NMR Analysis of Site-Specific Ligand Binding in Oligomeric Proteins. Dynamic Studies on the Interaction of Riboflavin Synthase with Trifluoromethyl-Substituted Intermediates†

Johannes Scheuring,‡,§ Markus Fischer,‡ Mark Cushman,|| John Lee,⊥ Adelbert Bacher,‡ and Hartmut Oschkinat*,▽

Department of Organic Chemistry and Biochemistry, Technical University of Munich, Lichtenbergstrasse 4, D-85747 Garching, Federal Republic of Germany, Department of Medicinal Chemistry and Pharmacognosy, Purdue University, West Lafayette, Indiana 47907, Department of Biochemistry, Boyd Graduate Studies Research Center, University of Georgia, Athens, Georgia 30602, and Forschungsinstitut für Molekulare Pharmakologie, Alfred-Kowalke-Strasse 4, D-10315 Berlin, Federal Republic of Germany

Received January 16, 1996; Revised Manuscript Received April 18, 1996⊗

ABSTRACT: The binding of small ligands to symmetrical oligomeric proteins may lead to a number of different partially ligated intermediates but should finally yield a symmetrical fully ligated enzyme/ligand complex. In the case of the trimeric protein, riboflavin synthase, some ligands form an unexpected protein/ligand complex, even in the presence of a large excess of ligand. Three different bound forms were observed by ^{19}F NMR spectroscopy, and Scatchard-type analysis suggested binding sites of similar affinities. NOESY analysis of the kinetic network revealed that the three bound states exchange with free ligand, but not with each other, thus suggesting that the trimeric enzyme could be asymmetrical. This information permits appropriate precautions to be taken during X-ray structure analysis of riboflavin synthase, which is in progress. Quantitative analysis of the NOESY spectra yielded different rate constants for the different binding sites. For comparison, the monomeric lumazine protein was investigated as an example of a case with simple two-site exchange. For such systems, all kinetic parameters including k_{on} and the dissociation constant can be determined from the NOESY spectrum. The data show that NMR spectroscopy can produce qualitative and quantitative information in cases of nonequivalent binding sites in oligomeric proteins if isolated NMR signals of the different forms can be observed. The technique is not limited to ^{19}F as reporter nucleus.

Enterobacteriaceae are devoid of an uptake system for riboflavin and are therefore dependent on the *de novo* synthesis of the vitamin and its respective cofactor derivatives. This opens the possibility to utilize inhibitors of riboflavin biosynthesis for antimicrobial chemotherapy. The situation is reminiscent of the inhibition of tetrahydrofolate biosynthesis in the absence of efficient folate uptake by therapeutic agents of the sulfonamide and trimethoprim type. However, inhibitors of riboflavin biosynthesis with *in vivo* activity are not yet available, although compounds with high inhibitory activity as monitored by *in vitro* enzyme assays have been reported (Harvey & Plaut, 1966; Al-Hassan et al., 1980; Otto & Bacher, 1981).

A potential target for antimicrobial inhibitors is the terminal step of riboflavin biosynthesis, i.e., the dismutation of 6,7-dimethyl-8-ribityllumazine (**1**) yielding riboflavin (**4**) and 5-amino-6-ribitylamino-2,4(1*H*,3*H*)pyrimidinedione (**5**) (Figure 1). The reaction is catalyzed by riboflavin synthase, an enzyme which has been observed in a wide variety of microorganisms and plants [for reviews see Bacher (1990), Bacher et al. (1993), Plaut (1971), and Plaut et al. (1974)]. A mechanism for this reaction (Figure 1) was proposed by Plaut, Wood, and their co-workers (Beach & Plaut, 1970; Plaut & Beach, 1976; Paterson & Wood, 1969, 1972). It requires the simultaneous binding of two substrate molecules which serve as donor and acceptor, respectively, of a four-carbon moiety which is transferred in the dismutation reaction. It has been proposed that the reaction is initiated by the attack of a nucleophile at the donor site, yielding the intermediate **3**. The nature of the nucleophile is unknown. Indirect evidence suggests that it could be a water molecule (Cushman et al., 1993). After deprotonation of the 7-methyl group of the acceptor molecule, the resulting anion **2** could attack C-6 of **3**.

Riboflavin synthase from *Bacillus subtilis* is a trimer of identical 23 kDa subunits (Bacher et al., 1980). The primary structure has been determined by Edman degradation (Schott et al., 1990) and by DNA sequencing of the cognate gene (Mironov et al., 1989; Perkins et al., 1991). The open reading frame predicted 13 amino acids at the C-terminus which were not observed by Edman sequencing of the protein.

The sequence of riboflavin synthase is characterized by a high degree of internal sequence homology suggesting that

† This work was supported by grants from the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie, and the European Community (EC ERBCHRXCT 930166).

‡ Technical University of Munich.

§ Present address: Albert Einstein College of Medicine, Department of Biochemistry, 1300 Morris Park Avenue, Bronx, NY 10461.

|| Purdue University.

⊥ University of Georgia.

▽ Forschungsinstitut für Molekulare Pharmakologie.

⊗ Abstract published in *Advance ACS Abstracts,* June 15, 1996.

[1] Abbreviations: NOESY, nuclear Overhauser enhancement; TPPI, time-proportional phase incrementation; IPTG, isopropyl thiogalactoside; NMR, nuclear magnetic resonance; EDTA, ethylenediaminetetraacetic acid disodium salt; SDS, sodium dodecyl sulfate; DEAE, diethylaminoethyl.

S0006-2960(96)00091-8 CCC: $12.00

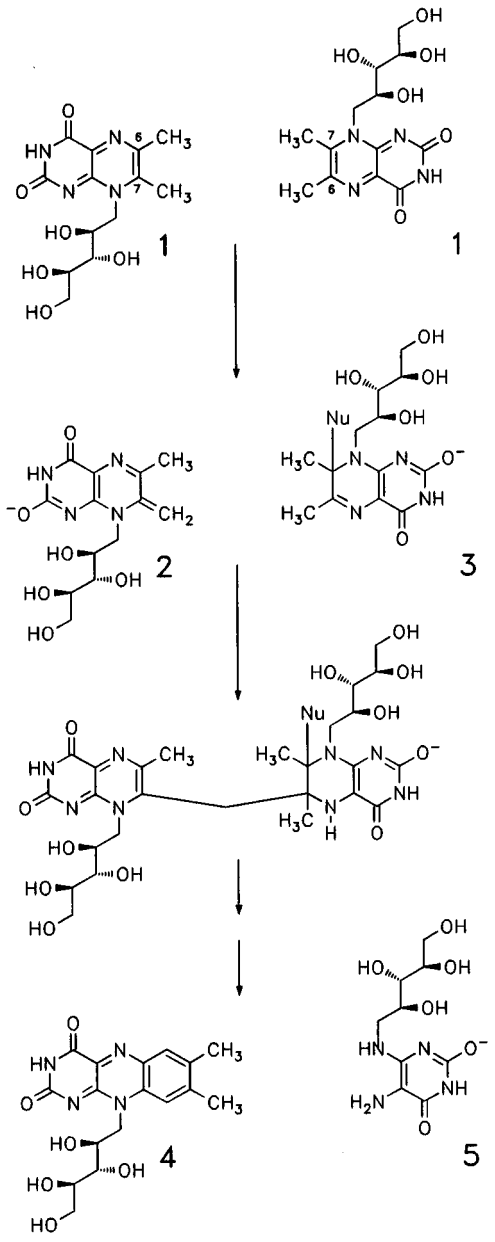

FIGURE 1: Hypothetical reaction mechanism of riboflavin synthase as proposed by Beach and Plaut (1970) and Paterson and Wood (1972). "Nu," unknown nucleophile.

the N-terminal and C-terminal halves of the protein form two structurally similar domains (Schott et al., 1990). Since the enzyme catalyzes a dismutation reaction, it appears plausible that each of the two homologous domains could provide the binding site for one of the two substrate molecules which must be present in close spatial proximity.

