Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic Chemistry 31 (2003) 278–287
BIOORGANIC CHEMISTRY
www.elsevier.com/locate/bioorg

Examination of a reaction intermediate in the active site of riboflavin synthase

Ya-Jun Zheng,[a,*] Douglas B. Jordan,[a,1] and Der-Ing Liao[b]

[a] *DuPont Agricultural Products, Stine-Haskell Research Center, 1094 Elkton Road, Post Office Box 30, Newark, DE 19714, USA*
[b] *DuPont Central Research and Development, Experimental Station, Wilmington, DE 19880, USA*

Received 20 December 2002

Abstract

The riboflavin synthase catalyzed reaction proceeds through a pentacyclic intermediate of undetermined stereochemistry. Calculations at the B3LYP/6-31G(d) level of theory indicate that the *trans* pentacyclic structure is favored over the *cis* by 3.3 kcal/mol. A model of the the *trans,* but not the *cis*, pentacycle in the enzyme active site shows good fitness and the availability of highly conserved protein residues for catalytic interactions. The model of the *trans* intermediate complements the model of the two substrates in the active site and allows for a hypothetical mechanism of the roles of specific protein residues in catalysis to be proposed.

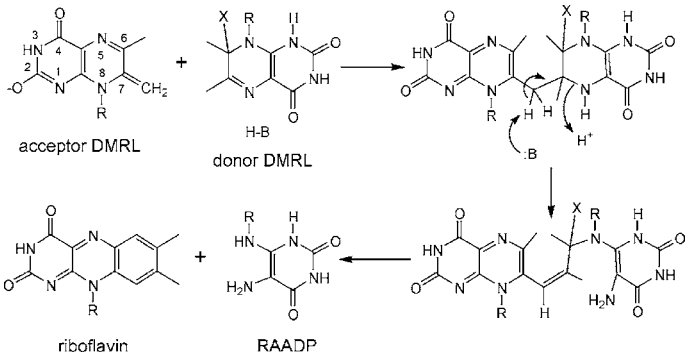
Keywords: Riboflavin synthase; Riboflavin biosynthesis; Catalytic mechanism; Enzyme mechanism; Stereochemistry; Modeling; X-ray crystallography; Structure-based design

1. Introduction

Riboflavin biosynthesis is an appealing venue for the design of antimicrobial agents owing, in part, to the requirement of a functional pathway by certain pathogens and the absence of the pathway in their mammalian hosts. Inhibitor design

* Corresponding author. Fax: +302-366-5738.
E-mail addresses: ya-jun.zheng@usa.dupont.com (Y.-J. Zheng), jordandbenz@yahoo.com (D.B. Jordan).
[1] Also corresponding author.

0045-2068/03/$ - see front matter
doi:10.1016/S0045-2068(03)00029-4

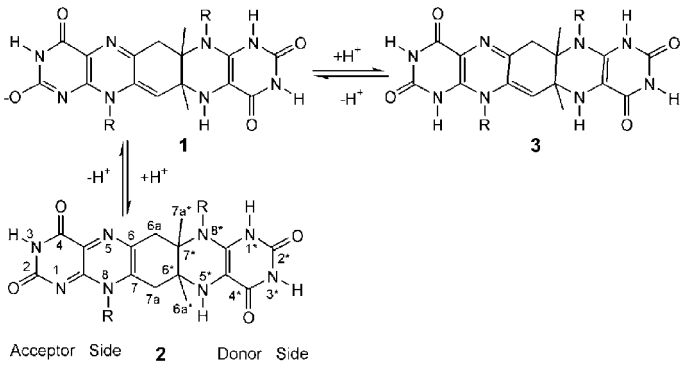


Scheme 1. An abbreviated version of proposed catalytic mechanisms of RS [10,11]. X is a nucleophile; R, a ribityl group; and B, a general acid/base. The mechanism has been generally accepted until recently; it does not accommodate the newly discovered pentacyclic intermediate.

