



Available online at www.sciencedirect.com



# A Structure-based Model of the Reaction Catalyzed by Lumazine Synthase from *Aquifex aeolicus*

### Xiaofeng Zhang<sup>1,2</sup>, Winfried Meining<sup>1,2</sup>, Mark Cushman<sup>3</sup>, Ilka Haase<sup>4</sup> Markus Fischer<sup>4</sup>, Adelbert Bacher<sup>4</sup> and Rudolf Ladenstein<sup>1\*</sup>

<sup>1</sup>Karolinska Institute, NOVUM Centre for Structural Biochemistry, S-14157 Huddinge, Sweden

<sup>2</sup>Södertörns Högskola, S-14157 Huddinge, Sweden

<sup>3</sup>Department of Medicinal Chemistry and Molecular Pharmacology, School of Pharmacy and Pharmacal Sciences, Purdue University West Lafayette, IN 47907, USA

<sup>4</sup>Lehrstuhl für Organische Chemie und Biochemie Technische Universität München, Lichtenbergstr. 4 D-85747 Garching, Germany 6,7-Dimethyl-8-ribityllumazine is the biosynthetic precursor of riboflavin, which, as a coenzyme, plays a vital role in the electron transfer process for energy production in all cellular organisms. The enzymes involved in lumazine biosynthesis have been studied in considerable detail. However, the conclusive mechanism of the reaction catalyzed by lumazine synthase has remained unclear. Here, we report four crystal structures of the enzyme from the hyperthermophilic bacterium Aquifex aeolicus in complex with different inhibitor compounds. The structures were refined at resolutions of 1.72 Å, 1.85 Å, 2.05 Å and 2.2 Å, respectively. The inhibitors have been designed in order to mimic the substrate, the putative reaction intermediates and the final product. Structural comparisons of the native enzyme and the inhibitor complexes as well as the kinetic data of singlesite mutants of lumazine synthase from Bacillus subtilis showed that several highly conserved residues at the active site, namely Phe22, His88, Arg127, Lys135 and Glu138 are most likely involved in catalysis. A structural model of the catalytic process, which illustrates binding of substrates, enantiomer specificity, proton abstraction/donation, inorganic phosphate elimination, formation of the Schiff base and cyclization is proposed.

© 2003 Elsevier Science Ltd. All rights reserved

\*Corresponding author

*Keywords:* lumazine synthase; *Aquifex aeolicus;* inhibitor complex; vitamin biosynthesis; catalytic mechanism

#### Introduction

The penultimate step of the riboflavin biosynthesis in the hyperthermophilic bacterium *Aquifex aeolicus* is catalyzed by lumazine synthase (LS), which is an assembly of 60 subunits arranged in a capsid with T = 1 icosahedral symmetry. In the wild-type *Bacillus subtilis* enzyme, 60 LS subunits enclose a trimer of riboflavin synthase (RS), which catalyzes the dismutation of 6,7dimethyl-8-ribityllumazine (**3**) resulting in the formation of riboflavin (**4**) and 5-amino-6-ribityl-

E-mail address of the corresponding author: rudolf.ladenstein@csb.ki.se amino-2,4(1*H*,3*H*)pyrimidinedione (1),<sup>1</sup> (see Figure 1). (see

The three-dimensional structures of several lumazine synthases from bacteria, yeast and plants have been studied by X-ray crystallography.<sup>2–8</sup> However, the precise mechanism of the catalytic reaction has remained unclear. Most of the existing functional and kinetic data have been obtained from studies on lumazine synthase from *B. subtilis.*<sup>9</sup> Because of high structural homology (48% sequence identity), these results can be considered as valid for *A. aeolicus* LS (LSAQ), too, and will be reviewed in the following section.

The active site, which is located at the interface of two adjacent subunits in close proximity to the inner surface of the icosahedral capsid, exists in 60 copies with the same environment.<sup>2,7</sup> The enzymatic function of the LS had remained unknown for several decades. The chemical nature of the four-carbon precursor of the pyrazine ring of 6,7-dimethyl-8-ribityllumazine (**3**) was not understood until Volk & Bacher<sup>10,11</sup> identified this compound as

X.Z. & W.M. contributed equally to this work. Abbreviations used: LS, lumazine synthase; RS, riboflavin synthase; RNOP, 5-nitroso-6-ribityl-amino-2,4(1*H*,3*H*)pyrimidinedione; LSAQ, *A. aeolicus* LS; RPP, 5-(6-D-ribitylamino-2,4(1*H*,3*H*)pyrimidinedione-5-yl)-1pentyl-phosphonic acid; RDL, 7-dioxo-5*H*-8ribitylaminolumazine; RPL, 3-(7-hydroxy-8ribityllumazine-6-yl)propionic acid.



Figure 1. Biosynthesis of ribo-flavin.

(3*S*)-3,4-dihydroxy-2-butanone-4-phosphate (2; Figure 1), and established the enzyme-catalyzed formation of 6,7-dimethyl-8-ribityl-lumazine (3) from the butanone-phosphate (2) and the pyrimidinedione (1).

Kinetic and spectroscopic studies by Kis et al.12 have revealed a number of mechanistic details of the reaction catalyzed by LS. It has been shown that both enantiomers of the butanone-phosphate (2) can serve as substrates for LS; however, the reaction rate of the natural S-enantiomer was about sixfold higher than that for the R-enantiomer. The  $K_{\rm m}$  value for substrate **2** (130  $\mu$ M) exceeded the  $K_{\rm m}$  value for the pyrimidine substrate 1 (5  $\mu$ M) by more than an order of magnitude, which suggested an ordered bi-bi kinetic mechanism, but conclusive evidence has not been given. Strict regio-specificity was detected for the enzyme-catalyzed condensation of the carbohydrate-phosphate with the pyrimidinedione. The methyl protons of substrate 2 exchange spontaneously with the solvent as shown by NMR analysis.<sup>13,14</sup> In the absence of the pyrimidine substrate 1 the enzyme did not react with the carbohydrate substrate 2. More specifically, the enzyme complex did not catalyze the exchange of the proton at C-3 of substrate 2 with solvent water, nor did it act as a racemase. It was thus concluded that the initial step of the enzyme-catalyzed reaction requires the presence of the pyrimidine substrate 1 in the active site. A reaction mechanism initiated by the formation of a Schiff base upon the reaction of the 5-amino group of the pyrimidine substrate 1 with the carbonyl group of substrate 2 has been proposed. It has been assumed that this step would be followed by proton abstraction and phosphate elimination with the resulting double bond, thus being in favorable conjugation with the pyrimidine system. The enolate intermediate could then tautomerize under formation of a carbonyl group, which then could be attacked by the 6-amino group, resulting in ring closure. Finally, the release of water would terminate the reaction, resulting in energetically favorable conjugation in the heterocyclic double ring system of product **3**.

