Biosynthesis of Tetrahydrofolate in Plants: Crystal Structure of 7,8-Dihydroneopterin Aldolase from Arabidopsis thaliana Reveals a Novel Adolase Class

Stefanie Bauer1*, Ann-Kathrin Schott2, Victoria Illarionova3 Adelbert Bacher2, Robert Huber1 and Markus Fischer2*

1Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany
2Lehrstuhl für Organische Chemie und Biochemie Technische Universität München, Lichtenbergstr. 4 D-85747 Garching, Germany
3Institute for Biophysics Krasnoyarsk 660036, Russia

Dihydroneopterin aldolase (DHNA) catalyses a retroaldol reaction yielding 6-hydroxymethyl-7,8-dihydropterin, a biosynthetic precursor of the vitamin, tetrahydrofolate. The enzyme is a potential target for antimicrobial and anti-parasite chemotherapy. A gene specifying a dihydroneopterin aldolase from Arabidopsis thaliana was expressed in a recombinant Escherichia coli strain. The recombinant protein was purified to apparent homogeneity and crystallised using polyethylenglycol as the precipitating agent. The crystal structure was solved by X-ray diffraction analysis at 2.2 Å resolution. The enzyme forms a D4-symmetric homo-octamer. Each polypeptide chain is folded into a single domain comprising an antiparallel four-stranded β-sheet and two long α-helices. Four monomers are arranged in a tetrameric ring, and two of these rings form a hollow cylinder. Well defined purine derivatives are found at all eight topologically equivalent active sites. The subunit fold of the enzyme is related to substructures of dihydroneopterin triphosphate epimerase, GTP cyclohydrolase I, and pyruvoyltetrahydropterin synthase, which are all involved in the biosynthesis of pteridine type cofactors, and to urate oxidase, although some members of that superfamily have no detectable sequence similarity. Due to structural and mechanistical differences of DHNA in comparison with class I and class II aldolases, a new aldolase class is proposed.

Introduction

Tetrahydrofolate and its derivatives are essential cofactors of one-carbon metabolism that are required for the biosynthesis of purines, thymidylate, serine and methionine in a wide variety of organisms; they are also required for the formulation of methionyl-tRNA.1,2 Whereas plants and many microorganisms obtain folate coenzymes by de novo synthesis, vertebrates depend on nutritional sources.3 Insufficient supply of the vitamin is conducive to anaemia in adults and to neural tube malformation in human embryos.4 The biosynthesis of tetrahydrofolate has been studied in some detail (Figure 1).5 The first committed step catalysed by GTP cyclohydrolase I is a mechanistically complex ring expansion reaction affording dihydroneopterin triphosphate (I),5–10 A pyrophosphatase and a phosphatase have been proposed to convert dihydroneopterin triphosphate into 7,8-dihydroneopterin (3) in two consecutive steps,11 but the details are still incompletely understood.12 The conversion of 7,8-dihydroneopterin into 6-hydroxymethyl-7,8-dihydropterin (4) and glycolaldehyde is catalysed by 7,8-dihydroneopterin aldolase (DHNA),13–18 The enzyme product is converted into dihydrofolate by the consecutive action of dihydropteroate synthase and dihydrofolate reductase,5,19–22 which
are both important chemotherapeutic targets for antibacterial and antiparasitic agents. More specifically, dihydropteroate synthase is inhibited by sulphonamides, the first synthetic antimicrobial and antiparasitic drugs with a broad sphere of action. Dihydrofolate reductase is inhibited by trimethoprim, which acts against a variety of bacterial pathogens. Inhibitors of the early steps of folate biosynthesis are not available but might be a useful addition to the therapeutic arsenal.

DHNA has been purified from cell extracts of various bacterial species. Hennig et al. described the crystal structure of the enzyme from Staphylococcus aureus and suggested a reaction mechanism. In analogy with the reaction catalysed by rabbit muscle aldolase, the DHNA mechanism has been interpreted as a retroaldol reaction.

In general, aldolases catalyse carbon–carbon bond formation and cleavage. They are attractive as synthetic catalysts due to their stereospecificity. They are divided into two major classes, based on the mode of stabilisation of the reaction intermediates. Class I aldolases are present in all groups of living organisms, from prokaryotes to eukaryotes, and are characterised by the formation of a Schiff base with the substrate. Class II aldolases occur only in prokaryotes and lower eukaryotes such as yeasts, algae and fungi and are dependent on divalent metal ions.

