

Biosynthesis of Riboflavin: Structure and Properties of 2,5-Diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate Reductase of *Methanocaldococcus jannaschii*

Lorenz Chatwell^{1*†}, Tobias Krojer^{1†}, Alexander Fidler²
Werner Römisch², Wolfgang Eisenreich², Adelbert Bacher²
Robert Huber¹ and Markus Fischer^{2*}

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

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

The pyrimidine reductase of the riboflavin biosynthetic pathway (MjaRED) specified by the open reading frame MJ0671 of *Methanocaldococcus jannaschii* was expressed in *Escherichia coli* using a synthetic gene. The synthetic open reading frame that was optimized for expression in *E. coli* directed the synthesis of abundant amounts of the enzyme with an apparent subunit mass of 25 kDa. The enzyme was purified to apparent homogeneity and was shown to catalyze the conversion of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate into 2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5'-phosphate at a rate of 0.8 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ at pH 8.0 and at 30 °C. The protein is a homodimer as shown by sedimentation equilibrium analysis and sediments at an apparent velocity of 3.5 S. The structure of the enzyme in complex with the cofactor nicotinamide adenine dinucleotide phosphate was determined by X-ray crystallography at a resolution of 2.5 Å. The folding pattern resembles that of dihydrofolate reductase with the *Thermotoga maritima* ortholog as the most similar structure. The substrate, 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate, was modeled into the putative active site. The model suggests the transfer of the *pro*-R hydrogen of C-4 of NADPH to C-1' of the substrate.

© 2006 Published by Elsevier Ltd.

Keywords: archaea; *Methanocaldococcus jannaschii*; riboflavin biosynthesis; pyrimidine nucleotide reductase; x-ray structure

*Corresponding authors

† L.C. and T.K. equally contributed to this work.

Present addresses: A. Fidler, Roche Diagnostics GmbH, Pharma Research Penzberg, TR-TNP 2, Process Research Biochemistry, Nonnenwald 2, D-82372 Penzberg, Germany; T. Krojer, IMP (Research Institute of Molecular Pathology), Dr. Bohr-Gasse 7, 1030 Vienna, Austria.

Abbreviations used: Mja, *Methanocaldococcus jannaschii*; r.m.s.d., root mean square deviation; SeMet, selenomethionine; Cymal-2, cyclohexyl-ethyl- β -D-maltoside; MjaRED, 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate reductase of *M. jannaschii*; DHFR, dihydrofolate reductase; ECDHFR, dihydrofolate reductase of *Escherichia coli*; TMDHFR, dihydrofolate reductase of *Thermotoga maritima*; DAROPP, 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate; DARIPP, 2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5'-phosphate; NADP, nicotinamide adenine dinucleotide phosphate.

E-mail addresses of the corresponding authors: chatwell@wzw.tum.de; markus.fischer@ch.tum.de

Introduction

Flavoenzymes are characterized by an extraordinary chemical versatility. They can catalyze redox processes involving one and two-electron transitions but also a variety of non-redox reactions such as photorepair of thymidine dimers in photodamaged DNA and the dehydration of certain non-activated organic substrates.^{1,2} More recently, they have also been shown to act as chromophores in blue light photoreceptors in plants and fungi.³ Moreover, flavoproteins are involved in numerous other physiological processes involving bioluminescence and circadian time-keeping.^{4–7}

The biosynthesis of riboflavin has been studied in some detail (for reviews see Fischer and Bacher^{8,9}). The convergent biosynthetic pathway (Figure 1) requires one molecule of GTP (1) and two molecules

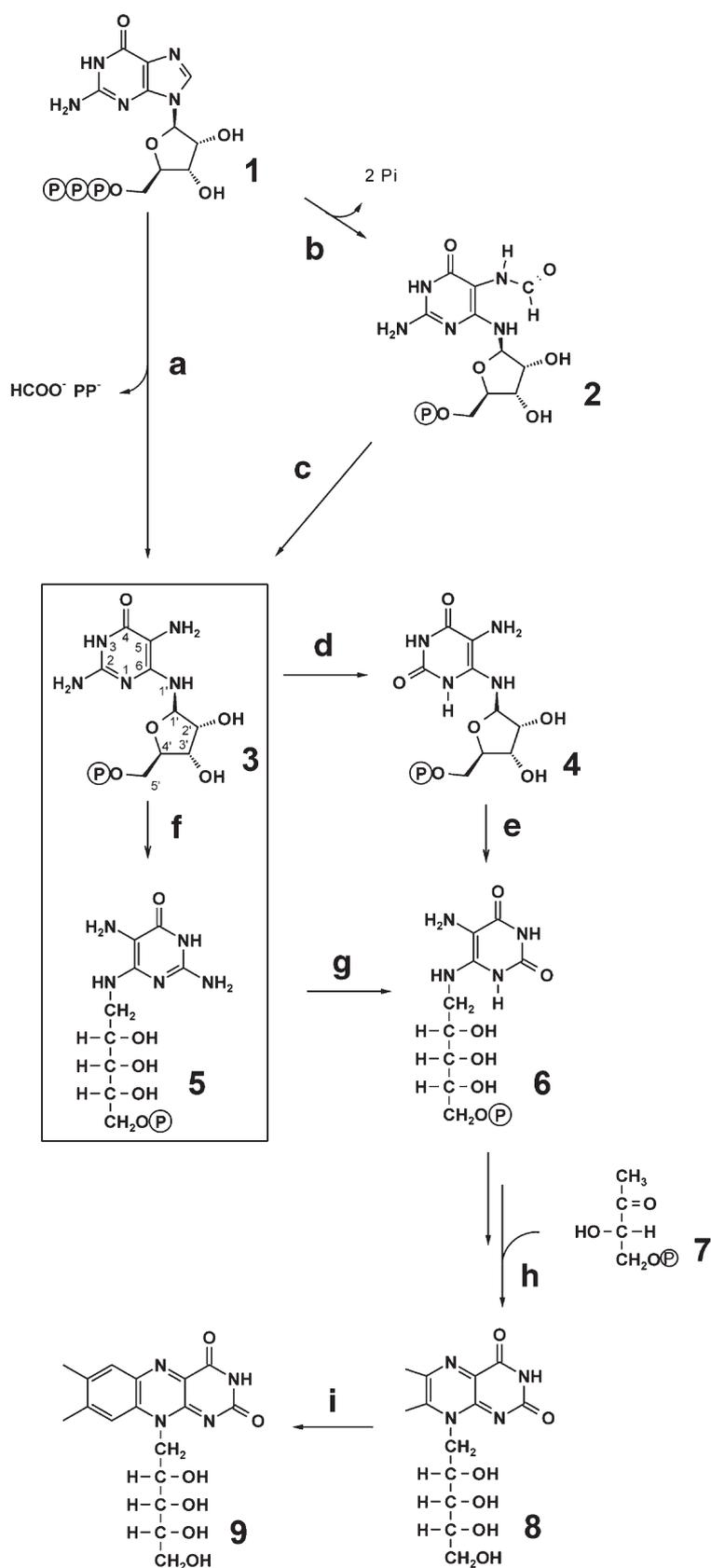


Figure 1. Biosynthesis of riboflavin. GTP cyclohydrolase II (**a**), GTP cyclohydrolase III (**b**), unknown enzyme (**c**), deaminase (**d**, **g**), reductase (**e**, **f**), lumazine synthase (**h**), riboflavin synthase (**i**). The biosynthetic pathway proceeds *via d* and *e* in Eubacteria and plants and *via f* and *g* in yeasts.^{15,16,73} 1, GTP; 2, 2-amino-5-formylamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate; 3, 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate; 4, 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate; 5, 2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5'-phosphate; 6, 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate; 7, 3,4-dihydroxy-2-butanone 4-phosphate; 8, 6,7-dimethyl-8-ribityllumazine; 9, riboflavin. The reaction (**f**) catalyzed by MJRED is highlighted by a box.

of ribulose 5-phosphate per molecule of the vitamin. In eubacteria, plants and yeasts, the first committed step is catalyzed by GTP cyclohydrolase II (Figure 1,

a) and involves the release of formate from the imidazole ring and the release of pyrophosphate from the side-chain of the nucleotide precursor

affording 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (3).

Early *in vivo* work with riboflavin deficient yeast mutants indicated that the ribosyl moiety of 3 is reductively transformed into the ribityl side-chain of 5.^{10,11} This sequence of reactions was later confirmed by enzymatic *in vitro* studies with *Ashbya gossypii* and *Saccharomyces cerevisiae*.^{12–14} On the other hand, the deamination step (Figure 1, d) was shown to precede the side-chain reduction (Figure 1, e) in eubacteria¹⁵ and in higher plants.¹⁶ Early studies

showed the basic pathway of riboflavin biosynthesis in Archaea to be similar to that of Eubacteria and fungi by *in vivo* studies with ¹³C-labeled precursors.¹⁷

5-Amino-6-ribitylamino-2,4(1H,3H)-pyrimidine-dione 5'-phosphate (6) needs to be dephosphorylated prior to the enzymatic formation of 6,7-dimethyl-8-ribityllumazine (8), but the hypothetical phosphatase has not yet been described.

The pyrimidine derivative 6 affords 6,7-dimethyl-8-ribityllumazine (8) by condensation with 3,4-