The structures of the LS active site and the binding sites of substrates 1 and 2 have been analyzed earlier in complexes of LS from *B. subtilis*, *Saccharo*myces cerevisiae, Spinacea oleracea and Magnaporthe grisea with the substrate-analogs 5-nitroso-6-ribitylamino-2,4(1H,3H)pyrimidinedione (RNOP; Figure  $2(c))^{2}$ 5-nitro-6-ribityl-amino-2,4(1H,3H)pyrimidinedione (RNO2P; Figure 2(c))<sup>3,4,15</sup> and 5-(6-D-ribitylamino-2,4(1H,3H)pyrimidinedione-5-yl)-1-pentylphosphonic acid (RPP; Figure 2(d)),<sup>6</sup> respectively. These studies indicated that the side-chain of highly conserved Arg127 is involved in binding of the phosphate group of substrate 2. The pyrimidine ring of substrate 1 was found in offset stacking aromatic interaction with the ring system of a phenylalanine or alternatively a tryptophane indole ring system (Phe22/Trp22). Several hydrogen bond interactions to the pyrimidine system and the ribityl chain hydroxyl groups of substrate 1 are formed upon binding. They have been analyzed in S. cerevisiae LS and described in detail.<sup>6</sup> The structure of an empty LS active site has been studied and described in LS from Brucella abortus<sup>5</sup> and A. aeolicus.<sup>7</sup> The observed conformational changes, obtained from comparisons with the occupied active sites, have suggested an inducedfit mechanism of substrate binding in LS. Several side-chains undergo specific conformational changes upon binding of the pyrimidine substrate; among the residues the highly conserved His88 (93% among 59 amino acid sequences; for an amino acid sequence alignment see Supplementary



**Figure 2.** Substrate and product analogues: (a) 6,7-dioxo-5*H*-8-ribityl aminolumazine (RDL); (b) 3-(7hydroxy-8-ribityllumazine-6-yl)propionic acid (RPL); (c) 6-ribitylamino-5-nitroso/nitro-2,4(1*H*,3*H*)pyrimidindione (R = NO RNOP,  $R = NO_2 RNO2P$ ); (d) 5-(6-D-ribitylamino-2,4(1*H*,3*H*)pyrimidinedione-5-yl)pentyl-1-phosphonic acid (RPP).

Material) suggests a critical involvement in catalysis. The aromatic ring of Phe22 adapts to the binding of the pyrimidine system by a swing movement, which results in a stacking aromatic interaction and proper orientation of the pyrimidine.

In this work we have determined and refined at high resolution the crystal structures of four complexes of A. aeolicus LS with substrate and product analogous inhibitors, among them for the first time an analogue of the product 6,7-dimethyl-8ribityl-lumazine (3). We analyze and describe in detail the binding modes of these inhibitors. By taking into account the constraints imposed by the active site structure, the functional data from single-site mutants of the highly homologous B. subtilis LS<sup>16</sup> and the conservation of distinct amino acid side-chains in the active site, a structural model of the catalytic process, which illustrates binding of both substrates, enantiomer specificity, proton abstraction/donation, inorganic phosphate elimination, formation of the Schiff base and cyclization will be proposed.

#### **Results and Discussion**

## Overall structure and quality of the refined models

The structures of icosahedral lumazine synthase from *A. aeolicus* (LSAQ) in complex with the inhibitors 6,7-dioxo-5*H*-8-ribitylaminolumazine (RDL; Figure 2(a)), 3-(7-hydroxy-8-ribityllumazine-6yl)propionic acid (RPL; Figure 2(b)), 6-ribitylamino-5-nitroso-2,4(1*H*,3*H*)pyrimidindione (RNOP; Figure 2(c)), and 5-(6-D-ribitylamino-2,4(1*H*,3*H*) pyrimidinedione-5-yl)pentyl-1-phosphonic acid (RPP; Figure 2(d)) were solved by molecular replacement using the native structure of LSAQ<sup>7</sup> as a search model. The native enzyme, as well as all enzyme–inhibitor complexes, was crystallized in space group l23. The unit cell contains 24 asymmetric units, each comprising one pentamer. Thus, one unit cell contains two complete icosahedral capsids with their local 3-fold axes in line with the unit cell body diagonal. The five subunits within the crystal asymmetric unit have exactly the same packing environment and were thus refined using

strict non-crystallographic 5-fold symmetry. A total of 5% reflections from each intensity data set with indexes corresponding to the native data set<sup>17</sup> were set aside for cross-validation. The crystal structures of the complexes were refined to a resolution of 1.75 Å (LSAQ-RDL), 1.82 Å (LSAQ-RPL), 2.05 Å (LSAQ-RNOP) and 2.2 Å (LSAQ-RPP), respectively. As summarized in Table 1, the final models have the conventional crystallographic Rfactors of 14.2% (LSAQ-RDL), 17.4% (LSAQ-RPL), 15.5% (LSAQ-RNOP) and 16.1% (LSAQ-RPP), respectively. The resulting final free R factors<sup>17</sup> were 15.7% (LSAQ-RDL), 18.9% (LSAQ-RPL), 18.1% (LSAQ-RNOP) and 17.7% (LSAQ-RPP). After the last refinement cycle, the averaged *B*-factors of the inhibitor molecules were  $10.8 \text{ Å}^2$ (RDL), 23.7 Å<sup>2</sup> (RPL), 13.7 Å<sup>2</sup> (RNOP) and 17.0 Å<sup>2</sup> (RPP), respectively. The averaged *B*-factors of the refined complex models agree well with the statistically estimated B-factors obtained from Wilson plots (Table 1).<sup>18</sup>

The main-chain atoms were generally very well defined in all structures with very few exceptions. Electron densities for the side-chains of Lys70 in all four complex structures, Lys151 in LSAQ–

	-		
RDL	RPL	RNOP	RPP
48.22-1.75	48.14-1.82	48.14-2.05	48.13-2.2
14.2	17.4	15.5	16.1
15.7	18.9	18.1	17.7
96.2	96.1	95.9	95.9
3.8	3.2	3.6	4.1
0	0.8	0.5	0
0.014	0.015	0.020	0.016
1.6	1.509	1.8	1.7
22.9	19.1	23.3	23.8
1.57	1.70	1.69	1.39
5885 (15)	5885 (15)	5885 (25)	5885 (45)
115	135	100	135
730	585	515	410
13.8	21.1	17.7	19.7
11.7	17.8	16.7	16.7
12.5	19.0	17.9	18.0
10.8	22.6	13.7	17.0
27.0	32.5	29.2	26.8
	RDL 48.22–1.75 14.2 15.7 96.2 3.8 0 0.014 1.6 22.9 1.57 5885 (15) 115 730 13.8 11.7 12.5 10.8 27.0	RDLRPL $48.22-1.75$ $48.14-1.82$ $14.2$ $17.4$ $15.7$ $18.9$ $96.2$ $96.1$ $3.8$ $3.2$ $0$ $0.8$ $0.014$ $0.015$ $1.6$ $1.509$ $22.9$ $19.1$ $1.57$ $1.70$ $5885(15)$ $5885(15)$ $115$ $135$ $730$ $585$ $13.8$ $21.1$ $11.7$ $17.8$ $12.5$ $19.0$ $10.8$ $22.6$ $27.0$ $32.5$	RDLRPLRNOP $48.22-1.75$ $48.14-1.82$ $48.14-2.05$ $14.2$ $17.4$ $15.5$ $15.7$ $18.9$ $18.1$ $96.2$ $96.1$ $95.9$ $3.8$ $3.2$ $3.6$ $0$ $0.8$ $0.5$ $0.014$ $0.015$ $0.020$ $1.6$ $1.509$ $1.8$ $22.9$ $19.1$ $23.3$ $1.57$ $1.70$ $1.69$ $5885(15)$ $5885(25)$ $115$ $135$ $100$ $730$ $585$ $515$ $13.8$ $21.1$ $17.7$ $11.7$ $17.8$ $16.7$ $12.5$ $19.0$ $17.9$ $10.8$ $22.6$ $13.7$ $27.0$ $32.5$ $29.2$

