Understanding Functional Divergence in Proteins by Studying Intragenomic Homologues†

James E. Spoonamore‡ and Vahe Bandarian*,‡,§

Department of Biochemistry and Molecular Biophysics and Department of Chemistry, The University of Arizona, Tucson, Arizona 85721

Received November 13, 2007

ABSTRACT: Studies of intragenomic homologues in bacterial genomes can provide valuable insights into functional divergence. Three GTP cyclohydrolase II homologues in the *Streptomyces coelicolor* genome have been shown to catalyze two related but distinct reactions [Spoonamore, J. E., Dahlgran, A. L., Jacobsen, N. E., and Bandarian, V. (2006) *Biochemistry 45*, 12144−12155]. Two of the homologues, SCO 1441 and 2687, convert GTP to 2,5-diamino-6-ribosylamino-4(3***H***)-pyrimidinone 5′-phosphate (APy); one of the homologues (SCO 6655) produces 2-amino-5-formylamino-6-ribosylamino-4(3***H***)-pyrimidinone 5′-phosphate (FAPy). We show herein that the differences in the fate of GTP in SCO 6655 relative to SCO 1441 and 2687 results from a single amino acid substitution in the active site of the protein: a Tyr residue in the active sites of SCO 1441 and SCO 2687 is replaced with a Met in SCO 6655. Site-directed interchange of this residue in the three *S. coelicolor* intragenomic homologues is necessary and sufficient for interconversion of catalytic function which, except for SCO 1441, occurs with little loss of catalytic efficiency. Furthermore, we show that of 14 additional site-directed variants at this position of SCO 6655, His confers catalytic efficiency within 1 order of magnitude of that of the wild type and supports conversion of GTP to both FAPy and APy. The results demonstrate a clear set of mutational events that permit GCH II to produce either FAPy or APy. These results highlight a mechanism whereby functional divergence can be achieved in enzymes that catalyze multistep transformations.

The mechanisms that underlie divergence of catalytic functions are complex, but it is generally accepted that Nature can borrow mechanistic strategies from proteins that catalyze related reactions and that mutation and natural selection lead to the emergence of the new catalysts (*1*, *2*). Intragenomic duplications are a common feature of many bacterial genomes (*3*) where the tendency to retain these genes, in light of the propensity for genome compaction in bacteria, suggests that the proteins that are encoded by them play essential roles for the organism. Intragenomic duplicates are a rich source of proteins from which one can construct case studies on divergence. Currently, there are multiple ongoing structural genomics and genome sequencing projects that are generating large amounts of protein sequence similarity-based annotations; however, a distinct danger of reliance on the annotations alone is that subtle changes in active site residues can potentially have drastic effects on activities of proteins (*4*). An approach utilizing bioinformatics as well as biochemical and structural data on intragenomic duplicates can reconstruct the mutational event(s) that leads to altered enzymatic activities. Any general principles that are gleaned from these studies should be applicable to proteins that catalyze similar reactions in any organism.

Intragenomic duplicates account for as much as 41% of the genes of *Streptomyces coelicolor* A3(2); of the ∼8000 proteins encoded by the genome of *S. coelicolor*, ∼3000 open reading frames (orf's) represent intragenomic homologues (*3*, *5*). We have become interested in three *S. coelicolor* genes, *sco1441*, *sco2687*, and *sco6655*, which encode homologues of GTP[1] cyclohydrolase II (GCH II) in this organism. GCH II catalyzes the first step in the biosynthesis of riboflavin. Each of the proteins encoded by these orfs has substantial sequence identity (>40% over the GCH II coding regions) with the *Escherichia coli* protein (B1277), which has been studied extensively (*6*−*10*). However, the three GCH II homologues of *S. coelicolor* are located in disparate regions of the chromosome, and genomic contexts suggest that only one of these proteins (SCO 1441) can be involved in the biosynthesis of riboflavin. The other two are in novel genome contexts; we have hypothesized that of the two remaining proteins, one (SCO 2687) may be involved in the biosynthesis of toxoflavin, which is produced from GTP in

† Support from the National Institutes of Health (NIH) (GM 72623) to V.B. is gratefully acknowledged. J.E.S. was supported by a NIH predoctoral training grant (GM 08804). In addition, the research of V.B. is supported (in part) by a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund.

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biophysics, 1041 E. Lowell St., Tucson, AZ 85721. Phone: (520) 626-0389. Fax: (520) 621-1697. E-mail: vahe@email.arizona.edu.