The Arabidopsis thaliana genome comprises three genes with similarity to bacterial DHNA genes. Here, we describe the three-dimensional structure of a 7,8-dihydroneopterin aldolase specified by the folb1 gene of A. thaliana, the first structure of a folate biosynthetic enzyme from the plant kingdom. The enzyme is a potential herbicide target.

Due to structural and mechanistical differences of DHNA in comparison with class I and class II aldolases, a new aldolase class is proposed.

Results and Discussion

Preparation of pseudomature dihydroneopterin aldolase

Sequence comparison suggested that the putative folb1 and folb2 genes located on chromosome 3 (gene At3g11750 and gene At5g62980) and the putative folb3 gene located on chromosome 5 of A. thaliana (gene At3g21730) specify dihydroneopterin aldolase isoenzymes. The putative catalytic domains of these enzymes show close sequence similarity; 64 amino acid residues (43%) are identical, and 23 amino acid residues (15%) are similar.

The gene segment specifying the putative catalytic domain of the folb1 gene (without a leader sequence, AtfolB1-D) was amplified by PCR and cloned into a plasmid vector. The sequence of the amplified cDNA segment was identical with that predicted earlier on the basis of genomic sequencing. The sequence has been deposited in GenBank (accession number AY507667).

An E. coli strain carrying the folb1 gene under the control of a T5 promoter and lac operator produced abundant amounts of a 14 kDa polypeptide (about
30% of cell protein) as judged by SDS gel electrophoresis. The recombinant protein was purified to apparent homogeneity by chromatographic procedures and appeared homogeneous as judged by SDS electrophoresis. Mass spectrometry afforded a relative mass of 14,128 Da, in good agreement with a predicted value of 14,129 Da. Partial Edman degradation afforded the N-terminal sequence MDKLILKGLKF, in agreement with the DNA sequence. Notably, the initial methionine residue had not been removed by post-translational processing in the bacterial host strain.

The recombinant enzyme catalysed the formation of 6-hydroxymethyl-7,8-dihydropterin from dihydroneopterin at a rate of 0.32 μmol mg⁻¹ min⁻¹ (pH 8, 30 °C) and from dihydromonapterin at a rate of 0.31 μmol mg⁻¹ min⁻¹ (pH 8, 30 °C). The enzyme also catalysed the reversible transformation of dihydroneopterin (1) into dihydromonapterin (2) (Figure 2). The catalytic rate was not affected by the addition of EDTA.

Structure determination

The recombinant protein was crystallised from a solution containing 0.1 M Tris hydrochloride (pH 7.2), 21% polyethylene glycol (PEG 2000 MME), and 115 mM C-HEGA-II (cyclohexylbutanoyl-N-hydroxyethylglucamide). Diffraction data were collected to a resolution of 2.2 Å. The crystals belonged to space group P1 with the unit cell constants a = 63.5 Å, b = 84.2 Å, c = 89.1 Å, α = 90.14°, β = 89.9°, γ = 76.2° and had a solvent content of 42%. Table 1 gives a summary of the data collection and phasing statistics. The crystal structure was solved by molecular replacement using the model of Staphylococcus aureus dihydroneopterin aldolase (PDB-ID: 1DHN) in which side-chains had been clamped. The rotational search showed peaks of height 9.0 and 8.7 for the two octamers in the asymmetric unit compared to 3.9 for the next highest peaks. The cross-translation search had a peak of 12.6 compared with 4.8 for the next.

The monomer of the recombinant A. thaliana DHNA consists of 126 residues that were well defined in all monomers of the asymmetric unit with the exception of the last five residues. The last three residues were disordered in all subunits;

Table 1. Statistics for data collection and phasing

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P1</td>
</tr>
<tr>
<td>Unit cell constants</td>
<td>a = 63.5 Å, b = 84.2 Å, c = 89.1 Å, α = 90.14°, β = 89.9°, γ = 76.2°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.0–2.2</td>
</tr>
<tr>
<td>Last shell (Å)</td>
<td>2.20–2.28</td>
</tr>
<tr>
<td>Measured reflections</td>
<td>240,705</td>
</tr>
<tr>
<td>Independent reflections</td>
<td>86,882</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>2.7</td>
</tr>
<tr>
<td>Rmerge overall (%)</td>
<td>9.0</td>
</tr>
<tr>
<td>Outermost shell (%)</td>
<td>30.0</td>
</tr>
<tr>
<td>Completeness overall (%)</td>
<td>95.5</td>
</tr>
<tr>
<td>Outermost shell (%)</td>
<td>92.6</td>
</tr>
</tbody>
</table>