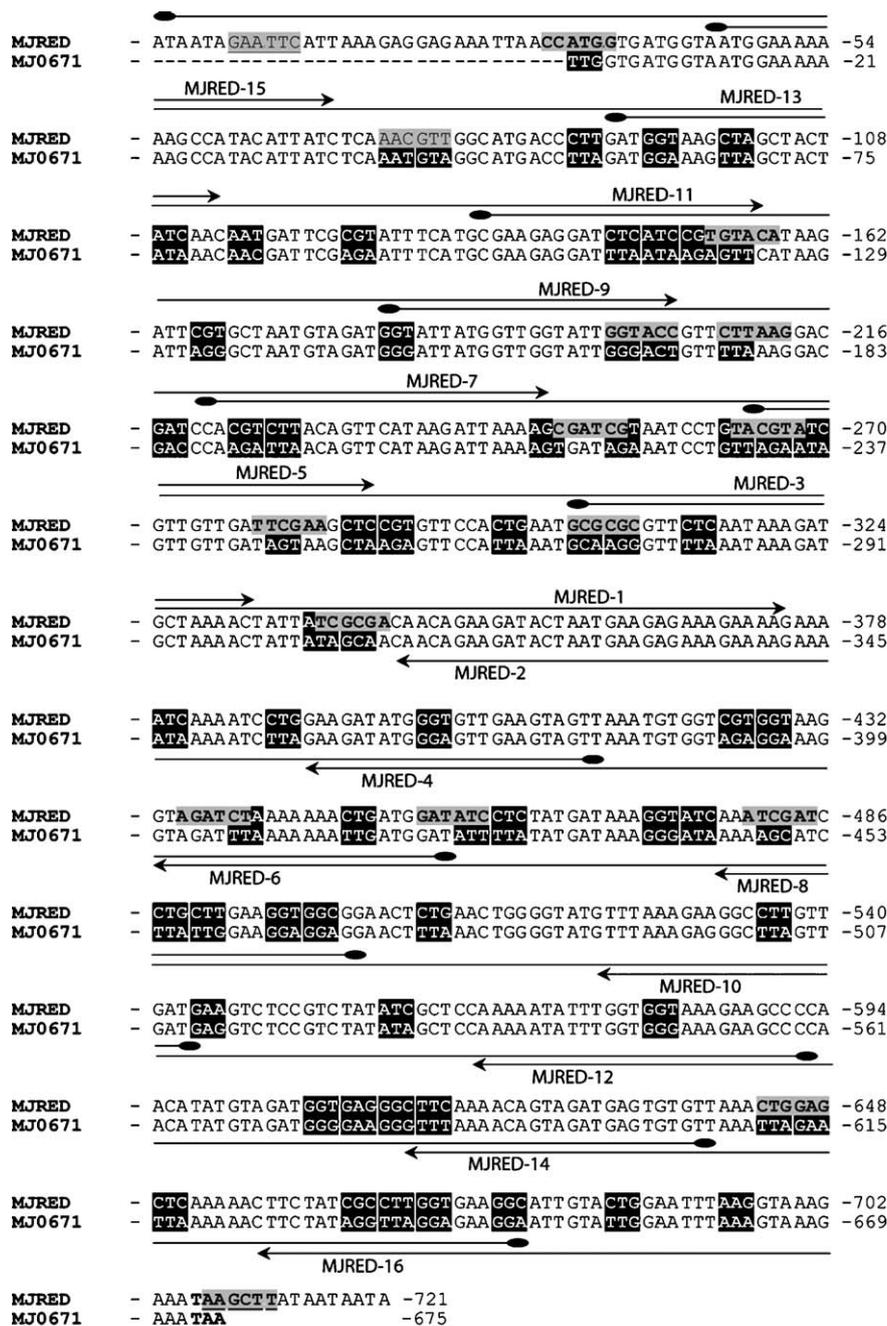


Figure 2. Construction of a synthetic gene for MjaRED. Alignment of the wild-type DNA sequence (MJ0671) and the synthetic DNA sequence (MJRED) with 5' and 3' overhangs including the synthetic EcoRI and BamHI sites. Changed codons are shaded in black. New single restriction sites are shaded in grey. Oligonucleotides used as forward primers are drawn above and reverse primers below the aligned DNA sequences.

dihydroxy-2-butanone 4-phosphate (7), which is obtained from ribulose 5-phosphate by a skeletal rearrangement. The final step of the biosynthetic pathway involves an unusual dismutation of the pteridine derivative 8 affording riboflavin (9) and 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (6).

It was also shown that the dephosphorylated compound 6 serves as an intermediate in the biosynthetic pathways of the 5-deazaflavin derivative, coenzyme F₄₂₀.¹⁸

Archaea have no detectable homologs of GTP cyclohydrolases II. Instead, they use a GTP cyclohydrolase of a different type (GTP cyclohydrolase III, Figure 1, b) that produces 2-amino-5-formylamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (2).¹⁹ The hypothetical enzyme (Figure 1, c) catalyzing the deformylation of 2 yielding 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (3) is still unknown.

The open reading frame MJ0671 of *Methanocaldococcus jannaschii* has been reported recently to specify an enzyme that converts 3 into 5 (Figure 1, f).²⁰

Archaea were shown to possess lumazine synthases and 3,4-dihydroxy-2-butanone 4-phosphate synthases closely resembling those of eubacteria, fungi and plants.^{21–24} Unusual riboflavin synthases devoid of similarity with the enzymes from Eubacteria have been cloned from *Methanobacterium thermoautotrophicum* and *M. jannaschii*^{25–27} and have been shown to display significant sequence similarity with 6,7-dimethyl-8-ribityllumazine synthases which catalyze the penultimate step of riboflavin biosynthesis.^{26,28}

Since the termination of this work, the structure of the RibG protein of *Bacillus subtilis* has been published.²⁹ The reductase domain of that protein catalyzes the reduction of 4 affording 6 (Figure 1). It should be noted that RibG protein and MjaRED use different substrates (4 and 3, respectively) and generate different products (6 and 5, respectively). On the other hand, both enzymes catalyze the reduction of the ribose side-chain of their respective substrate, and their sequences and three-dimensional structures are similar.

Here we report the three-dimensional structure of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate reductase of *M. jannaschii* (MjaRED). The enzyme from the Archaeobacterium *M. jannaschii* closely resembles 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate reductases found in fungi.

Results

Cloning, expression and biochemical characterization

The hypothetical open reading frame MJ0671 of *M. jannaschii* predicts a protein of 224 amino acid residues. Because MJ0671 contains numerous codons that are poorly translated in *E. coli*, we

designed a synthetic gene that was optimized for the *E. coli* codon usage. Approximately 31% (69 of 224) of the codons were replaced, and 14 singular restriction sites were introduced (Figure 2). The DNA sequence was assembled from 16 synthetic oligonucleotides by a sequence of eight PCR steps. The synthetic gene was transcribed efficiently in a recombinant *E. coli* strain, affording approximately 30% of cellular protein.

The recombinant protein was purified to apparent homogeneity. The N-terminal sequence of the recombinant protein was verified by partial Edman degradation affording the sequence motif VMVMEKKPYI. Notably, the N-terminal methionine had been removed by post-translational modification.

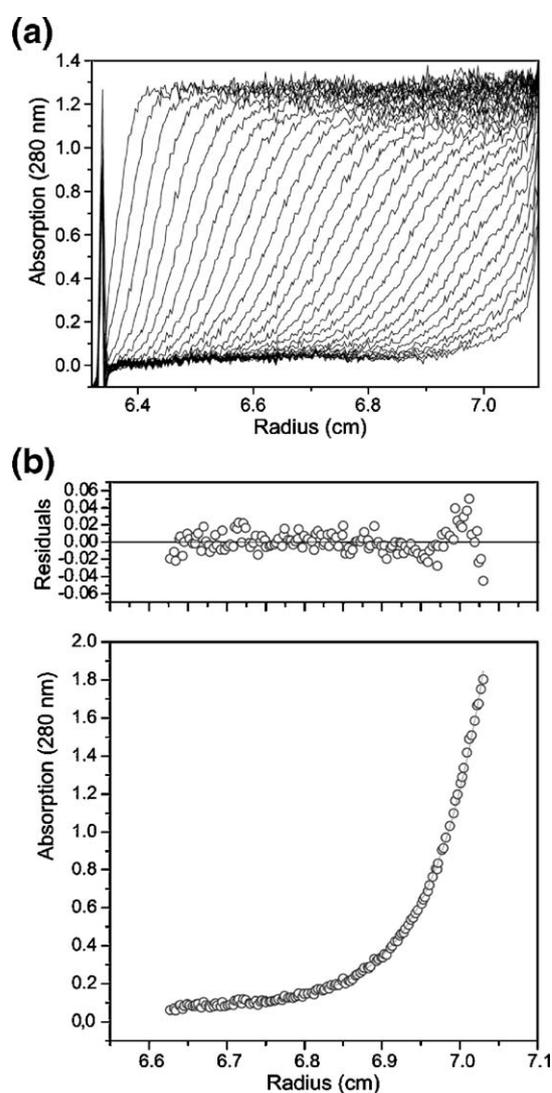


Figure 3. (a) Boundary sedimentation of MjaRED. Conditions used were: 20 °C, 50 mM Tris hydrochloride (pH 7.0), 5 mM sodium chloride, 5 mM dithiothreitol and 3.0 mg of protein per ml. The sample was scanned at intervals of 5 min. (b) Sedimentation equilibrium of MjaRED. Conditions used were: 100 mM potassium phosphate (pH 5.0), 300 mM sodium chloride and 0.4 mg of protein per ml at 10,000 rpm and 4 °C. Residuals are shown at the top.

Electrospray mass spectrometry afforded a peak corresponding to a molecular mass of 24,906, well in line with the cleavage of the initial methionine (calculated mass of the full-length protein, 25,037; calculated mass without initial methionine, 24,906).

The recombinant protein was shown to catalyze the formation of 2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5'-phosphate (5) from 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (3) at a rate of 0.8 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ at pH 8.0 and at 30 °C as shown by ^{13}C NMR spectroscopy.

The enzyme sediments as a single, symmetrical boundary with an apparent velocity of 3.5 S at 20 °C (Figure 3(a)). Sedimentation equilibrium afforded a relative mass of 50 kDa, which is close to the mass predicted for a homodimer (Figure 3(b)).

Structural features

MjaRED was crystallized as described in Materials and Methods. X-ray data processing and phasing statistics are given in Table 1. The crystals belonged to space group $P4_32_12$, with the asymmetric unit containing six 25.5 kDa monomers. These are arranged as two non-physiological trimers (Figure 4). Trimerization occurs principally *via* the secondary structure element αC leading to a symmetrical structure. The two trimers are related by a translation of approximately 0.5/0.5/0.49. The average buried surface area between MjaRED monomers is 300–350 \AA^2 . Interestingly, the contacts between monomers were mediated by the detergent Cymal-2 (cyclohexyl-ethyl- β -D-maltoside), which was essential for crystallization (Figure 4). Most of the crystal contacts as well as the contacts within the asymmetric unit are provided by this detergent. The trimerization site is mediated by two Cymal-2 molecules per monomer, which are well defined by strong electron densities. The carbohydrate tail of Cymal-2 is mainly surrounded by polar and hydrophobic residues from helices C, but also from βH and βI . Some residues contacting the detergent tail are isoleucine 44, 12, and 118, leucine 212 and 89, and phenylalanine 209. The solvent-exposed disaccharide group is hydrogen bonded to several residues in all helices of neighbouring molecules, but especially the tyrosine residues 10 and 210 play a major role. Monomer interactions in the crystal give no hint for the dimer formation of MjaRED in solution as observed by analytical ultracentrifugation (Figure 3).

Interpretable main chain electron density was observed for 219 of the total number of 224 residues of each MjaRED monomer. Also observed in the asymmetric unit are 274 ordered water molecules and one NADP bound per monomer. This ligand was not added for crystallization but remained bound during all preparation steps. It appears most likely that the bound NADP did not remain in the reduced state, but a clear distinction on the basis of structural arguments, due to the limitation of the resolution, is not possible. After refinement, 88.7% of residues are within the most favored region for main chain phi-psi angles on a Ramachandran plot³⁰ and

have an average protein *B*-factor of 53.7 \AA^2 (Table 2). The r.m.s. deviations between the C-alpha atoms of the six subunits range from 0.13 \AA to 0.28 \AA , indicating that all molecules are virtually identical.

The overall structure

Each monomer is composed of a mixed α/β structure³¹ comprising a nine-stranded β -sheet (strands $\beta\text{A} - \beta\text{I}$). The sheet is flanked by four α -helices (A, B, D, and E), two helices on either side of the sheet and an additional one (C) on the top side. Seven more or less extended loops provide connectivity within the secondary structure elements (Figure 5). The substrate and coenzyme binding functions of MjaRED are carried out by non-contiguous segments of the amino acid sequence rather than by separate domains.

