGPSΔ(1−26) superimposed on that of wild-type sIGPS in complex with rCdRP (*24*). The backbone conformation of loop β1α1 with P57 in sIGPSΔ(1−26) differs slightly from that of the wild-type enzyme due to its intrinsic flexibility as described in Figure 3. In contrast, the backbones of loops β2α2, β3α3, and β6α6 are essentially identical in both structures. The side chain of F89 remains unchanged upon removal of helix α0. However, the phenyl ring of F112 is rotated by approximately 30° around χ^2, and the side chain of L184 has shifted toward the solvent. As shown in Table 2, deletion of residues 1−9 of sIGPS does not alter k_{cat} appreciably but increases K_M^{CdRP} significantly. Since residues 1−7 are distant from the substrate binding site, it is likely that the loss of W8 (and posssibly L9) is mainly responsible for the decreased substrate binding affinity of sIGPSΔ(1−9). The

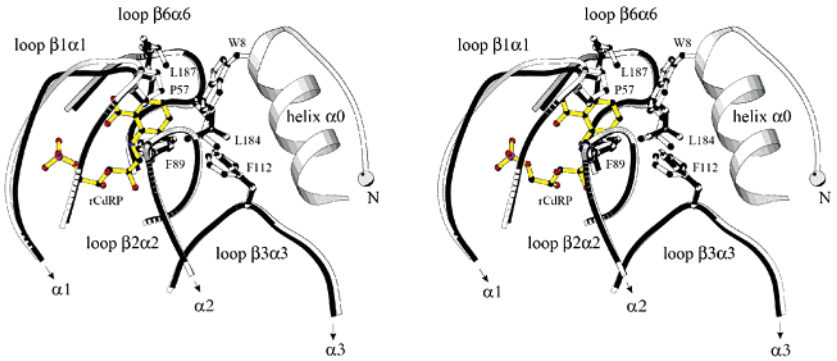


FIGURE 4: Ball-and-stick representation of the hydrophobic pocket at the active site of sIGPS. The structures of wild-type sIGPS with bound rCdRP (carbon atoms in black, oxygen atoms in red, nitrogen atom in blue, phosphorus atom in magenta, and bonds between atoms in yellow) (PDB entry code 1LBF) and sIGPSΔ(1−26) (PDB entry code 2c3z) are superimposed. The drawing was produced with MOLSCRIPT (*27*).

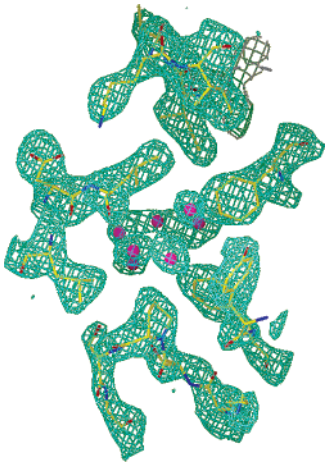


FIGURE 5: $2F_o - F_c$ electron density map contoured at 1.0 σ viewing the hydrophobic pocket at the active site of sIGPSΔ(1−26). The water-filled cavity (magenta spheres) that arises by the cleavage of the N-terminal extension is in the wild-type enzyme partially occupied by the hydrophobic side chains of residues W8 and L9 at the N-terminus of helix α0. The figure was produced with "O" (*28*).

further increased K_M^{CdRP} of sIGPSΔ(1−26) might be attributed to the deletion of L12 and V13, which are not an integral part of the hydrophobic (r)CdRP binding pocket but are located close enough to influence substrate affinity. In the absence of either substrate or product molecules, the active site of sIGPSΔ(1−26) forms an hydrophobic channel that is filled with crystallographically well-defined water molecules (Figure 5). These water molecules might compete with the substrate for binding to the active site and thus also contribute to the high K_M value for CdRP. In the wild-type enzyme this cavity is covered by helix α0, and some water molecules that are closer to the protein surface are displaced by the side chains of W8 and L9.