*Rmerge = Σ||I-I(h,k,l) - <I(h,k,l)>||/ΣI(h,k,l).*
Asn121 and Thr122 were disordered to various extents in different subunits. Most of the side-chains were clearly defined by their electron density map, except for some located at the surface of the protein. The final model of the DHNA octamer comprises 975 amino acid residues with an $R_{cryst}$ of 19.5% and an $R_{free}$ of 25.8% (Table 2).

A well defined guanine molecule is found in all subunits bound at the putative active site that had not been added but carried through all preparation and purification steps.

### Structural overview

The polypeptide chain of 7,8-dihydroneopterin aldolase of *A. thaliana* is folded into a single domain comprising a four-stranded antiparallel $\beta$-sheet ($\beta_1$, $\beta_2$, $\beta_3$ and $\beta_4$) and two long $\alpha$-helices, $\alpha_4$ and $\alpha_5$ (Figure 3). The loops connecting strands $\beta_1$ and $\beta_2$ and strand $\beta_3$ and helix $\alpha_4$ contain two further short helical segments called $\alpha_1$ (residues 19–24) and $\alpha_2$ (residues 38–44). A short $\beta_\parallel$-helical part $\alpha_3$ (residues 47–49) is inserted between $\alpha_2$ and the long helix $\alpha_4$. The four-stranded $\beta$-sheet is composed of residues 3–13, 26–35, 92–99 and 112–119. Two long $\alpha$-helices ($\alpha_4$, residues 53–65, and $\alpha_5$, residues 72–86) both lie on the same side of the $\beta$-sheet. The monomer has an ellipsoidal shape with dimensions 50 Å × 25 Å × 25 Å, and its $\beta$baab fold shows topological similarity to substructures of other tetrahydrobiopterin biosynthetic enzymes (GTP cyclohydrolase I, pyruvoyl tetrahydropterin synthase, 7,8-dihydroneopterin triphosphate epimerase; for details see below). The hydrophobic core of the molecule is formed by non-polar residues of the two long $\alpha$-helices and of the $\beta$-sheet. The structure of *A. thaliana* DHNA shows high similarity to that of *S. aureus* DHNA (Figure 4). Both structures can be superimposed with a root mean squared deviation (rmsd) fit between 117 C$^\alpha$ positions of 0.72 Å.

DHNA is a $D_4$-symmetric homooctamer with the shape of a hollow cylinder assembled from two ring-shaped tetramers. The 16 antiparallel $\beta$-strands of four symmetry-related polypeptide chains in a tetramer are arranged to form a barrel with a continuous $\beta$-sheet hydrogen bonding network, surrounded by a ring of helices (Figure 5). Six backbone hydrogen bonds from the N-terminal strand $\beta_1$ (residues 3–11) to the C-terminal strand $\beta_4$ (residues 107–113) as well as a salt bridge between residues Asp2 and Arg118 stabilise the interaction of adjacent subunits in the tetramer. Further stabilisation is achieved by a hydrophobic cluster formed by the side-chains of residues Leu24, Tyr12, Leu104 and Ala103, whereas Leu104, Ala103 and Leu24 interact with Tyr12 and Leu24 of the adjacent subunit.