‡ Department of Biochemistry and Molecular Biophysics.

§ Department of Chemistry.

[1] Abbreviations: APy, 2,5-diamino-6-ribosylamino-4(3***H***)-pyrimidinone 5′-phosphate; GCH, GTP cyclohydrolase; GTP, guanosine 5′-triphosphate; FAPy, 2-amino-5-formylamino-6-ribosylamino-4(3***H***)-pyrimidinone 5′-phosphate; Tris, 2-amino-2-(hydroxymethyl)propane-1,3-diol; ICP-OES, inductively coupled plasma optical emission spectroscopy.

10.1021/bi702263z CCC: $40.75
Published on Web 02/19/2008

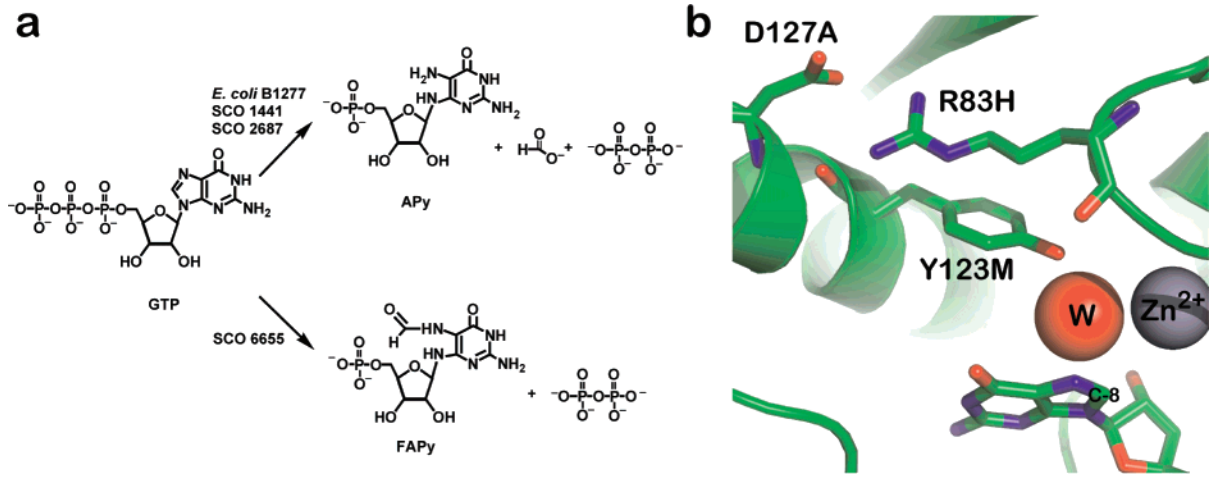

FIGURE 1: (a) Reactions catalyzed by three GCH II homologues of *S. coelicolor* and *E. coli* GCH II (B1277) and (b) active site of the GCH II protein from *E. coli*. The X-ray crystal structure of *E. coli* GCH II (PDB entry 2bz0) showing residues that are conserved in all enzymes that catalyze conversion of GTP to APy, but different in SCO 6655. Gly209 was not visible in the structure. Residue numbers refer to the SCO 2687 sequence. Panel b was generated with Pymol (*33*).

a series of steps that are likely initiated by GCH II (*11*−*14*). The biological role of the cluster that bears the third GCH II homologue (SCO 6655), however, has remained elusive. Intriguingly, despite substantial degrees of sequence identity between these proteins, one catalyzes a related but distinct transformation.

