doi:10.1016/j.jmb.2004.04.034 J. Mol. Biol. (2004) 339, 967–979

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

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

Stefanie Bauer[1]*, Ann-Kathrin Schott[2], Victoria Illarionova[3] Adelbert Bacher[2], Robert Huber[1] and Markus Fischer[2]*

[1]*Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany*

[2]*Lehrstuhl für Organische Chemie und Biochemie Technische Universität München, Lichtenbergstr. 4 D-85747 Garching, Germany*

[3]*Institute for Biophysics Krasnoyarsk 660036, Russia*

**Corresponding authors*

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 D_4-symmetric homooctamer. 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.

Keywords: tetrahydrofolate biosynthesis; aldolase classes; retroaldol reaction; purin binding; Schiff base

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 formylation 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 (**1**).[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

Abbreviation used: DHNA, dihydroneopterin aldolase.

E-mail addresses of the corresponding authors: stbauer@biochem.mpg.de; markus.fischer@ch.tum.de

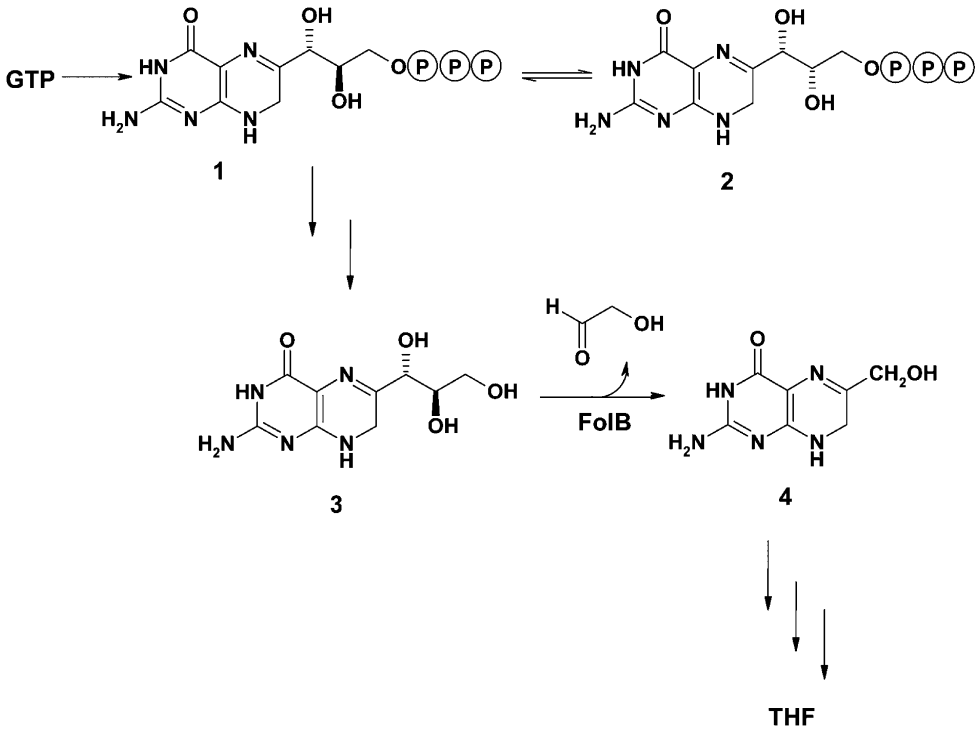

Figure 1. Pteridine biosynthesis pathway: **1**, 7,8-dihydro-D-neopterin 3′-triphosphate; **2**, 7,8-dihydro-L-monapterin 3′-triphosphate; **3**, 7,8-dihydro-D-neopterin; **4**, 6-hydroxymethyl-7,8-dihydropterin.

are both important chemotherapeutic targets for antibacterial and antiparasitic agents.[23,24] 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.[14,20,25–28] Hennig *et al*. described the crystal structure of the enzyme from *Staphylococcus aureus* and suggested a reaction mechanism.[29] In analogy with the reaction catalysed by rabbit muscle aldolase, the DHNA mechanism has been interpreted as a retroaldol reaction.[14,16,18,29]

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.[30] Class II aldolases occur only in prokaryotes and lower eukaryotes such as yeasts, algae and fungi and are dependent on divalent metal ions.[31–34]

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-ΔE) 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

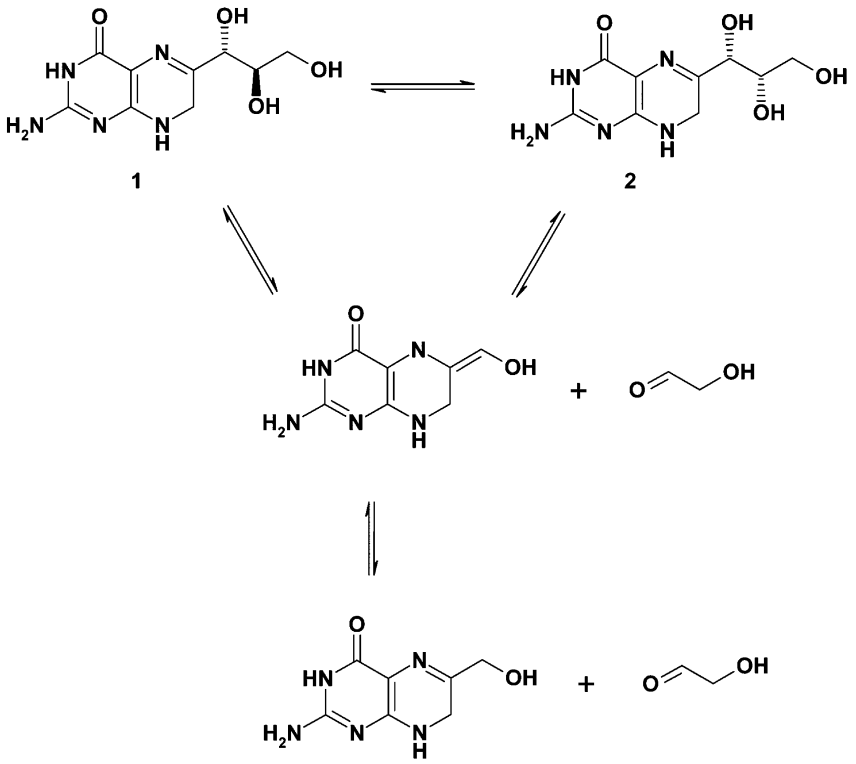

Figure 2. Reaction catalysed by 7,8-dihydroneopterin aldolase.