The crystal structure shows that MjaRED resembles the typical structural features of the oxidoreductase family of dihydrofolate reductases (DHFR). The Dali algorithm³² used to search for structural homologues identified the DHFR from *Thermotoga maritima* as the most similar structure in

Table 1. X-ray data processing and phasing statistics

	Peak	Inflection point	High remote
<i>Crystal data</i>			
Space group	Mercury substituted crystals		
Resolution (\AA) ^a	14,22		
Unit cell (\AA)	20–2.7 (2.8–2.7)		
	$a = 135.6, b = 135.6, c = 208.8,$		
	$\alpha = 90.0, \beta = 90.0, \gamma = 90.0$		
Wavelength (\AA)	1.006	1.009	0.950
$\langle I \rangle / \langle \sigma \langle I \rangle \rangle$	29.5(5.0)	28.0(3.2)	29.3(5.2)
Completeness (%)	98.8(90.4)	98.3(90.3)	99.5(94.2)
Measured reflections	305,981	304,691	325,107
Unique reflections	53,783	53,597	56,776
R_{sym} (%) ^b	0.056(0.471)	0.060(0.552)	0.057(0.340)
<i>Phasing procedure</i>			
Phasing power	1.36/1.70	1.64/1.20	0/1.63
iso/ano ^c			
R_{cullis} iso/ano ^d	0.53/0.73	0.52/0.84	0/0.73
Figure of merit		0.50	
<i>Refinement</i>			
Resolution (\AA)	20.00–2.7		
Number of protein atoms/solvent molecules			4447/–
rmsd bond length (\AA)/angles($^\circ$)/bonded B_s (\AA^2)			0.0109/1.59/1.63
Number of reflections $R_{\text{work}}/R_{\text{free}}$			50,999/2656
Ramachandran: most favored/disallowed			84.7/0
$R_{\text{cryst}}/R_{\text{free}}$ (%) ^e			37.3/39.2
Average <i>B</i> -value (\AA^2) overall			46.4

^a Last resolution shell is given in parentheses.

^b R_{sym} is the unweighted *R*-value on *I* between symmetry mates.

^c Phasing power is the mean value of the heavy-atom derivative structure factor amplitude divided by the residual lack of closure error.

^d $R_{\text{cullis}} = \sum_{\text{hkl}} | |F_{\text{ph}}| - |F_{\text{p}}| - F_{\text{h,calc}} | / \sum_{\text{hkl}} |F_{\text{ph}} - F_{\text{p}}|$.

^e $R_{\text{cryst}} = \sum_{\text{hkl}} | |F_{\text{obs}}(\text{hkl})| - k |F_{\text{calc}}(\text{hkl})| | / \sum_{\text{hkl}} |F_{\text{obs}}(\text{hkl})|$ for the working set of reflections; R_{free} is the *R*-value for 5% of the reflections excluded from refinement.

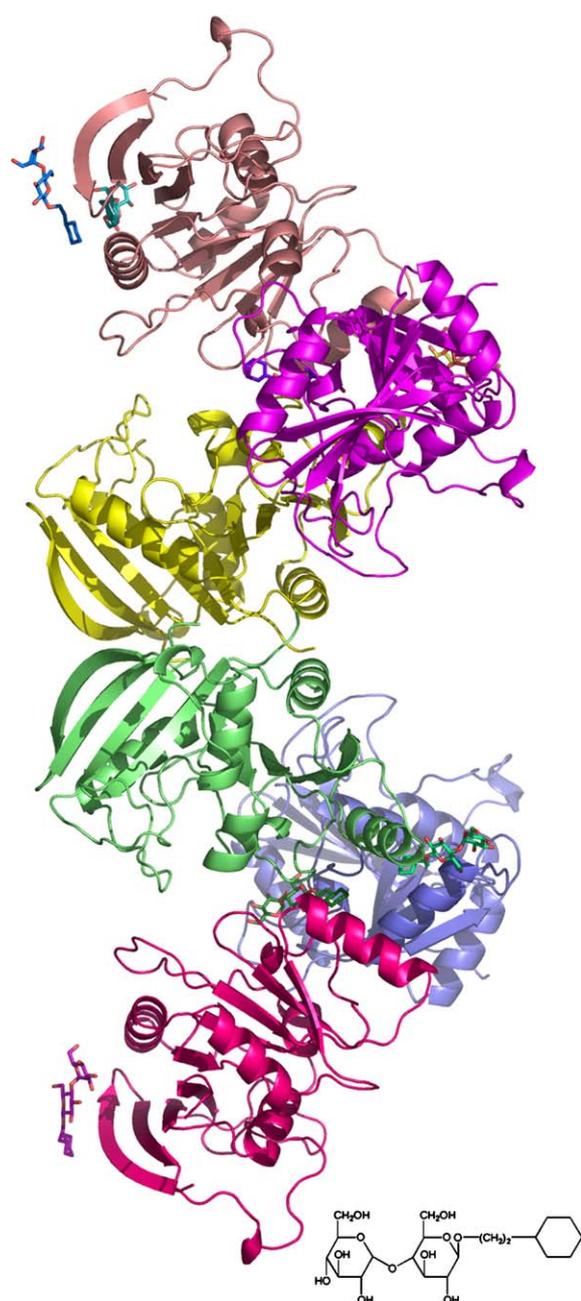


Figure 4. Ribbon view of the packaging arrangement of MjaRED with bound Cymal-2 molecules. The Cymal-2 molecules are drawn as ball-and-stick models. The chemical formula of Cymal-2 (cyclohexyl-ethyl- β -D-maltoside) is given below.

the database (PDB entry 1D1G) (Z score 18.6 Å, superposition rmsd of 2.2 Å across the alpha carbons). This enzyme shows 22% sequence identity with MjaRED (Figure 6(a) and (b)). The structural similarity of the MjaRED monomer to the TMDHFR³³ noted above does not extend to the packing architecture of the functional oligomers. The DHFR of *T. maritima* crystallizes as a homodimer, in accordance with the solution state of MjaRED (Figure 3(b)). MjaRED is considerably

larger than the members of the DHFR family (224 residues compared to 168 (TMDHFR) and 159 (ECDHFR)^{33,34}) and most of the large structural deviations are due to insertions (Figure 6(a) and (b)).

Based on comparison to the DHFR of *E. coli* the overall structure of MjaRED can be divided into two subdomains. The adenosine binding subdomain³⁵ comprises strands B, C, D, and F and the helices B, D, and E. The loop subdomain comprises strands A, G, H and I and the helix A, but about 50% of this subdomain consists of loops. The gap separating the two subdomains is formed by a break in the hydrogen-bonding pattern between strands A and F. The nicotinamide ring, coplanar with the sheet, spans part of this gap (Figure 5).

The cofactor binding site

MjaRED and DHFR both use NADPH as cofactor and possess common structural characteristics

Table 2. Crystal data and final refinement statistics

	Peak	Inflection point	High remote
<i>Crystal data</i>			
Space group	P4 ₃ 2 ₁ 2		
Resolution (Å) ^a	30–2.5 (2.56–2.5)		
Unit cell (Å)	a=136.9, b=136.9, c=213.6, α=90.0, β=90.0, γ=90.0		
Wavelength (Å)	0.9790	0.9795	0.9567
<I>/<σ (I)>	17.6(2.0)	16.4(1.7)	15.3(1.3)
Completeness (%)	97.9 (83.4)	98.3(83.9)	99.4 (94.2)
Measured reflections	491,750	498,508	510,658
Unique reflections	131,576	132,214	134,041
R _{sym} (%) ^b	0.078(0.549)	0.084 (0.634)	0.089 (0.804)
<i>Phasing procedure</i>			
Phasing power iso/ano ^c	0/1.68	0.83/0.78	0.79/0.65
R _{cullis} iso/ano ^d	0/0.72	0.74/0.82	0.81/0.87
Figure of merit	0.47		
<i>Refinement</i>			
Resolution (Å)	25.00–2.6		
Number of protein atoms/solvent molecules	10,260/274		
rmsd bond length (Å)/angles(°)/bonded Bs(Å ²)	0.007/1.403/1.430		
Number of reflections R _{work} /R _{free}	127,183/6681		
Ramachandran: most favored/disallowed	88.7/0		
R _{cryst} /R _{free} (%) ^e	23.0/25.3		
Average B-value (Å ²) overall	55.9		
Average B-value (Å ²) protein	53.7		
Average B-value (Å ²) NADPH	83.2		
Average B-value (Å ²) solvent	84.3		

^a Last resolution shell is given in parentheses.

^b R_{sym} is the unweighted R-value on I between symmetry mates.

^c Phasing power is the mean value of the heavy-atom derivative structure factor amplitude divided by the residual lack of closure error.

^d $R_{cullis} = \sum_{hkl} | |F_{ph}| - |F_p| - F_{h,calc} | / \sum_{hkl} |F_{ph} - F_p|$.

^e $R_{cryst} = \sum_{hkl} | |F_{obs} (hkl)| - k |F_{calc} (hkl)| | / \sum_{hkl} |F_{obs} (hkl)|$ for the working set of reflections. R_{free} is the R-value for 5% of the reflections excluded from refinement.

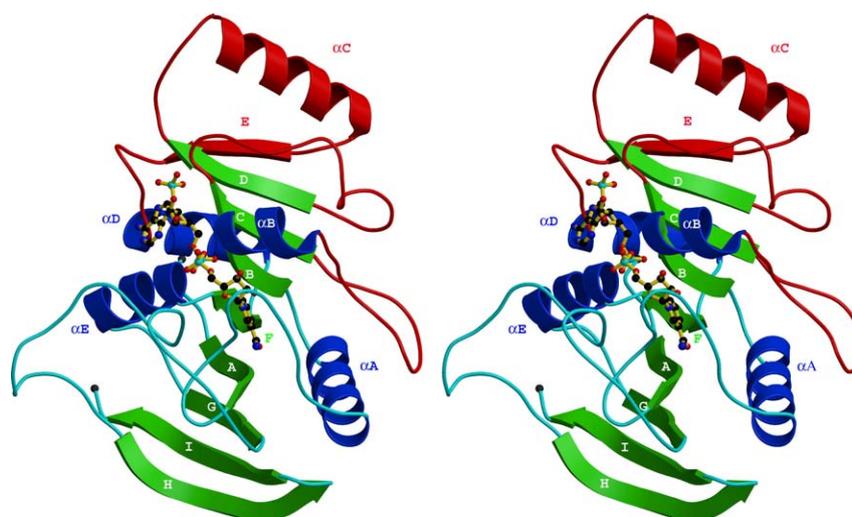


Figure 5. Stereo ribbon views of the overall structure of MjaRED with bound NADP. Alpha helices are colored in blue and beta strands in green. Structural elements not present in TMDHFR are colored in red. The cofactor is drawn as a ball-and-stick model.