GCH II catalyzes the conversion of GTP to 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5′-phosphate (APy) (Figure 1a). GCH II also has a phosphatase activity and converts GTP to GMP during abortive turnover cycles (*9*, *10*, *15*). While GMP is not a substrate, the phosphatase activity is thought to be the first chemical step of the transformation of GTP to APy (*9*, *10*). While 8-oxo-GTP is not a substrate, it is dephosphorylated by GCH II (*16*). SCO 6655 catalyzes the conversion of GTP to 2-amino-5-formylamino-6-ribosylamino-4(3***H***)-pyrimidinone 5′-phosphate (FAPy) (Figure 1a); by contrast, SCO 1441 and 2687 produce APy. Previous site-directed mutagenesis studies have shown that a methionine residue at position 120 of SCO 6655 dictates the fate of GTP: replacement of the Met with a Tyr residue, which is found in the analogous position of other GCH II proteins (Y326 and Y123 in SCO 1441 and SCO 2687, respectively), leads to acquisition of GTP → APy activity in SCO 6655 (*11*). The X-ray crystal structure of the *E. coli* protein shows that the corresponding residue (Y105) is located in the active site of the protein 4.1 Å from N-7 of the substrate (Figure 1b) and forms a hydrogen bond to a water molecule that coordinates to the active site zinc (*17*). This water molecule has been postulated to be involved in the reaction catalyzed by the protein. The phenotype of the site-directed variant of SCO 6655 is remarkable because it suggests that this sequence signature could permit one to detect other GCH II homologues that catalyze the conversion of GTP to FAPy instead of the canonical reaction (GTP → APy). Moreover, if the reciprocal effect can be shown in the two GTP cyclohydrolase II proteins of *S. coelicolor* that catalyze the canonical enzymatic transformations, one would have identified a clear mutational event that is both necessary and sufficient for functional divergence in this family.

Changes in SCO 6655 residues that are conserved in other GCH II proteins are not limited to the Tyr−Met exchange. In this study, we have examined three additional residues that are conserved in GCH II proteins generally but different in SCO 6655. The data show that these residues do not substantially alter the course of turnover, while the Tyr−Met interchange accounts completely for the change in the fate of GTP. Furthermore, an exhaustive site-directed mutagenesis study shows that only two amino acids, Tyr and His, best support the canonical GCH II reaction (GTP → APy). The results confirm that, indeed, this residue serves as a switch in these proteins and provides a signature sequence that can be used for assignment of catalytic function.

MATERIALS AND METHODS

Preparation of Site-Directed Variants of GCH II Homologues of S. coelicolor. The variants were prepared using mutagenic primers listed in Table 1 of the Supporting Information and a QuickChange mutagenesis kit (Stratagene) essentially as described by the manufacturer except that 10% (v/v) dimethyl sulfoxide was included to facilitate melting of the high-G+C *S. coelicolor* DNA. The presence of the desired mutation was confirmed by DNA sequencing (University of Michigan DNA Sequencing Core Facility, Ann Arbor, MI). All PCR templates were based on pET 29 or pET 28 vectors for expression of native or His_6-tagged protein, respectively.

Overexpression and Purification of Wild-Type and Site-Directed Variants of GCH II from S. coelicolor. The proteins were overexpressed and quantified as described previously (*11*). Metal analysis was conducted by Garrett-Callahan Co. (Burlingame, CA). Histidine affinity tagged variants of SCO 2687 were purified over Ni^{2+}-charged Hi-Trap Chelating HP columns (GE Healthcare) as described by the manufacturer. SCO 1441, SCO 2687, and their variants as well as wild-type SCO 6655 and the Met120Ala, -Gly, -His, -Ile, -Lys, and -Trp variants were purified as described previously (*11*). Other SCO 6655 variants (Met120Asn, -Asp, -Cys, -Gln, -Glu, -Ser, and -Thr) were purified from 2 g of wet cell paste as follows. The cleared cell lysate was treated first with 25% ammonium sulfate and centrifuged at 12000*g* for 15 min. The pellet was discarded, and the supernatants were brought to 40% saturation with ammonium sulfate and centrifuged

Table 1: Steady-State Kinetic Parameters and Product Distributions Observed with His_6-Tagged Variants of SCO 2687

variant	$K_{M(GTP)}$ (μM)	k_{cat} (min^{-1})	$k_{cat}/K_{M(GTP)}$ (M^{-1} min^{-1})	ring open		[ring open]/[GMP]
				% FAPy	% APy[a]	
Arg83His	ND[b]	~0.02[c]	ND[b]	5	95	2
Tyr123Met	92 ± 8	3.8 ± 0.1	(4.1 ± 0.4) × 10^4	100	(<0.1)	20
Asp127Ala	85 ± 8	1.3 ± 0.1	(1.6 ± 0.2) × 10^4	0.2	99.8	14
Gly209Asp	230 ± 40	2.8 ± 0.2	(1.2 ± 0.2) × 10^4	0.5	99.5	13

[a] Parentheses indicate species not observed; percentages are based on the lower limit of detection. [b] Not determined. [c] Enzyme activity decreases with turnover; k_{cat} based on observation of 2.8 μM APy produced in 30 min with 1 mM GTP and 5 μM variant.

at 12000*g* for 15 min. The resulting pellet was suspended in a small volume of 20 mM Tris-HCl (pH 8) containing 0.5 mM DTT and applied to a 1 mL HiTrap Q-Sepharose column (GE Healthcare), washed with the loading buffer, and eluted with a 0 to 0.5 M linear gradient of KCl in 20 mM Tris-HCl (pH 8.0) containing 0.5 mM DTT. Fractions containing protein were pooled and exchanged into 20 mM Tris-HCl (pH 8.0) buffer containing 0.5 mM DTT with centrifugal concentrators (YM-10, Amicon). Protein was frozen in liquid nitrogen and stored at −80 °C until it was used.