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% polyethyleneglycol (PEG 2000 MME), and 115 mM C-HEGA-11 (cyclohexylbutanoyl-*N*-hydroxyethylglucamide). Diffraction data were collected to a resolution of 2.2 Å. The crystals belonged to space group *P*1 with the unit cell constants $a = 63.5$ Å, $b = 84.2$ Å, $c = 89.1$ Å, $\alpha = 90.14°$, $\beta = 89.9°$, $\gamma = 76.2°$ 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

Space group	*P*1
Unit cell	
a	63.55
b	84.22
c	89.05
α	90.14
β	89.99
γ	76.17
Resolution range (Å)	20.0–2.2
Last shell (Å)	2.20–2.28
Measured reflections	240,705
Independent reflections	86,882
Multiplicity	2.7
R_{merge} overall (%)[a]	9.0
Outermost shell (%)	30.0
Completeness overall (%)	95.5
Outermost shell (%)	92.6

[a] $R_{merge} = \Sigma_{hkl}[(\Sigma_i|I_i - \langle I\rangle|)/\Sigma_i I_i]$.

Table 2. Refinement statistics

Resolution range	20–2.2
Reflections (working set)	82,949
Reflections (test set)	3837
R_{cryst}[a] (%)	19.5
R_{free}[a] (%)	25.8
Non-hydrogen protein atoms	15,312
Solvent molecules	929
Ligand atoms	224
Average *B*-value (Å^2) overall	22.5
Average *B*-value (Å^2) protein	22.3
Average *B*-value (Å^2) ligand	19.6
Average *B*-value (Å^2) solvent	24.8
rmsd from standard groups	
Bond length (Å)	0.007
Bond angle (°)	1.32
rmsd from local symmetry-related subunits (Å)	0.33

[a] $R_{cryst} = \sum_{hkl \subset W} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl \subset W} |F_{obs}|$, $R_{free} = \sum_{hkl \subset T} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl \subset T} |F_{obs}|$

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*-factor 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 β-sheet (β1, β2, β3 and β4) and two long α-helices, α4 and α5 (Figure 3). The loops connecting strands β_1 and β_2 and strand β_2 and helix α4 contain two further short helical segments called α1 (residues 19–24) and α2 (residues 38–44). A short 3_{10}-helical part α3 (residues 47–49) is inserted between α2 and the long helix α4. The four-stranded β-sheet is composed of residues 3–13, 26–35, 92–99 and 112–119. Two long α-helices (α4, residues 53–65, and α5, residues 72–86) both lie on the same side of the β-sheet. The monomer has an ellipsoidal shape with dimensions 50 Å × 25 Å × 25 Å, and its ββααββ 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 α-helices and of the β-sheet. The structure of *A. thaliana* DHNA shows high similarity to that of *S. aureus* DHNA[29] (Figure 4). Both structures can be superimposed with a root mean squared deviation (rmsd) fit between 117 C^α 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 β-strands of four symmetry-related polypeptide chains in a tetramer are arranged to form a barrel with a continuous β-sheet hydrogen bonding network, surrounded by a ring of helices (Figure 5). Six backbone hydrogen bonds from the N-terminal strand β1 (residues 3–11) to the C-terminal strand β4 (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

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

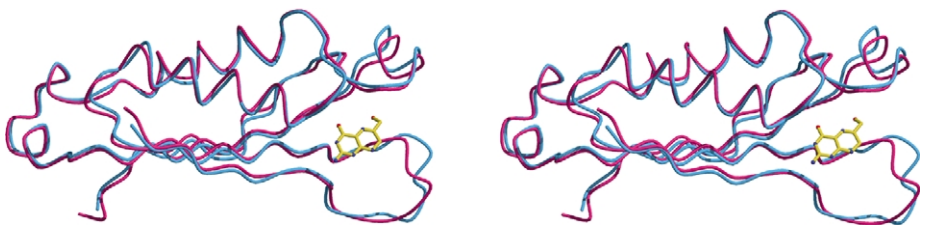

Figure 4. Superposition of the *A. thaliana* 7,8-dihydroneopterin aldolase structure (red) with that of *S. aureus* 7,8-dihydroneopterin aldolase (blue).