efforts are facilitated by the availability of X-ray structures for three of the pathway enzymes [1–5]. Riboflavin synthase (RS) catalyzes the terminal step of riboflavin biosynthesis, dismutation of two molecules of 6,7-dimethyl-8-(1′-D-ribityl)-lumazine (DMRL) to yield riboflavin and 4-ribitylamino-5-amino-2,6-dihydroxypyrimidine (RAADP) [6–11]. Overall, a 4-carbon unit is transferred from one DMRL molecule (the donor DMRL that becomes RAADP) to another (the acceptor DMRL that becomes riboflavin). Reactions, on and off the enzyme, follow the same regiochemistry with a head-to-tail orientation of the two 4-carbon moieties from which the xylene ring of riboflavin is derived. The reaction mechanism, proposed over 30 years ago (Scheme 1) [8–11], has been generally accepted until recently.

Two recent developments further the prospects for inhibitor design and understanding the mechanism of RS: the determination of the X-ray structure of RS [5] and the discovery of a pentacyclic intermediate on the RS reaction pathway [12]. Although the X-ray structure of the *Escherichia coli* enzyme was determined in the absence of substrate-like ligands, the beta barrels of RS closely resemble those of 4 flavoproteins [5]. By overlaying the beta barrels of the individual flavoproteins onto those of RS and transferring the DMRL portion of their flavin cofactors to RS, a consistent view of the binding mode for DMRL in the RS active site was generated (Fig. 1A) [5]. In the model, the two DMRL molecules fit well in the enzyme active site with only minor clashes. It suggests the identities of the donor DMRL and acceptor DMRL, and has conserved[2] amino acid residues in the vicinity of the substrates to promote the reaction. A NMR structure of the N-terminal domain of the *E. coli* RS complexed with riboflavin substantiates the binding mode of DMRL predicted by the model of Fig. 1A [13]. When incubated with DMRL, a site-directed mutant of the *E. coli* RS accumulates a pentacyclic intermediate derived from two

[2] Comparisons of RS sequences refer to the alignment of 15 species of RS reported previously [5]. The terms "highly conserved" and "conserved" are used to indicate that the residues are identical in the sequence alignment.