Spectrophotometric Assays of Wild-Type and Site-Directed Variants of GCH II. Reaction mixtures (0.8 mL) contained 0.1 M Tris-HCl (pH 8), 5 mM $MgCl_2$, 0.5 mM DTT, and wild-type or variant GCH II. Reactions were initiated by addition of GTP, and rates were computed using the following difference extinction coefficients: $\Delta\epsilon_{299} = 9040$ M^{-1} cm^{-1} and $\Delta\epsilon_{275} = 4700$ M^{-1} cm^{-1} for APy and FAPy, respectively. Fitting to the Michaelis−Menten equation was carried out with the statistical package R (*18*).

HPLC Analysis of APy, FAPy, and GMP Produced by Wild-Type and Site-Directed GCH II Variants of S. coelicolor. The reaction mixtures (0.1 mL) contained 0.1 M Tris-HCl (pH 8), 5 mM $MgCl_2$, 0.5 mM DTT, and 0.1 mM GTP. Reactions were generally initiated by addition of 5−25 μM enzyme. With the largely inactive Tyr326Phe SCO 1441 variant, the final concentration of protein was 70 μM. Incubations were carried out for 30 min at 25 °C. Reactions were quenched by addition of 0.2 mL of acetonitrile followed by vigorous vortexing for 30 s, followed by incubation for 5 min at room temperature. Precipitated protein was removed by centrifugation for 5 min at 20000*g*. An aliquot of the supernatant (20 μL) was analyzed on an Agilent 1100 HPLC system equipped with a diode array detector. Separations were performed using an Agilent Zorbax 5 μm SAX column (4.6 mm × 150 mm) equilibrated with a buffer solution that contained 1% solvent A [0.75 M ammonium phosphate (pH 4.1) with 2% (v/v) acetonitrile] and 99% solvent B (water) (*19*). The gradient program was as follows: 1% A for 5 min, followed by a gradient from 1 to 25% A over 20 min and a gradient from 25 to 70% A over 5 min. APy, FAPy, and GMP had retention times of 4.5, 20.0, and 21.4 min, respectively. The lower limits for detection for each of the compounds (in an injection volume of 20 μL) were determined to be 20 nM for FAPy and GMP and 50 nM for APy using samples of known concentrations. Each of the components was quantified by integrating the area corresponding to the compound at the appropriate wavelengths and bandpass for the compound (275 and 16 nm for APy and FAPy and 254 and 16 nm for GTP, respectively) and comparing the areas to standard curves constructed using purified material.

Partial Proteolysis of Tyr326Met SCO 1441. Tryptic proteolysis of SCO 1441 and Tyr326Met SCO 1441 was performed at a 500:1 ratio of GCH II protein to porcine trypsin (Sigma). Digestions were carried out in 0.1 M Tris-HCl (pH 8.0) and 5 mM $MgCl_2$ at a protein concentration of 0.5 mg/mL. Samples (5 μg) were withdrawn 0, 5, 15, 30, 60, 90, and 120 min after the addition of trypsin, mixed with SDS−PAGE sample buffer, and immediately placed in a 95 °C water bath for 5 min. The samples were kept on ice until they were analyzed on a 12% SDS−PAGE gel. Bands were visualized with Brilliant Blue R dye (Sigma).