(Figure 6(a)). The main chain conformational differences are relatively modest and the general mode of cosubstrate binding is fairly well conserved. Nevertheless there are several deviations. The position of the adenine site is shifted by 1.5 Å as compared to the one bound in the DHFR of *T. maritima* and the specific interatomic contacts are different. These are mainly van der Waals and hydrophobic interactions as described for DHFR.³³

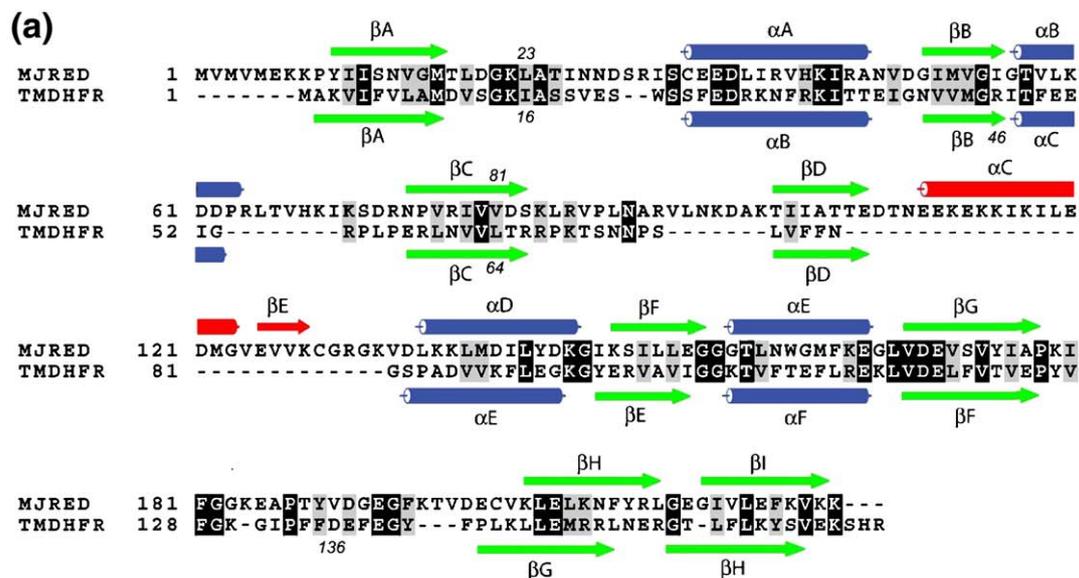
The strictly conserved Leu64 (TMDHFR) is replaced by Val81 in MjaRED (Figure 6(a)). The side-chain of Ser83 and the backbone of Val134 and Gly162 form the polar walls. Only the backbone oxygen of Val134 builds up a hydrogen bond with the adenine. As the distances of the other polar residues are approximately 3.3 Å to the cofactor, they are best characterized as polar interactions.

The strands β B and β F both terminate abruptly at the pyrophosphate binding site where each chain turns sharply at a glycine residue (Gly54 and Gly155, respectively) to begin the following helices, α B and α E, respectively (Figure 5(a)). The pyrophosphate bridge is extensively hydrogen bonded to backbone amino groups of the N termini of the alpha helices F and E (residues 55–57 and 155–159, respectively). In addition to its contacts with the protein backbone, the dinucleotide pyrophosphate also makes a hydrogen bond with the side-chain of Thr57. Lys60 and the amino group of Gly157 are linked to the phosphates by bridging solvent molecules.

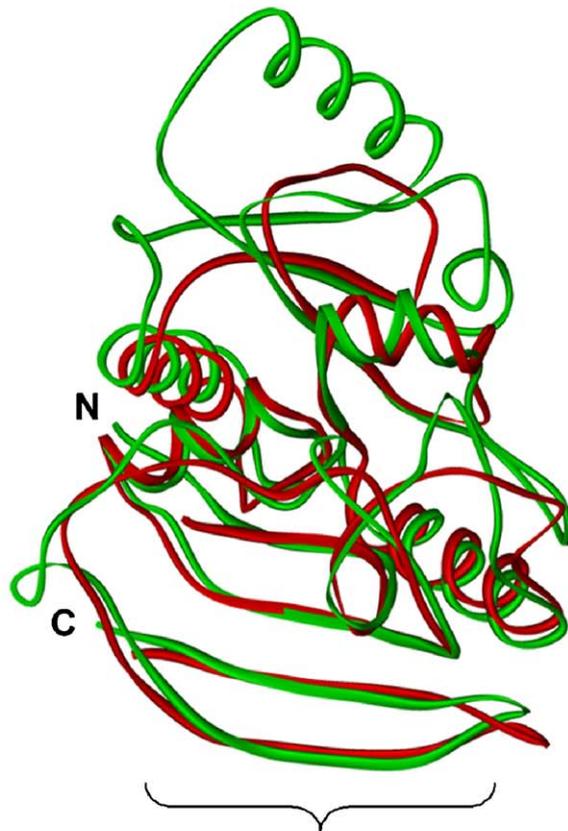
The basic amino acid residue at the position of Arg46 (TMDHFR, strictly conserved in the DHFR family) is not present in MjaRED (Figure 6(a)); it is replaced by a residue that does not participate in the binding. Although the pyrophosphate of the cofactor is shifted approximately 1.5 Å, the binding site in MjaRED is qualitatively similar to that found in the DHFRs.

Both ribosyl groups of the cofactor occupy relatively solvent-exposed regions of the binding cleft. Binding of the NMN-ribose is mediated by hydrogen bonds to the side-chain oxygen atoms of Leu23 and Asp61 and indirectly by bridging solvent, to the side-chains of Gly157 and Lys60 (Figure 7). Whereas this part of the cofactor occupies a similar position in the TMDHFR and MjaRED, the strictly invariant residue Ile16 (TMDHFR) of the DHFR family members is replaced by a Leu23. Asp136 (TMDHFR), conserved in most of the DHFR family members and responsible for tight binding, is missing in MjaRED and other members of this reductase family. Thus, the binding of the AMN-ribose appears to be much weaker and less specific in MjaRED compared to DHFR. The AMN 2'-phosphate is fixed by four hydrogen bonds and salt-bridges to protein side-chains (Arg86, Lys84 and Ser83). Together with the pyrophosphate bridge, these interactions mainly determine the position and orientation of AMN-ribose.

The nicotinamide-binding pocket is formed by residues of the two central strands of the pleated sheet, β A and β F, the alpha helix α E and the convoluted loop following β A (residues 18–35) (Figure 7). The conformation of this loop is stabilized by hydrogen bonds within the loop and by hydrogen bonds with a second loop following β G. The backbone of Gly16 contributes hydrogen bonds with the carboxamide and amino group of the nicotinamide ring. The carboxamide is also hydrogen bonded by the backbone of Leu23. These interactions determine the position of the nicotinamide carboxamide. In comparison to DHFR, the cleft around the nicotinamide ring is much wider because Trp23 (TMDHFR) and Phe107 (TMDHFR) are replaced by Ser30 and Asn160, respectively. The temperature factors for the cofactor are significantly increased by comparison with the protein (Table 2),



(b)



protein-protein interaction of TMDHFR dimer

Figure 6. (a) Structure-based sequence alignment of MjaRED and TMDHFR (PDB entry code 1D1G³³). Secondary structure elements (α -helices, rods; β -strands, arrows) corresponding to MjaRED are drawn above the sequences, those corresponding to TMDHFR are drawn below. (b) Structural alignment of MjaRED (green) and TMDHFR (red; PDB entry code 1D1G³³).

most likely due to a lower occupancy. We conclude from our studies that although the secondary structure elements in the superimposed structure of DHFR and MjaRED share a partly common fold (Figure 6), the mode of binding is different. Additionally, the orientation of the adenine ring

relative to the AMN-ribose, which is, to the best of our knowledge, in *anti*-conformation in all known structures of the EC1 group, is rotated approximately 180° around the bond connecting the AMN-ribose to the adenine ring and is therefore present in *syn*.

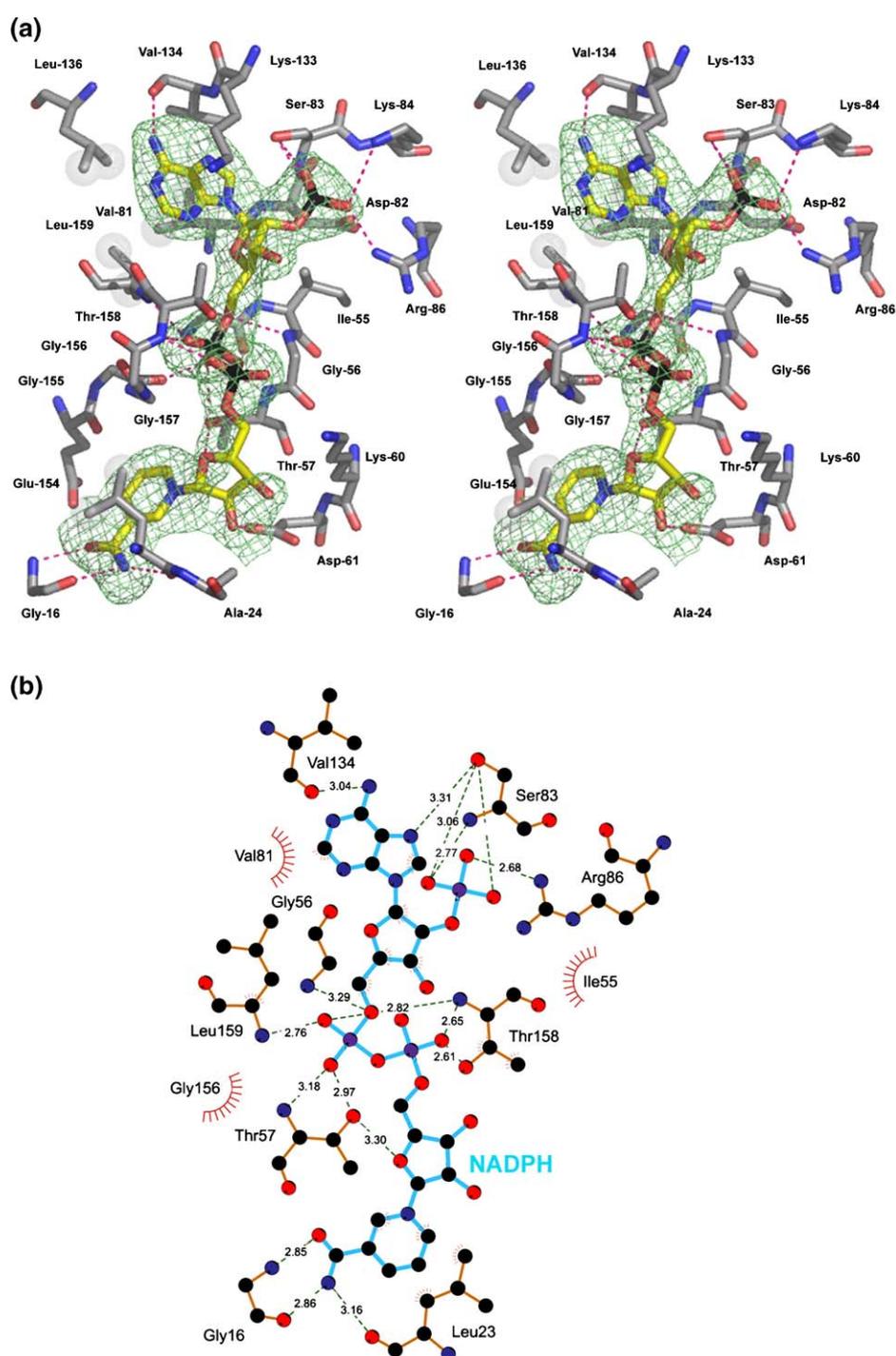


Figure 7. The cofactor-binding site of MjaRED. (a) Residues of the protein were selected that are located within a 4 Å sphere around the NADP cofactor. The initial $F_o - F_c$ difference electron density map covering the cofactor is contoured at 2.5σ . The map was calculated without the cofactor molecule. Possible hydrogen bonds and hydrophobic interactions are represented by broken lines and small spheres, respectively. (b) Schematic diagram of MjaRED interactions with the cofactor. Hydrogen bonds are presented as broken lines. Distances are given in angstroms. Spheres indicate hydrophobic contacts between the cofactor and the surrounding residues.