Table 1. Refinement statistics for A. aeolicus lumazine synthase in complex with inhibitors

RNOP and LSAQ–RPP as well as Lys7 in LSAQ– RPP were not well defined or missing. The sidechains of residue Lys7 and Lys135 in LSAQ–RDL and the side-chain of Lys7 in LSAQ–RPL were found to adopt alternate conformations. The inhibitor compounds in all complex structures could be unambiguously located with the aid of  $2F_o - F_c$ electron density maps (Figure 3). The water molecules were well identified in the final structures with the averaged *B*-factors of 27.0 Å<sup>2</sup> (LSAQ– RDL), 32.3 Å<sup>2</sup> (LSAQ–RPL), 29.2 Å<sup>2</sup> (LSAQ– RNOP) and 26.8 Å<sup>2</sup> (LSAQ–RPP), respectively (for detailed information, see Table 1). The overall folding of the enzyme in complexes with inhibitors is identical to that of the native structure.<sup>7</sup> The subunit is built up by a central  $\beta$ sheet flanked by two  $\alpha$ -helices on one side and three  $\alpha$ -helices on the other side. This structural arrangement has also been found in icosahedral LSs of spinach<sup>4</sup> and *B. subtilis*<sup>2,3,15</sup> as well as the pentameric LSs.<sup>4,6,8</sup> In the complex structures, a phosphate ion is situated at the non-crystallographic 5-fold channel, where, in the native structure, a water molecule has been observed instead. Lys98 side-chains from five surrounding subunits extend towards the central phosphate ion forming salt



**Figure 3.** The  $2F_o - F_c$  electron densitiy ( $\sigma = 1.5$ ) around the active site of *A. aeolicus* lumazine synthase in complex with the product analogue 6,7-dioxo-5*H*-8-ribitylamino-lumazine (RDL). The average *B*-factor of water molecules, indicated by red spheres, is below 21 Å<sup>2</sup>.



bridges blocking the 5-fold channels. However, in the empty LSAQ capsid, the lysine side-chains are bent away from the 5-fold axis forming an open channel along the 5-fold.

Compared with the structure of empty LSAQ, no significant global changes were observed in the enzyme–inhibitor complex structures. The r.m.s. deviation between C<sup> $\alpha$ </sup> atoms of the empty LSAQ and the complexes are 0.13 Å (LSAQ–RDL), 0.21 Å (LSAQ–RPL), 0.19 Å (LSAQ–RNOP) and 0.19 Å (LSAQ–RPP), respectively. Ramachandran plots showed that all residues of each enzyme–inhibitor complex were within the favored and allowed regions.

#### Architecture of the active site

The active site of lumazine synthase is situated at the subunit interface close to the inner wall of the icosahedral capsid. The overall conformation of the active site in all enzyme-inhibitor complexes is well conserved with respect to the active site of the native enzyme. Five active site pockets within the crystallographic asymmetric unit are topologically equivalent. The secondary structural elements constructing the active sites are  $\beta$ -turns (residues 21-24, 54-58 and 81-92) from one subunit and a  $\beta\text{-strand}$  (residue 127'-142') as well as an  $\alpha$ -helix (residue 113'-116') from the neighboring subunit. As shown in the surface representation of the inhibitor-binding site in LSAQ-RPL (Figure 4), the ribityl side-chain of the inhibitor is embedded in a surface depression, which is less accessible to solvent than the ring system and the phosphonyl chain. The hetero-aromatic ring system of the inhibitors is located in a hydrophobic pocket formed by Phe22, Val80, Leu81, Phe89 and Ile92. The phosphate-binding region of the active site pocket is an open pit, which is fully exposed to the solvent. On binding of the inhibitor com**Figure 4**. The active site of *A. aeolicus* lumazine synthase in complex with the reaction intermediate analogue 3-(7-hydroxy-8-ribityllumazine-6-yl)propionic acid (RPL). The solvent-accessible surface was calculated with a water probe radius of 1.4 Å. The two adjacent subunits constituting the active site are shown in different colors (salmon and wheat).

pounds, the solvent accessible surfaces of the active site of the complexes were reduced by 354.4  $\text{\AA}^2$  (LS–RDL), 344.4  $\text{\AA}^2$  (LS–RPL), 368.8  $\text{\AA}^2$  (LS–RNOP) and 369.7  $\text{\AA}^2$  (LS–RPP), respectively.

Although the overall structure of the active site in the complexes is similar to that of the native enzyme, several notable conformational differences were observed. The binding of substrate analogous inhibitors to LS appears to follow an induced fit mechanism:

- 1. The phenyl ring of Phe22 swings into an orientation parallel to the hetero-aromatic ring system of the inhibitors in all complexes, whereas in the native enzyme it is tilted away by more than 30° relative to this orientation (Figure 5). With an inter-ring distance of about 3.4 Å, the offset-stacked geometry suggests that the  $\pi-\sigma$  interaction<sup>19</sup> between the ring systems of the inhibitors and Phe22 helps to gain stabilization energy. Adopting different side-chain conformations, Phe22 acts like a "gate" controlling the path between the active site cavity and the solvent environment. Position 22 is occupied by either a phenylalanine or a tryptophan residue with only one exception (in this case a serine) among 59 compared amino acid sequences.
- 2. The imidazole group of His 88 is twisted away from the inhibitor by about 0.7 Å in all the complex structures, except LSAQ–RNOP, in which the atom N<sup>81</sup> on the imidazole ring of His88 forms a hydrogen bond with the oxygen atom of the inhibitor's nitroso group. The amino acid sequence alignment shows that position 88 is occupied exclusively by a positively charged residue, i.e. either a histidine (93%) or an arginine (7%).



Figure 5. The superimposed active site structures of lumazine synthase from A. aeolicus in complex with 6,7-dioxo-5H-8-ribitylaminolumazine (RDL, ligands wheat), 3-(7-hydroxy-8-ribityllumazine-6-yl)propionic acid (RPL, green), 6-ribitylamino-5-nitroso-2,4(1H,3H) pyrimidindione (RNOP, blue), and 5-(6-D-ribitylamino-2,4(1*H*,3*H*)pyrimidinedione-5-yl)pentyl-1-phosphonic acid (RPP, yellow) as well as the native enzyme (red) with an empty active site. Note the alternate side-chain locations of Lys135' and Glu138' in the complex structures, the adaptation movements of His88 and the tilted phenyl ring of Phe22 in the structure of the native enzyme.