RESULTS

Identification of Conserved GCH II Residues that Are Different in SCO 6655. To determine which residues can account for the differences in activity between canonical GCH II proteins and SCO 6655, a BLAST (*20*, *21*) search of the NCBI protein database was carried out using *E. coli* GCH II as the query sequence. A multiple-sequence alignment of the 73 highest-scoring sequences (see Table 2 of the Supporting Information) was constructed and aligned with reference to the sequence of SCO 6655 using ClustalW (*22*). The alignment revealed (data not shown) four residues (SCO 2687 numbering), Arg83, Tyr123, Asp127, and Gly209, which are absolutely conserved in the GCH II proteins (such as SCO 2687) that catalyze the conversion of GTP to APy, but different in SCO 6655 which converts GTP to FAPy. The effect of interchanging these residues in SCO 2687 with those found in SCO 6655 was examined. To facilitate purification of the proteins, each of the variants was prepared in an N-terminally His_6-tagged form and purified by Ni affinity chromatography.

With the exception of the Arg83His variant of SCO 2687, which appeared substantially impaired in overall activity, each of the other proteins appeared to be at most ~10-fold less efficient than the wild-type protein (Table 1). In addition to determining the kinetic parameters of each variant, we determined the distribution of the three potential reaction products by HPLC. Two of the products result from hydrolysis of the C-8−N-7 bond of GTP (GTP → FAPy) or release of formate (GTP → APy). Each of the variants also produces substantial quantities of a third product, GMP; although the mechanistic significance of production of GMP is not understood, dephosphorylation of GTP by GCH II has been reported (*9*, *10*, *15*). The HPLC conditions permitted simultaneous monitoring of each product in the mixture. Interestingly, the analysis reveals that one of the variants, Tyr123Met, produces FAPy exclusively (see Table 1), whereas the other three variants did not alter the product distribution. This observation mirrors that obtained with the Met120Tyr variant of SCO 6655 (*11*), in which the analogous interchange led to an enzyme that produced APy from GTP. The interchange of activities of SCO 6655 and 2687 by the

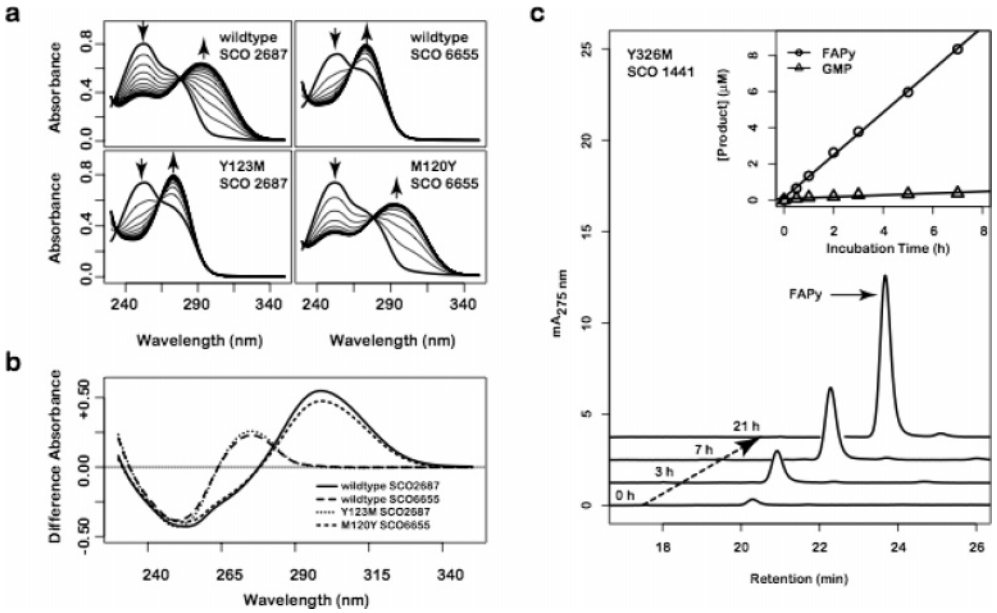

FIGURE 2: One mutation can alter the course of turnover of SCO 6655, SCO 2687, and SCO 1441 with GTP. Panel a shows spectra obtained with wild-type SCO 2687 and SCO 6655 (top) and Tyr123Met or Met120Tyr variants of SCO 2687 and SCO 6655 (bottom) at 2.3 min intervals after addition of GTP. The final spectrum with each protein was recorded at 30 min. In each sample, the absorbance due to GTP (253 nm) disappears concomitant with appearance of a new feature at 299 nm (SCO 2687 and Met120Tyr SCO 6655) or 274 nm (Tyr123Met SCO 2687 and SCO 6655). Panel b shows the difference spectra that are obtained by subtracting the spectrum obtained immediately after addition of GTP and after reaction was complete. Panel c shows the region of the HPLC traces that show formation of FAPy with Tyr326Met SCO 1441. The inset shows that formation of FAPy and GMP was linear over the course of 7 h.