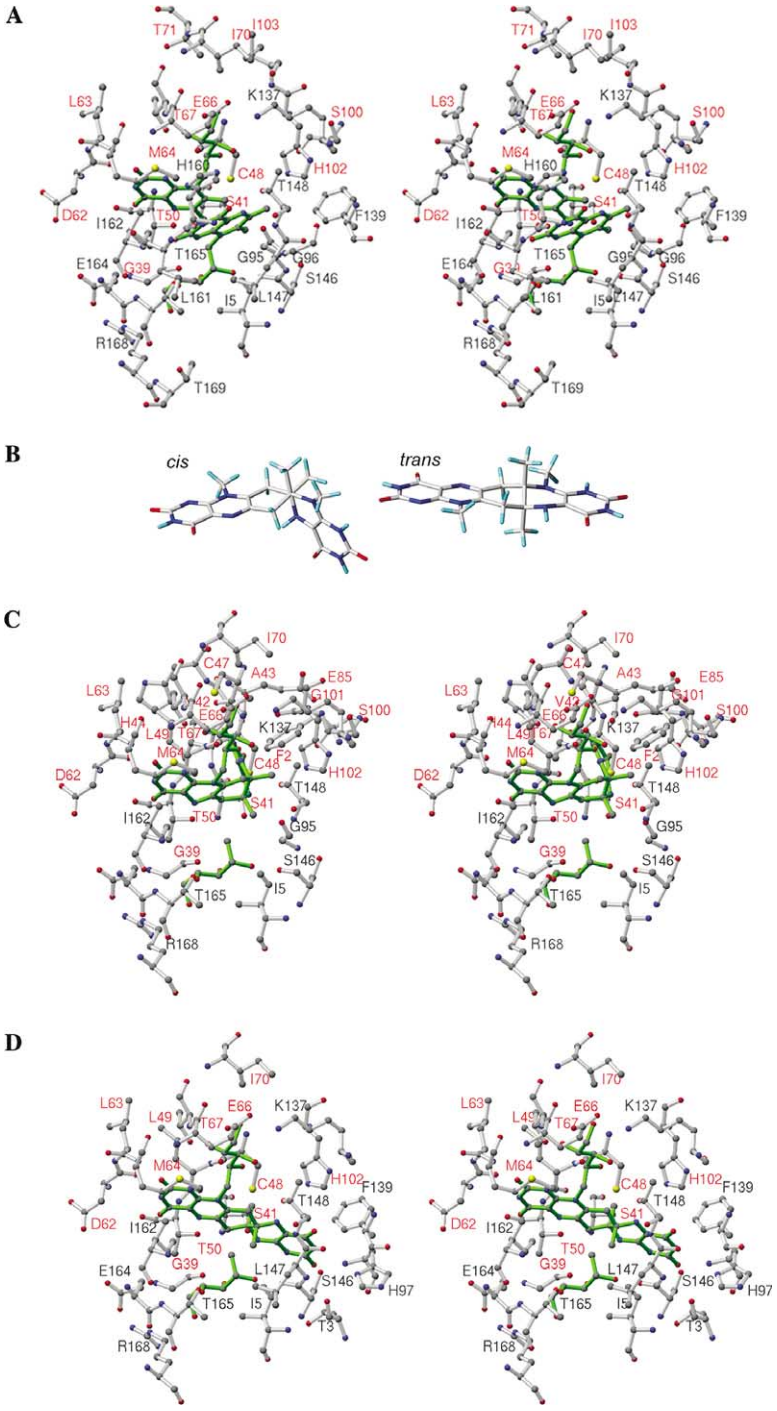


Scheme 2. Pentacyclic intermediate and its tautomers [12]. R is a ribityl group. The acceptor side atoms (the left half of the molecule as drawn) are numbered according to the convention for DMRL; the donor side (the right half) atoms are likewise numbered with the addition of an asterisk.

DMRL molecules joined in a head-to-tail fashion (Scheme 2) [12]. The intermediate was rigorously shown to belong within the catalytic cycle of the native RS, thus requiring a revision to Scheme 1.

Stereochemistry of the pentacyclic intermediate has not been determined experimentally. Recently, a crystal structure of RS from *Schizosaccharomyces pombe* complexed with the inhibitor, 6-carboxyethyl-7-oxo-8-ribityllumazine, was determined [14]. Unlike the original structure of the unliganded, native enzyme, in the crystal of the enzyme–inhibitor complex, RS exists in a catalytically incompetent monomeric state (instead of the native homotrimer) without a formed active site. After docking the pentacyclic intermediate into the formed active site of the first model (Fig. 1A) [5], it was shown that, after allowing the protein to move using molecular dynamics, the pentacyclic intermediate with a *cis* conformation can be accommo-

Fig. 1. Substrate and reaction intermediate models. (A) Stereoview model of two DMRL molecules in the active site of RS [5]. Residues within 4 Å of the substrates are shown. DMRL bonds are in green. Residues labeled in black belong to the C-terminal beta barrel. Residues labeled in red belong to the N-terminal beta barrel. (B) Calculated models of the pentacyclic intermediate. The ribityl groups were truncated to methyl groups. (C) Stereoview model of the *cis* pentacyclic intermediate in the active site of RS. The N barrel is modeled as containing the acceptor DMRL and the C barrel is modeled as containing the donor DMRL. The acceptor DMRL is kept in place. The ring of the donor DMRL is allowed to move with its ribityl group kept in place, causing the bond to be broken between the ring system and the ribityl group (distance = 4.6 Å); the broken bond between the ring of the donor DMRL and its ribityl group is apparent. Residues within 4 Å of the intermediate are shown. Coloring is the same as that in A. The protein orientation is the same as in A. (D) Stereoview model of the *trans* pentacyclic intermediate in the active site of RS. The N barrel is modeled as containing the acceptor DMRL and the C barrel is modeled as containing the donor DMRL. The acceptor DMRL is kept in place. The ring of the donor DMRL is allowed to move with its ribityl group kept in place, causing the bond to be broken between the ring system and the ribityl group (distance = 2.7 Å); the broken bond between the ring of the donor DMRL and its ribityl group is apparent. This is the best model. Residues within 4 Å of the intermediate are shown. Coloring is the same as that in A and C. The protein orientation is the same as in A and C.