In order to find suitable inhibitors, Cushman and coworkers synthesized a series of trifluoromethyl analogs of putative reaction pathway intermediates of riboflavin synthase substrate (Cushman et al., 1991, 1992, 1993). For the present study, we have used the hydrate of 6,7-bis(trifluoromethyl)-8-ribityllumazine (epimer A, compound **6**, Figure 2), 6-(trifluoromethyl)-7-oxo-8-ribityllumazine (**7**), and 6-(trifluoromethyl)-7-methyl-8-ribityllumazine (**8**). Owing to the presence of the trifluoromethyl groups in **6**, the compound forms a highly stable hydrate with a chiral center at C-7. The epimers of **6** can be separated by HPLC, but the absolute configuration at C-7 is still unknown. As shown earlier, epimer A (but not epimer B) binds to riboflavin synthase

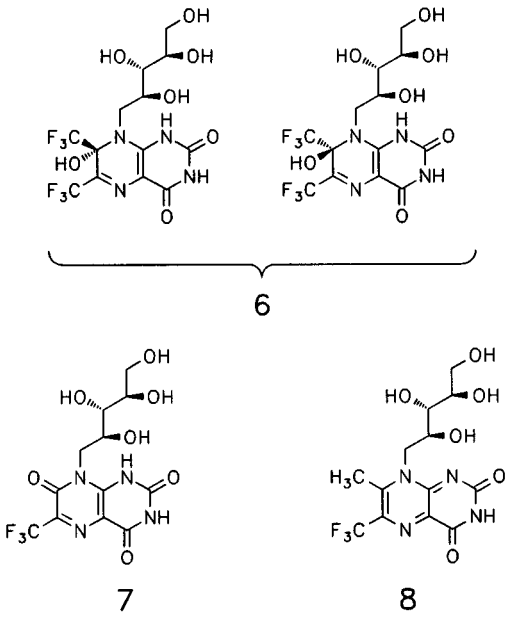

FIGURE 2: Structures of trifluoromethyl lumazines used in this study.

with a stoichiometry of one molecule per subunit (Cushman et al., 1991) whereas **7** binds with a stoichiometry of two molecules per subunit (Cushman et al., 1991, 1992). ^{19}F NMR experiments showed multiple signals for the bound ligands in each case. A titration of ligand **6** (epimer A) to riboflavin synthase showed a concerted increase of three signals due to bound forms (Cushman et al., 1992). Each of the three signals corresponds to approximately one ligand per trimeric enzyme as determined by signal integration. This could imply a non-symmetric oligomeric protein or different modes of binding for each ligand in its binding site.

This situation can be analyzed by quantitative evaluation of NOESY spectra using the relaxation matrix formalism (Johnston et al., 1989) which yields the topology of the kinetic network and, depending on the situation, provides an estimation of rate constants. The rigorous mathematical treatment is necessary to distinguish between true rate constants and effects due to "spin diffusion". We expect that this more detailed NMR analysis, in conjunction with the expected X-ray structure, will lead to an understanding of the unexpected ligand binding features of riboflavin synthase observed in the earlier, one-dimensional NMR experiments (Cushman et al., 1991, 1992, 1993). Moreover, the knowledge of the topology of the kinetic network could help to avoid potential pitfalls in the course of X-ray structure analysis, i.e., to distinguish between "pseudo-symmetry" and true symmetry. Intuitively, one would expect C_3 symmetry for a homotrimer; however, a more or less distorted trimer (i.e., a pseudosymmetric trimer) could still appear "positive" in a Patterson search for elements describing 3-fold symmetry.

It has recently been shown that riboflavin synthase forms part of a larger protein family. Other members of this family are lumazine protein and yellow fluorescent protein which serve as phototransducers in bacterial bioluminescence (O'Kane et al., 1991; O'Kane & Prasher, 1992). Both proteins show a high degree of sequence homology with riboflavin synthase. In contrast to the trimeric riboflavin synthase, the phototransducer proteins are monomers. Lumazine protein binds one molecule of the riboflavin synthase

Table 1: Bacterial Strains and Plasmids

strain or plasmid	relevant characteristics	source
E. coli strains		
XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lac1^{q}ZΔM15 Tn10(tet^r)]	Stratagene
M15[pRep4]	LacZM15,Str^R, Kan^R	Hoffmann LaRoche AG
plasmids		
pBlueScriptII SK−	high-copynumber phagemit vector	Stratagene
pNCO113	expression vector	A. van Loon
pRF2	pBR322 with 10 kb fragment containing the rib-operon	A. van Loon
pα215	pNCO113 with wild type ribB gene	this study
pα202	pNCO113 with C-terminal truncated ribB gene (202 aa)	this study
pα197	pNCO113 with C-terminal truncated ribB gene (197 aa)	this study

substrate **1** (O'Kane & Lee, 1985; Lee et al., 1991), which serves as the chromophore for bioluminescence emission. The protein also binds the fluorinated compounds **7** and **8** with an equimolar stoichiometry (Scheuring et al., 1994a,b). For this reason, lumazine protein was chosen as a test case, because a pure two-site exchange yields the parameters k_{off} and k_{on}[E] directly from the NOESY experiment when the concentration of free enzyme is known, whereas the latter value is more difficult to determine in a system showing multisite exchange.

In the course of an inhibitor design project it is desirable to minimize the dissociation rate. This parameter may be determined from ^{19}F 2D NOESY spectra, even in the case of multisite exchange.

EXPERIMENTAL PROCEDURES

Materials. The preparation of 6,7-dimethyl-8-ribityllumazine (Plaut & Harvey, 1971), 6,7-bis(trifluoromethyl)-8-ribityllumazine hydrate (epimer A of **6**), 6-(trifluoromethyl)-7-oxo-8-ribityllumazine (**7**), and 6-(trifluoromethyl)-7-methyl-8-ribityllumazine (**8**) (Cushman et al., 1991, 1992) has been described.

Restriction enzymes were from New England Biolabs (Schwalbach, Germany) and Pharmacia Biotech (Freiburg, Germany). T_4 DNA ligase was from Gibco BRL (Eggenstein, Germany). Oligonucleotides were synthesized by MWG Biotech (Ebersberg, Germany). DNA fragments were purified with the Geneclean II Kit from Bio 101 (San Diego, CA). Nucleobond AX100 columns were from Macherey und Nagel (Düren, Germany), isopropyl β-D-thiogalactopyranoside (IPTG) was from Bissendorf Biochemicals (Hannover, Germany), DEAE-cellulose DE52 was from Whatman (Maidstone, England), CM-Sephadex C50 was from Pharmacia (Freiburg, Germany), lysozyme was from Sigma (Munich, Germany), and bovine pancreas DNAse I (grade II) was from Boehringer (Mannheim, Germany).

Strains and Plasmids. Bacterial strains and plasmids used in this study are summarized in Table 1.

Restriction Enzyme Digestion of DNA. DNA was digested at 37 °C with restriction enzymes in reaction buffers specified by the supplier. The digested DNA was analyzed by horizontal electrophoresis in 0.8%−2% agarose gels.