Table 2: Zinc Content, Steady-State Kinetic Parameters, and Product Distributions Observed with Wild-Type and Met, Tyr, and Phe Site-Directed Variants of GCH II Intragenomic Homologues of *S. coelicolor*

protein	variant	no. of Zn^{2+} per monomer[a]	$K_{M(GTP)}$ (μM)	k_{cat} (min^{-1})	$k_{cat}/K_{M(GTP)}$ (M^{-1} min^{-1})	ring open		[ring open]/[GMP]
						FAPy %	APy %[b]	
SCO 1441	wild type	0.9	63 ± 8	2.5 ± 0.1	$(4.0 \pm 0.6) \times 10^4$	<0.1	100	24
	Tyr326Met	1.1	ND[c]	0.081 ± 0.001	ND[c]	100	<0.1	19
	Tyr326Phe	0.9	ND[c]	0.020 ± 0.001	ND[c]	100	(<1.7)	0.94
SCO 2687	wild type	1.0	24 ± 3	3.4 ± 0.2	$(1.4 \pm 0.2) \times 10^5$	<0.1	100	32
	Tyr123Met	1.2	48 ± 4	5.8 ± 0.2	$(1.2 \pm 0.1) \times 10^5$	100	(<0.05)	24
	Tyr123Phe	1.2	ND[c]	0.053 ± 0.004	ND[c]	100	(<0.8)	0.96
SCO 6655	wild type	1.1	56 ± 9	7.2 ± 0.5	$(1.3 \pm 0.2) \times 10^5$	100	(<0.05)	10
	Met120Tyr	0.8	8 ± 1	1.4 ± 0.1	$(1.9 \pm 0.3) \times 10^5$	1	99	16
	Met120Phe	0.9	19 ± 4	0.78 ± 0.05	$(4.1 \pm 0.9) \times 10^4$	100	(<0.5)	0.67

[a] Determined by inductively coupled plasma optical emission spectroscopy analysis of samples by Garrett-Callahan Co. [b] Parentheses indicate species not observed; percentages are based on the lower limit of detection. [c] Not determined.

single site-directed mutation was examined further to determine if this residue is a universal switch that controls the fate of GTP.

Universal Switch in GCH II. The properties of native Tyr326Met, Tyr123Met, and Met120Tyr variants of recombinant SCO 1441, SCO 2687, and SCO 6655, respectively, were examined. The site-directed variants were expressed in the native state and purified to homogeneity by ion exchange and hydrophobic interaction chromatographic steps. ICP-OES analysis indicated that as expected, each protein preparation contained 0.8−1.2 mol of Zn^{2+}/mol of monomer.

The site-directed variants were assayed for turnover with GTP as the substrate, and the results were compared to the values of the corresponding wild-type proteins. Representative data obtained with these proteins are shown in Figure 2 and are summarized in Table 2. Figure 2a shows the UV−visible spectral changes that are observed with the wild type, Tyr123Met SCO 2687, and Met120Tyr SCO 6655. The spectral changes and the resulting difference spectra (Figure 2b) are consistent with switching the reactivity of the protein, for instance, from GTP → APy in wild-type SCO 2687 to GTP → FAPy in the variant. The spectral changes and difference spectra for Met120Tyr SCO 6655 are consistent with our previous report (*11*) in which we showed that the variant gains the ability to produce APy from GTP. Although the Tyr326Phe SCO 1441 variant is substantially impaired in overall turnover, it produces FAPy and GMP, as assessed by HPLC analysis (Figure 2c). To eliminate the possibility that the changes in activities of these proteins resulted from nonconservative mutations (Y → M or M → Y), we also examined the effect of introducing Tyr326Phe, Tyr123Phe, and Met120Phe into the analogous position of SCO 1441, SCO 2687, and SCO 6655, respectively. Spectral changes and difference spectra for the Phe variants of SCO 6655 and SCO 2687 are also consistent with the GTP → FAPy reactivity (Figure 1 of the Supporting Information), which would be predicted. As with the Tyr → Met variant, the Tyr → Phe variant of SCO 1441 was also catalytically impaired. The fidelity of conversion of GTP to either FAPy or APy was assessed by HPLC analysis of the reaction mixtures as