The interaction of the N-terminal extension with the $(\beta\alpha)_8$-barrel core of sIGPS has been analyzed on the basis of the wild-type structure. Most of these interactions are formed by hydrogen bonds, which are listed in Table 4. In addition, salt bridges are found between R3 and D165, K6 and E185, and R19 and D111 as well as R26 and D128. From these ion pairs R3/D165 and K6/E185 are observed in only one out of three different sIGPS crystal forms (*19*) and can therefore be classified as weak, transient interactions. In contrast, the salt bridges R19/D111 and R26/D128 seem to be highly specific and strongly supported by the molecular context (for a detailed description see refs *9* and *19*). Since structural differences between the $(\beta\alpha)_8$-barrel cores of wild-type and truncated sIGPS are restricted to minor conformational rearrangements (Figure 3), it is apparent that the disruption of the hydrogen bonds and salt bridges between the N-terminal extension and the core are crucial for the decreased stability of sIGPSΔ(1−26) compared to sIGPS (Table 3).

CONCLUSIONS

The comparison of thermal inactivation and steady-state enzyme kinetic data of sIGPS and tIGPS with sIGPSΔ(1−26) and tIGPSΔ(1−25) shows that the N-terminal helix α0 is important for both resistance to irreversible thermal inactivation and substrate binding. These findings can be rationalized on the basis of the structures of sIGPS and sIGPSΔ(1−26). Numerous stabilizing contacts between side chains of the N-terminal extension and the canonical $(\beta\alpha)_8$ barrel, as well as key residues involved in binding CdRP, are lost in the truncated protein. The intrinsic enzymatic function of IGPS, however, does not require the extension, which proves that all catalytic residues are provided by the barrel core of the enzyme. This notion is in keeping with the hypothesis that TIM-barrel enzymes involved in the tryptophan biosynthesis have diverged from a common ancestor during evolution (*4*).

ACKNOWLEDGMENT

The authors thank Halina Szadkowski and Jeannette Ueckert for excellent technical assistance and Rainer Deutzmann for N-terminal protein sequencing. Thanks also to Astrid Pappenberger, Miriam Deuss, and André Fischer for critical discussions and to Marion Strieder for preparing Figures 2A and 3A.

REFERENCES

1. Gerlt, J. A., and Raushel, F. M. (2003) Evolution of function in $(\beta/\alpha)_8$-barrel enzymes, *Curr. Opin. Chem. Biol. 7*, 252−264.
2. Wise, E. L., and Rayment, I. (2004) Understanding the importance of protein structure to nature's routes for divergent evolution in TIM barrel enzymes, *Acc. Chem. Res. 37*, 149−158.
3. Sterner, R., and Höcker, B. (2005) Catalytic versatility, stability, and evolution of the $(\beta\alpha)_8$-barrel enzyme fold, *Chem. Rev. 105*, 4038−4055.
4. Henn-Sax, M., Höcker, B., Wilmanns, M., and Sterner, R. (2001) Divergent evolution of $(\beta\alpha)_8$-barrel enzymes, *Biol. Chem. 382*, 1315−1320.
5. Gerlt, J. A., Babbitt, P. C., and Rayment, I. (2005) Divergent evolution in the enolase superfamily: the interplay of mechanism and specificity, *Arch. Biochem. Biophys. 433*, 59−70.
6. Copley, R. R., and Bork, P. (2000) Homology among $(\beta\alpha)_8$ barrels: implications for the evolution of metabolic pathways, *J. Mol. Biol. 303*, 627−641.
7. Wilmanns, M., Hyde, C. C., Davies, D. R., Kirschner, K., and Jansonius, J. N. (1991) Structural conservation in parallel β/α-barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis, *Biochemistry 30*, 9161−9169.
8. Stehlin, C., Dahm, A., and Kirschner, K. (1997) Deletion mutagenesis as a test of evolutionary relatedness of indoleglycerol phosphate synthase with other TIM barrel enzymes, *FEBS Lett. 403*, 268−272.
9. Hennig, M., Darimont, B., Sterner, R., Kirschner, K., and Jansonius, J. N. (1995) 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability, *Structure 3*, 1295−1306.
10. Merz, A., Knöchel, T., Jansonius, J. N., and Kirschner, K. (1999) The Hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilized by mutational disruption of two solvent-exposed salt bridges, *J. Mol. Biol. 288*, 753−763.
11. Beismann-Driemeyer, S., and Sterner, R. (2001) Imidazole glycerol phosphate synthase from *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex, *J. Biol. Chem. 276*, 20387−20396.
12. Merz, A., Knöchel, T., Jansonius, J. N., and Kirschner, K. (1999) The hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilized by mutational disruption of two solvent-exposed salt bridges, *J. Mol. Biol. 288*, 753−763.
13. Leopoldseder, S., Claren, J., Jürgens, C., and Sterner, R. (2004) Interconverting the catalytic activities of $(\beta\alpha)_8$-barrel enzymes from different metabolic pathways: sequence requirements and molecular analysis, *J. Mol. Biol. 337*, 871−879.
14. Hommel, U., Eberhard, M., and Kirschner, K. (1995) Phosphoribosyl anthranilate isomerase catalyzes a reversible Amadori reaction, *Biochemistry 34*, 5429−5439.