Plasmid Construction. The construction of expression plasmids for riboflavin synthase is shown in Figure 3. Briefly, the plasmid pRF_2 was cleaved with *Hin*dIII, and a 1.3 kb fragment containing the ribB gene was subcloned into the vector BlueScriptII SK−, which had also been prepared with *Hin*dIII. The resulting plasmid pBlueI was cut with *Afl*II and *Xho*I and was ligated with a synthetic oligonucleotide B to introduce a suitable recognition site at the

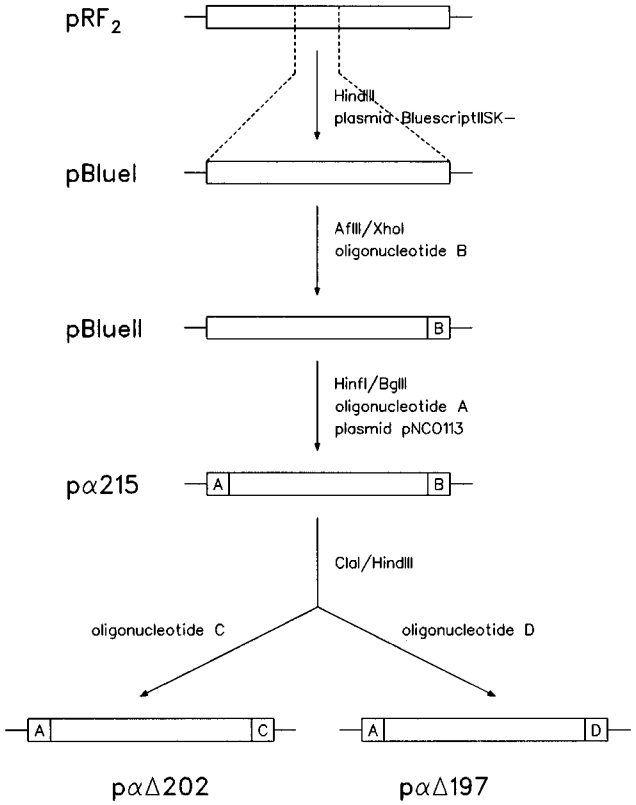

FIGURE 3: Construction of expression clones for the cloning of the riboflavin synthase gene (ribB). The sequences of synthetic oligonucleotides used are shown in Table 2. The scheme is not drawn to scale.

C-terminal end of the riboflavin synthase gene. The resulting plasmid pBlueII was cleaved with *Hin*fI and *Bgl*II, and the resulting 0.6 kb fragment together with a synthetic oligonucleotide A was cloned into a *Nco*I−*Bam*HI-digested pNCO113 expression vector yielding the plasmid pα215. In order to generate C-terminally truncated riboflavin synthases, pα215 DNA was cleaved with *Cla*I and *Hin*dIII and was subsequently ligated with synthetic oligonucleotides C or D, resulting in the respective plasmids pαΔ202 and pαΔ197. The ligated plasmids were transformed into XL1-Blue cells. Transformants were selected on LB plates containing ampicillin (150 μg/mL).

All constructs were monitored by restriction analysis and by DNA sequencing. After verification, the plasmids were transformed into the *Escherichia coli* M15 [pREP4] host strain (Stüber et al., 1990) carrying the pREP4 repressor plasmid for the overexpression of lac repressor protein. Kanamycin (25 μg/mL) and ampicillin (150 μg/mL) were added to secure the maintenance of both plasmids in the host strain.

Table 2: Oligonucleotides Used for the Construction of Expression Plasmids and for Sequencing

designation	sequence	overhangs	internal recognition sites
A	5' C**ATG**TTCACCGGCATCATCGAAACTGGTACCATCG 3' 3' AAGTGGCCGTAGTAGCTTTGACCATGGTAGCTTA 5'	*Nco*I-*Hin*fI	*Kpn*I
B	5' TTAAGCGAAAACGGCTTT**TAG**AGATCTGCGGCCGCC 3' CGCTTTTGCCGAAAATCTCTAGACGCCGGCGGAGCT	*Afl*II-*Xho*I	*Bgl*II
C	5' CGATTCCTCCACAAAGCTAACGAAAACAAAACCCAG**TAG**A 3' TAAGGAGGTGTTTCGATTGCTTTTGTTTTGGCTCATCTTCGA	*Cla*I-*Hin*dIII	–
D	5' CGATTCCTCCACAAAGCTAAC**TAG**ATATCGCGA 3' TAAGGAGGTGTTTCGATTGATCTATAGCGCTTCGA	*Cla*I-*Hin*dIII	*Eco*RV
P1	5' GTGAGCGGATAACAATTTCACAC 3'	–	–
P2	5' CTGGATCTATCAACAGGAGTC 3'	–	–

Table 3: Summary of Parameters for NMR Experiments

sample		concentration (μM)			
protein	ligand	protein	ligand	mixing time (ms)	T (°C)
riboflavin synthase	**6**	380	1390	0; 15; 30; 50; 70[a] 0; 5; 10; 15; 20[b]	29
riboflavin synthase	**7**	250	1140	0; 10; 15; 20	24
lumazine protein	**8**	165	345	0; 50; 100; 150; 200; 250	4
lumazine protein	**7**	180	310	0; 100; 200; 300	24

[a] Spectra used for the generation of the plot shown in Figure 7. [b] Spectra used for quantitative analysis.

In the expression plasmids, the riboflavin synthase genes are under the control of the T5 promotor and the lac operator. Protein expression was induced by the addition of IPTG to the lac operator.

DNA Sequencing. Sequencing was performed by the Sanger dideoxy chain termination method (Sanger et al., 1977) using a model 373A DNA sequencer from Applied Biosystems (Foster City, CA). Plasmid DNA was isolated from cultures (50 mL) of XL1-Blue strains, grown overnight in LB medium containing ampicillin (150 μg/mL), using the lysozyme/Triton X-100 method (Davis et al., 1986) and nucleobond AX-100 columns from Macherey and Nagel (Düren, Germany). Plasmids were sequenced with forward (P1) and reverse primers (P2) as specified in Table 2.

Isolation of Riboflavin Synthase. Recombinant *E. coli* strains were grown in LB medium containing ampicillin (150 μg/mL) and kanamycin (25 μg/mL). At an OD_{600nm} of about 0.7−0.9, IPTG was added to a final concentration of 2 mM. After an additional incubation for 5 h, the cells were harvested by centrifugation (Sorvall GS3 rotor, 5000 rpm, 15 min, 4 °C). Cells were washed twice with 0.9% NaCl and frozen at −20 °C.

The cell mass was thawed in 50 mM Na/K phosphate buffer, pH 8.0, containing 1 mg of lysozyme and 0.1 mg of DNase per g. The suspension was incubated for 1.5 h at 37 °C. It was centrifuged (Sorvall SS-34 rotor, 15 000 rpm, 15 min, 4 °C), and the supernatant was placed on a column of DEAE cellulose (2 × 15 cm). The column was developed with 50 mM phosphate buffer, pH. 8.0. The enzyme was not retarded on the column under these conditions. The riboflavin synthase fraction was adjusted to pH 5.5 by the addition of 0.1 M H_3PO_4, and the solution was applied to a column of CM-Sephadex (2 × 15 cm, equilibrated with 50 mM phosphate, pH 5.5). The protein was eluted from the column by a pH gradient from 5.5 to 7.0 (Schott et al., 1990). Riboflavin synthase activity was eluted at pH 6.8. Fractions were collected and concentrated by ultracentrifugation (Kontron TFT.70 rotor, 55 000 rpm, 20 h, 4 °C). Protein concentrations of pure preparations were estimated photometrically using an absorbance coefficient of $E^{1\%}_{280nm,1cm} = 3.0$ (Bacher et al., 1980).