Active site architecture

Despite the lack of significant sequence identity and the lack of relation in substrate type and substrate procession, the crystal structure shows

that MjaRED resembles the typical structural features of the oxidoreductase family of dihydrofolate reductases.

The shape of the active site is a cavity of approximately $9 \text{ \AA} \times 14 \text{ \AA} \times 18 \text{ \AA}$, which is fairly

large for the substrate. The cavity is built up by the beta sheets A, B and F, whereas the walls comprise helix A and two loops (Figure 5).

Unfortunately, the substrate 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (DAROPP, Figure 1, 3) could not be used for cocrystallization or soaking experiments due to its insufficient stability and inhibitors were not available. To approach this problem, we generated a model of the substrate DAROPP using the Dundee PRODRG Server and modelled the substrate into the active site of the protein with the automated docking program AutoDock³⁶ (Figure 8). Ten trials were performed and the best solution was picked to perform the geometrical energy minimization calculation.

Based on our model, the polar groups of the pyrimidinone ring of DAROPP form several hydrogen bonds to main-chain and side-chain atoms of the protein (Figure 8). The side-chain amide of Asn14 interacts with the position 5 amino group of the pyrimidine ring of DAROPP (3.14 Å). Further interactions are observed between the side-chain carboxyl group of Glu154 with the position 5 group (2.44 Å) and with N-1' (4.11 Å) of DAROPP and with the N-atoms of the guanidine group of Arg45 (N^δ, 2.8 Å; N^ε, 2.8 Å). The backbone oxygen of Glu154 interacts with OH-2' (3.2 Å). Arg45 N^ε also interacts with the OH-3' (3.03 Å). Further interactions are observed between the N-1 and N-3 of DAROPP and the hydroxyl group of Ser33 (3.61 and 4.25 Å, respectively).

Arg45 side-chain amino groups (N^ε) interact with the bridging phosphate oxygen (3.33 Å) as well as with the terminal phosphate oxygen atoms (2.73–3.52 Å). These phosphate oxygen atoms are also interacting with backbone amino groups of Arg64 (2.86 and 4.00 Å, respectively) and Leu65 (2.75 and 3.69 Å). Another interaction is observed between the imidazole nitrogen of His68 and one of the terminal phosphate oxygen atoms of DAROPP (4.24 Å). Notably, binding of the phosphorylated substrate is essential for the catalytic activity of the enzyme; riboflavin type reductases are not able to utilize the dephosphorylated form of the substrate.³⁷

Arg45 and Glu154 residues seem to be essential for all reductases of the riboflavin pathway. In addition, Ser33 is also conserved, either a serine or a threonine residue is found at this position (Figure 9).

Further interactions are seen between the carboxamide oxygen of the nicotinamide ring and the position 5 amino amino group (4.09 Å), C-6 (3.42 Å), and N-1' (2.59 Å) of the pyrimidine ring of the substrate.

The amino acid residue Asp37 is conserved over all archaeal reductases of the riboflavin pathway known till now. In all known RIB7 proteins of fungi, which catalyze the same type of reduction, a highly conserved glutamic acid is found at this position. In our model the closest distance between the side-chain carboxyl group and the N-3 of the

pyrimidine ring of the substrate is 3.89 Å. In contrast, this residue is not found in eubacterial type reductases, e.g. the RibG protein of *B. subtilis* (Figure 9). Interestingly, this residue superimposes very well with Asp28 (TMDHFR) which is strictly conserved and mechanistically crucial in all DHFR family members (Figure 10(c)).

Oligomerization

As described above, the Dali algorithm³² identified the dimeric DHFR from *T. maritima* as the most similar 3D-structure (Z score of 16.6). As stabilization by oligomerization has been observed for a number of hyperthermophilic proteins it was presumed that oligomerization could be a factor for the high thermal stability.³³ In contrast to the homodimeric solution state of native MjaRED (Figure 3) we found the enzyme to be pseudo-trimeric in the crystal structure (Figure 4).

Despite the low sequence identity of 22% between the TMDHFR and MjaRED and the extra secondary structure elements of MjaRED (α C and β E; Figure 5 (a) and (b)), both proteins show a remarkable degree of homology within the dimerization site of TMDHFR (Figure 6(b)). Moreover, the calculation of the electrostatic surface potential with GRASP showed that the surface of MjaRED in this part of the protein is mainly of hydrophobic nature. Thus we used the program FTDock³⁸ to model the initial complex and subsequently refined the solution with MultiDock.³⁹ Although we did not impose any restraints on the software, the program suite provided a solution that was in good agreement with the structure of TMDHFR (Supplementary Data Figure 1S).

Discussion

The conversion of 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (Figure 1, 3) into 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate (Figure 1, 6) had been shown earlier to proceed *via* different intermediates in fungi and bacteria.^{14,37,40} In yeasts and in Archaea, the reaction sequence is initiated by the reductive conversion of the ribosyl side-chain of 3 into the ribityl side-chain of 2,5-diamino-6-ribitylamino-4(3*H*)-pyrimidinone 5'-phosphate (5); subsequent deamination of the pyrimidine moiety yields 6. In yeasts, the reactions are catalyzed by independent proteins; in Archaea the enzyme catalyzing the deamination step (G) is still unknown. In bacteria, on the other hand, the reaction sequence begins with the deamination of 3 yielding 5-amino-6-ribosylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate (3). That reaction and the subsequent side-chain reduction of 4 are catalyzed by bifunctional fusion proteins consisting of 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate deaminase (D) and 5-amino-6-ribosylamino-2,4(1*H*,3*H*)-

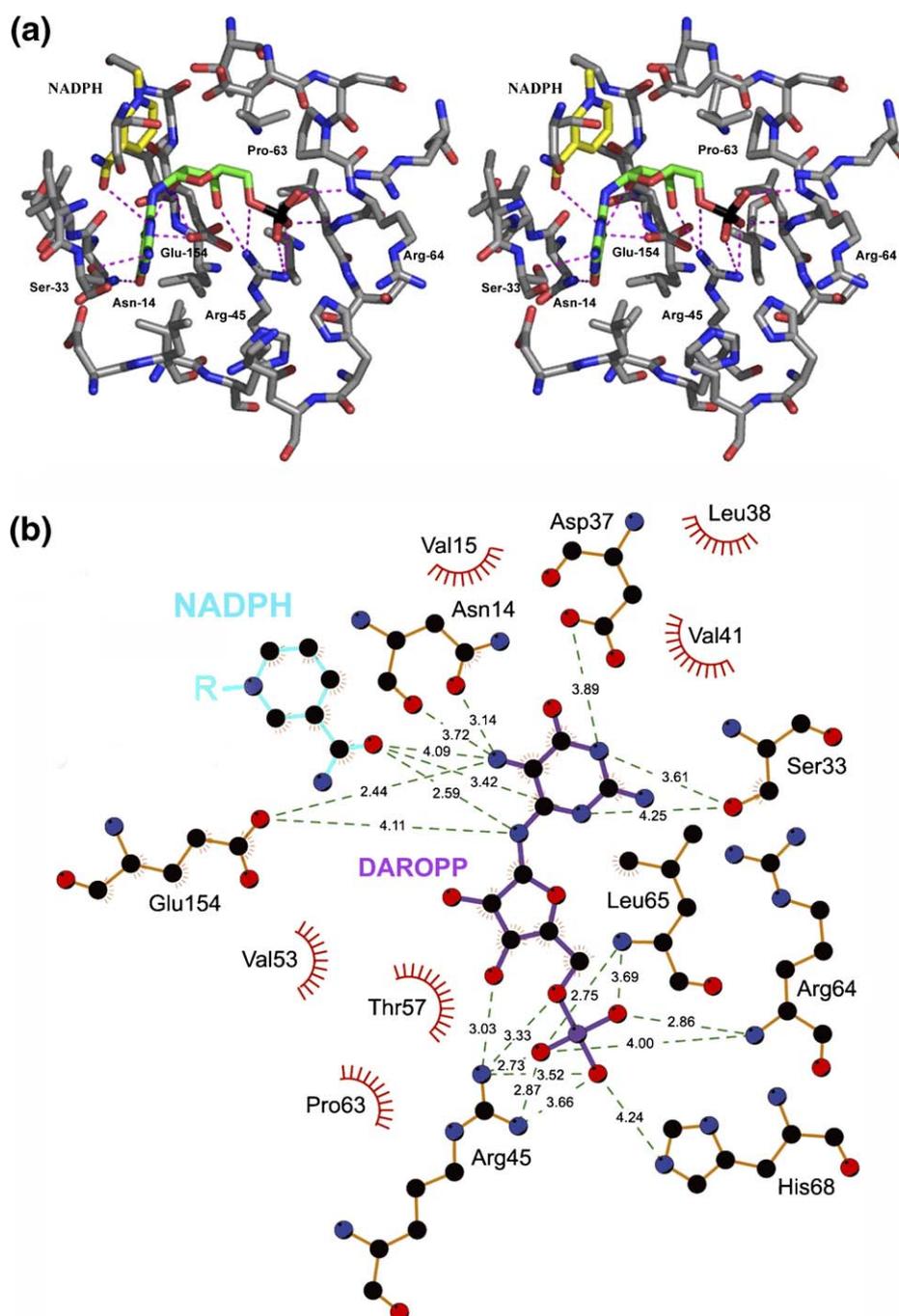


Figure 8. (a) Stereo view of the active site of MjaRED with the modelled substrate. The image shows all residues that are located within a 4 Å sphere around the modelled substrate. Carbon atoms of the protein and of the substrate are colored in gray and green, respectively. Carbon atoms of the nicotinamide moiety of the NADP cofactor are colored in yellow. Possible hydrogen bond interactions are represented by broken lines. (b) Schematic diagram of MjaRED interactions with the substrate. Hydrogen bonds are presented as broken lines. Distances are given in angstroms. Spheres indicate hydrophobic contacts between the cofactor and the surrounding residues.

pyrimidinedione 5'-phosphate reductase (E) domains in the majority of completely sequenced eubacteria.¹⁵

Previously, it could be shown by *in vivo* studies using the Ascomycete *A. gossypii* that the biosynthesis of riboflavin involves the introduction of a hydrogen atom derived from NADPH into the 1' position of the ribityl side-chain.^{10,41} On the other

hand, no detailed information has been provided for biosynthesis of riboflavin in Archaea.

Mechanistically, it is believed that the substrate of 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate reductases is reduced in its imine form⁴² (Figure 10(b)).