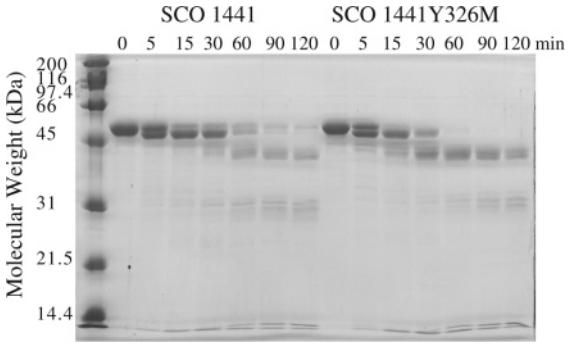

FIGURE 3: Partial tryptic proteolysis of Tyr326Met SCO 1441. Each lane contained a sample (5 μg) that was withdrawn from the digestion mixture at the indicated time. The predicted molecular mass of a SCO 1441 monomer is 46.1 kDa.

described above for the SCO 2687 variants. Substrate (0.1 mM) was incubated with enzyme, and the product distribution (GMP, FAPy, and APy) was determined by HPLC. The HPLC traces of these reactions (data not shown) reveal that in each case, one observes the ring open product that would be expected on the basis of whether a Tyr (GTP → APy) or a Met or Phe (GTP → FAPy) is present in the protein.

The results presented thus far clearly establish that a single residue in each of the GCH II enzymes of *S. coelicolor* controls the fate of GTP. The Met → Tyr and Tyr → Met variants of SCO 6655 and SCO 2687 appear to function essentially as well as their wild-type counterparts. The kinetic parameters for wild-type and variant GCH II proteins are summarized in Table 2. We have been unable to measure K_M values for the Tyr → Phe variants of SCO 1441 and SCO 2687 and the Tyr326Met variant of SCO 1441 due to the substantially reduced catalytic activity. The HPLC analysis of the reaction products, however, shows a clear interchange of GTP → FAPy and GTP → APy functionalities of the three intragenomic GCH II homologues of *S. coelicolor*, which correlates with the presence of Met or Phe (GTP → FAPy) or Tyr (GTP → APy).

Molecular interpretation of the reduced k_{cat}/K_M for SCO 1441 must await additional studies, such as a comparison of the phenotypes of the wild-type and variant proteins in the context of the GCH II domains alone. Partial proteolysis experiments with Tyr326Met SCO 1441, however, support a model where in the variants there is a change in the overall conformational dynamics of the protein. Figure 3 shows a partial tryptic digest comparing the susceptibility of wild-type and Tyr326Met variants of SCO 1441 to that of trypsin. While a fraction of the wild-type enzyme remains undigested >30 min after initiation of the proteolysis reaction, the variant protein is completely proteolyzed within 15 min. Similar results were obtained in proteinase K digest comparisons of wild-type and Tyr326Met SCO 1441 (Figure 2 of the Supporting Information). Therefore, it is possible that the dramatic loss of activity observed with this protein reflects changes in the conformational states of this protein relative to that of the wild type.

Activities of Other M120X Variants of SCO 6655. The residue occupying position 120 of SCO 6655 appears to be a key determinant of the product that will be produced by the protein from GTP. A Tyr residue supports conversion of GTP to APy, whereas the presence of a Met residue leads to formation of FAPy. To determine if other residues could also substitute, several additional Met120 SCO 6655 variants (Ala, Asn, Asp, Cys, Gln, Glu, Gly, Ile, Lys, Ser, Thr, and Trp) were prepared and assessed by UV−visible spectrophotometry and HPLC for activity and product distribution. The variants have been classified into five groups on the basis of how robust the GTP → FAPy activity of the variant was relative to that of the GTP → APy reaction (Table 3). The wild-type (M120) and Tyr-containing proteins constitute two of the five groups and represent "end points" for the possible phenotypes. We have also classified the Met120Ala, -Gly, -Phe, and -Trp variants in the first group since each of these produces FAPy only (the % APy values reported in Table 3 represent our limit of detection). Members of the second group, which is composed of Met120Asn, -Ile, -Ser, and -Gln variants, retain ≥10% of the wild-type catalytic efficiency but gain the ability to produce small quantities of APy (0.1−0.2% of the total ring open product). The third group contains Met120Cys, -Thr, -Glu, and -Lys variants, which are generally ∼10−100-fold less active than the wild-type protein, but ∼1−10% of the ring open product is FAPy. Members of the fourth group, which contains Met120His and -Asp, produce substantially more FAPy (∼50%); although the Asp variant is substantially impaired in catalysis, the catalytic efficiency of the His variant is within ∼10-fold of that of the wild-type protein. The last class contains the Met120Tyr variant, which produces APy almost exclusively.