Enzyme Assay. Assay mixtures contained 0.1 M phosphate buffer, pH 7.4, 10 mM sodium sulfite, 10 mM EDTA, 0.6 mM 6,7-dimethyl-8-ribityllumazine, and protein. They were incubated at 37 °C. At the start and at the end of incubation, aliquots were retrieved, and protein was precipitated by the addition of 15% trichloroacetic acid. The amount of riboflavin formed was determined photometrically from the increase of optical density at 470 nm (= 9100 M^{-1} cm^{-1}) (Plaut & Harvey, 1971; Bacher, 1986).

Polyacrylamide Gel Electrophoresis. Sodium dodecyl sulfate polyacrylamide gel electrophoresis was performed as described by Laemmli (1970). Molecular weight standards were supplied by Sigma (Munich, Germany).

Isolation of Lumazine Protein. Lumazine protein was isolated from *Photobacterium phosphoreum* as described earlier (O'Kane & Lee, 1985). The apoprotein was prepared by published procedures (Lee et al., 1992).

NMR Spectroscopy. Samples of riboflavin synthase contained 170 mM phosphate, pH 6.8, and 10 mM Na_3SO_3. Samples of lumazine protein contained 100 mM phosphate, pH 7.0, and 1 mM mercaptoethanol. The concentrations of proteins and ligands are summarized in Table 3. ^{19}F NMR spectra were recorded at 338 MHz or 470 MHz using AM360 and AM500 NMR instruments from Bruker Instruments (Karlsruhe, Germany).

NOESY spectra (Jeener et al., 1979) were recorded with different mixing times using the TPPI method to achieve quadrature detection. Recovery delays were 2−4 s. The number of increments was between 150 and 200. The number of scans was 32 per increment. Experimental data for the individual series of NOESY experiments are also summarized in Table 3. The spectra were transformed using a Gaussian shape function in F2 and a shifted sine bell function in F1.

Quantitative Evaluation of NOESY Spectra. In the so-called slow-exchange regime, kinetic parameters such as rate and equilibrium constants may be obtained from the NOESY spectra provided that first-order reactions are involved. In the study of enzyme/ligand systems, this condition is fulfilled when the ligands do not take part in a chemical reaction. An enzyme/ligand exchange system of the type E + L $\rightleftharpoons$ EL can then be rewritten as

$$\mathrm{L} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}[\mathrm{E}]}{\rightleftharpoons}} \mathrm{EL}$$

using the pseudo-first-order rate constant k_{on}[E] where [E] is the concentration of the free enzyme. This exchange equilibrium determines the normalized peak amplitudes (written as matrix **A**) in a NOESY spectrum according to Macura and Ernst (1980) and Johnston et al. (1986):

$$\mathbf{A} = \mathrm{e}^{-\mathbf{R}\tau_{\mathrm{m}}} \tag{1}$$

where τ_m is the NOESY mixing time and **R** is the exchange matrix, here defined as

$$\mathbf{R} = \begin{bmatrix} \varrho_{\mathrm{A}} & \sigma_{\mathrm{BA}} \\ \sigma_{\mathrm{AB}} & \varrho_{\mathrm{B}} \end{bmatrix} + \begin{bmatrix} +k_{\mathrm{on}}[\mathrm{E}] & -k_{\mathrm{off}} \\ -k_{\mathrm{on}}[\mathrm{E}] & +k_{\mathrm{off}} \end{bmatrix} \tag{2}$$

The quantities and are cross- and auto-relaxation rates defined in the usual manner (Macura & Ernst, 1980) and are assumed to be zero in the following discussion. In practical cases they can be considered to be zero when relaxation occurs on much longer time scales than chemical exchange. The matrix of normalized peak amplitudes is given by

$$\mathbf{A} = \begin{bmatrix} a_{11}/A_0 & a_{21}/B_0 \\ a_{12}/A_0 & a_{22}/B_0 \end{bmatrix} \tag{3}$$

The a_{ij} are the peak amplitudes of the NOESY spectrum, and the A_0 and B_0 correspond to the populations of the individual components involved in the kinetic network. They can be obtained from the diagonal peak amplitudes of a NOESY spectrum measured with zero mixing time. The transition probabilities may be computed by calculating the logarithm of the normalized spectrum (Johnston et al., 1986). Although k_{off} is obtained directly, values for k_{on} can only be obtained if either the concentration of the free enzyme is measured or the dissociation constant of the system is known.

Multisite exchange can be treated in the same manner, but the details depend on the type of kinetic network. A quantitative analysis would allow one to obtain proper values for k_{off}, but there are now different intermediates consisting of the trimeric enzyme with zero, one, or two ligands. The rates are then determined by a convolution of the particular on-rates and the concentrations of the different adducts or intermediates that may occur.

If the multiple signals in the ^{19}F NMR spectra arise from ligand binding to one binding site with a subsequent change in its mode of binding, only one second-order reaction occurs:

$$\mathrm{L} \underset{k_{\mathrm{off},1}}{\overset{k_{\mathrm{on},1}}{\rightleftharpoons}} \mathrm{EL}_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{EL}_2$$

This topology should be detectable by NOESY, and it should be analyzable as a two-site exchange. In this case, cross-peaks between the different bound forms should occur in the spectrum at short mixing times. If a spectrum is recorded with longer mixing times and the transition probabilities are computed according to equation 1, non-zero elements in the matrix **R** should be obtained between the different bound forms.

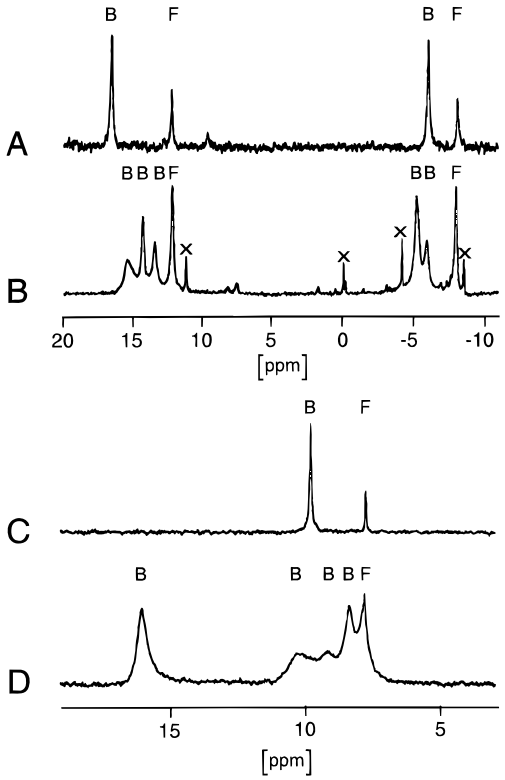

FIGURE 4: 338 MHz ^{19}F NMR spectra of trifluoromethyl-substituted lumazines bound to riboflavin synthase from *B. subtilis* and lumazine protein from *P. phosphoreum*. A, **6** and lumazine protein; B, **6** and riboflavin synthase; C, **7** and lumazine protein; D, **7** and riboflavin synthase. Signals of protein bound ligands are designated **B**, the free ligand is designated **F**. The signals of spurious contaminants (×) account for less than 2% of the total integral.