C-1' of the modelled MjaRED substrate occupies approximately the same position as the reactive C-

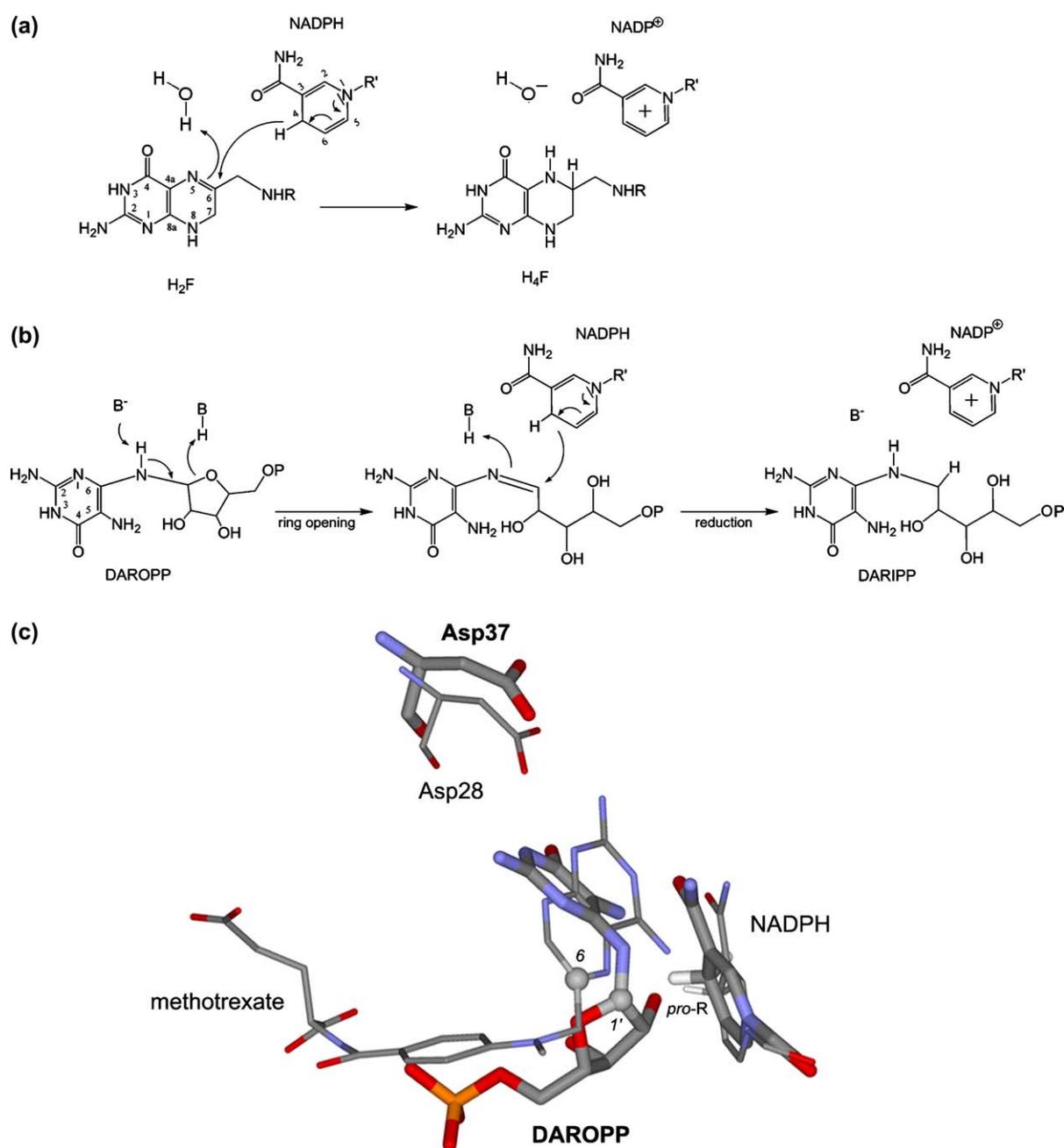


Figure 10. Reaction mechanisms and stereochemistry. (a) Reaction catalyzed by DHFR. The educt is H_2F (dihydrofolate), the product is H_4F (tetrahydrofolate), and the cofactor is NADPH/NADP⁺.⁷⁴ (b) Proposed reaction mechanism for reduction of the ribosylamino to the ribitylamino moiety in riboflavin biosynthesis. The educt is DAROPP (2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate), the product is DARIPP (2,5-diamino-6-ribitylamino-4(3H)-pyrimidinone 5'-phosphate), and the cofactor is NADPH/NADP⁺.⁴² (c) Structural alignment of MjaRED (thick bonds) and TMDHFR (PDB entry code 1D1G,³³ thin bonds) cofactor (NADPH) and substrate (DAROPP) or substrate analog molecule (methotrexate), respectively. The well aligning residues Asp28 (TMDHFR) and Asp37 (MjaRED) are drawn in the same manner. The *pro-R* hydrogen atoms at C-4 of both NADPH molecules are highlighted.

molecule. Although no water molecule could be observed in the X-ray structure representing the Michaelis complex of the *E. coli* enzyme, energy minimization calculations revealed that water could be placed in hydrogen bonding distance of N-5 with only minor conformational changes in DHFR.⁴⁹

No ionizable group in the vicinity of N-1' of the modelled DAROPP could be found in the proposed active site of MjaRED. In analogy to the DHFR reaction, a water molecule could also act as a direct proton donor of N-1'. Structural arguments like the similar architecture of the active site, the similar orientation and position of the reactive molecules

and the same stereospecificity of the hydride transfer suggest that the mechanism of riboflavin type reductases is rather similar to DHFR.

Materials and Methods

Materials

Restriction enzymes were obtained from New England Biolabs (Schwalbach, Germany). T4 DNA ligase was from Gibco BRL (Karlsruhe, Germany). EXT DNA polymerase and Taq polymerase were from Finnzymes (Epsø, Finland). Oligonucleotides were synthesized by Thermo Electron GmbH (Ulm, Germany). A Plasmid Miniprep Kit from PEQLab (Erlangen, Germany) was used for plasmid DNA isolation and purification. DNA fragments and PCR amplicates were purified with a Gel Extraction Kit or Cycle Pure Kit from PEQLab. Casein hydrolyzate and yeast extract were from Gibco BRL, and isopropyl β -D-thiogalactopyranoside was from Biomol (Hamburg, Germany). The detergent Cymal-2 was obtained from Hampton Research (Aliso Viejo, CA).

Strains and plasmids

Bacterial strains and plasmids used in this study are summarized in Table 3. Bacteria were grown at 37 °C in LB medium containing 170 mg/l ampicillin and 15 mg/l kanamycin where appropriate.

Preparation of a synthetic gene

The partially complementary oligonucleotides MjaRED-1 and MjaRED-2 were annealed and treated with DNA polymerase. The resulting 111 bp segment was elongated by a series of seven PCR amplifications using pairwise combinations of oligonucleotides according to Table 4 (Figure 2). The resulting 721 bp DNA fragment was digested with EcoRI and HindIII and ligated into the plasmid pNCO113, which had been treated with the same restriction endonucleases affording the expression plasmid pNCO-MjaRED-syn (Table 3).

Transformation of bacterial cells

Ligation mixtures were transformed into *E. coli* XL1-Blue cells. Transformants were selected on LB solid

Table 3. Bacterial strains and plasmids used in this study

Strain/plasmid	Relevant characteristics	Source
<i>E. coli</i> strains XL1-Blue	<i>recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac[F', proAB, lacI^qZΔM15, Tn10(tet^r)]</i>	71
M15[pREP4]	<i>lac, ara, gal, mtl, recA⁺, uvr⁺, Str^R, (pREP4: Kan^R, lacI)</i>	72
Expression plasmids pNCO113	<i>E. coli</i> expression vector	72
pNCO-MjaRED-syn	Gene MJ0671 with optimized codon usage for <i>E. coli</i>	This study

medium supplemented with ampicillin (170 mg/l). The plasmids were re-isolated and analyzed by restriction analysis and by DNA sequencing. The expression plasmid was then transformed into *E. coli* M15 [pREP4] cells carrying the pREP4 repressor plasmid for the overexpression 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-MjaRED-syn was grown in LB medium containing ampicillin (170 mg/l) and kanamycin (15 mg/l) at 37 °C with shaking overnight. Erlenmeyer flasks containing 500 ml of medium were then inoculated at a ratio of 1:50 and were incubated at 37 °C with shaking. At an absorbance of 0.6 (600 nm), isopropyl- β -thiogalactoside was added to a final concentration of 2 mM, and incubation was continued for 4 h. Cells were harvested by centrifugation and stored at -20 °C. Frozen cell mass (5 g) was thawed in 20 mM potassium phosphate containing 2 mM dithiothreitol (DTT) (pH 7). The suspension was ultrasonically treated and centrifuged. The supernatant was brought to 70 °C. After 5 min, the mixture was cooled to 10 °C and centrifuged (15,000 rpm (Sorvall SS34), 20 min). The supernatant was placed on top of a HA Macrorep 45 μ m column (45 ml, Amersham Biosciences), which had been equilibrated with 20 mM potassium phosphate (pH 7.0). The column was developed with a gradient from 20 mM to 1 M potassium phosphate (pH 7.0). Fractions were combined and concentrated by ultrafiltration. The supernatant was placed on top of a Superdex-200 column (26 cm \times 60 cm, Amersham Biosciences), which was then developed with 20 mM Tris hydrochloride (pH 7.8), containing 100 mM potassium chloride and 5 mM dithiothreitol. Fractions were combined and concentrated by ultrafiltration using Amicon 10 kDa membranes.

Analytical ultracentrifugation

Experiments were performed with an analytical ultracentrifuge Optima XL-A from Beckman Instruments (Palo Alto, CA) equipped with absorbance optics. Aluminum double sector cells equipped with quartz windows were used throughout. Sedimentation equilibrium experiments were performed with solutions containing 100 mM potassium phosphate (pH 5.0), 300 mM sodium chloride and 0.4 mg of protein per ml at 10,000 rpm (Beckman AN-60Ti) and 4 °C. Boundary sedimentation experiments were performed at 59,000 rpm and 20 °C using a solution containing 50 mM Tris hydrochloride (pH 7.0), 100 mM sodium chloride, 5 mM dithiothreitol and 3.0 mg of protein per ml. The partial specific volume was estimated from the amino acid composition yielding a value of 0.7414 ml/g.⁵⁰

Assay of reductase activity

Assay mixtures contained 100 mM Tris hydrochloride (pH 8.2), 10 mM MgCl₂, 10 mM DTT, 10% ²H₂O, 5 mM NADH, 0.5 mM ATP, 5 mM PEP and 5 mM [¹³C]GTP. The reductase substrate was generated by pre-incubation of this mixture with GTP cyclohydrolase II, guanylate kinase and pyruvate kinase at 37 °C. The last two enzymes were used to recycle GMP, which is a byproduct of GTP cyclohydrolase II. The substrate mixture was controlled by ¹³C NMR spectroscopy. After initiation of the reaction by

Table 4. Oligonucleotides used for plasmid construction

Designation	Sequence (5' to 3') ^a
MJRED-1	<u>g</u> cgcgcttcaataaagatgctaaaactattatcgcgacaacagaagactaatgaagagaagaaag
MJRED-2	aactactcaacaccatactccaggattttgattttctttcttcttcaatgatctctctgtg
MJRED-3	acgtatcgtttgattcgaagctccggttccactgaatgcgcgcttcaataaagatgctaaaac
MJRED-4	atccatcagttttaaagatctactaccacgaccatttaactactcaacaccatactctc
MJRED-5	ccagcttcaagtcataaagattaaaagcagatcgtaactctgtagtattggttgattcgaagctc
MJRED-6	cgccacctcaagcaggatcgattgatacctttatcatagaggatccatcagttttaaagatctac
MJRED-7	ggtattatgggttgattggtacggttcttaaggacgatccagcttcaagtcataaagattaaaag
MJRED-8	catcaacaaggcctctttaaacaatcccagttcagagttccgacacccaagcaggatcgattg
MJRED-9	gcgaagaggatctacccgtacataagattcgtgtaagttagatggtattggttgattggtacc
MJRED-10	gggctctttaccaccaaatattttggagcgatagacggagacctcaacaaggcctctttaaac
MJRED-11	gatggtagctagctactatcaacaatgattcgcgtattcatgcaagaggatctaccggtgac
MJRED-12	aacacactctactgtttgaagcctcaccactacatagttgggctcttaccaccaaatattttg
MJRED-13	aatggaaaaaagccatacattatctcaaacgttggcatgacctgtagtgtaagctagctactcaac
MJRED-14	gcctcactaggcgatagaagttttcagttccagtttaacacactctactgtttgaag
MJRED-15	ataatagaattcattaagaggagaaattaacatggtgatgtaagtgaaaaaagccatacattac
MJRED-16	tattattatagcttattctttaccftaaattccagtaaatgcctcactaggcgatagaag

^a Recognition sites used for cloning are underlined.

addition of reductase ¹³C NMR spectra were recorded at intervals of 10 min.