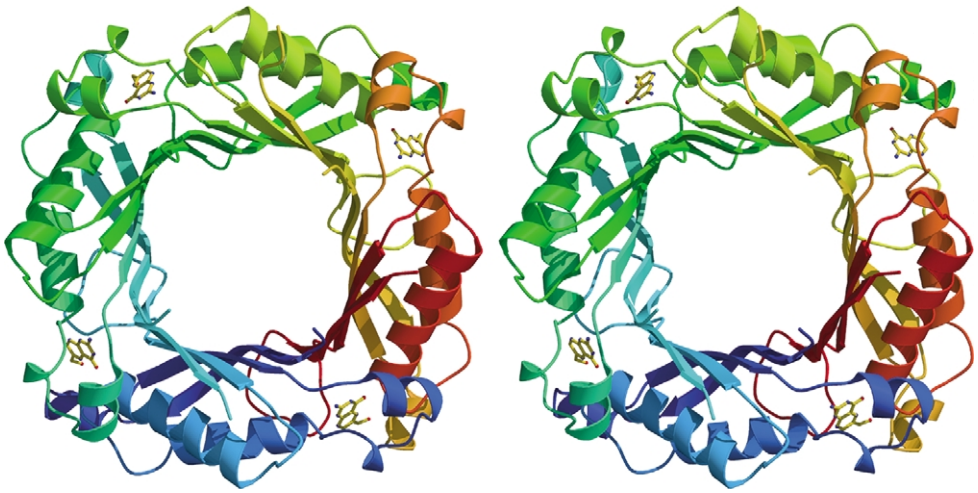

Figure 5. Stereo drawing of a ribbon presentation of the DHNA tetramer with the modelled product bound to the four active sites.

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 Å^2 corresponding to 29.2% of the surface of each monomer is buried upon tetramerisation.

The octamer is formed by a C_2 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 Å^2 (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[35], 6-pyruvoyl tetrahydropterin synthase (PTPS)[36,37], the C-terminal domain of GTP cyclohydrolase I (residues 94–238)[38] and both domains of urate oxidase (UO1 and UO2)[39]. 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-dihydropterin 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

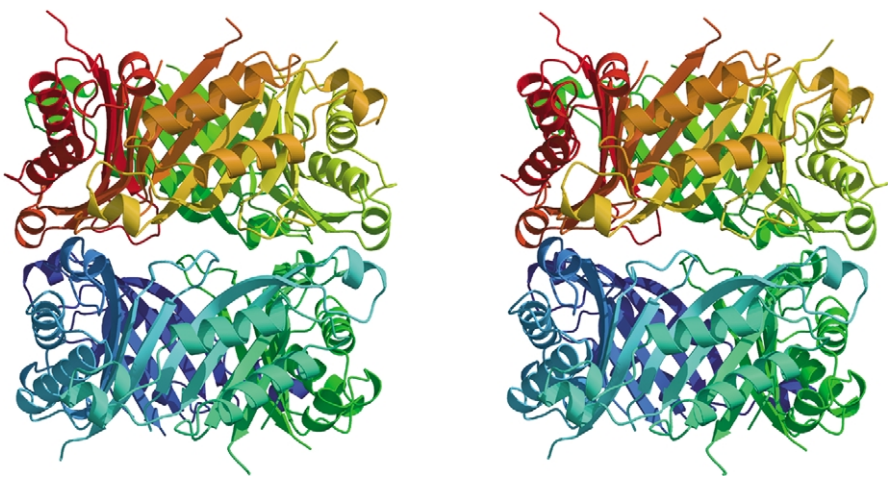

Figure 6. Stereo drawing of the DHNA octamer without bound product.

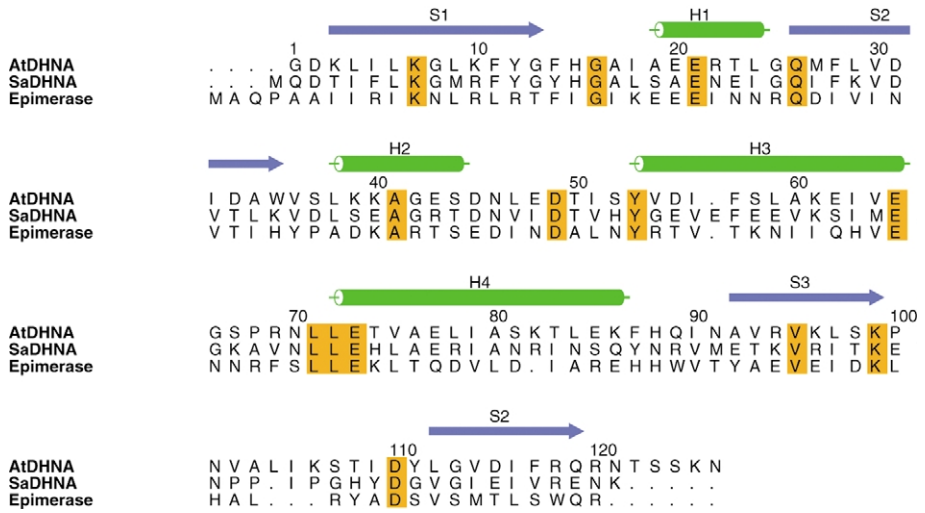

Figure 7. Secondary structure assignment of *A. thaliana* DHNA and sequence alignment with *S. aureus* DHNA and *E. coli* 7,8-dihydroneopterin triphosphate epimerase. Secondary structure elements (green, α-helices; blue, β-strands) found in *A. thaliana* DHNA are shown above the sequences. Identical residues (orange background) are indicated. The Figure was prepared with ALSCRIPT.[66]

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-Dihydroneopterin 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