3. Arg127, which is involved in forming the ionic contact with the phosphate ion, is highly conserved throughout the compared species. Arg127 forms a salt bridge with

Glu126. The sequence comparisons with other species indicate a highly conserved negatively charged residue (Glu/Asp) at this position. Together with Lys131 and His132, these charged residues form a tetrad extending over the subunit interface (Figure 6). Interestingly, this tetrad constructs a pocket with another charged tetrad from the neighboring pentamer related by the crystallographic 2fold. The active site cavity is located beside the tetrad with Phe22 facing the 2-fold. The 2-fold pocket and the symmetry-related tetrads as well as the inhibitor compounds of LSAQ-RDL are shown in a surface representation (Figure 6). In the pentameric enzyme from S. cerevisiae,<sup>6</sup> a four-residue extension is located at the loop region between positions 129 and 130. The computer modeling study has shown that this extra insertion would clash with residues 23-26 from the adjacent pentamer when the S. cerevisiae LS pentamers were assembled into an icosahedral capsids.6

In the structure of LSAQ–RDL, the side-chain of Lys135 was observed to adopt two conformations, i.e. either bending into the active site or stretching towards the boundary of the substrate-binding pocket (Figure 5). Interestingly, two side-chain conformations of Glu138 were observed in all complex structures of *A. aeolicus* lumazine synthase as well as in the structure of the native enzyme. It seems that the conformation of Glu138 shifts simultaneously with Lys135. Therefore, Lys135 can establish ionic interactions with the Glu residue in both positions. The amino acid sequence alignment shows that in 88% of the compared sequences, position 135 is occupied by Lys, Arg, or His, in most cases a lysine. Position 138 is in 71% of the com-



Figure 6. Solvent-accessible surface representation (water probe r = 1.4 Å) of the 2-fold opening. The product analogue 6,7-dioxo-5H-8-ribitylaminolumazine (RDL) is shown by space-filling models and four charged residues, which form a tetrad, are shown by stick models. The color codes are wheat and blue for subunits from one pentamer, salmon and green for subunits from the neighboring pentamer.



**Figure 7**. Schematic drawings of the interactions between the enzyme and the ligands: (a) 6,7-dioxo-5*H*-8-ribityl-aminolumazine (RDL); (b) 3-(7-hydroxy-8-ribityllumazine-6-yl)propionic acid (RPL); (c) 6-ribitylamino-5-nitroso-2,4(1*H*, 3*H*)pyrimidindione (RNOP); (d) 5-(6-D-ribitylamino-2,4(1*H*,3*H*)pyrimidinedione-5-yl)pentyl-1-phosphonic acid (RPP).

pared species occupied by either an aspartic acid or a glutamic acid.

#### Inhibitor-binding mode

Inhibitor compounds were designed in order to mimic several states of the catalytic reaction. Compound RNOP resembles substrate **1**. RPP is similar to the presumed transition state intermediate after a connection to the ring is established by nucleophilic attack. RPL mimics the intermediate with the phosphate carrying side-chain in the proper orientation for ring closure. RDL represents a product analogue of the reaction. A considerable amount of hydrogen bonds were observed between the enzyme and the inhibitor molecules. Some of them were found in all complex structures in spite of different binding ligands. Schematic drawings of the inhibitor binding topology are shown in Figure 7 and the distances between the inhibitors and the active site residues are listed in Table 2.

As a common binding feature of all enzymeinhibitor complexes, the ribityl side-chains of the ligands interacts with a series of residues of the enzyme *via* main-chain and side-chain contacts. The hydroxyl group at position 2' binds to the main-chain nitrogen atom of Ser56 and Trp57; 3'-OH is in contact with the main-chain nitrogen

**Table 2.** Interactions between inhibitor compounds and the enzyme

	Inhibitor atom	Protein atom	Distance between protein atoms and inhibitor atoms (Å)			
		_	RDL	RPL	RNOP	RPP
1	2-O	Ser56_N	2.9	3.0	2.9	2.9
		Ser56_O <sup>γ</sup>	2.9	2.8	2.8	2.7
		Val80_O	3.5	3.4	3.5	3.7
	N3	Val80_O	2.8	2.7	2.8	2.9
	4-O	Val80_O	3.7	3.7	3.6	3.6
		Asn23_N <sup>82</sup>	3.6	3.5	3.4	3.2
		Ile82_N	3.2	3.1	3.2	3.2
		Ile82_O	3.7	3.6	3.8	3.9
	5-NO	$His 88_N^{\delta 1}$	-	-	3.3	-
	6-O	$His 88_N^{\delta 1}$	3.4	-	-	-
	7-O	$His 88_N^{\delta 1}$	3.6	3.7	-	-
		Lys135_N <sup>c</sup>	$2.5^{*}$	2.8	-	-
2	2′-O	Ser56_N	3.3	3.3	3.2	3.3
		Trp57_N	3.1	3.1	3.1	3.2
	3'-O	Gly55_N	3.6	3.6	3.6	3.5
		Glu58_O <sup>ε1</sup>	2.7	2.7	2.7	2.7
		Glu58_O <sup>2</sup>	3.4	3.5	3.5	3.4
	4'-O	Phe113_0	2.8	2.8	2.9	2.9
	-/ ->	Lys135_N <sup>c</sup>	3.0*	2.9	_	2.8
	5'-0	Gly58_O <sup>e2</sup>	2.9	2.7	2.6	2.6
		PheII3_N	2.9	2.9	2.9	3.0
		PheII3_O	3.7	3.6	3.6	3.6
3	O1	$His88_N^{\delta 1}$	_	_	3.8	-
		Gly84_N	4.0	4.0	3.9	_
	O2(O2P)	Ala85_N	3.3	3.4	3.3	3.5
		Thr86_N	2.9	2.9	2.8	3.0
		Thr86_O	3.6	3.5	3.6	3.6
	O(O1D)	$1 \text{ nr86}_{-}\text{O}^{1}$	2.6	2.6	2.6	2.5
	03(01P)	$1 \text{ nr 86}_0^{\mu}$	3.8	3.8	3.7	3.9
		$\text{Arg}_{127}$ NU 12	2.8	2.9	2.7	2.7
		Arg127_INHZ	3.6	3.6	3.0	3.5
	O4(O3P)	Alab5_IN $\Lambda = 127$ N <sup>K</sup>	3.3	3.0	3.3	2.8
		Arg127_IN <sup>®</sup>	3.9	3.8 2.9	3.7	3.8 29
		Lyc125 NS	2.9	2.0	2.9	2.0
4	13-01	$C_{11}138 \Omega^{\epsilon 1}$	5.9	27	_	_
т	10-01	$Glu138 O^{\epsilon 2}$	_	35	_	_
	13-02	$Glu138 O^{\epsilon 1}$	_	3.6	_	_
	10-02	Lys135 $N^{\zeta}$	_	3.0	_	_
				0.0		

Distances within 4.0 Å are listed. (\*) Distances measured for the alternate conformation of Lys135. (–) Atom does not exist or distance longer than 4.0 Å. Atomic interactions between protein atoms and: (1) atoms on the ring systems of inhibitor compounds; (2) atoms on the ribityl side-chain; (3) the phosphate ion; (4) atoms on the propionic acid side chain of the inhibitor compound RPL.

atom of Gly55 and Glu58-O<sup>e1</sup>/O<sup>e2</sup>; 4'-OH forms hydrogen bonds with the main-chain nitrogen atom of Phe113' and N<sup>c</sup> of Lys135'; 5'-OH interacts with Gly58-O<sup>e2</sup> as well as the main-chain nitrogen and oxygen atoms of Phe113'. The phenyl ring of Phe113' and the indole ring of Trp57 adopt an offset-stacking configuration with an inter-planar distance of about 4.5 Å. The conservation of aromatic residues at these two positions is 66% and 81%, respectively. Similar binding schemes were observed in enzyme from *B. subtilis* LS with the presence of inhibitor RNOP<sup>2</sup> as well as in the structure of *S. cerevisiae* lumazine synthase in complex with inhibitor RPP.<sup>6</sup> The central part of all inhibitors is formed by a hetero-aromatic system, which is either a pyrimidinedione ring in inhibitor RNOP and RPP or a lumazine ring in inhibitor RDL and RPL. Similar interactions between the enzyme and the core ring systems of all inhibitors were observed. The carbonyl-oxygen 2-O on the ring is in contact with Ser56-N, Ser56-O<sup> $\gamma$ </sup> and Val80-O of the enzyme. Hydrogen bonds were formed between carbonyl-oxygen 4-O and the enzyme atoms Asn23-N<sup> $\delta$ 2</sup>, Val80-O, Ile82-N and Ile82-O, between imido N3 atom of the ligands and the protein atom Val80-O.