¹³C spectra were acquired with a DRX 500 spectrometer from Bruker Instruments, Karlsruhe, Germany, at transmitter frequencies of 500.13 and 125.76 MHz, respectively.

Preparation of a selenomethionine protein

Recombinant SeMet-protein was produced with *E. coli* strain M15 [pREP4] containing the vector pNCO-MjaRED-syn grown in New Minimal Medium⁵¹ as described.⁵² Mass spectrometry confirmed an average replacement of four to five Met positions by SeMet of the possible seven (data not shown).

Crystallization, crystal transformation and structure solution

Both native and SeMet-substituted proteins were crystallized at 18 °C using the sitting-drop vapour diffusion method. Equal volumes of protein solution (10 mg/ml) in Tris hydrochloride (pH 7.8) containing 100 mM KCl, 5 mM DTT and 12 mM Cymal-2 were mixed with a reservoir solution containing 100 mM Tris hydrochloride (pH 7.5), and 1.1 M sodium citrate. Tetrakis-hexagonal-shaped crystals appeared within a week and were shown to belong to space group $P4_12_12$ or $P4_32_12$ with cell parameters $a=b=136.9$ Å, $c=213.7$ Å, suggesting six monomers per asymmetric unit. Native and SeMet-substituted crystals diffracted to 2.5 Å. Data were collected on beamline BW6 at the DESY (Hamburg, Germany) and on beamline ID14-4 at the ESRF (Grenoble, France). Data were processed using DENZO and SCALEPACK.⁵³ Prior to flash-cooling, crystals were transferred into 6 µl per-fluoropolyether and immediately frozen in the cryostream. Data were processed using DENZO and SCALEPACK.⁵³ Unfortunately, native and SeMet crystals exhibited pseudo-body centering, which prevented structure solution by either multiple isomorphous replacement (MIR) or multiple wavelength anomalous dispersion (MAD) methods.

Transformation of the crystals with a free-mounting system (Kiefersauer, Proteros Biostructures, Germany) resulted in crystals that appeared to be body-centered with space group $I4_122$ or $I422$. Before transformation, these crystals had been incubated for 24 h in a reservoir solution containing 10 mM mersalyl acid (2-[N-(3-hydro-

xymercuri-2-methoxypropyl)-carbonyl]-phenoxyacetic acid). A complete MAD experiment of the mercury substituted crystals was performed at beamline BW6 at the DESY (Hamburg, Germany) and seven mercury sites per asymmetric unit were identified with SHELXD using space group $I4_122$. No solution was found in $I422$. The resulting electron density was readily interpretable for two molecules of the expected three in the asymmetric unit. The third molecule showed only diffuse density, but could be reconstructed by application of a 3-fold symmetry operator. The subsequent refinement procedure halted at an R_{free} of around 39.2% (Table 1). An anomalous difference Fourier map of the SeMet data set calculated at this stage of the refinement showed 20 selenium sites of the expected 42. A re-evaluation of the data in space group $P4_32_12$ revealed that the body-centering is obviously not complete as the ratio of intensities of odd and even $h+k+l$ reflections amounts to about 10% for higher resolution reflections. Fortunately, we were now able to solve the structure of the non-transformed SeMet in space group $P4_32_12$ crystals by molecular replacement using the program PHASER⁵⁴ and the MjaRED trimer as a search model. Other maximal non-isomorphous sub-groups for space group $I4_122$ gave no solution. PHASER placed two MjaRED trimers into the asymmetric unit. After rigid-body refinement of the individual monomers, we located the selenium positions by anomalous difference Fourier techniques. Refinement of heavy-atom parameters and phase calculation was done with SHARP.⁵⁵ The resulting electron density map was modified and improved by solvent flattening with SHARP. Table 1 gives a summary of the data collection and phasing statistics.

Model building and refinement

Model building was carried out with O.⁵⁶ Energy-restrained crystallographic refinement of MjaRED (SeMet) in space group $P4_32_12$ was carried out with maximum likelihood algorithms implemented in CNS,⁵⁷ using the protein parameters of Engh and Huber. The initial model was subjected to a rigid body and positional refinement. Bulk solvent, overall anisotropic B -factor corrections and non-crystallographic restraints were introduced depending on the behavior of the R_{free} index. Refinement proceeded in several cycles, which were interrupted for manual rebuilding. All refinements were performed using the maximum-likelihood target with Hendrickson-Lattman coefficients.

After the addition of solvent molecules, the refinement converged at an R -factor of 23.0% (R_{free} of 25.3%, Table 2).

Graphical representation

Graphical representations were prepared using the programs MOLSCRIPT,⁵⁸ BOBSCRIPT,⁵⁹ RASTER3D,⁶⁰ GRASP,⁶¹ ALSRIPT⁶² and Pymol.⁶³ Secondary structure elements were assigned with STRIDE.⁶⁴ Final Ramachandran plot analysis with the program PROCHECK³⁰ resulted in 88.7% residues in the most favored regions, 10.8% residues in the additionally allowed regions, 0.5% residues in the generously allowed regions and no residues in the disallowed regions. The program LSQMAN⁶⁵ was used for the superposition of the different 3D structures.

Molecular modelling

The Dundee PRODRG Server⁶⁶ was used to generate a molecular model of the substrate 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (Figure 1, 3). Automated docking simulations were conducted using the AutoDock 3.06 program suite.³⁶ The calculations were performed using the Lamarckian Genetic Algorithm (LGA) as described in the corresponding manual. The modelling Program FTDock was used for dimer interaction studies. A final monomer model was duplicated and the starting position of the second monomer was randomized. Additional preparations were performed following the FTDock manual.

Miscellaneous

DNA was sequenced using the method described by Sanger⁶⁷ by the custom sequencing service of GATC Biotech (Konstanz, Germany). N-terminal protein sequencing was performed by the automated Edman method using a 471A Protein Sequencer (Perkin-Elmer). Protein concentration was determined by published procedures.⁶⁸ Electrospray mass spectrometry experiments were performed as described by Mann and Wilm⁶⁹ using a triple quadrupole ion spray mass spectrometer API365 (SciEx, Thornhill, ON, Canada). SDS polyacrylamide gel electrophoresis using 16% (w/v) polyacrylamide gels was performed as described by Laemmli with molecular mass standards provided by Sigma.⁷⁰

Protein Data Bank accession code

The coordinates were deposited at the RCSB Protein Data Bank under the accession number 2AZN.

Acknowledgements

We thank Richard Feicht for expert help with protein preparation. This work was supported by the Fonds der Chemischen Industrie and the Hans-Fischer-Gesellschaft e.V. We also thank G. P. Bourenkov and G. Kachalova for assistance during data collection at beamline BW6 (DESY, Hamburg) and J. McCarthy and G. Leonard for support during measurements at beamline ID14-4 (ESRF, Grenoble).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.04.045

References

- Imada, Y., Iida, H., Ono, S. & Murahashi, S. (2003). Flavin catalyzed oxidations of sulfides and amines with molecular oxygen. *J. Am. Chem. Soc.* **125**, 2868–2869.
- Sancar, A. (1994). Structure and function of DNA photolyase. *Biochemistry*, **33**, 2–9.
- Briggs, W. R., Beck, C. F., Cashmore, A. R., Christie, J.M., Hughes, J. Jarillo, J. A. *et al.* (2001). The phototropin family of photoreceptors. *Plant Cell*, **13**, 993–997.
- Lee, J. (1993). Lumazine protein and the excitation mechanism in bacterial bioluminescence. *Biophys. Chem.* **48**, 149–158.
- Meighen, E. A. (1993). Bacterial bioluminescence: organization, regulation, and application of the lux genes. *FASEB J.* **7**, 1016–1022.
- Sancar, A. (2000). Cryptochrome: The second photoactive pigment in the eye and its role in circadian photoreception. *Annu. Rev. Biochem.* **69**, 31–67.
- Sancar, A. (2004). Regulation of the mammalian circadian clock by cryptochrome. *J. Biol. Chem.* **279**, 34079–34082.
- Fischer, M. & Bacher, A. (2005). Biosynthesis of flavocoenzymes. *Nat. Prod. Rep.* **22**, 324–350.
- Fischer, M. & Bacher, A. (2006). Biosynthesis of vitamin B2 in plants. *Physiol. Plant.* **126**, 304–318.
- Bacher, A. & Lingens, F. (1968). Accumulation of 2,5-diamino-6-hydroxy-4-(ribitylamino)pyrimidine in a riboflavine-deficient mutant of *Saccharomyces cerevisiae*. *Angew. Chem., Int. Ed. Engl.* **7**, 219–220.
- Bacher, A. & Lingens, F. (1969). Structure of the purine precursor in riboflavine biosynthesis. *Angew. Chem., Int. Ed. Engl.* **8**, 371–372.
- Hollander, I. & Brown, G. M. (1979). Biosynthesis of riboflavin: Reductase and deaminase of *Ashbya gossypii*. *Biochem. Biophys. Res. Commun.* **89**, 759–763.
- Klein, G. & Bacher, A. (1980). Biosynthesis of riboflavin. Enzymatic deamination of 2,5-diamino-6-hydroxy-4-ribitylamino-pyrimidine 5'-phosphate. *Z. Naturforsch.* **35b**, 482–484.
- Nielsen, P. & Bacher, A. (1981). Biosynthesis of riboflavin. Characterization of the product of the deaminase. *Biochim. Biophys. Acta*, **662**, 312–317.
- Richter, G., Fischer, M., Krieger, C., Eberhardt, S., Lüttgen, H., Gerstenschläger, I. & Bacher, A. (1997). Biosynthesis of riboflavin: characterization of the bifunctional deaminase-reductase of *Escherichia coli* and *Bacillus subtilis*. *J. Bacteriol.* **179**, 2022–2028.
- Fischer, M., Römisch, W., Saller, S., Illarionov, B., Richter, G., Rohdich, F. *et al.* (2004). Evolution of vitamin B2 biosynthesis: structural and functional similarity between pyrimidine deaminases of eubacterial and plant origin. *J. Biol. Chem.* **279**, 36299–36308.
- Eisenreich, W., Schwarzkopf, B. & Bacher, A. (1991). Biosynthesis of nucleotides, flavins, and deazaflavins in *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* **266**, 9622–9631.
- Reuke, B., Korn, S., Eisenreich, W. & Bacher, A. (1992).