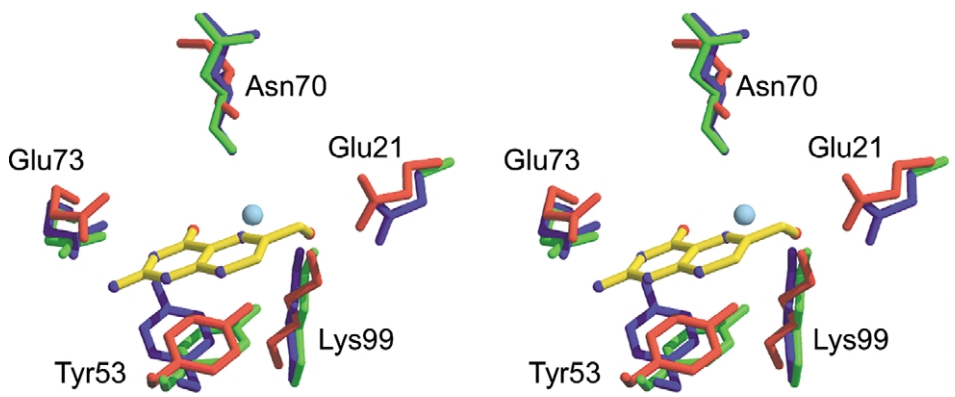

Figure 8. Stereo drawing of the superposition of catalytic active residues of *A. thaliana* DHNA (green), *S. aureus* DHNA (blue), and *E. coli* 7,8-dihydroneopterin triphosphate epimerase (red) with the modelled reaction product 6-hydroxymethyl-7,8-dihydropterin and the bound water molecule. In 7,8-dihydroneopterin triphosphate epimerase Asn70 is changed to a serine.

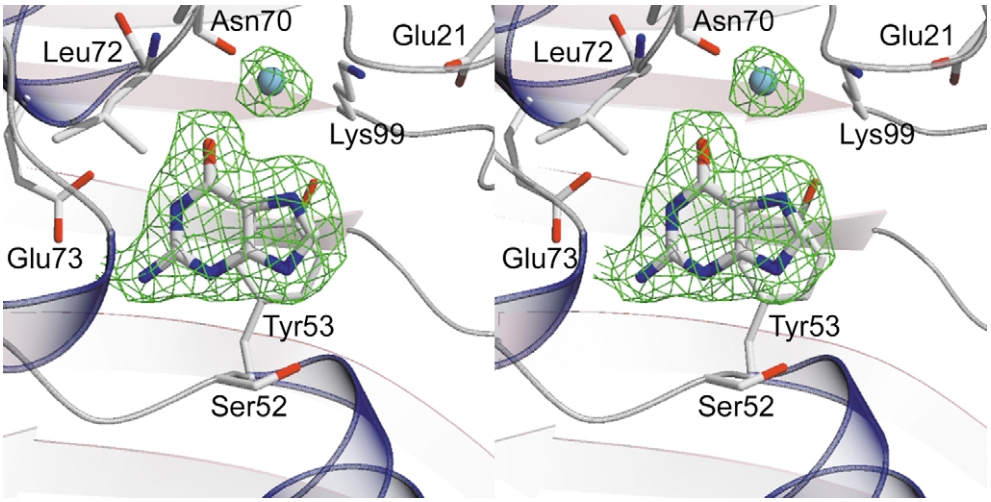

Figure 9. Stereo drawing of the *A. thaliana* DHNA active site with electron density for the bound guanine and the water molecule. Residues that form hydrophobic or hydrogen bond interactions are shown. Ser52 and Tyr53 belong to the adjacent subunit. The final $2F_o - F_c$ electron density map covering the ligand and the water molecule is contoured at 1 sigma.

(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-dihydroneopterin is bound consists of Val36, Leu38, Tyr53 and Val54 of one subunit, and Leu71 and Leu72 of the adjacent subunit. In detail, several hydrogen bonds to the modelled 7,8-dihydroneopterin 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 carbonyl group of Asn70 and the ε-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.

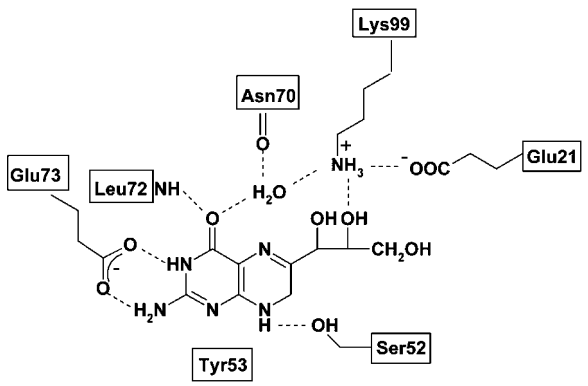

Figure 10. Schematic presentation of substrate interactions.

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),[40,41] transaldolase,[42] KDPG aldolase,[43] type I 3-dehydroquinase,[44] D-2-desoxyribose-5-phosphate aldolase (DRPA),[45] *N*-acetylneuraminate lyase[46] and dihydrodipicolinate synthease (DHDPS)[47] are known. The subunits of these enzymes all show classical TIM β/α-folds with a C-terminal region involved in substrate specifity. 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,[48] 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[42] and Asp187 and Asp188 in DHDPS.[47] In *E. coli* KDPG aldolase Glu45 is

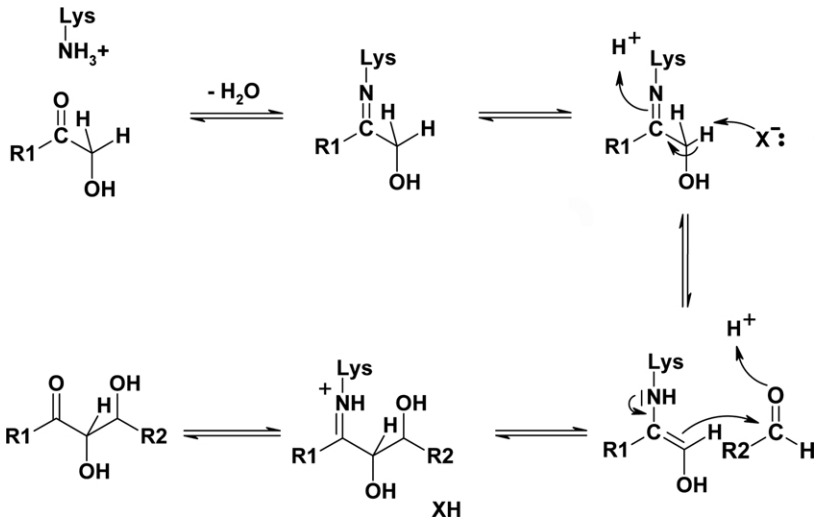

Figure 11. Catalytic mechanisms of class I aldolases with compound X as nucleophilic group.