As observed in other enzyme–inhibitor complex structures, a phosphate ion is bound to the active site *via* hydrogen bond networks though contacts with Gly84-N, Ala85-N, Thr86-N/O/O<sup> $\gamma$ 1/O<sup> $\gamma$ 2</sup>. The highly conserved Arg127 contributes its N<sup>\*</sup>/N<sup> $\gamma$ </sup> atom for an ionic contact with the phosphate O1.</sup>

In the structure of LSAQ–RDL, the 6- and 7keto-oxygens on the lumazine ring form hydrogen bonds with the  $N^{\delta 1}$  atom of His88. In the LS–RPL structure, the carboxyl oxygen atoms on the propionic acid side-chain build a salt bridge with the  $N^{\epsilon}$  atom of Lys135' and these oxygen atoms also interact with Glu138'-O<sup>ε1</sup>/O<sup>ε2</sup> of the enzyme.

More than ten water molecules forming hydrogen bond networks with the enzyme and the ligands are located within the active site region. Several water-inhibitor hydrogen bond interactions were observed in all complex structures. As shown in Figure 3, Wat1 establishes a hydrogen bond to the 2-O atom at the ring of the inhibitor, Wat2 interacts with the phosphate oxygen atoms and Wat3 is in contact with the ribityl 3'-oxygen atom. A main-chain hydrogen bond interaction between Thr117-N and Wat4 was observed in all complex structures. On binding of the inhibitors, several water molecules involved in forming hydrogen bonds with the enzyme residues in the native structure, are replaced by oxygen atoms of the inhibitor compounds, namely, 2-O and 4-O on the pyrimidine ring and 2'-, 3'- as well as 5'- oxygen atoms on the ribityl chain.

#### Catalytic reaction mechanism

By comparing inhibitor binding data presented here and considering structural constraints implied by the geometry of the active site, a catalytic mechanism can be suggested, which represents an extended and refined version of the mechanism proposed earlier by Kis et al.<sup>12</sup> This mechanism consists of the following steps: substrate binding, nucleophilic attack, formation of a Schiff base intermediate, phosphate elimination and ring closure (Figure 8). The involvement of critical residues (i.e. Phe22, Arg127, His88, and the couple Lys135/ Glu138) will be considered and discussed on the basis of structural proximity and activity measurements on mutants of the highly homologous lumazine synthase from *B. subtilis*. An animation of this mechanism can be found in Supplementary Material.



Figure 8. Hypothetical mechanism for the biosynthesis of lumazine suggested by Kis et al.<sup>12</sup>

#### Substrate binding

All inhibitors used so far in crystallographic studies of the LS–inhibitor complexes have both the ribityl side-chain and a pyrimidine ring system in common. The binding mode of this part of the inhibitors is, in all cases, very similar. The ribityl side-chain establishes a number of main-chain and side-chain contacts to the neighboring residues *via* its hydroxyl groups and the ring system is in all cases nearly coplanar with either a phenylalanine or a tryptophane residue. Thus, it is assumed that the corresponding portions of the substrate will bind in a similar way and be located in a similar position throughout the reaction cycle.

The binding site of substrate **2** has been located and defined on the basis of the binding position of the phosphate group, which is in ionic interaction with Arg127,<sup>2,3</sup> and the known conformation of the pentyl-phosphonate group of the bound inhibitor RPP.<sup>6</sup> A binding model was generated by superimposing substrate **2** with the pentyl-phosphonic acid side-chain of RPP. The model shows that substrate **2** fits well into the active site pocket in the presence of substrate **1**. The 2-carbonyl group and the 3-hydroxyl group of substrate **2** could form hydrogen bonds with surrounding residues Ile82, Gly84, His88 and Ala89 (Figure 9(a)).

Arg127 (present in 93% of all sequences) is responsible for the binding of the phosphate group of substrate **2**. It is replaced by a lysine residue in a number of lumazine synthases. The mutation of Arg127 to Ala, Leu, Ser, Thr in LS from *B. subtilis* results in complete inactivation of the enzyme.<sup>16</sup> However, the mutant Arg127His, which obviously can be involved in the binding of the phosphate group of substrate **2**, shows 62% residual activity in comparison to the wild-type enzyme.<sup>16</sup>

Substrate 2 does not racemise, nor does it exchange protons at the C-3 position in the absence of substrate 1.<sup>12</sup> Moreover the binding affinity of substrate 2 is lower compared with substrate 1. Thus, it is assumed that substrate 2 binds to the active site while substrate 1 is already present. The obvious binding site for substrate 2 is the region around Arg127 including residues Ile82, Gly84, His88 and Ala89.

#### Nucleophilic attack

The NMR studies by Volk & Bacher<sup>14</sup> suggested that the reaction is initiated by a nucleophilic attack on the carbonyl group of substrate **2** by the 5-amino group of the pyrimidine substrate **1**. This attack can proceed while the phosphate group is bound to the phosphate-binding site and the ribityl side-chain is attached to the ribityl-binding site at the distal end of the active site. The position and binding mode of inhibitor RPP in complex with LSAQ and LS from yeast indicate that this attack can proceed without strain.<sup>6</sup>



Figure 9 (legend opposite)

The computer-generated model in Figure 9(a) shows that by fixing the phosphate group of substrate 2 at the binding site, the intermediate can be orientated in a way that the carbonyl C atom (C2) is geometrically accessible to the 5-amino group of substrate 1. The reaction would then start by a nucleophilic attack. There is no clear evidence for the involvement of certain active site residues in the nucleophilic attack; however, His88 could play a role in proton transfer. As consequence of the high nucleophilicity of the amino group of substrate 1, it seems to be, apart from orientation and proximity effects, unnecessary for the enzyme to be involved in direct chemical catalysis. The nucleophilic attack of the C2 atom, followed by the proton rearrangements, would lead to intermediate 5 (shown in Figure 9(b)). The intermediate **5** was modeled into the active site using the complex structure of LS–RPP as a scaffold.

His88 is in an appropriate position and is therefore a strong candidate for an involvement in these proton transfer steps. His88 is highly conserved in all known lumazine synthase sequences and is strongly sensitive to mutation: the mutants His88Ala, His88Phe, His88Lys were found to be not completely inactive;<sup>16</sup> however, their activities are below 10% of the activity of the wild-type enzyme. The residual activity of the His88 mutants can be explained by the fact that there is still a possibility for proton transfer from solvent without the influence of a catalytic group, resulting in reduced reaction rate. His88 might be involved in other steps of the reaction, too, which will be discussed below.



Figure 9 (contd.)

**Figure 9**. Proposed mechanism of lumazine biosynthesis shown by computer generated models (The chiral carbon atom originally from C3 of substrate **2** is marked with an asterisk.): (a) both substrates are bound to the active site, the 5-amino group of substrate **1** is in favored orientation for the nucleophilic attack; (b) the reorientation of the intermediate phosphate side-chain is aided by the coordinated conformational change of Lys135 and Glu138; (c) the reorientation of the intermediate phosphate side-chain allows the proton transfer from His88; (d) the Schiff base is formed in *cis* configuration allowing for conjugation of the double bond with the pyrimidine ring; (e) phosphate ion elimination followed by the nucleophilic attack of the 6-amino group, the side-chain rotations of Lys135 and Glu138 are assumed to push the phosphate ion away; (f) the formation of the second ring, His88-N<sup>®</sup> can serve as a protonator for the 15-hydroxyl oxygen; (g) the formation of lumazine. An animation of the proposed mechanism is provided in Supplementary Material.