A
T71
I70
I103
L63
E66
T67
K137
S100
M64
H160
C48
H102
T148
D62
T50
S41
F139
I162
T165
G95
T96
E164
G39
S146
L161
I5
L147
R168
T169
B
cis
trans
C
I70
C47
A43
E85
L63
V42
S101
S100
H44
E66
L49
T67
K137
M64
C48
F2
H102
D62
T148
S41
I162
T50
G95
G39
S146
T165
I5
R168
D
I70
L63
L49
E66
K137
T67
M64
C48
H102
F139
D62
T148
S41
I162
T50
G39
E164
L147
S146
H97
T165
I5
T3
R168

dated by the enzyme active site [14]; the *trans* configuration was not examined. However, the roles of active-site residues in catalysis are not satisfactorily accounted for in this proposal. Here we report a theoretical examination of the structure of the pentacyclic intermediate, its fitness in the active site of RS, and the potential roles of active site residues in the catalytic mechanism. Our model of the *trans* intermediate in the RS active site is consistent with available experimental data and provides a more satisfying account for the roles of active-site residues in catalysis than the models of the *cis* intermediate.

2. Materials and methods

We used a hybrid density functional theory method at the B3LYP/6-31G(d) level of theory [15,16] to investigate the structure and conformation of the pentacyclic intermediate, truncating the ribityl groups to methyl groups for the calculations. Quantum mechanical calculations were performed using the *Gaussian 98* program [17]. Solvent effects were estimated using a consistent continuum solvent model as implemented in Jaguar [18]; solvent effects were examined at two dielectric constants (2.02 for cyclohexane and 10.65 for 1,2-dichloroethane).

Docking studies on the pentacyclic intermediate in the active site were accomplished using the molecular modeling software *Sybyl* 6.8 [19]. It is well known that, given enough flexibility in molecular dynamics simulations (or molecular docking/scoring), even poor-binding ligands can be made to fit well in enzyme binding pockets. Appropriate treatments of protein flexibility, solvation effects, and electrostatic interactions are some of the critical factors in molecular dynamics simulations. Since RS is a trimeric protein, appropriate molecular dynamics simulation would be rather difficult. Instead of performing a restrained molecular dynamics simulation, we take a different approach. To avoid inappropriate distortion of the protein structure that could arise from using molecular dynamics, in our models, we hold the protein structure fixed to its X-ray coordinates for assessing how different isomers of the pentacyclic intermediate bind. Even though both protein and substrate are normally expected to undergo some conformational changes during catalysis, our model for RS nevertheless will be a reasonable starting point for examining the catalytic mechanism. The pentacyclic intermediate was manually docked into the active site by overlaying the DMRL portion of the ring system of the intermediate onto one of the DMRL rings of Fig. 1A; this was accomplished by using the Fit Atoms routine within *Sybyl* 6.8.

3. Results

The intermediate has two chiral centers and four possible stereoisomers (*RR*, *SS*, *RS*, and *SR*). *RR* and *SS* are mirror images with equivalent energies as are *RS* and *SR*. In the calculated structures (shown in Fig. 1B), the pentacyclic ring of the *trans* isomer (*RR* or *SS*) of **2** is 3.3 kcal/mol lower in energy than the *cis* isomer (*RS* or

SR) of **2**. To examine how the protein environment may change the relative energy difference, continuum solvation calculations were carried out using the gas phase geometries. The calculated energy differences (*cis* minus *trans*) are 2.8 and 2.2 kcal/mol in hexane ($\varepsilon = 2.02$) and dichloroethane ($\varepsilon = 10.65$), respectively. Although the *trans* configuration is favored over the *cis* in all the calculations, the energy differences are not large enough to exclude the *cis* configuration from consideration as a viable possibility for being the reaction intermediate in the enzyme-catalyzed reaction.