There are a number of potential sources for experimental errors in the evaluation of the NOESY spectra. At short mixing times, the cross-peaks are small, hence integration errors in the normalized peak amplitudes are large. At longer mixing times, spin−lattice relaxation effects may obscure the computation of transition probabilities. For these reasons, a series of NOESY spectra with a number of different mixing times was recorded in each case.

RESULTS

The lumazines **6** and **7** (Figure 2) are structural analogs of intermediates **3** and **2**, respectively, in the hypothetical reaction pathway of riboflavin synthase (Figure 1). They were used as probes for the mode of binding of ligands to riboflavin synthase.

Figure 4 shows the ^{19}F NMR spectra of the lumazines **6** (epimer A) and **7** in the presence of the monomeric lumazine protein or the trimeric riboflavin synthase (wild type protein from *B. subtilis*). In each case, a signal is observed for each trifluoromethyl group of the respective ligand in the free solution state. The addition of lumazine protein gives rise to one additional signal for each respective trifluoromethyl group of **6** and **7** in their respective bound states (Figure 4A,C). The addition of riboflavin synthase leads to a more complex situation in which multiple signals are observed for the trifluoromethyl groups of the protein-bound ligands (Figure 4B,D). In particular, the lumazine hydrate **6** (epimer

A), which is bound with a stoichiometry of one ligand molecule per protein subunit, shows sets of bound signals (Figure 4B) from each of the two different trifluoromethyl groups (Scheuring et al., 1994a,b; Cushman et al., 1991, 1992). Both groups contain signals from more than one bound form. The best resolved pattern is observed at 338 MHz for the downfield-shifted 6-fluoromethyl group, which shows three different signals from bound forms whose line widths differ. The number of three bound forms is confirmed by three well-resolved $^{19}F-^{19}F$ NOE between the two respective trifluoromethyl groups of each bound form. This also agrees with the results obtained by equilibrium dialysis (Cushman et al., 1991). Lumazine **7**, which has only one trifluoromethyl group, shows a similar pattern around 8−10 ppm. There is an additional signal around 16 ppm which counts for more than two trifluoromethyl groups and which we assign to the second set of three ligands binding to riboflavin synthase (Cushman et al., 1992, 1993). Importantly, riboflavin synthase binds two substrate molecules per monomer, which is necessary for the dismutation reaction to take place.

Three possible reasons for the observation of several bound forms need to be considered: (i) the preparation of riboflavin synthase could be heterogeneous; (ii) different conformations of the bound ligands in each of the binding sites on the NMR time scale could be present; (iii) the trimeric ligand−enzyme complex could be asymmetrical. These three possibilities are addressed in the following sections.

Cloning and Expression of C-Terminally Truncated Forms of Riboflavin Synthase. We considered first the possibility that the multiple signals of the fluorolumazines bound to riboflavin synthase could be due to heterogeneity of the protein preparation (Cushman et al., 1992). The riboflavin synthase gene of *B. subtilis* has a length of 645 base pairs, predicting a peptide of 215 amino acids (Perkins et al., 1991; Mironov et al., 1989). On the other hand, Edman sequencing of the protein yielded only 202 amino acid residues (Schott et al., 1990). Attempts to determine the peptide mass by ion spray or laser desorption mass spectrometry were unsuccessful. Microheterogeneity at the N-terminus was ruled out on the basis of Edman sequencing, but C-terminal microheterogeneity resulting from partial proteolysis appeared possible (Schott et al., 1990).

In order to resolve this problem, a plasmid for hyperexpression of the *B. subtilis* gene in *E. coli* was constructed (Figure 3). The vector pNCO113 was used, and the gene was expressed under control of a T5 promoter and lac operator (Stüber et al., 1990). Clones were obtained for the full-length gene and for genes with C-terminal deletion of 13 amino acids (designated Δ202) and of 18 amino acids (designated Δ197), respectively. All vector constructs were checked by dideoxynucleotide sequencing. The proteins could be expressed to a level around 30% in *E. coli*. They were purified by DEAE cellulose and CM-Sephadex chromatography as described under Methods. SDS polyacrylamide gel electrophoresis confirmed the expected mass differences (Figure 5). All recombinant proteins showed full enzymatic activity of about 45 000 nmol mg^{-1} h^{-1}.

Figure 6A shows ^{19}F NMR spectra of **7** bound to riboflavin synthase isolated from the derepressed *B. subtilis* mutant H94. Experiments with the recombinant proteins expressed in *E. coli* are shown in Figure 6B−D. Multiple NMR signals for the trifluoromethyl group of the bound ligand are

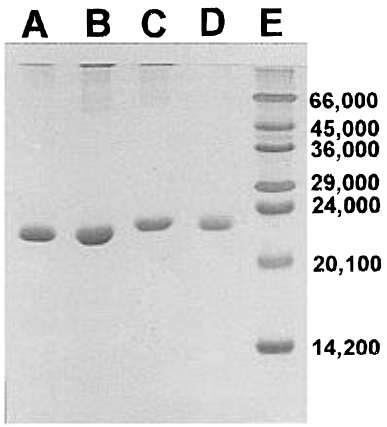

FIGURE 5: SDS−polyacrylamide gel electrophoresis of purified riboflavin synthases. Lane A, truncated enzyme Δ202; lane B, truncated enzyme Δ197; lane C, recombinant full-length protein; lane D, enzyme isolated from *B. subtilis* mutant H94; lane E, marker proteins. Molecular weights are indicated.

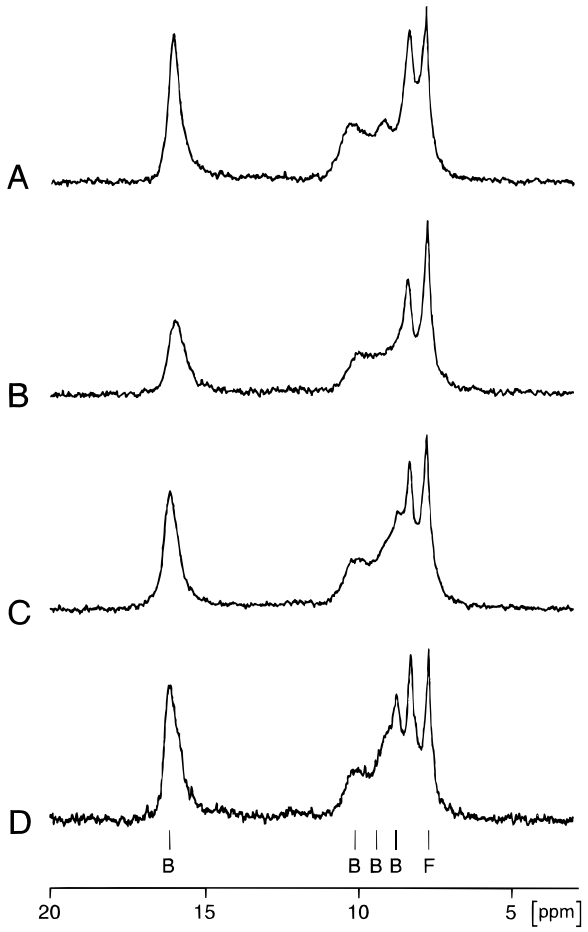

FIGURE 6: ^{19}F NMR spectra of **7** and riboflavin synthase. Spectrum A, enzyme isolated from *B. subtilis*; spectrum B, full-length recombinant enzyme expressed in *E. coli*; spectrum C, truncated protein 202, expressed in *E. coli*; spectrum D, truncated protein 197. B, enzyme-bound **8**; F, free ligand. Spectra were measured at 338 MHz in 170 mM phosphate, pH 6.8 at 24 °C.

observed with all proteins. A relatively sharp signal at 7.8 ppm increases with decreasing length of the construct. The intensities of the other signals remain essentially the same. It can be concluded that microheterogeneity is not the source for multiple signals observed in the spectra of the protein isolated from *B. subtilis* as there are only minor differences between the spectra in Figure 6.