15. Eberhard, M., Tsai-Pflugfelder, M., Bolewska, K., Hommel, U., and Kirschner, K. (1995) Indoleglycerol phosphate synthase-phosphoribosyl anthranilate isomerase: comparison of the bifunctional enzyme from *Escherichia coli* with engineered monofunctional domains, *Biochemistry 34*, 5419−5428.
16. Sterner, R., Kleemann, G. R., Szadkowski, H., Lustig, A., Hennig, M., and Kirschner, K. (1996) Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer, *Protein Sci. 5*, 2000−2008.
17. The CCP4 suite (1994) Programs for protein crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760−763.
18. Brünger, A. T. (1992) X-PLOR Version 3.1. A system for X-ray crystallography and NMR, Yale University Press, New Haven, CT.
19. Knöchel, T. R., Hennig, M., Merz, A., Darimont, B., Kirschner, K., and Jansonius, J. N. (1996) The crystal structure of indole-3-glycerol phosphate synthase from the hyperthermophilic archaeon *Sulfolobus solfataricus* in three different crystal forms: effects of ionic strength, *J. Mol. Biol. 262*, 502−515.
20. Knöchel, T., Pappenberger, A., Jansonius, J. N., and Kirschner, K. (2002) The crystal structure of indoleglycerol-phosphate synthase from *Thermotoga maritima*. Kinetic stabilization by salt bridges, *J. Biol. Chem. 277*, 8626−8634.
21. Merz, A., Yee, M. C., Szadkowski, H., Pappenberger, G., Crameri, A., Stemmer, W. P., Yanofsky, C., and Kirschner, K. (2000) Improving the catalytic activity of a thermophilic enzyme at low temperatures, *Biochemistry 39*, 880−889.
22. Andreotti, G., Cubellis, M. V., Palo, M. D., Fessas, D., Sannia, G., and Marino, G. (1997) Stability of a thermophilic TIM-barrel enzyme: indole-3-glycerol phosphate synthase from the thermophilic archaeon *Sulfolobus solfataricus*, *Biochem. J. 323*, 259−264.
23. Forsyth, W. R., and Matthews, C. R. (2002) Folding mechanism of indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*: a test of the conservation of folding mechanisms hypothesis in $(\beta\alpha)_8$ barrels, *J. Mol. Biol. 320*, 1119−1133.
24. Hennig, M., Darimont, B. D., Jansonius, J. N., and Kirschner, K. (2002) The catalytic mechanism of indole-3-glycerol phosphate synthase: crystal structures of complexes of the enzyme from *Sulfolobus solfataricus* with substrate analogue, substrate, and product, *J. Mol. Biol. 319*, 757−766.
25. Darimont, B., Stehlin, C., Szadkowski, H., and Kirschner, K. (1998) Mutational analysis of the active site of indoleglycerol phosphate synthase from *Escherichia coli*, *Protein Sci. 7*, 1221−1232.
26. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
27. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
28. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.

BI051640N