Notably, the *trans* form of tautomer **3** is 8.3 kcal/mol higher in energy than the *trans* **2**. This energy difference is smaller than that for other keto-enol tautomerizations (e.g., ~13 kcal/mol) [20], suggesting that the C7a–H is more acidic than a C–H group adjacent to a typical carbonyl group. Indeed, a pK_a of ~8.5 has been determined for the analogous C7a methyl group of DMRL [8,11].

RS is a homotrimer of 23 kDa subunits. The monomer comprises two similar beta barrels (the N barrel formed from N-terminal residues and the C barrel formed from C-terminal residues) and a C-terminal α helix [5]. Trimer contacts are mainly through the C-terminal α helices. The active site shares the N barrel of one subunit and the C barrel of another subunit. Only one active site can be formed at a time with two subunits merging; the other two potential active sites remain unformed with their subunits far apart. The formed active site defined by the X-ray structure [5] was used for modeling DMRL in Fig. 1A and below for modeling the pentacyclic intermediate.

There are four straight-forward ways to model the intermediate into the active site of RS using the model of Fig. 1A without assuming the acceptor and donor DMRL identities and without moving the protein residues: overlay the DMRL portion of the ring system of the intermediate (either in the donor or acceptor orientation) onto one of the DMRL rings of Fig. 1A at a time (2 ways) while retaining the connectivity of one of the DMRL ribityl groups to the intermediate at a time (2 ways); i.e., the acceptor DMRL could be either of the two DMRL molecules of Fig. 1A (2 ways) and either the donor or acceptor DMRL rings of Fig. 1A would need to move to form the intermediate leaving its ribityl group behind (2 ways).

In its four models, the *cis* pentacyclic isomer is not accepted well by the RS active site—it overlaps with protein backbone atoms and highly conserved residues of RS are remote from the reaction site (e.g., see Fig. 1C).[3] Inspection of the models of the *trans* isomer reveals that there is only one that has a satisfactory fitting within the RS active site; the other 3 models show considerable overlaps with the protein and have the conserved residues out of place for reactivity.[3] In the best *trans* model (Fig. 1D), the acceptor DMRL belongs to the N barrel and retains positioning of its DMRL ribityl group; the donor DMRL belongs to the C barrel, and its ribityl group must move by about 1 Å to accommodate the formation of the intermediate. Such ribityl movement could occur readily as no protein atoms are in its way. In proceeding from the model of the two DMRL molecules in Fig. 1A to that of the *trans* pentacyclic

[3] Coordinates for all models discussed in this work can be obtained from the corresponding authors.

intermediate in Fig. 1D, it appears that minor clashes might occur with the protein: i.e., even though this part of the protein has a good deal of open space, some movement of Ser146, Leu147, and Thr148 is needed for the ring system of donor DMRL to swing from its position in Fig. 1A to its new position in Fig. 1D as part of the pentacyclic intermediate. Such movement of the residues is not necessarily a requirement for the conversion of the two reactants to the pentacyclic intermediate on the enzyme, given that Figs. 1A and D are models approximating the two states.