proposed to act as a general base.[49,50] In both *N*-acetylneuraminate 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-dihydroneopterin and D-erythro-dihydroneopterin as substrates for the formation of 6-hydroxymethyl-7,8-dihydroneopterin, but it can also catalyse the epimerisation of C2′ of dihydroneopterin and dihydromonapterin at appreciable velocity. A possible reaction mechanism is shown in Figure 12. The proton at the position 2′ hydroxyl group is believed to be abstracted by Lys99, thus initiating the cleavage of the C1′–C2′ carbon bound under formation of the enamine type intermediate **2**, which can be converted to the carbinol type product (**3**) by tautomerisation. Epimerisation can be described as a sequence of retroaldol cleavage followed by aldol addition, which proceeds in a non-stereospecific mode. As shown earlier, the catalytic mechanisms of DHNA and 7,8-dihydroneopterin triphosphate epimerase are very similar,[16] except that the carbinol type intermediate cannot be released by the epimerase.

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.[7] 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 transimination). 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 β/α-folds with the C-terminal region of the barrel involved in substrate specificity while DHNA shows a ββααββ fold with elliptical shape and the active site located at the interface of two adjacent subunits. The DHNA ββααββ fold also occurs in 7,8-dihydroneopterin triphosphate epimerase, GTP cyclohydrolase I, 6-pyruvoyl tetrahydropterin synthase and urate oxidase, but with the exception of 7,8-dihydroneopterin 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-Dihydroneopterin and dihydromonapterin were purchased from Dr B. Schircks (Jona, Switzerland).

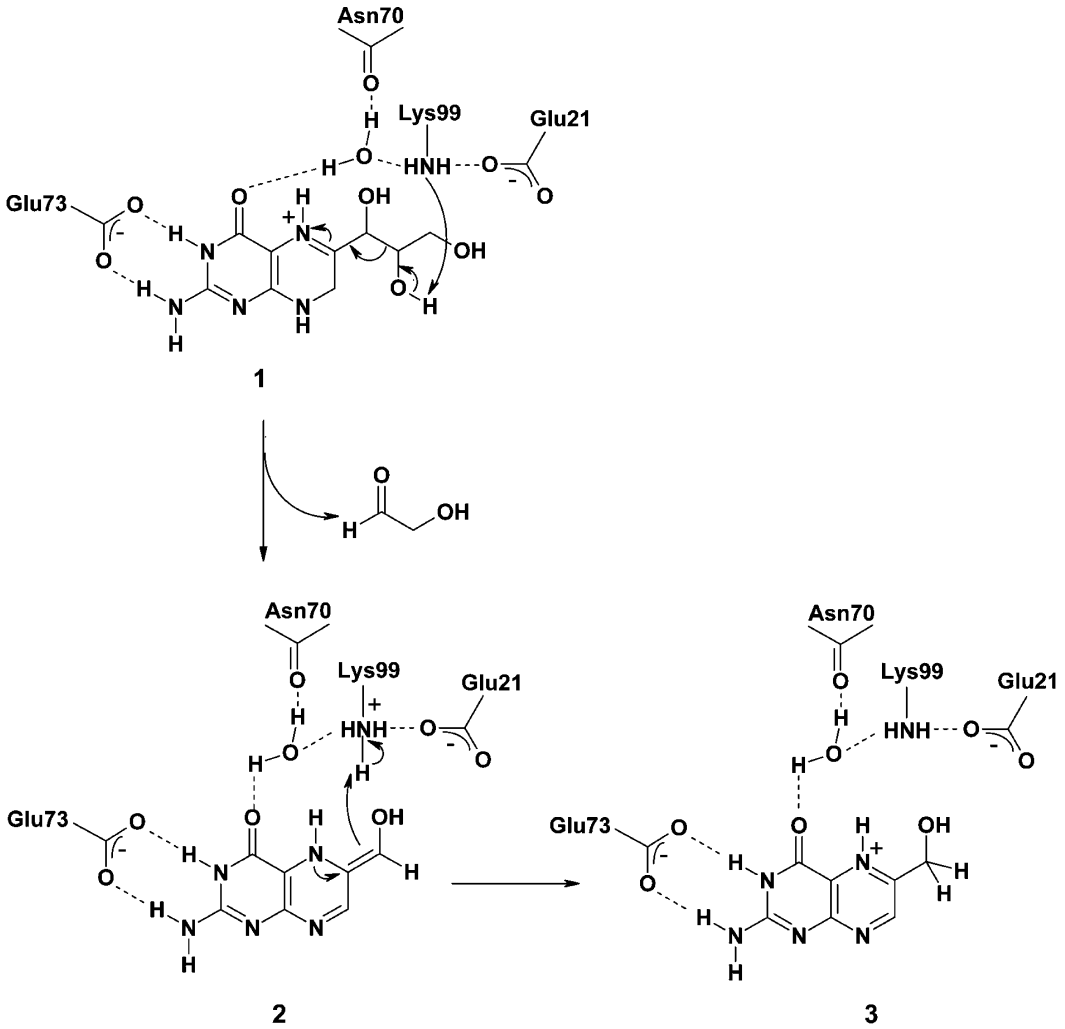

Figure 12. Proposed catalytic mechanism of *A. thaliana* DHNA.