Figure 9 (legend on previous page)

#### Formation of the Schiff base

Kis et al.<sup>12</sup> have proposed the formation of the Schiff base intermediate (6). According to this hypothesis, the elimination of a water molecule would lead to the formation of the double bond between N8 and C9, which, due to the different arrangements of chemical groups with respect to the double bond, can have either cis or trans configuration. For the pyrazine ring closure, which would lead to the formation of the lumazine ring, only the *cis* isomer of the Schiff base allows a sufficient proximity of C10 and N7. Thus, a *cis* configuration must be either accomplished by a suitable orientation of the butyl phosphate side-chain before double bond formation or by a transformation from trans to cis configuration after the double bond formation.

The complex of the inhibitor compound RPP with LSAQ can serve as model for the binding mode of the reaction intermediate 5 with the phosphate group bound to Arg127'. If an intermediate were protonated at O10 followed by the water elimination, the spatial arrangement of the chain would only lead to a trans configuration of the Schiff base. Therefore, either a rotation around the N8–C9  $\sigma$ -bond before the Schiff base formation or an isomerization step about the N8–C9  $\pi$ -bond after the Schiff base formation is required. Zheng et al. calculated energy barriers for the isomerization in the range of 19–20 kcal/mol using a semiempirical approach and argued that the cis-/transisomerization step could be the rate-limiting step of the catalysis.<sup>20</sup> Here we propose an enzymeassisted reorientation of the butyl chain before the double bond formation.

As mentioned earlier, the side-chain of Lys135' in the LS–RDL structure adopts alternate conformations (Figure 5) and the Lys135' side-chains in other complex structures were all found to occupy either of these two positions. Furthermore, the side-chain conformation of Glu138', which forms a salt bridge with Lys135', alternates in a cooperative manner with the conformational change of Lys135'.

Shown in Figure 9(b), adopting one of the multiple conformations, Lys135' is extended towards the phosphate group of the intermediate 5. Thus, a salt bridge is formed between the O1 atom of the phosphate group and Lys135'-N<sup>€</sup>. Accordingly, the phosphate moiety could be re-orientated by rotation around the  $\sigma$ -bonds. It is therefore suggested that along with the side-chain shift of Lys135' and Glu138', the phosphate moiety of intermediate 5 could be "pulled" away from the original binding site forming a new conformation (Figure 9(c)). His88  $N^{\delta 1}$  could then serve as a proton donor in the subsequent water elimination, which leads to the formation of the Schiff base 6 in a *cis* configuration (Figure 9(d)). The  $N^{\delta 1}$  atom of His88 could be re-protonated by surrounding solvent molecules.

The mutation Lys135Glu inactivates the enzyme almost completely (5% residual activity), which

could be due to the impairment of the presumed phosphate shifting mechanism and a local destabilization of the structure by two neighboring repulsive charges (Glu135, Glu138).

## Proton abstraction and elimination of phosphate

It has been suggested that the formation of the Schiff base intermediate **6** is followed by the proton abstraction from C10 and the elimination of the phosphate group. The resulting double bond would be favorable due to its conjugation with the pyrimidine  $\pi$  system.<sup>12</sup>

The enol intermediate 7 (Figure 9(e)) could form spontaneously or after proton abstraction from C10 under the involvement of water, accompanied by the elimination of the phosphate group. The phosphate ion could either be released from the active site or bind to the phosphate-binding site and thus compete with the incoming reaction substrate for the next reaction cycle.

After tautomerization, the resulting carbonyl group in compound **8** could then be directly attacked by the 6-amino group leading to the closure of the second ring of compound **9**. Favorable for conjugation with the pyrimidine  $\pi$  system, the double bonds of the intermediate **7** could be constrained to the same plane. Moreover, the phenyl ring of Phe22, which is in stacking orientation with respect to the pyrimidine ring, could also assist the conjugation *via* the  $\pi$ - $\sigma$  interaction.

As expected, the replacement of Phe22 has some impact on the maximum reaction velocity. The mutations to aliphatic (Phe22Val) or polar sidechains (Phe22Asp, Phe22Ser) reduce the  $V_{\text{max}}$  value to 22%, 12% and 40%, respectively, of the maximum velocity of the wild-type enzyme. The polar replacements generally increase the  $K_{\text{m}}$  value for both substrates. The  $\pi-\sigma$  interaction in between Phe22 and the pyrimidine system presumably contributes to the substrate-binding energy.

#### **Ring closure**

The nucleophilic attack by the 6-amino group would lead to a bicyclic intermediate 9, on which the 10-hydroxyl group could be oriented in two directions. The 10-hydroxyl group of intermediate 9 from the S-2 enantiomer is oriented towards His88, which is at an appropriate distance to serve as a proton donor (Figure 9(f)). Alternatively, if the R-2 enantiomer had served as the substrate, the 10-hydroxy group would be oriented to the opposite side of the intermediate ring system. Hence the observed distance (5 Å) between the N<sup>ε</sup> atom of His88 and the 10-hydroxyl oxygen atom would not allow direct proton transfer.<sup>21,22</sup> This could be the reason for the sixfold higher reaction rate for the natural S-enantiomer.<sup>12</sup> After the 10hydroxyl group has taken up the proton provided by His88, it would be eliminated in the form of a water molecule. Deprotonation of N1 followed by double bond rearrangement leads then to the formation of the pyrazine ring, which finalizes the enzyme-catalyzed reaction cycle. The release of the newly formed 6,7-dimethyl-8-ribityllumazine molecule from the binding pocket may be facilitated by a swing-out movement of the phenyl ring of Phe22 to the position as observed in the native structure (Figure 9(g)).

#### **Materials and Methods**

#### Materials

RDL was prepared as described;<sup>23</sup> RPL was synthesized by Cushman *et al.* (M.C., unpublished results); RNOP was synthesized with published procedures;<sup>24</sup> RPP was prepared as described by Cushman *et al.*<sup>25</sup>

#### Molecular biological and enzymological methods

LS from *A. aeolicus* was prepared as described.<sup>7</sup> The reaction mixtures were incubated at 50 °C. The protein concentration was determined by measuring the absorbance at 280 nm ( $\epsilon_{280 \text{ nm},1 \text{ cm}} = 0.855 \text{ ml/mg}$ ).

#### Crystallization, co-crystallization and crystal soaking

The purified protein was concentrated to 15 mg/ml and stored in a solution containing 0.5 mM EDTA, 0.5 mM sodium sulfite and 50 mM potassium phosphate (pH 7.0). Crystallization and co-crystallization with various inhibitors were performed using the sitting-drop vapor diffusion method at room temperature.

Crystals of the native protein were obtained by mixing 2  $\mu$ l of protein solution (15 mg/ml) and 2  $\mu$ l of reservoir solution containing 4% polyethylene glycol ( $M_r = 400 \text{ Da}, \text{ w/v}$ ), 0.3 M lithium sulfate and 0.1 M Mops (pH 6.5). Crystals were observed after two weeks.