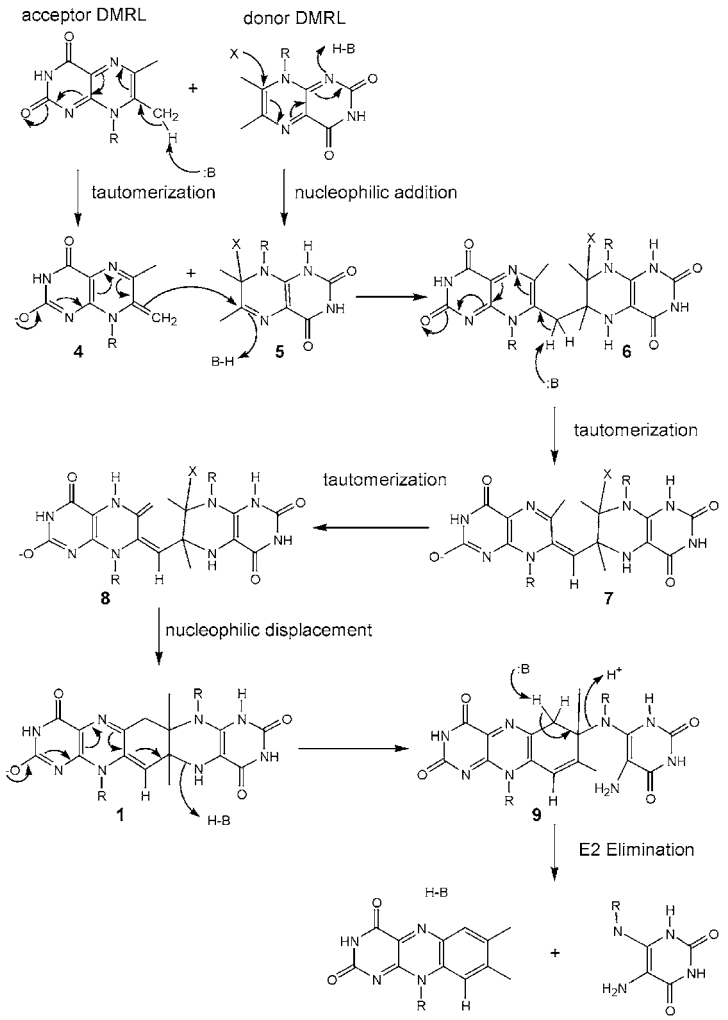
Importantly, the model of Fig. 1D is in accord with the tentative assignment of the donor DMRL and acceptor DMRL made previously on the basis of their orientation in the modeled substrate complex where the methyl groups of the (acceptor) DMRL belonging to the N barrel are in position to attack the ring system of the (donor) DMRL belonging to the C barrel as shown in Fig. 1A [5]. Also, the model of Fig. 1D has protein residues that are highly conserved among biological species in the vicinity of the intermediate, [2] including Ser41, Cys48, His97, His102, Ser146, and Thr148. Thr50 is also in the vicinity; among species, it is either a Thr or Ser. Both the *SS* and the *RR* stereoisomers fit well in the model of Fig. 1D, but only the *SS* stereoisomer of the *trans* intermediate satisfies the required stereochemistry of the nucleophilic displacement reaction (vide infra). The *SS* stereoisomer is shown in Fig. 1D.

4. Discussion

Immediate implications arising from our structural modeling are that the *trans*, but not the *cis*, penatcyclic structure is viable as the intermediate; only one of the models of the *trans* isomer is viable when considering the accessibility of the intermediate to conserved residues for participation in catalysis; and the N barrel constitutes the acceptor binding site (its DMRL becomes riboflavin) and the C barrel constitutes the donor binding site (its DMRL becomes RAADP). Thus, Fig. 1A approximates the Michaelis complex, preceding the junction of two DMRL molecules, and Fig. 1D approximates the bound pentacyclic intermediate following the junction.

Having the models of Figs. 1A and D in place allows us to propose a detailed molecular mechanism (Scheme 3). Although the scheme itself is similar to that proposed previously [12], our elaboration on the detailed roles of specific amino acid residues does constitute a new proposal. According to our hypothesis, the acceptor DMRL is first deprotonated by the His102/Thr148 dyad (B:).[4] The developing negative charge of the resulting **4** is stabilized by a hydrogen bonding interaction between the C2 carbonyl O and a backbone amide NH group formed between Leu63 and Met64 (not conserved residues). The donor molecule is activated to **5** by nucleophilic attack at

[4] In all cases where the dyad is proposed as a general acid/base, the hydroxyl group of Thr148 is thought to interact with the substrate and intermediates because of its proximity as modeled ($\sim$3 Å between heavy atoms).



Scheme 3. Hypothetical mechanism of the RS catalytic cycle. X is either Cys48 or a water; R, ribityl group; and B, a general acid/base.