The tetramer is torus-shaped with a 60 Å

### Table 2. Refinement statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range</td>
<td>20–2.2 Å</td>
</tr>
<tr>
<td>Reflections (working set)</td>
<td>82,949 reflections</td>
</tr>
<tr>
<td>Reflections (test set)</td>
<td>3857 reflections</td>
</tr>
<tr>
<td>$R_{cryst}$ (%)</td>
<td>19.5%</td>
</tr>
<tr>
<td>$R_{free}$ (%)</td>
<td>25.8%</td>
</tr>
<tr>
<td>Non-hydrogen protein atoms</td>
<td>15,312</td>
</tr>
<tr>
<td>Solvent molecules</td>
<td>925</td>
</tr>
<tr>
<td>Ligand atoms</td>
<td>224</td>
</tr>
<tr>
<td>Average $B$-value ($\AA^2$) overall</td>
<td>22.5 Å$^2$</td>
</tr>
<tr>
<td>Average $B$-value ($\AA^2$) protein</td>
<td>22.3 Å$^2$</td>
</tr>
<tr>
<td>Average $B$-value ($\AA^2$) ligand</td>
<td>19.6 Å$^2$</td>
</tr>
<tr>
<td>Average $B$-value ($\AA^2$) solvent</td>
<td>24.8 Å$^2$</td>
</tr>
<tr>
<td>rmsd from standard groups</td>
<td>0.33 Å</td>
</tr>
<tr>
<td>Bond length ($\AA$)</td>
<td>0.007 Å</td>
</tr>
<tr>
<td>Bond angle (°)</td>
<td>1.32 °</td>
</tr>
</tbody>
</table>

* $R_{cryst} = \sum_{\text{hkl}, W_h} |F_{\text{obs}}|-|F_{\text{calc}}|/\sum_{\text{hkl}, W_h} |F_{\text{obs}}|$

$R_{free} = \sum_{\text{hkl}, T} |F_{\text{obs}}|-|F_{\text{calc}}|/\sum_{\text{hkl}, T} |F_{\text{obs}}|$

---

**Figure 3.** Stereo drawing of the *A. thaliana* 7,8-dihydroneopterin aldolase structure with modelled product.

**Figure 4.** Superposition of the *A. thaliana* 7,8-dihydroneopterin aldolase structure (red) with that of *S. aureus* 7,8-dihydroneopterin aldolase (blue).
diameter and a height of 25 Å and it encloses a hydrophilic pore that is 30 Å in diameter. An accessible surface area of approximately 2136 Å² corresponding to 29.2% of the surface of each monomer is buried upon tetramerisation.

The octamer is formed by a C₂ symmetric arrangement of two tetramers in a head-to-head fashion to form a hollow cylinder (Figure 6). Stabilisation of this assembly is achieved by hydrophobic side-chain interactions. Here, only 690 Å² (corresponding to 9.5% of the monomer surface) are buried upon oligomerisation sharing a dominance of the intra-tetramer over the inter-tetramer interactions in the octamer.

Structural comparison

The DHNA fold is topologically similar to 7,8-dihydroneopterin triphosphate epimerase, 6-pyruvoyl tetrahydropterin synthase (PTPS), the C-terminal domain of GTP cyclohydrolase I (residues 94–238) and both domains of urate oxidase (UO1 and UO2). These enzymes have been jointly addressed as a superfamily of tetrahydrobiopterin biosynthesis enzymes (although urate oxidase is not involved in pteridine biosynthesis). Dihydroneopterin triphosphate epimerase catalyses the reversible epimerization of carbon 2’ affording dihydromonapterin triphosphate from dihydroneopterin triphosphate, GTP cyclohydrolase I catalyses the conversion of guanosine triphosphate to dihydroneopterin triphosphate, PTPS converts dihydroneopterin triphosphate into 6-pyruvoyl tetrahydrobiopterin, and urate oxidase catalyses the conversion of uric acid into allantoin in the purine degradation pathway. They all bind purine or pterin derivatives in topologically similar sites between adjacent subunits.

Only two members of the superfamily, namely DHNA and 7,8-dihydroneopterin triphosphate epimerase, share more than 20% identical amino acid residues (Figure 7). Six out of seven catalytic site residues are highly conserved between DHNA and epimerase; in fact, the only significant difference in the active sites of these enzymes is

![Figure 5](image1.png) Stereo drawing of a ribbon presentation of the DHNA tetramer with the modelled product bound to the four active sites.

![Figure 6](image2.png) Stereo drawing of the DHNA octamer without bound product.
the putative triphosphate binding site of the epimerase. The high similarity between the binding mode of DHNA from *A. thaliana* and *S. aureus* and epimerase from *E. coli* is shown in Figure 8.

No significant sequence similarity exists between the other members of the superfamily. In all members of the superfamily, the main secondary structure elements are conserved, but several insertions and deletions involving not only loops but also short secondary structure elements in the regions outside the four strands and the two helices are found. The only amino acid residue that is common to all members except domain 2 of urate oxidase of the superfamily is a glutamate residue (Glu73 in DHNA), which serves as anchor for the pyrimidine ring of each respective substrate. In domain 2 of urate oxidase it is a glutamine.