In order to obtain crystals of enzyme–inhibitor complexes, different techniques were employed. The protein in complex with RDL and RNOP were co-crystallized by adding 1  $\mu$ l of saturated inhibitor–water solution (ca 1 mM) to 2  $\mu$ l of protein solution (15 mg/ml) and 3  $\mu$ l of a solution, which contained 0.9 M sodium–potassium tartrate and 0.1 M Hepes (pH 7.5). Large crystals were obtained after ten days. For complexes of protein with inhibitor compound RPL and RPP, crystals were preThe Catalytic Mechanism of Lumazine Synthase

pared by soaking the native crystals over night in the reservoir buffer containing 2 mM inhibitors.

#### Data collection and processing

The crystals were transferred to the corresponding crystallization solutions containing 20% (v/v) glycerol as cryo-protectant. The crystals were soaked in the solution for several minutes before mounting.

Data collection was performed at HASYLAB, Beam Line X11 (DESY, EMBL Hamburg Outstation). For each enzyme–inhibitor complex, a data set covering the full acquisitive resolution range as well as a redundant data set covering the low-resolution range up to 3.5 Å, were collected. The reflection data were evaluated, merged and scaled using the program package HKL.<sup>26</sup> Data collection and evaluation statistics are shown in Table 3.

#### Crystallographic refinement and model building

The positions of five subunits were first optimized by rigid-body refinement using the program CNS<sup>27</sup> with constrained non-crystallographic 5-fold symmetry. Subsequently, simulated annealing by torsion angle dynamics was performed. Starting at a temperature of 5000 K, the system was slowly cooled down to 293 K at a rate of 50 K per cycle.<sup>27</sup>

The molecular models of inhibitor RDL and RPL were generated using the program O and the geometric parameters were optimized with X-PLOR  $^{\rm 28}$  The models of inhibitor compound RNOP and RPP were taken from the refined structure of LS from *B. subtilis*<sup>2</sup> and LS from S. cerevisiae,<sup>6</sup> respectively. The initial atomic temperature factors of the complex models were set to an overall value given by the Wilson plot.18 The initial density maps were of remarkable quality, which allowed unambiguous positioning of the ligands. All main-chain atoms and more than 90% of the side-chain atoms were visible in the initial maps. Electron densities for most of the missing atoms appeared in later refinement cycles. Occupancies of residues with alternate locations were refined with trial values starting from 20%/80% until 80%/20% with an incremental/decremental rate of 10%. The values, which gave the best temperature factors for both conformations, were set for the subsequent refinement cycle. Dictionaries for crystallographic refinement were generated using the program REFMAC5.<sup>29</sup> Libraries needed for model building and simulated annealing were prepared by HIC-Up.3

Table 3. X-ray data evaluation statis	stics
---------------------------------------	-------

5				
	LS-RDL	LS-RPL	LS-RNOP	LS-RPP
Cell dimensions <i>a</i> (Å)	180.6	180.1	180.1	180.1
Resolution range (Å)	48.22-1.75	48.14-1.82	48.14 - 2.05	48.13-2.20
Highest resolution shell (Å)	1.78-1.75	1.85-1.83	2.09 - 2.05	2.24 - 2.20
Number of observed reflections	735,750	619,465	480,544	377,239
Number of unique reflections	97,916	86,400	60,538	49,238
Completeness (%)	99.9	97.6	99.4	100
Overall $I/\sigma$	25.8	40.1	20.9	20.6
Last shell $I/\sigma$	4.5	5.4	5.0	5.0
Overall R <sub>merge</sub> (%)	6.2	5.2	8.7	9.0
Last shell $R_{merge}$ (%)	27.1	27.6	29.7	30.0

All data sets were collected using X-rays with a wavelength of 0.8482 Å. All crystals belong to space group I23.  $R_{\text{merge}} = \sum_i |I_i - \langle I_i \rangle| / \sum_i |\langle I_i \rangle|$ , where  $I_i$  is the scaled intensity of the *i*th observation and  $\langle I \rangle$  is the mean intensity for that reflection.

The molecular masks for map averaging were calculated using MAMA.<sup>31</sup> After each macro cycle of the crystallographic refinement, NCS operators and the masks were optimized with the program IMP.<sup>31</sup> The electron density maps were averaged applying the non-crystallographic 5-fold symmetry using the program AVE.<sup>32</sup>

Model building was performed using the program O.<sup>33</sup> REFMAC5 was used throughout crystallographic refinement. Water molecules were determined from the  $2F_o - F_c$  and the  $F_o - F_c$  electron density maps with a sigma level of 1.3 using the program Arp/Warp.<sup>34</sup> NCS relations among water molecules were detected using the program Sortwater.<sup>35</sup> Water molecules with *B*-factors larger than 50 Å<sup>2</sup> were rejected from the refined structures.

Solvent accessible surface calculations were performed using the program CNS.<sup>27</sup>

#### Amino acid sequence alignment of the lumazine synthase, quality analysis and graphical representation of the model

The amino acid sequences of lumazine synthase from 48 species were obtained from Swiss-Prot (date 6 May 2002) and the alignment was performed using ClustalW.<sup>36</sup> The refined structures were checked using PROCHECK<sup>37</sup> and WhatIf.<sup>38</sup> The r.m.s. deviations between main-chain atoms of the models were calculated using the program RMSPDB<sup>†</sup>. The r.m.s. deviations of bond lengths, angles, dihedral angles and improper dihedrals were calculated using CNS. The program Pymol<sup>‡</sup> was used for the graphic representations of the enzyme–inhibitor complex structures.

#### Protein Data Bank accession codes

The structure coordinates and structure factors have been deposited at the RCSB Protein Data Bank. The accession codes of the complexes are 1NQU (LSAQ– RDL), 1NQX (LSAQ–RPL), 1NQV (LSAQ–RNOP) and 1NQW (LSAQ–RPP), respectively. The coordinates and the structure factors will be released on 14–02–2004.

#### Acknowledgements

This work was supported by grants from the Swedish Natural Science Research Council, NIH grant GM1469, the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

#### References

- Bacher, A., Baur, R., Eggers, U., Harders, H. D., Otto, M. K. & Schnepple, H. (1980). Riboflavin synthases of *Bacillus subtilis*. Purification and properties. *J. Biol. Chem.* 255, 632–637.
- Ladenstein, R., Schneider, M., Huber, R., Bartunik, H. D., Wilson, K., Schott, K. & Bacher, A. (1988).

†Kleywegt, G. J. (1992–2001), http://xray.bmc.uu.se/ usf/rmspdb\_man.html Heavy riboflavin synthase from *Bacillus subtilis*. Crystal structure analysis of the icosahedral  $\beta_{60}$  capsid at 3.3 Å resolution. *J. Mol. Biol.* **203**, 1045–1070.