C6 by the thiol of Cys48 with concomitant protonation at N1*; the proton comes from the solvent. Although it may appear from Fig. 1A that the acceptor DMRL might shield the SH group of Cys48 from attacking the ring system of the donor DMRL, the model of Fig. 1A has enough space for the SH group to approach the ring system of the donor DMRL without hindrance from the acceptor DMRL. Nucleophilic attack by **4** at C6* of the **5** imine is activated by the nearby, protonated His102/Thr148 dyad (B–H), providing **6** and returning His102 to the neutral form. Two consecutive tautomerizations follow next. The first (**6**–**7**) is initiated by the His102/Thr148 dyad (B:) and assisted by the hydrogen bond between the C2–O

and the backbone amide of Leu63–Met64. Thr50 is the obvious active site residue available to promote the second tautomerization yielding **8**. Intermediate **1** is formed from **8** through an intramolecular nucleophilic displacement of the Cys48 at C7*. Breakdown of **1**–**9** to riboflavin and RAADP requires protonation of N5* and N8*. Breaking the C6*···N5* bond is assisted by the protonated His102/Thr148 dyad, while breaking the C7*···N8* bond is via an E2 elimination reaction initiated by Cys48 and assisted by Ser41.

The proposed catalytic mechanism has two deprotonations of C7a. The C7a methyl group has a pK_a of ~8.5 [8,11], and the His102/Thr148 dyad is a suitable base for such deprotonations. It also specifies a S_N2 nucleophilic displacement and an E2 elimination. Both reactions have specific stereoelectronic requirements. It should be noted that the nucleophilic displacement could be either a S_N2 or S_N1 (e.g., elimination of X to give the imminium ion followed by a nucleophilic addition) process; there is no evidence to exclude either possibility. In either case, the nucleophilic displacement (or addition) occurs via an inversion of configuration at the carbon center (C7*). The approaching C6a attacks C7* from the opposite side of Cys48, suggesting that, in the resulting intermediate **1**, the C7a* methyl should be on the same side as Cys48. As to the E2 elimination, anti-elimination is normally favored, and this means that the leaving amine should be on the opposite side of the ring as the base (Cys48). Our model of the *SS* stereoisomer in the RS active site (Fig. 1D) does imply that both requirements are met.

The proposed role for Ser41 accounts for the Ser41Ala mutation causing accumulation of the pentacyclic intermediate: Ser41 does not appear to have a role in the formation of the pentacyclic intermediate so mutation to Ala should not impede its formation; yet, the hydroxyl group of Ser41 is considered critical for the breakdown of the intermediate in the direction of riboflavin formation. The mutation erodes V_{max} for riboflavin formation by 7000-fold [21]. The His102Asn mutation erodes V_{max} for riboflavin formation by 260-fold [21]. In our model, the imidazole group of the highly conserved His102 is proposed to have general acid/base roles in the formation and breakdown of the pentacyclic intermediate in the forward reaction trajectory (acting through Thr148 of the dyad). Our model has the ring of the highly conserved Phe2 serving to hold the thiol group of Cys48 in a reactive position; the truncation mutant, lacking Met1 and Phe2, has no measurable catalytic activity [5,21]. The Cys48Ala mutant is catalytically inactive and the Cys48Ser mutant is 5-fold less active than the wild-type enzyme [14], consistent with the view that the native residue serves as the nucleophile for initiating the catalytic cycle (X of Schemes 1 and 3). Mutations of Thr148 and Thr50 have not been examined.

A salient feature of the proposed RS mechanism is that the early stages of the reaction cannot commence until two subunits merge, bringing together residues from the C barrel and the N barrel to form the active site: e.g., the His102/Thr148 dyad is composed of His102 from the N barrel and Thr148 from the C barrel, and Cys48 from the N barrel acts upon the donor DMRL of the C barrel. In this way the catalyst prevents generation of reactive species of DMRL until the acceptor DMRL and the donor DMRL are close to one another for making riboflavin.