- Biosynthetic precursors of deazaflavins. *J. Bacteriol.* **174**, 4042–4049.
19. Graham, D. E., Xu, H. & White, R. H. (2002). A member of a new class of GTP cyclohydrolases produces formylaminopyrimidine nucleotide monophosphates. *Biochemistry*, **41**, 15074–15084.
 20. Graupner, M., Xu, H. & White, R. H. (2002). The pyrimidine nucleotide reductase step in riboflavin and F(420) biosynthesis in archaea proceeds by the eukaryotic route to riboflavin. *J. Bacteriol.* **184**, 1952–1957.
 21. Fischer, M., Römisch, W., Schiffmann, S., Kelly, M., Oschkinat, H. Steinbacher, S. *et al.* (2002). Biosynthesis of riboflavin in archaea. Studies on the mechanism of 3,4-dihydroxy-2-butanone-4-phosphate synthase of *Methanococcus jannaschii*. *J. Biol. Chem.* **277**, 41410–41416.
 22. Steinbacher, S., Schiffmann, S., Bacher, A. & Fischer, M. (2004). Metal sites in 3,4-dihydroxy-2-butanone 4-phosphate synthase from *Methanococcus jannaschii* in complex with the substrate ribulose 5-phosphate. *Acta Crystallog. sect. D*, **60**, 1338–1340.
 23. Steinbacher, S., Schiffmann, S., Richter, G., Huber, R., Bacher, A. & Fischer, M. (2003). Structure of 3,4-dihydroxy-2-butanone 4-phosphate synthase from *Methanococcus jannaschii* in complex with divalent metal ions and the substrate ribulose 5-phosphate: implications for the catalytic mechanism. *J. Biol. Chem.* **278**, 42256–42265.
 24. Haase, I., Mörtl, S., Köhler, P., Bacher, A. & Fischer, M. (2003). Biosynthesis of riboflavin in archaea. 6,7-dimethyl-8-ribityllumazine synthase of *Methanococcus jannaschii*. *Eur. J. Biochem.* **270**, 1025–1032.
 25. Eberhardt, S., Korn, S., Lottspeich, F. & Bacher, A. (1997). Biosynthesis of riboflavin: An unusual riboflavin synthase of *Methanobacterium thermoautotrophicum*. *J. Bacteriol.* **179**, 2938–2943.
 26. Fischer, M., Schott, A. K., Römisch, W., Ramsperger, A., Augustin, M., Fidler, A. *et al.* (2004). Evolution of vitamin B2 biosynthesis. A novel class of riboflavin synthase in Archaea. *J. Mol. Biol.* **343**, 267–278.
 27. Illarionov, B., Eisenreich, W., Schramek, N., Bacher, A. & Fischer, M. (2005). Biosynthesis of vitamin B2. Diastereomeric reaction intermediates of archaeal and non-archaeal riboflavin synthases. *J. Biol. Chem.* **280**, 28541–28546.
 28. Ramsperger, A., Augustin, M., Schott, A. K., Gerhardt, S., Krojer, T. Eisenreich, W. *et al.* (2006). Crystal structure of an archaeal pentameric riboflavin synthase in complex with a substrate analog inhibitor: stereochemical implications. *J. Biol. Chem.* **281**, 1224–1232.
 29. Chen, S. C., Chang, Y. C., Lin, C. H. & Liaw, S. H. (2006). Crystal structure of a bifunctional deaminase and reductase from *Bacillus subtilis* involved in riboflavin biosynthesis. *J. Biol. Chem.* **281**, 7605–7613.
 30. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). Procheck- a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
 31. Levitt, M. & Chothia, C. (1976). Structural patterns in globular proteins. *Nature*, **261**, 552–558.
 32. Holm, L. & Sander, C. (1994). Searching protein-structure databases has come of age. *Proteins: Struct. Funct. Genet.* **19**, 165–173.
 33. Dams, T., Auerbach, G., Bader, G., Jacob, U., Ploom, T., Huber, R. & Jaenicke, R. (2000). The crystal structure of dihydrofolate reductase from *Thermotoga maritima*: molecular features of thermostability. *J. Mol. Biol.* **297**, 659–672.
 34. Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C. & Kraut, J. (1982). Crystal-structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate-reductase refined at 1.7 Å resolution 0.1. General features and binding of methotrexate. *J. Biol. Chem.* **257**, 3650–3662.
 35. Schnell, J. R., Dyson, H. J. & Wright, P. E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 119–140.
 36. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662.
 37. Burrows, R. B. & Brown, G. M. (1978). Presence in *Escherichia coli* of a deaminase and a reductase involved in biosynthesis of riboflavin. *J. Bacteriol.* **136**, 657–667.
 38. Gabb, H. A., Jackson, R. M. & Sternberg, M. J. E. (1997). Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* **272**, 106–120.
 39. Jackson, R. M., Gabb, H. A. & Sternberg, M. J. E. (1998). Rapid refinement of protein interfaces incorporating solvation: Application to the docking problem. *J. Mol. Biol.* **276**, 265–285.
 40. Bacher, A. & Lings, F. (1970). Biosynthesis of riboflavin. Formation of 2,5-diamino-6-hydroxy-4-(1'-D-ribitylamino)pyrimidine in a riboflavin auxotroph. *J. Biol. Chem.* **245**, 4647–4652.
 41. Bacher, A., Baur, R., Oltmanns, O. & Lings, F. (1969). Biosynthesis of riboflavin. Mutants accumulating 6-hydroxy-2,4,5-triaminopyrimidine. *FEBS Letters*, **5**, 316–318.
 42. Keller, P. J., Le Van, Q., Kim, S. U., Bown, D. H., Chen, H. C., Kohnle, A. *et al.* (1988). Biosynthesis of riboflavin: mechanism of formation of the ribitylamino linkage. *Biochemistry*, **27**, 1117–1120.
 43. Bystroff, C., Oatley, S. J. & Kraut, J. (1990). Crystal structures of *Escherichia coli* dihydrofolate reductase: the NADP+holoenzyme and the folate. NADP+ternary complex. Substrate binding and a model for the transition state. *Biochemistry*, **29**, 3263–3277.
 44. Charlton, P. A., Young, D. W., Birdsall, B., Feeney, J. & Roberts, C. K. (1985). Stereochemistry of reduction of the vitamin folic acid by dihydrofolate reductase. *J. Chem. Soc. Perkin Trans.* **1**, 1349–1353.
 45. Miller, G. P. & Benkovic, S. J. (1998). Strength of an interloop hydrogen bond determines the kinetic pathway in catalysis by *Escherichia coli* dihydrofolate reductase. *Biochemistry*, **37**, 6336–6342.
 46. Sawaya, M. R. & Kraut, J. (1997). Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: Crystallographic evidence. *Biochemistry*, **36**, 586–603.
 47. Bajorath, J., Kraut, J., Li, Z. Q., Kitson, D. H. & Hagler, A. T. (1991). Theoretical studies on the dihydrofolate reductase mechanism: Electronic polarization of bound substrates. *Proc. Natl Acad. Sci. USA*, **88**, 6423–6426.
 48. Bajorath, J., Kitson, D. H., Kraut, J. & Hagler, A. T. (1991). The electrostatic potential of *Escherichia coli* dihydrofolate reductase. *Proteins: Struct. Funct. Genet.* **11**, 1–12.
 49. Shrimpton, P. & Allemann, R. K. (2002). Role of water

- in the catalytic cycle of *Escherichia coli* dihydrofolate reductase. *Protein Sci.* **11**, 1442–1451.
50. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., & Horton, J. C., eds), pp. 90–125, Royal Society of Chemistry, Cambridge.
 51. Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur. J. Biochem.* **230**, 788–796.
 52. Budisa, N., Karnbrock, W., Steinbacher, S., Humm, A., Prade, L., Neufeind, T. *et al.* (1997). Bioincorporation of telluromethionine into proteins: a promising new approach for X-ray structure analysis of proteins. *J. Mol. Biol.* **270**, 616–623.
 53. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A*, vol. 276. (pp. 307–326) San Diego, CA: Academic Press.
 54. Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallog. sect. D*, **60**, 432–438.
 55. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In: *Macromolecular Crystallography, Pt A*, vol. 276. (pp. 472–494) Academic Press, San Diego, CA.
 56. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
 57. Brünger, A. T., Adams, P. D., Clore, G. M., De Lano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D* **54**, 905–921.
 58. Kraulis, P. J. (1991). Molscript- a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
 59. Esnouf, R. M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallog. sect. D*, **55**, 938–940.
 60. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. In: *Macromolecular Crystallography, Pt B*, vol. 277, pp. 505–524, Academic Press, San Diego.
 61. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). Grasp-graphical representation and analysis of surface-properties. *Biophys. J.* **64**, A166.
 62. Barton, G. J. (1993). Alscript- a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.
 63. Delano, W. L. (2002). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA.
 64. Frishman, D. & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Struct. Funct. Genet.* **23**, 566–579.
 65. Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallog. sect. D*, **50**, 178–185.
 66. Schuttelkopf, A. W. & van Aalten, D. M. F. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallog. sect. D*, **60**, 1355–1363.
 67. Sanger, F., Niklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
 68. Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
 69. Baur, S., Becker, J., Li, Z., Niegemann, E., Wehner, E., Wolter, R. & Brendel, M. (1995). Sequence analysis of a 5.6 kb fragment of chromosome II from *Saccharomyces cerevisiae* reveals two new open reading frames next to CDC28. *Yeast*, **11**, 681–689.
 70. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
 71. Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). XL-blue: a high efficiency plasmid transforming recA *Escherichia coli* strain with β -galactosidase selection. *BioTechniques*, **5**, 376–379.
 72. Stueber, D., Matile, H. & Garotta, G. (). System for high level production in *E. coli* and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies and structure function analysis. In: *Immunological Methods IV* (Lefkowitz, I & Pernis P., eds), pp.121–125, Academic Press, Orlando, FL.
 73. Shavlovskii, G. M., Logvinenko, E. M., Sibirnyi, A. A., Fedorovich, D. V. & Zakal'skii, A. E. (1981). Activity of the enzyme of the 2d step of flavinogenesis, 2,5-diamino-6-hydroxy-4-ribosylaminopyrimidine-5'-phosphate reductase, in *Pichia guilliermondii* yeasts. *Mikrobiologiya*, **50**, 1008–1011.
 74. Benkovic, S. J. & Hammes-Schiffer, S. (2003). A perspective on enzyme catalysis. *Science*, **301**, 1196–1202.

Edited by I. Wilson

(Received 21 February 2006; received in revised form 19 April 2006; accepted 20 April 2006)