**Active site architecture**

The eight topologically equivalent substrate binding sites are all located at the interfaces between adjacent subunits in the tetramer assemblies. Well defined electron density resembling guanine is found in all eight active sites. The ligand is obviously carried over from the preparation, but has not been characterised further (Figure 9). The localisation of the pyrimidine moiety corresponds to the general binding mode observed in other folate biosynthesis enzymes. As the reaction substrate and product also comprise a pyrimidine moiety, we suggest the same binding mode for these ligands in DHNA. 7,8-Dihydropteridine has been modelled on the basis of the bound guanine. One subunit contributes residues from α1 (Glu21) and the beginning of α4 (Leu72 and Glu73), from the end of β3

![Figure 8](image-url)
(Lys99) and the loop connecting α3 and α4 (Asn70), the other subunit contributes one residue at the beginning of α3 (Tyr53) and one residue at the loop connecting α2 and α3 (Ser52) whereas Lys99, Glu21 and Asn70 are involved in the hypothetical side-chain binding or in fixing of bound water molecules. The hydrophobic pocket to which the pyrimidine ring 7,8-dihydronicotinamide is bound consists of Val36, Leu58, Tyr53 and Val54 of one subunit, and Leu71 and Leu72 of the adjacent subunit. In detail, several hydrogen bonds to the modelled 7,8-dihydronicotinamide are formed that involve main-chain and side-chain interactions (Figure 10). The N1H group and the amino group of the pterin ring form hydrogen bonds with the carboxylate group of Glu73, respectively. The carbonyl oxygen O2 and the amide group of Leu72 form a hydrogen bond as well as the N10H group with the hydroxyl group of Ser52. A water molecule is part of a hydrogen bonding network involving carbonyl oxygen O2 of the pterin ring, the carboxyl group of Asn70 and the α-amino group of Lys99. The latter is in close contact to the carboxylate group of Glu21 and the hydroxyl group of the product side-chain.

**Reaction mechanism of DHNA in comparison with class I aldolase mechanism**

Aldolases, in general, can be divided into two major classes, based on the mode of stabilisation of the reaction intermediates. Class I aldolases are characterised by the formation of a Schiff base, whereas class II aldolases require a metal cofactor and are also designated metalloaldolases. DHNA from *A. thaliana* is not dependent on divalent metal ions and is hence not a member of the class II aldolase family.

Class I aldolases form a family of enzymes that are not related in amino acid sequence and that catalyse a large variety of aldol cleavage/condensation reactions. A common feature of the members of this enzyme class is the formation of a covalent intermediate (Schiff base) between a lysine residue of the enzyme and a carbonyl carbon atom of the substrate. A number of enzymes catalysing aldol cleavage and/or condensation belong to the class I aldolase family. Several enzymes have been studied by X-ray crystallography, and the three-dimensional structures of fructose-1,6-bisphosphate aldolase (FBPA), transaldolase, KDPG aldolase, type I 3-dehydroquinase, p-2-desoxyribose-5-phosphate aldolase (DRPA), N-acetyl neuraminate lyase and dihydrodipicolinate synthase (DHDPS) are known. The subunits of these enzymes all show classical TIM β/α-folds with a C-terminal region involved in substrate specificity. In FBPA, Asp33 was originally hypothesised to protonate the incipient carbinolamine hydroxyl (C2) with Tyr363 abstracting the proton α to the carbonyl (C3 of DHAP). However, in a recent study of human muscle aldolase, Glu187 is now proposed as the catalytic acid and Asp33 as the catalytic base by analogy with the corresponding Glu96 and Asp17 of transaldolase B and Asp187 and Asp188 in DHDPS. In *E. coli* KDPG aldolase Glu45 is
proposed to act as a general base. In both N-acetyleneuraminic acid lyase and type I 3-dehydroquinase a histidine residue (His79 and His143) might assume the latter role.

Superposition of type I aldolase enzymes (data not shown) shows that, with the exception of transaldolase, all reactive lysine residues are not only on the same strand, but also in the same location with respect to the centre of the barrel. A similar constellation of key catalytic residues appears to be present in type I aldolase binding sites, but their precise configuration differs from enzyme to enzyme, suggesting that the aldolase mechanism does not require a unique geometric arrangement of catalytic residues. The reaction mechanism of type I aldolases is shown in Figure 11.