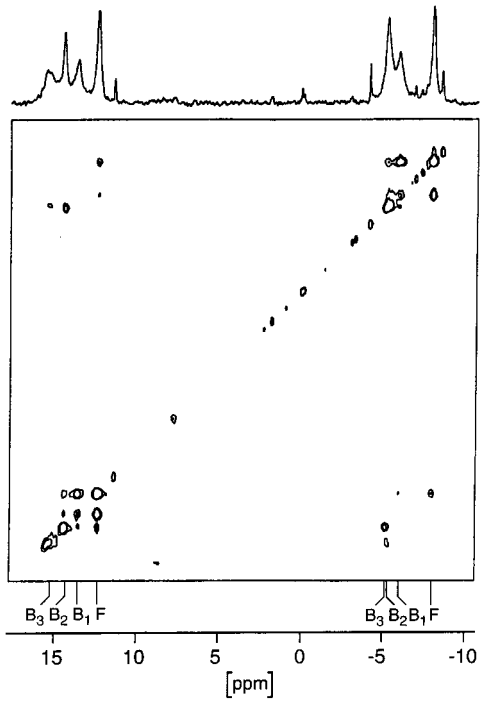

FIGURE 7: 338 MHz ^{19}F 2D NOESY spectrum of **6** and riboflavin synthase from *B. subtilis*. The sample contained 1.3 mM **6** and riboflavin synthase from *B. subtilis* (10 mg/mL) in 170 mM phosphate containing 10 mM Na_2SO_3, pH 6.8. The following parameters were as follows: mixing time, 200 ms; number of increments, 145; 256 scans per increment; relaxation delay, 2 s; 90° pulse, 5.8 μs; data set, 0.5K per increment. F, free ligand; B_1, B_2, B_3, bound ligands.

Topology of the Kinetic Matrix of Riboflavin Synthase and Ligands. The two remaining reasons for the signal patterns observed in the spectra of riboflavin synthase and the ligands **6** and **7** are either different binding modes (i.e., with different conformation or topology) for each ligand with different associated lifetimes or three structurally different binding sites. It is assumed that the knowledge of the topological network could help to understand the situation.

The effects of the chemical exchange are most apparent in the NOESY spectrum of the enzyme in the presence of **6** (epimer A) which was recorded with a mixing time of 200 ms (Figure 7). Two signal groups are observed for each of the trifluoromethyl groups. Cross-peaks of different amplitudes are observed between the free species and the bound species designated B1 and B2. At this long mixing time, signals are also observed between bound forms. One of the bound forms designated B3 shows no signal due to chemical exchange, reflecting strong binding with regard to the NMR time scale. Apart from the exchange cross-peaks, there are NOE peaks between the 6- and 7-trifluoromethyl signals of the different bound forms and of the free form in this spectrum (upper left and lower right quadrants). These peaks underline the presence of three dominantly populated bound forms, of which two peaks are superimposed in the high-field signal group.

The detection of sufficiently rapid exchange between the free form and two bound forms of **6** made it possible to determine the topology of the kinetic network. For quantitative analysis, a series of spectra was recorded with mixing times of 0, 15, 30, 50 and 70 ms. In these spectra, no NOE between the 6- and 7-trifluoromethyl groups was observed. Only cross-peaks within each respective signal group occurred, including exchange signals between bound forms at mixing times of 50 and 70 ms. No cross-peaks were detected between the free and the third bound form. The signal of the latter showed a considerably broader line shape, indicating relatively tight binding and hence, small off-rates. A plot of signal intensities at different mixing times is shown in Figure 8. The nature of the exchange signals between the bound forms was investigated by relaxation matrix analysis. The resonances at low field (Figure 7) which showed the best signal separation were used for the analysis. In case of the spectrum recorded with $\tau_m = 50$ ms, the following relaxation matrix was obtained:

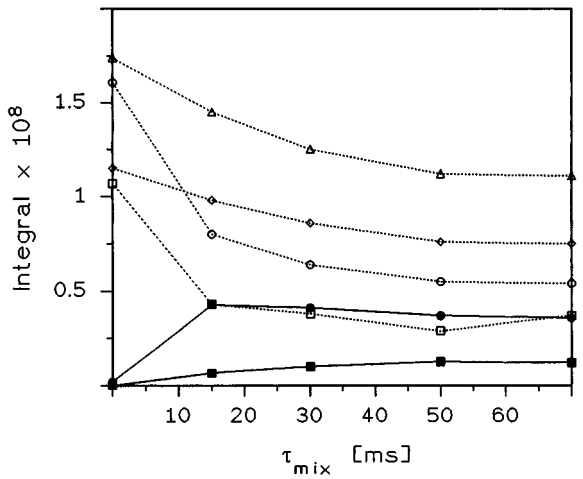

FIGURE 8: Buildup and relaxation of NOE in a sample containing riboflavin synthase and the lumazine **6** as specified in Table 3. Data from 2D ^{19}F NOESY experiments. Relaxation of the diagonal signals (- - -), free ligand; ○, bound ligand at 13.6 ppm; □, bound ligand at 14.4 ppm; △, bound ligand at 15.2 ppm. Building up curves of the exchange signal: ●, crosspeak between free ligand (12.3 ppm) and bound ligand at 13.6 ppm; ■, cross-peak between free and bound ligand at 14.4 ppm.

$$\mathbf{R} = \begin{bmatrix} 38.9 & -39.0 & -7.8 \\ -25.9 & 46.8 & 0.2 \\ -8.7 & 0.86 & 56.1 \end{bmatrix}$$

The results show that the cross-peak between the two bound forms (Figure 7) does not reflect chemical exchange between the two bound forms because only small (and even positive) values are obtained in the positions $^3/_2$ and $^2/_3$ of the matrix which contain the constants for the exchange between B_1 and B_2. Analysis of NOESY data recorded at other mixing times yielded similar results. They indicate the presence of a kinetic network in which two bound forms exchange predominantly with the free form and not with each other (Figure 9).

Measurement of Kinetic Constants. The binding of ligands **7** and **8** to the lumazine protein was investigated first as an example of a two-site exchange. The steric demands of the two ligands are similar. Similar values for the respective on- and off-rates were found (Table 4). The rate constants yielded dissociation constants similar to those observed by titration (Scheuring et al., 1994a,b). Since the titration experiment with **7** had been performed at 4 °C, it is not surprising that the NOE measurements performed at a higher temperature yielded a somewhat larger value for K_D.

As discussed above, riboflavin synthase shows multiple signals with either ligands **6** or **7**. Two of the three binding sites observed for ligand **6** show detectable exchange signals

Table 4: Kinetic Constants for Binding of Ligands to Lumazine Protein

ligand	concentration of free protein (mM)	T (°C)	τ_m (s)	k_{on} (mol s^{-1} L^{-1})	k_{off} (s^{-1})	K_D (M)	K_D (M, from titration)[a]
7	10	24	0.10	4.2×10^4	0.44	1.0×10^{-5}	3.0×10^{-6} [b]
			0.20	5.4×10^4	0.52	0.96×10^{-5}	
			0.30	5.2×10^4	0.55	1.6×10^{-5}	
8	20	4	0.10	3.4×10^4	1.11	3.3×10^{-5}	3.0×10^{-5}
			0.15	4.6×10^4	2.06	4.5×10^{-5}	
			0.20	10.3×10^4	3.30	3.2×10^{-5}	
			0.25	4.0×10^4	1.36	3.4×10^{-5}	

[a] Values determined by titration monitored by ^{19}F NMR from Scheuring et al. (1994). [b] Value determined at 4 °C.