- Ritsert, K., Huber, R., Turk, D., Ladenstein, R., Schmidt-Bäse, K. & Bacher, A. (1995). Studies on the lumazine synthase/riboflavin synthase complex of *Bacillus subtilis*: crystal structure analysis of reconstituted, icosahedral β-subunit capsids with bound substrate analogue inhibitor at 2.4 Å resolution. *J. Mol. Biol.* 253, 151–167.
- Persson, K., Schneider, G., Jordan, D. B., Viitanen, P. V. & Sandalova, T. (1999). Crystal structure analysis of a pentameric fungal and an icosahedral plant lumazine synthase reveals the structural basis for differences in assembly. *Protein Sci.* 8, 2355–2365.
- Braden, B. C., Velikovsky, C. A., Cauerhff, A. A., Polikarpov, I. & Goldbaum, F. A. (2000). Divergence in macromolecular assembly: X-ray crystallographic structure analysis of lumazine synthase from *Brucella abortus*. J. Mol. Biol. 297, 1031–1036.
- Meining, W., Mörtl, S., Fischer, M., Cushman, M., Bacher, A. & Ladenstein, R. (2000). The atomic structure of pentameric lumazine synthase from *Saccharomyces cerevisiae* at 1.85 Å resolution reveals the binding mode of a phosphonate intermediate analogue. J. Mol. Biol. 299, 181–197.
- Zhang, X., Meining, W., Fischer, M., Bacher, A. & Ladenstein, R. (2001). X-ray structure analysis and crystallographic refinement of lumazine synthase from the hyperthermophile *Aquifex aeolicus* at 1.6 Å resolution: determinants of thermostability revealed from structural comparisons. *J. Mol. Biol.* **306**, 1099–1114.
- Gerhardt, S., Haase, I., Steinbacher, S., Kaiser, J. T., Cushman, M., Bacher, A. *et al.* (2002). The structural basis of riboflavin binding to *Schizosaccharomyces pombe* 6,7-dimethyl-8-ribityllumazine synthase. *J. Mol. Biol.* **318**, 1317–1329.
- Bacher, A. & Ladenstein, R. (1990). The lumazine synthase/ribofloavin synthase complex of *Bacillus* subtilis. In *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., ed.), vol. I, pp. 293–316, Chemical Rubber & Co, Boca Raton, FL.
- Volk, R. & Bacher, A. (1991). Studies on the 4-carbon precursor in the biosynthesis of riboflavin. Purification and properties of L-3,4-dihydroxy-2-butanone-4-phosphate synthase. *J. Biol. Chem.* 265, 19479–19485.
- Bacher, A. (1991). Biosynthesis of flavins. In Chemistry and Biochemistry of Flavoenzymes (Müller, F., ed.), vol. I, pp. 215–259, Chemical Rubber & Co, Boca Raton, FL.
- Kis, K., Volk, R. & Bacher, A. (1995). Biosynthesis of riboflavin. Studies on the reaction mechanism of 6,7dimethyl-8-ribityllumazine synthase. *Biochemistry*, 34, 2883–2892.
- Beach, R. L. & Plaut, G. W. E. (1970). Stereospecificity of the enzymic synthesis of the *o*-xylene ring of riboflavin. *J. Am. Chem. Soc.* 92, 2913–2916.
- Volk, R. & Bacher, A. (1991). Biosynthesis of riboflavin. Studies on the mechanism of L-3,4-dihydroxy-2butanone 4-phosphate synthase. *J. Biol. Chem.* 266, 20610–20618.
- Ladenstein, R., Ritsert, K., Huber, R., Richter, G. & Bacher, A. (1994). The lumazine synthase/riboflavin synthase complex of *Bacillus subtilis*. X-ray structure analysis of hollow reconstituted β-subunit capsids. *Eur. J. Biochem.* 223, 1007–1017.

<sup>‡</sup>DeLano, W. L. (2002), http://www.pymol.org

- 16. Fischer, M., Haase, I., Kis, K., Meining, W., Ladenstein, R., Cushman, M. et al. (2003). Enzyme catalysis via control of activation entropy. Site directed mutagenesis of 6,7-dimethyl-8-ribityllumazine synthase. . Mol. Biol. **326**, 783–793.
- 17. Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature, 355, 472-475.
- 18. Wilson, A. J. C. (1949). The probability distribution of X-ray intensities. Acta Crystallog. 2, 318-321.
- 19. Hunter, C. A. & Sanders, J. K. M. (1990). The nature of  $\pi$ - $\pi$  interactions. J. Am. Chem. Soc. **112**, 5525–5534.
- 20. Zheng, Y.-J., Viitanen, P. V. & Jordan, D. B. (2000). Rate limitations in the lumazine synthase mechanism. Bioorg. Chem. 27, 89-97.
- 21. Goetz, J. M., Poliks, B., Studelska, D. R., Fischer, M., Kugelbrey, K., Bacher, A. et al. (1999). Investigation of the binding of fluorolumazines to the 1-MDa capsid of lumazine synthase by <sup>15</sup>N{<sup>19</sup>F} REDOR NMR. J. Am. Chem. Soc. 121, 7500-7508.
- 22. Scheuring, J., Kugelbrey, K., Weinkauf, S., Cushman, M., Bacher, A. & Fischer, M. (2001). <sup>19</sup>F NMR ligand perturbation studies on 6,7-bis(trifluoromethyl)-8ribityllumazine-7-hydrates and lumazine the synthase complex of Bacillus subtilis. Site-directed mutagenesis changes the mechanism and the stereoselectivity of the catalyzed haloform-type reaction. J. Org. Chem. **66**, 3811–3819.
- 23. Al-Hassan, S. S., Kulick, R. J., Livingstone, D. B., Suckling, C. J. & Wood, H. C. S. (1980). Specific enzyme inhibitors in vitamin biosynthesis. Part 3. The synthesis and inhibitory properties of some substrates and transition state analogs of riboflavin synthase. J. Chem. Soc. Perkin Trans. 1, 2645-2656.
- 24. Plaut, G. W. E. (1971). Metabolism of water soluble vitamins. The biosynthesis of riboflavin. In Comprehensive Biochemistry (Florkin, M. & Stotz, E. H., eds), pp. 11-45, Elsevier, Amsterdam.
- 25. Cushman, M., Mihalic, J. T., Kis, K. & Bacher, A. (1999). Design, synthesis, and biological evaluation of homologous phosphonic acids and sulfonic acids as inhibitors of lumazine synthase. J. Org. Chem. 64, 3838-3845.
- 26. Otwinowski, Z. & Minor, W. (1997). Processing of Xray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307-326.
- 27. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallog. sect. D, 54, 905–921.
- 28. Brünger, A. T. (1992). X-PLOR Manual Version, 3.865 edit., Yale University, New Haven, CT. 29. Murshudov, G. N., Vagin, A. A. & Dodson, E. J.
- (1997). Refinement of macromolecular structures by

the maximum-likelihood method. Acta Crystallog. sect. D, 53, 240-255.

- 30. Kleywegt, G. J. & Jones, T. A. (1998). Databases in protein crystallography. Acta Crystallog. sect. D, 54, 1119–1131.
- 31. Kleywegt, G. J. & Read, R. J. (1997). Not your average density. Structure, 5, 1557-1569.
- 32. Kleywegt, G. J. & Jones, T. A. (1994). Halloween ...masks and bones. In From First Map to Final Model (Bailey, S., ed.), pp. 59-66, SERC, Daresbury Laboratory, Warrington.
- 33. Jones, A. T., Zou, J. Y., Cowtan, J. Y. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in the model. Acta Crystallog. sect. A, 47, 110-119.
- 34. Lamzin, V. S. & Wilson, K. S. (1993). Automated refinement of protein models. Acta Crystallog. sect. D, 49, 129-149.
- 35. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallog. sect. D, 50, 760-763.
- 36. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22, 4673-4680.
- 37. Laskowski, R. A. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallog. 26, 283-291.
- 38. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. 8, 52-56.

Edited by R. Huber

(Received 16 October 2002; received in revised form 29 January 2003; accepted 31 January 2003)

DIRECT® SCIENCE www.sciencedirect.com

Supplementary Material for this paper comprising the amino acid sequence alignment, and an animation of the reaction mechanism is available on Science Direct