As no crystal structure of a substrate or product complex of A. thaliana DHNA exists, the following reaction mechanism is based on the comparison of active site residues and binding modes of other folate biosynthesis enzymes, especially of dihydroneopterin triphosphate epimerase and of S. aureus DHNA. A. thaliana DHNA can use L-threo-dihydronopterin and D-erythro-dihydronopterin as substrates for the formation of 6-hydroxy-

The substrate (1) of DHNA can be viewed as an intramolecular Schiff base, where the Schiff base motif forms part of the dihydropyrazine ring of the dihydropterin system. In fact, this Schiff base motif in dihydroneopterin arises biosynthetically by condensation of a carbonyl motif with an amino group in the multistep reaction catalysed by GTP cyclohydrolase I. Hence, the formation of an intermolecular Schiff base with a lysine residue of the enzyme is not required (although it should be mentioned that the pyrazine of the substrate could be opened hydrolytically, and the resulting carbonyl group could then form a Schiff base with the enzyme; this reaction sequence would constitute a reversible transamination). The origin of the ligand present in the enzyme crystals is at present unknown.

From structural aspects, all class I aldolases and class II whose structures have been determined show classical TIM \( \beta/\alpha \)-folds with the C-terminal region of the barrel involved in substrate specificity while DHNA shows a \( \beta\alpha\beta\beta \) fold with elliptical shape and the active site located at the interface of two adjacent subunits. The DHNA \( \beta\alpha\beta\beta \) fold also occurs in 7,8-dihydronopterin triphosphate epimerase, GTP cyclohydrolase I, 6-pyruvyl tetrahydropterin synthase and urate oxidase, but with the exception of 7,8-dihydronopterin triphosphate epimerase, these enzymes do not catalyse aldolase type reactions.

As a result of the different reaction mechanism and a completely different overall fold of DHNA it is proposed that DHNAs present a new class of aldolase enzymes, both mechanistically as well as structurally.

**Material and Methods**

**Materials**

D-Dihydronopterin and dihydromonapterin were purchased from Dr B. Schircks (Jona, Switzerland).
Oligonucleotides were custom synthesised by Interactiva, Ulm, Germany. Restriction enzymes and Vent DNA polymerase were purchased from New England Biolabs, Schwalbach, Germany. T4 DNA ligase and reverse transcriptase (SuperScriptII) were from Gibco BRL, Eggenstein, Germany. Taq Polymerase was from Finnzyme, Espoo, Finland. DNA fragments were purified with the QIAquick PCR Purification Kit from Qiagen, Hilden, Germany.

Microorganisms and plasmids

Microorganisms and plasmids used in this study are summarised in Table 3.

Preparation of cDNA

The isolation of RNA from *A. thaliana* and the preparation of cDNA were performed as described.\(^{33}\)

Construction of an expression plasmid

The hypothetical open reading frame (accession number AAF23191, without leader sequence) was amplified by PCR using *A. thaliana* cDNA as template and the oligonucleotides ATfolB1-ΔLS-1 and ATfolB1-2 (Table 4) as primers. The amplificate was digested with EcoRI and BamHI and was ligated into the expression vector pNCO113, which had been digested with the same

<table>
<thead>
<tr>
<th>Table 3. Bacterial strains and plasmids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain or plasmid</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>E. coli strain</td>
</tr>
<tr>
<td>Plasmids</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
enzymes. The resulting plasmid pNCO-AtfolB1-ΔE was transformed into *E. coli* XL-1 Blue cells by published procedures.52 Transformants were selected on LB agar plates supplemented with ampicillin (170 mg/l).

The pNCO113 type plasmid reisolated from XL-1 Blue cells was transformed into *E. coli* M15 [pREP4] cells53 carrying the pREPl repressor plasmid, which directs the synthesis of lac repressor protein. Kanamycin (15 mg/l) and ampicillin (170 mg/l) were added to secure the maintenance of both plasmids in the host strain.