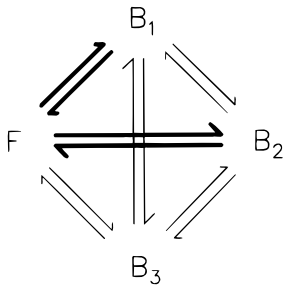

FIGURE 9: Kinetic network for binding of **6** to riboflavin synthase. F, free ligand; B_1, B_2, B_3, ligand bound to the different subunits corresponding to the three signals in the ^{19}F NMR spectrum. The observed exchanges between free and bound state are designated by bold arrows, the other possible exchanges which are not observed in the NOESY spectrum are designated by the thin arrows.

with the free ligand. Within the context of slow chemical exchange on the chemical shift time scale, one site exchanges slowly on the time scale of the NOESY experiment, one shows an intermediate rate, and one exchanges rapidly. Average values for k_{off} in the range of 55 and 8 s^{-1} are obtained for the two binding sites with lower affinity (Table 5). The values obtained at different mixing times seem to vary systematically, getting smaller as τ_m increases. The reason for this can be seen in other relaxation pathways which now become active such as cross relaxation between the trifluoromethyl groups.

An interesting change in the values for k_{off} is observed for the riboflavin synthase and ligand **7**. Now one of the two sites exchanges even slower, i.e., no cross-peaks are detected, for the other site a larger off-rate (75 s^{-1}) is observed. The exchange of a trifluoromethyl- and a hydroxyl group with an oxo group enhances the binding to one site but yields a larger value for k_{off} for the binding site that shows the weakest binding strength.

DISCUSSION

Riboflavin synthase of *B. subtilis* is a homotrimer of 23 kDa subunits. Cushman and his co-workers (Cushman et al., 1991, 1992, 1993) synthesized several trifluoromethyl lumazines with close structural similarity to presumed reaction pathway intermediates (Plaut et al., 1974; Paterson & Wood, 1972). Ligand perturbation studies monitored by ^{19}F NMR unexpectedly showed the existence of three different binding modes for the interaction of each respective compound with riboflavin synthase (Cushman et al., 1991, 1992, 1993). An approximate 1:1:1 stoichiometry for these different binding modes at saturating ligand concentrations was deduced from ^{19}F NMR signal integrals. Moreover, binding isotherms obtained by ligand titration experiments monitored by ^{19}F NMR indicated that the dissociation constants for the different binding sites had similar values (Cushman et al., 1991, 1992, 1993).

Table 5: Kinetic Constants for Binding of Ligands to Riboflavin Synthase

ligand	T (°C)	τ_m (s)	$k_{off,1}$ (s^{-1})	$k_{off,2}$ (s^{-1})
6 (epimer A)	29	0.015	86.2	4.6
		0.030	50.5	9.5
		0.050	38.0	7.8
		0.070	17.5	2.1
7	24	0.005	57.9	
		0.010	75.7	
		0.015	80.3	
		0.020	155.0	

A more detailed analysis of these unexpected binding patterns was performed by two-dimensional NOESY spectroscopy. A quantitative analysis showed that the ligands bound in the different modes exchange with the free ligand but do not exchange with each other at significant rates. Moreover, the quantitative NOESY analysis showed that the k_{off} rates for dissociation of the ligands from the different binding modes are significantly different. Since the dissociation constants K_D are similar for each binding mode, it follows that the k_{on} rates must vary in parallel with the experimentally determined k_{off} rates.

In contrast to the trimeric riboflavin synthase, the monomeric lumazine protein shows only one binding mode for each of the trifluoromethyl analogs (Scheuring et al., 1994a,b). In this case, it is possible to determine both k_{on} and k_{off} rates for the ligand by quantitative NOESY experiments. The dissociation constants, calculated from the rate constants, agreed with the values obtained from binding isotherm measurements performed by ligand titration experiments monitored by ^{19}F NMR.

The following hypotheses could account for the multiplicity of binding modes in the homotrimeric riboflavin synthase. (i) Molecular heterogeneity of the protein due to post-translational modification could artifactually lead to the appearance of different ^{19}F NMR signals. (ii) The ligands could assume different conformations or topological positions in the active site of the enzyme. (iii) Despite the assembly from three identical subunits, the protein could assume an inherently asymmetrical structure conducive to three different ligand binding sites. These possibilities are discussed in more detail below.

(i) The sequence of the *ribB* gene of *B. subtilis* coding for riboflavin synthase predicts a protein of 215 amino acid residues (Perkins & Pero, 1991; Mironov et al., 1989). Thirteen amino acids at the C-terminus had not been observed by Edman degradation of the protein (Schott et al., 1990). It therefore appeared possible that the C-terminus of the protein could be removed by post-translational

processing *in vivo*. In this case, incomplete proteolytic processing could be conducive to protein heterogeneity, resulting in the different chemical shift values for bound ligands.

In order to rule out the possibility of C-terminal heterogeneity, recombinant proteins with partial C-terminal deletions were prepared and studied by ^{19}F NMR. The C-terminally truncated proteins gave slightly different protein perturbation spectra as compared to the wild type protein, but signal multiplicity was clearly present in each case. It should also be noted that numerous batches of the full-length protein isolated from *B. subtilis* and from a recombinant *E. coli* strain reproducibly showed the same perturbation spectra. On this basis, protein heterogeneity can be ruled out as the origin of signal multiplicity.

(ii) Riboflavin synthase catalyzes a dismutation reaction requiring the binding of two identical substrate molecules in close spatial proximity. In line with this hypothesis, it has been proposed that the high-field signals observed with the ligand **8** correspond to ligand molecules bound at the donor site, whereas the low-field signal shifted by 8 ppm from the signal of the free ligand represents the ligand molecule bound at the acceptor site. This would imply that the multiple signals at higher field represent three topologically different states of the ligand bound at the donor site. The results of the NOESY analysis would imply that these topologically different states are separated by energy barriers that are high in comparison with the energy barriers between each of the three bound states and the free ligand. Since the bound states are equally populated, the hypothesis further requires that the free enthalpy of binding is almost identical for each of the hypothetical topologically different ligand states. Whereas the hypothesis of multiple topological states of the ligand at the donor site can not be ruled out on the basis of the present evidence, the coincidence of these constraints appears rather unlikely.

(iii) We must then consider the possibility that the different binding modes reflect the presence of three nonequivalent binding sites in the trimeric riboflavin synthase. Numerous examples for half site reactivity in multimeric proteins have been reported [for review see Matthews and Bernhard, (1973) and Seydoux et al. (1974)]. The binding of a first ligand to an oligomeric protein may select one of several conformationally different oligomeric states, or may induce a conformational change at the symmetry-related binding site(s). In some cases, only one binding site becomes significantly occupied and hence, reactive (e.g., malate dehydrogenase) (Tsernoglu et al., 1972). The observed nonequivalent binding sites might be caused by ligand-induced structural changes or sterical hindrance at the other binding sites. However, this leaves open the question of why the trimeric enzyme does not become symmetrical upon binding of three ligands.