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, Epsoo, 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.[51]

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 3. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source
E. coli strain		
XL-1-Blue	recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac[F′, proAB, lacI^{q}ZΔM15, Tn10(tet^r)]	26
M15 [pREP4]	lac, ara, gal, mtl, recA$^+$, uvr$^+$, [pREP4: lacI, kan^r]	43
Plasmids		
pNCO113	Expression vector	29
pNCO-AtfolB1-ΔE	Vector expressing the aldolase of *A. thaliana* without a leader sequence	This study

Table 4. Oligonucleotides used in this study

Designation	Endonuclease	Sequence
ATfolB1-ΔE-1	EcoRI	5′ ataatagaattc attaaagaggag aaattaaccatg gacaaactgata ctaaaagggttg 3′
ATfolB1-2	BamHI	5′ tattatggatcc ttagttctttga actagtgtttcgctgtc 3′

Restriction sites are underlined.

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] cells[53] carrying the pREP4 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 ($A_{600\ nm}$), 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). The column was developed with a linear gradient of 0–1.0 M potassium chloride in 50 mM Tris hydrochloride (pH 8.3) (total volume, 500 ml). Fractions were combined, and ammonium sulphate was added to 60% saturation. The precipitate was harvested by centrifugation and dissolved in 50 mM Tris hydrochloride (pH 8.3), containing 70 mM potassium chloride. The solution was placed on top of a Superdex 200 column (2.6 cm × 60 cm), which was then developed with 50 mM Tris hydrochloride (pH 8.3), containing 70 mM potassium chloride at a flow rate of 3 ml min^{-1}. Fractions were combined and placed on top of a column of Red-Sepharose CL-6B (11 ml) that had been equilibrated with 50 mM Tris hydrochloride (pH 8.3). The column was developed with a linear gradient of 0–1.0 M potassium chloride (total volume, 80 ml). Fractions were combined and concentrated by ultrafiltration.

Estimation of protein concentration

Protein concentration was estimated by a modified Bradford procedure.[54]

Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis was performed with 16% (w/v) polyacrylamide gels by published procedures.[55] Molecular mass standards were supplied by Sigma (Munich, Germany).

Protein sequencing

Sequence determination was performed by the automated Edman method using a 471 A Protein Sequencer (Perkin Elmer).

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 (cyclohexyl-butanoyl-*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 ($Cu\ K_\alpha$) = 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.[56] The electron density was improved by solvent flattening using the program DM. The structure was solved by molecular replacement using the program AMORE[57] 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.[58] Refinement steps carried out consisted of conjugate gradient minimisation, simulated annealing, and *B*-factor refinement with the program CNS using the mlf target.[59] 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.[60] The refinement was carried out on the resulting model until convergence was reached at an *R*-factor of 19.5% and an R_{free} of 25.8%.

The Ramachandran plot[61] calculated with the program PROCHECK[62] 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

Stereochemical parameters were assessed throughout refinement with PROCHECK.[62] Secondary structure elements were assigned with STRIDE.[63] Structural Figures were prepared with MOLSCRIPT,[64] BOBSCRIPT[65] 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

1. Cossins, E. A. (1980). The Biochemistry of Plants (Davies, D. D., ed.), vol. 2, pp. 365, Academic Press, New York.
2. Cossins, E. A. (1997). Folates and one-carbon metabolism in plants and fungi. *Phytochemistry*, **45**, 437–452.
3. Young, D. W. (1986). The biosynthesis of the vitamins thiamin, riboflavin, and folic acid. *Nature Prod. Rep.* **3**, 395–419.
4. Feinleib, M., Beresford, S. A., Bowman, B. A., Mills, J. L., Rader, J. I., Selhub, J. & Yetley, E. A. (2001). Folate fortification for the prevention of birth defects: case study. *Am. J. Epidemiol.* **154**, 60–69.
5. Green, J. M., Nichols, B. P. & Matthews, R. G. (1996). *Escherichia coli* and *Salmonella typhimurium* (Neidhardt, F. C., ed.), vol. 1, pp. 665–673, American Society for Microbiology, Washington, DC.
6. Bracher, A., Schramek, N. & Bacher, A. (2001). Biosynthesis of pteridines. Stopped-flow kinetic analysis of GTP cyclohydrolase I. *Biochemistry*, **40**, 7896–7902.
7. Schramek, N., Bracher, A., Fischer, M., Auerbach, G., Nar, H., Huber, R. & Bacher, A. (2002). Reaction mechanism of GTP cyclohydrolase I: single turnover experiments using a kinetically competent reaction intermediate. *J. Mol. Biol.* **316**, 829–837.
8. Nar, H., Huber, R., Auerbach, G., Fischer, M., Hösl, C., Ritz, H. *et al.* (1995). Active site topology and reaction mechanism of GTP cyclohydrolase I. *Proc. Natl Acad. Sci. USA*, **92**, 12120–12125.
9. Yim, J. J. & Brown, G. M. (1976). Characteristics of guanosine triphosphate cyclohydrolase I purified from *Escherichia coli*. *J. Biol. Chem.* **251**, 5087–5094.
10. Fan, C. L. & Brown, G. M. (1976). Partial purification and properties of guanosine triphosphate cyclohydrolase from *Drosophila melanogaster*. *Biochem. Genet.* **14**, 259–270.
11. Suzuki, Y. & Brown, G. M. (1974). The biosynthesis of folic acid. XII. Purification and properties of dihydroneopterin triphosphate pyrophosphohydrolase. *J. Biol. Chem.* **249**, 2405–2410.
12. De Saizieu, A., Vankan, P. & von Loon, A. P. G. M. (1995). Enzymic characterization of *Bacillus subtilis* GTP cyclohydrolase I. Evidence for a chemical dephosphorylation of dihydroneopterin triphosphate. *Biochem. J.* **306**, 371–377.
13. Mathis, J. B. & Brown, G. M. (1980). Dihydroneopterin aldolase from *Escherichia coli*. *Methods Enzymol.* **66**, 556–560.
14. Mathis, J. B. & Brown, G. M. (1970). The biosynthesis of folic acid. XI. Purification and properties of dihydroneopterin aldolase. *J. Biol. Chem.* **245**, 3015–3025.
15. Zimmerman, M., Tolman, R. L., Morman, H., Graham, D. W. & Rogers, E. F. (1977). Inhibitors of folate biosynthesis. 1. Inhibition of dihydroneopterin aldolase by pteridine derivatives. *J. Med. Chem.* **20**, 1213–1215.
16. Hauβmann, C., Rohdich, F., Schmidt, E., Bacher, A. & Richter, G. (1998). Biosynthesis of pteridines in *Escherichia coli*. Structural and mechanistic similarity of dihydroneopterin triphosphate epimerase and dihydroneopterin aldolase. *J. Biol. Chem.* **273**, 17418–17424.
17. Deng, H., Callender, R. & Dale, G. E. (2000). A vibrational structure of 7,8-dihydrobiopterin bound to dihydroneopterin aldolase. *J. Biol. Chem.* **275**, 30139–30143.
18. Illarionova, V., Eisenreich, W., Fischer, M., Hauβ mann, C., Römisch, W., Richter, G. & Bacher, A. (2002). Biosynthesis of tetrahydrofolate: stereochemistry of dihydroneopterin aldolase. *J. Biol. Chem.* **277**, 28841–28847.
19. Ortiz, P. J. & Hotchkiss, R. D. (1966). The enzymatic synthesis of dihydrofolate and dihydropteroate in cell-free preparations from wild-type and sulfonamide-resistant pneumococcus. *Biochemistry*, **5**, 67–74.
20. Richey, D. P. & Brown, G. M. (1969). The biosynthesis of folic acid. IX. Purification and properties of the enzymes required for the formation of dihydropteroic acid. *J. Biol. Chem.* **244**, 1582–1592.
21. Richey, D. P. & Brown, G. M. (1970). A comparison of the effectiveness with which p-aminobenzoic acid and p-aminobenzoylglutamic acid are used as substrate by dihydropteroate synthetase from *Escherichia coli*. *Biochim. Biophys. Acta*, **222**, 237–239.
22. Suckling, C. J., Sweeney, J. R. & Wood, H. C. (1977). Dihydropteroate synthase: purification by affinity chroma-tography and mechanism of action. *J. Chem. Soc.* **4**, 439–442.
23. Achari, A., Somers, D. O., Champness, J. N., Bryant, P. K., Rosemond, J. & Stammers, D. K. (1997). Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. *Nature Struct. Biol.* **4**, 490–497.
24. Chio, L. C., Bolyard, L. A., Nasr, M. & Queener, S. F.