### Protein purification

The recombinant *E. coli* strain M15[pREP4]-pNCO-AtfolB1-ΔE was grown in LB medium containing ampicillin (170 mg/l) and kanamycin (20 mg/l) at 37°C with shaking overnight. Erlenmeyer flasks containing 500 ml of medium were inoculated at a ratio of 1:50 and were incubated at 37°C with shaking. At an absorbance of 0.6 (A600 nm), isopropylthiogalactoside was added to a final concentration of 2 mM, and incubation was continued for five hours. The cells were harvested by centrifugation, washed with 0.9% NaCl and stored at −20°C.

Frozen cell mass (6 g) was thawed in 35 ml of 50 mM Tris hydrochloride (pH 8.3). The suspension was subjected to ultrasonic treatment and was then centrifuged (20 minutes, 15,000 rpm, 4°C). The supernatant was placed on top of a Q-Sepharose column (90 ml), which had been equilibrated with 50 mM Tris hydrochloride (pH 8.3), containing 70 mM potassium chloride at a flow rate of 3 ml min⁻¹. The column was developed with a linear gradient of 0–1.0 M potassium chloride in 50 mM Tris hydrochloride (pH 8.3). The column was developed with a linear gradient of 0–1.0 M potassium chloride (total volume, 500 ml). Fractions were combined, and ammonium sulphate was added to 60% saturation. The precipitate was harvested by centrifugation, washed with 0.9% NaCl and stored at −20°C.

### Assay of DHNA

The enzyme assay was performed as described.16

### Electrospray mass spectrometry

Experiments were performed with a triple quadrupol ion spray mass spectrometer API365 (SciEx, Thornhill, Ontario, Canada).

### Crystallisation

Crystals of *A. thaliana* DHNA were obtained using the sitting-drop vapour-diffusion method by mixing 2 μl of a solution containing 11.5 mg of protein per ml with an equal volume of reservoir solution containing 0.1 M Tris hydrochloride (pH 7.2), 21% polyethylene glycol (PEG 2000 MME), and 115 mM C-HEGA-11 (cyclohexylbutanoyl-N-hydroxyethylglucamide). The drops were allowed to equilibrate over a reservoir of 0.3 ml of the precipitating buffer. Crystals grew to a maximum dimension of 80 μm × 80 μm × 80 μm in about 24 hours at 20°C.

### Data collection and structure solution

X-ray data were collected on a MARResearch345 image plate detector mounted on a Rigaku RU-200 rotating anode operated at 50 mA and 100 kV with λ (Cu Kα) = 1.542 Å. The data set was collected under cryogenic conditions at 100 K using an Oxford Cryostream. For cryo measurements, crystals were transferred into Perfluoropolyether. Diffraction intensities were integrated with DENZO and were scaled and merged using the HKL suite.63 The electron density was improved by solvent flattening using the program DM. The structure was solved by molecular replacement using the program AMORE57 with the *S. aureus* 7,8-dihydroneopterin aldolase octamer (PDB-ID: 1DHN) as search model.

### Model building and refinement

The 2.2 Å electron density map was of sufficient quality to permit unambiguous chain tracing for about 90% of the model in the first round of model building using the program MAIN.66 Refinement steps carried out consisted of conjugate gradient minimisation, simulated annealing, and B-factor refinement with the program CNS using the mlf target.69 For cross-validation,
a random test set of 5% of the total number of reflections was excluded from the refinement and used for the calculation of the free R-factor. The refinement was carried out on the resulting model until convergence was reached at an R-factor of 19.5% and an Rfree of 25.8%.

The Ramachandran plot calculated with the program PROCHECK showed no residues with angular values in forbidden areas; 90.8% of the non-glycine residues are in the most favoured regions, and 9.2% are in additionally allowed regions.

Analysis and graphical representation

Stereocchemical parameters were assessed throughout refinement with PROCHECK. Secondary structure elements were assigned with STRIDE. Structural Figures were prepared with MOLSCRIPT, BOBSCRIPT and Swiss-PDB Viewer.

Atomic coordinates

The coordinates have been deposited with the Protein Data Bank (accession code 1SQL) and will be released upon publication.

Acknowledgements

The authors thank Andrew D. Hanson for helpful discussions and Stefan Gerhardt for graphical advice. This work was supported by the Fonds der Chemischen Industrie and the Hans-Fischer-Gesellschaft e.V.

References


Edited by D. Rees

(Received 7 January 2004; received in revised form 7 April 2004; accepted 10 April 2004)