A second reason could be the presence of an inherently asymmetrical trimeric enzyme. On the basis of marked internal sequence similarity, it has been proposed that each subunit of riboflavin synthase should fold into two domains with close topological similarity (Schott et al., 1990). As the two homologous domains may provide relatively similar surfaces, contacts between monomers could be formed alternatively by the respective surface of either one of the domains. This model is counterintuitive, as a C_3 symmetrical trimer would be expected, but it cannot be ruled out on the basis of the present data.

The three-dimensional structure of riboflavin synthase of *B. subtilis* is unknown. Crystals have been obtained but were unsuitable for crystallographic analysis. More recently, crystals of the enzyme from *E. coli* have been obtained with good X-ray diffraction quality, and X-ray structure analysis is in progress. The ligand perturbation ^{19}F NMR spectra of the *E. coli* protein show the presence of different ligand binding modes in analogy to the *B. subtilis* protein. Patterson analysis of the X-ray data showed the presence of noncrystallographic 3-fold symmetry or pseudosymmetry.

An unequivocal interpretation of the ligand perturbation NMR data should become possible on the basis of X-ray structure data. However, it should also be noted that the NMR data are important for the progress of the X-ray structure analysis since they indicate that problems of pseudosymmetry (inherent or ligand-induced) may arise.

Nuclear magnetic resonance is a suitable method for the study of site specific ligand binding, as it is simultaneously sensitive to dynamic processes and to structural parameters. In particular, NOESY spectra (Jeener et al., 1979) can be evaluated quantitatively to characterize various binding sites by their individual rate constants and dissociation constants and to interpret the topology of the kinetic network (Johnston et al., 1986). The ^{19}F signals of the lumazine derivatives are suitable monitors for binding, as the fluorine NMR signals of the free and bound ligands are well resolved. Although the use of fluorine facilitates this study due to the simplicity of the spectra, it is generally possible to perform such investigations with selectively ^{15}N- or ^{13}C-labeled ligand molecules or even unlabeled ligands.

ACKNOWLEDGMENT

We thank Sabine Eberhardt for DNA sequencing and Angelika Kohnle for help with the preparation of the manuscript. We also thank Dr. A. van Loon, Hoffmann La Roche AG, Basle, for plasmids.

REFERENCES

Al-Hassan, S. S., Kulick, R. J., Livingstone, D. B., Suckling, C. J., Wood, H. C. S., Wrigglesworth, R., & Ferone, R. (1980) *J. Chem. Soc., Perkin Trans. 1* 2645−2656.

Bacher, A. (1986) *Methods Enzymol. 122*, 192−199.

Bacher, A. (1990) in *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., Ed.) Vol. 1, pp 215−259, CRC Press, Boca Raton, FL.

Bacher, A., Baur, R., Eggers, U., Harders, H. D., Otto, M. K., & Schnepple, H. (1980) *J. Biol. Chem. 255*, 632−637.

Bacher, A., Eisenreich, W., Kis, K., Ladenstein, R., Richter, G., Scheuring, J., & Weinkauf, S. (1993) in *Bioorganic Chemistry Frontiers* (Dugas, H., & Schmidtchen, F. P., Eds.) Vol. 3, pp 147−192, Springer, Berlin.

Beach, R. L., & Plaut, G. W. E. (1970) *J. Am. Chem. Soc. 92*, 2913−2916.

Cushman, M., Patrick, D. A., Bacher, A., & Scheuring, J. (1991) *J. Org. Chem. 56*, 4603−4608.

Cushman, M., Patel, H. H., Scheuring, J., & Bacher, A. (1992) *J. Org. Chem. 57*, 5630−5643.

Cushman, M., Patel, H. H., Scheuring, J., & Bacher, A. (1993) *J. Org. Chem. 58*, 4033−4042.

Davis, L. G., Dibner, M. D., & Battey, J. F. (1986) in *Basic Methods in Molecular Biology*, Elsevier, New York.

Harvey, R. A., & Plaut, G. W. E. (1966) *J. Biol. Chem. 241*, 2120−2136.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
Johnston, E. R., Dellno, M., & Hendrix, J. (1986) *J. Magn. Reson. 66*, 399−409.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Lee, J., Matheson, I. B. C., Müller, F., O'Kane, D. J., Vervoort, J., & Vissert, A. J. W. (1991) in *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., Ed.) Vol. 2, pp 109−151, CRC Press, Boca Raton, FL.
Lee, J., Gibson, B. O., O'Kane, D. J., Kohnle, A., & Bacher, A. (1992) *Eur. J. Biochem. 210*, 711−719.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95−117.
Matthews, B., & Bernhard, S. A. (1973) *Annu. Rev. Biophys. Bioeng. 2*, 257−317.
Mironov, V. W., Kraev, A. S., Chernov, B. K., Ul'yanov, A. V., Golova, Yu. B., Pozmogova, G. E., Simonova, M. L., Gordeev, V. K., Stepanov, A. I., & Skryabin, K. G. (1989) *Dokl. Akad. Nauk SSR 305*, 482.
O'Kane, D. J., & Lee, J. (1985) *Biochemistry 24*, 1467−1475.
O'Kane, D. J., & Prasher, D. C. (1992) *Mol. Microbiol. 6*, 443−449.
O'Kane, D. J., Woodward, B., Lee, J., & Prasher, D. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1100−1104.
Otto, M. K., & Bacher, A. (1981) *Eur. J. Biochem. 115*, 511−517.
Paterson, T., & Wood, H. C. S. (1969) *J. Chem. Soc., Chem. Commun.* 290−291.
Paterson, T., & Wood, H. C. S. (1972) *J. Chem. Soc., Perkin Trans. 8*, 1051−1056.
Perkins, B., Pero, J. G., & Sloma, A. (1991) Eur. Pat. Appl. 0405370.
Plaut, G. W. E. (1971) in *Comprehensive Biochemistry* (Florkin, M., & Stotz, E. H., Eds.) Vol. 21, pp 11−45, Elsevier, Amsterdam.
Plaut, G. W. E., & Harvey, R. A. (1971) *Methods Enzymol. 18B*, 515−538.
Plaut, G. W. E., & Beach, R. L. (1976) in *Flavins and Flavoproteins* (Singer, T. P., Ed.) pp 737−746, Elsevier, Amsterdam.
Plaut, G. W. E., Smith, C. M., & Alworth, W. L. (1974) *Annu. Rev. Biochem. 43*, 899−922.
Sanger, F., Niklen, S., & Coulson, A. R. (1977) *Proc. Nat. Acad. Sci. U.S.A. 74*, 5463−5467.
Scheuring, J., Lee, J., Cushman, M., Patel, H. H., Patrick, D. A., & Bacher, A. (1994a) *Biochemistry 33*, 7634−7640.
Scheuring, J., Cushman, M., Oschkinat, H., & Bacher, A. (1994b) in *Flavins and Flavoproteins* (Yagi, K., Ed.) pp 75−78, Walter de Gruyter, Berlin.
Schott, K., Kellermann, J., Lottspeich, F., & Bacher, A. (1990) *J. Biol. Chem. 265*, 4204−4209.
Seydoux, F., Malhotra, O. P., & Bernhard, S. A. (1974) *CRC Crit. Rev. Biochem. 2*, 227−257.
Stüber, D., Matile, H., & Garotta, G. (1990) in *Immunological Methods* (Lefkovits, I., Pernis, B., Eds.) Vol. 4, pp 121−152, Academic Press, Orlando, FL.
Tibbelin, G., Eberhardt, S., Bacher, A., & Ladenstein, R. (1996) *Proteins: Struct. Funct., Genet.* (in press).
Tsernoglou, D., Hill, E., & Banaszak, L. J. (1972) *J. Mol. Biol. 69*, 75−87.

BI9600916