† http://www.expasy.ch.spdv

(1996). Identification of a class of sulfonamides highly active against dihydropteroate synthase form *Toxoplasma gondii*, *Pneumocystis carinii*, and *Mycobacterium avium*. *Antimicrob. Agents Chemother.* **40**, 727–733.
25. Shiota, T. & Disraely, M. N. (1961). The enzymic synthesis of dihydrofolate from 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine and p-amino-benzoylglutamate by extracts of *Lactobacillus plantarum*. *Biochem. Biophys. Acta*, **52**, 467–473.
26. Shiota, T. & Palumbo, M. P. (1965). Enzymatic synthesis of the pteridine moiety of dihydrofolate from guanine nucleotides. *J. Biol. Chem.* **240**, 4449–4453.
27. Shiota, T., Baugh, C. M., Jackson, R. & Dillard, R. (1969). The enzymatic synthesis of hydroxymethyl-dihydropteridine pyrophosphate and dihydrofolate. *Biochemistry*, **8**, 5022–5028.
28. Jones, T. H. & Brown, G. M. (1967). The biosynthesis of folic acid. VII. Enzymatic synthesis of pteridines from guanosine triphosphate. *J. Biol. Chem.* **242**, 3989–3997.
29. Hennig, M., D'Arcy, A., Hampele, I. C., Page, M. G. P., Oefner, C. & Dale, G. E. (1998). Crystal structure and reaction mechanism of 7,8-dihydroneopterin aldolase from *Staphylococcus aureus*. *Nature Struct. Biol.* **5**, 357–362.
30. Gefflaut, T., Blonski, C., Perie, J. & Willson, M. (1995). Class I aldolases: substrate specificity, mechanism, inhibitors and structural aspects. *Prog. Biophys. Mol. Biol.* **63**, 301–340.
31. Gross, W., Bayer, M. G., Schnarrenberger, C., Gebhart, U. B., Maier, T. L. & Schenk, H. (1994). Two distinct aldoalses of class II type in the cyanoplasts and in the cytosol of the alga cyanophora pardoxa. *Plant Physiol.* **105**, 1393–1398.
32. Smith, G. M., Mildvan, A. S. & Harper, E. T. (1980). Nuclear relaxation studies of the interaction of substrates with a metalloaldolase from yeast. *Biochemistry*, **19**, 1248–1255.
33. Szwergold, B. S., Ugurbil, K. & Brown, T. R. (1995). Properties of fructose-1,6-bisphophate aldolase from *Escherichia coli*: an NMR analysis. *Arch. Biochem. Biophys.* **317**, 244–252.
34. Rutter, W. J. (1960). Aldolase. *Enzymes*, **5**, 341–372.
35. Ploom, T., Hauβmann, C., Hof, P., Steinbacher, S., Bacher, A., Richardson, J. & Huber, R. (1999). Crystal structure of 7,8-dihydroneopterin triphosphate epimerase. *Structure*, **7**, 509–516.
36. Nar, H., Huber, R., Heizmann, C. W., Thöny, B. & Bürgisser, D. (1994). Three-dimensional structure of 6-pyruvoyl tetrahydropterin synthase, an enzymee involved in tetrahydrobiopterin biosynthesis. *EMBO J.* **13**, 1255–1262.
37. Ploom, T., Thöny, B., Yim, J., Lee, S., Nar, H., Leimbacher, W. *et al.* (1999). Crystallographic and kinetic investiagtions on the mechanism of 6-pyruvoyl tetrahydropterin synthase. *J. Mol. Biol.* **286**, 851–860.
38. Nar, H., Huber, R., Meining, W., Schmmid, C., Weinkauf, S. & Bacher, A. (1995). Atomic structure of GTP cyclohydrolase I. *Structure*, **3**, 459–466.
39. Colloc'h, N., El Hajji, M., Bachet, B., L'Hermite, G., Schiltz, M., Prangé, T. *et al.* (1997). Crystal structure of the protein drug urate oxidase-inhibitor complex at 2.05 Å resolution. *Nature Struct. Biol.* **4**, 947–952.
40. Choi, K. H., Shi, J., Hopkins, C. E., Tolan, D. R. & Allen, K. N. (2001). Snapshots of catalysis: the structure of fructose-1,6-(bis)phosphate aldolase covalently bound to the substrate dihydroxyacetone phosphate. *Biochemistry*, **40**, 13868–13875.
41. Sygusch, J., Beaudry, D. & Allaire, M. (1987). Molecular architecture of rabbit skeletal muscle aldolase at 2.7 Å resolution. *Proc. Natl Acad. Sci. USA*, **84**, 7846–7850.
42. Jia, J., Huang, W., Schorken, U., Sahm, H., Sprenger, G. A., Lindqvist, Y. & Schneider, G. (1996). Crystal structure of transaldolase B from *Escherichia coli* suggests a cricular permutation of the α/β barrel within the class I aldolase family. *Structure*, **4**, 715–724.