From one perspective, it is appealing to model the *cis* penatcylic structure into the RS active site because the model of the two DMRL molecules in the RS active site has the two attacking methyl groups of the acceptor DMRL on the same side of the donor DMRL ring system (Fig. 1A). Thus, dimerization of the two DMRL molecules yielding the *cis* intermediate would require the least amount of movement among the two substrates. However, the *cis* intermediate does require considerable movement of the protein in order to fit within the active site. Also, because the dimerization is a two step process with respect to the attacking methyl groups, there is no requirement that the intermediate has *cis* stereochemistry. Indeed, in this work, we have demonstrated that the *trans* intermediate can be well accommodated by the RS active site without moving the protein. Furthermore, the model of the *trans* intermediate has conserved, active-site residues in good position for participating in catalysis, but the model of the *cis* intermediate does not. Although, the *cis* intermediate cannot be ruled out, its model in the RS active site is considerably less satisfying than that of the *trans* intermediate. Our proposed catalytic mechanism, based on the pentacyclic intermediate having the *trans* configuration, is consistent with all available experimental observations. From this framework, further studies can design active-site inhibitors and probe roles of active-site residues in catalysis.

References

[1] D.-I. Liao, Y.-J. Zheng, P.V. Viitanen, D.B. Jordan, Biochemistry 41 (2002) 1795–1806.
[2] R. Ladenstein, M. Schneider, R. Huber, H.D. Bartunik, K. Wilson, M. Schott, A. Bacher, J. Mol. Biol. 203 (1988) 1045–1070.
[3] K. Ritsert, R. Huber, D. Turk, R. Ladenstein, K. Schmidt-Base, A. Bacher, J. Mol. Biol. 253 (1995) 151–167.
[4] K. Persson, G. Schneider, D.B. Jordan, P.V. Vittanen, T. Sandalova, Protein Sci. 8 (1999) 2355–2365.
[5] D.-I. Liao, Z. Wawrzak, J.C. Calabrese, P.V. Viitanen, D.B. Jordan, Structure 9 (2001) 399–408.
[6] T. Rowan, H.C.S. Wood, Proc. Chem. Soc. London (1963) 21–22.
[7] T. Rowan, H.C.S. Wood, J. Chem. Soc. Commun. (1969) 452–458.
[8] T. Paterson, H.C.S. Wood, J. Chem. Soc. Chem. Commun. (1969) 290–291.
[9] R.L. Beach, G.W.E. Plaut, Tetrahedron Lett. (1969) 3489–3492.
[10] R.L. Beach, G.W.E. Plaut, J. Am. Chem. Soc. 92 (1970) 2913–2916.
[11] G.W.E. Plaut, R. Beach, T. Aogaichi, Biochemistry 9 (1970) 771–785.
[12] B. Illarionov, W. Eisenreich, A. Bacher, Proc. Natl. Acad. Sci. USA 98 (2001) 7224–7229.
[13] V. Truffault, M. Coles, T. Diercks, K. Abelmann, S. Eberhardt, H. Luttgen, A. Bacher, H. Kessler, J. Mol. Biol. 309 (2001) 949–960.
[14] S. Gerhardt, A.-K. Schott, N. Kairies, M. Cushman, B. Illarionov, W. Eisenreich, A. Bacher, R. Huber, S. Steinbacher, M. Fisher, Structure 10 (2002) 1371–1381.
[15] A.D. Becke, J. Chem. Phys. 98 (1993) 5648–5652.
[16] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B: Condens. Matter 37 (1988) 785–789.
[17] Gaussian 98, revision A.11. Gaussian, Inc., Pittsburgh, PA, 1998.
[18] Jaguar 4.1, Schrodinger, Inc., Portland, Oregon, 2000.
[19] Sybyl Molecular Modeling Software, version 6.8. Tripos, Inc., St. Louis, MO, 2001.
[20] J.A. Gerlt, P.G. Gassman, J. Am. Chem. Soc. 115 (1993) 11552–11568.
[21] B. Illarionov, K. Kemter, S. Eberhardt, G. Richter, M. Cushman, A. Bacher, J. Biol. Chem. 276 (2001) 11524–11530.