43. Mavridis, I. M. & Tulinsky, A. (1976). The folding and quaternary structure of trimeric 2-keto-3-deoxy-6-phosphogluconic aldolase at 3.5-A resolution. *Biochemistry*, **15**, 4410–4417.
44. Gourley, D. G., Shrive, A. K., Polikarpov, I., Krell, T., Coggins, J. R., Hawkins, A. R., Isaacs, N. W. & Sawyer, L. (1999). The two types of 3-dehydro-quinase have distinct structures but catalyze the same overall reaction. *Nature Struct. Biol.* **6**, 521–525.
45. Heine, A., DeSantis, G., Luz, J. G., Mitchell, M., Wong, C.-H. & Wilson, I. A. (2001). Observation of covalent intermediates in an enzyme mechanism at atomic resolution. *Science*, **294**, 369–374.
46. Izard, T., Lawrence, M. C., Malby, R. L., Lilley, G. G. & Colman, P. M. (1994). The three-dimensional structure of *N*-acetylneuraminate lyase from *Escherichia coli*. *Structure*, **2**, 361–369.
47. Mirwaldt, C., Korndörfer, I. & Huber, R. (1995). The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *J. Mol. Biol.* **246**, 227–239.
48. Dalby, A., Dauter, Z. & Litllechild, J. A. (1999). Crystal structure of human muscle aldolase complexed with fructose 1,6-bisphosphate: mechanistic implications. *Protein Sci.* **8**, 291–297.
49. Wymer, N., Buchanan, L. V., Henderson, D., Mehta, N., Botting, C. H., Pocivavsek, L. *et al.* (2001). Directed evolution of a new catalytic site in 2-keto-3-deoxy-6-phosphogluconate aldolase from *Escherichia coli*. *Structure*, **9**, 1–10.
50. Allard, J., Grochulski, P. & Sygusch, J. (2001). Covalent intermediate trapped in 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase structure at 1.95-Å resolution. *Proc. Natl Acad. Sci. USA*, **98**, 3679–3684.
51. Fischer, M., Haase, I., Feicht, R., Richter, G., Gerhardt, S., Changeux, J. P. *et al.* (2002). Biosynthesis of riboflavin 6,7-dimethyl-8-ribityllumazine synthase of *Schizosaccharomyces pombe*. *Eur. J. Biochem.* **269**, 519–526.
52. Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). XL1-BLUE: A high efficiency plasmid transforming reca *Escherichia coli* strain with beta-galactosidase selection. *BioTechniques*, **5**, 376–379.
53. Stüber, D., Matile, H. & Garotta, G. (1990). System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies, and structure-function analysis. Immunological Methods (Lefkovits, I. & Pernis, P., eds), vol. 4, pp. 121–152, Academic Press, Orlando, FL.
54. Read, S. M. & Northcote, D. H. (1981). Minimization of variation in the response to different proteins of the Coomassie blue G dye-binding assay for protein. *Anal. Biochem.* **116**, 53–64.
55. Laemmli, U. K. (1970). Cleavage of structural

proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
56. Otwinowski, Z. & Minor, W. (1997). Proceedings of X-ray diffraction data collected in oscillations mode. *Methods Enzymol.* **276**, 307–326.
57. Navaza, J. (1994). AMoRE: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.
58. Turk, D. Dissertation, Technische Universität (1992).
59. Brünger, A. T. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
60. Brünger, A. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
61. Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283–437.
62. Laskowski, R. A., McArthur, M. W., Moss, D. S. & Thronton, J. M. (1993). PROCHECK - a programm to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
63. Frishman, D. & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Struct. Funct. Genet.* **23**, 566–579.
64. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
65. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 132–134.
66. Barton, G. J. (1993). ALSCRIPT a tool to format multiple sequences alignments. *Protein Engng.* **6**, 37–40.

Edited by D. Rees

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