The data show that Tyr and Met alone appear to fully support overall turnover (GTP → APy or GTP → FAPy), without producing substantial quantities of GMP (6−9% of total turnover) or loss of catalytic efficiency $[(k_{cat}/K_M)^{WT}/(k_{cat}/K_M)^{Tyr} \sim 0.7]$.

DISCUSSION

GCH II proteins catalyze the cleavage of C-8 of GTP in two half-reactions. In the first, GTP is converted to FAPy, which undergoes hydrolysis to produce APy in the second half-reaction. Interestingly, SCO 6655 catalyzes only the first half-reaction; however, the second half-reaction is readily unmasked by interchange of a single residue (Met120) in this protein with the Tyr in the other GCH II intragenomic homologues of *S. coelicolor*. Many enzymatic reactions proceed with partial reactions whereby initial transformation of the substrate in one partial reaction is followed by additional transformations leading to formation of products. One may imagine that an enzyme that catalyzes two or more successive partial reactions with stable intermediates could have the capacity, at least in principle, to catalyze each half-reaction independently. This may represent a hitherto unrecognized pathway for evolution of new activities from existing protein folds, whereby silencing of one half-reaction could lead to emergence of a new catalyst that shares an identical fold with the original. Reversion(s) as required by evolutionary pressures could readily be accommodated in the existing scaffold.

We have identified four residues in enzymes that catalyze canonical GCH II chemistry (GTP → APy) that are different from those that appear in SCO 6655, which catalyzes the conversion of GTP to FAPy. The site-directed interchange of three of these residues with those found in SCO 6655 led to proteins that still produced APy; one of the variants

Table 3: Steady-State Kinetic Parameters and Product Distributions Observed with Native Recombinant M120 Variants of SCO 6655

SCO 6655	$K_{M(GTP)}$ (μM)	k_{cat} (min^{-1})	$k_{cat}/K_{M(GTP)}$ (M^{-1} min^{-1})	ring open		[ring open]/ [GMP]
				FAPy %	APy %[a]	
Met (wild type)	56 ± 9	7.2 ± 0.5	(1.3 ± 0.2) × 10^5	100	(<0.05)	10
Ala	59 ± 2	0.7 ± 0.1	(1.2 ± 0.5) × 10^4	100	(<0.1)	3.5
Gly	190 ± 30	0.95 ± 0.07	(5.0 ± 0.8) × 10^3	100	(<0.1)	2.2
Phe	19 ± 4	0.78 ± 0.05	(4.1 ± 0.8) × 10^4	100	(<0.5)	0.67
Trp	ND[b]	0.03 ± 0.003	ND[b]	100	(<0.3)	2.0
Asn	67 ± 4	1.8 ± 0.04	(2.7 ± 0.2) × 10^4	99.9	0.1	6.8
Ile	15 ± 0.1	1.3 ± 0.01	(8.7 ± 0.3) × 10^4	99.9	0.1	4.2
Ser	33 ± 9	1.4 ± 0.1	(4 ± 1) × 10^4	99.9	0.1	1.9
Gln	11 ± 2	2.0 ± 0.1	(1.8 ± 0.4) × 10^5	99.8	0.2	6.8
Cys	67 ± 1	1.1 ± 0.04	(1.6 ± 0.4) × 10^4	99	1	10.7
Thr	150 ± 75	0.34 ± 0.08[c]	(2 ± 1) × 10^3	96	4	3.9
Glu	ND[b]	0.053 ± 0.005[c]	ND[b]	92	8	0.01
Lys	8.9 ± 3	0.18 ± 0.01[c]	(2.0 ± 0.7) × 10^4	92	8	1.5
His	50 ± 9	0.18 ± 0.01[c]	(2.0 ± 0.7) × 10^4	53	47	0.44
Asp	ND[b]	0.007 ± 0.0005[c]	ND[b]	52	48	0.1
Tyr	8 ± 1	1.4 ± 0.1	(1.9 ± 0.3) × 10^5	1	99	16

[a] Parentheses indicate the species was not observed; the percentages are based on the lower limit of detection. [b] Not determined. [c] k_{cat} based on the sum of ring-opened products.

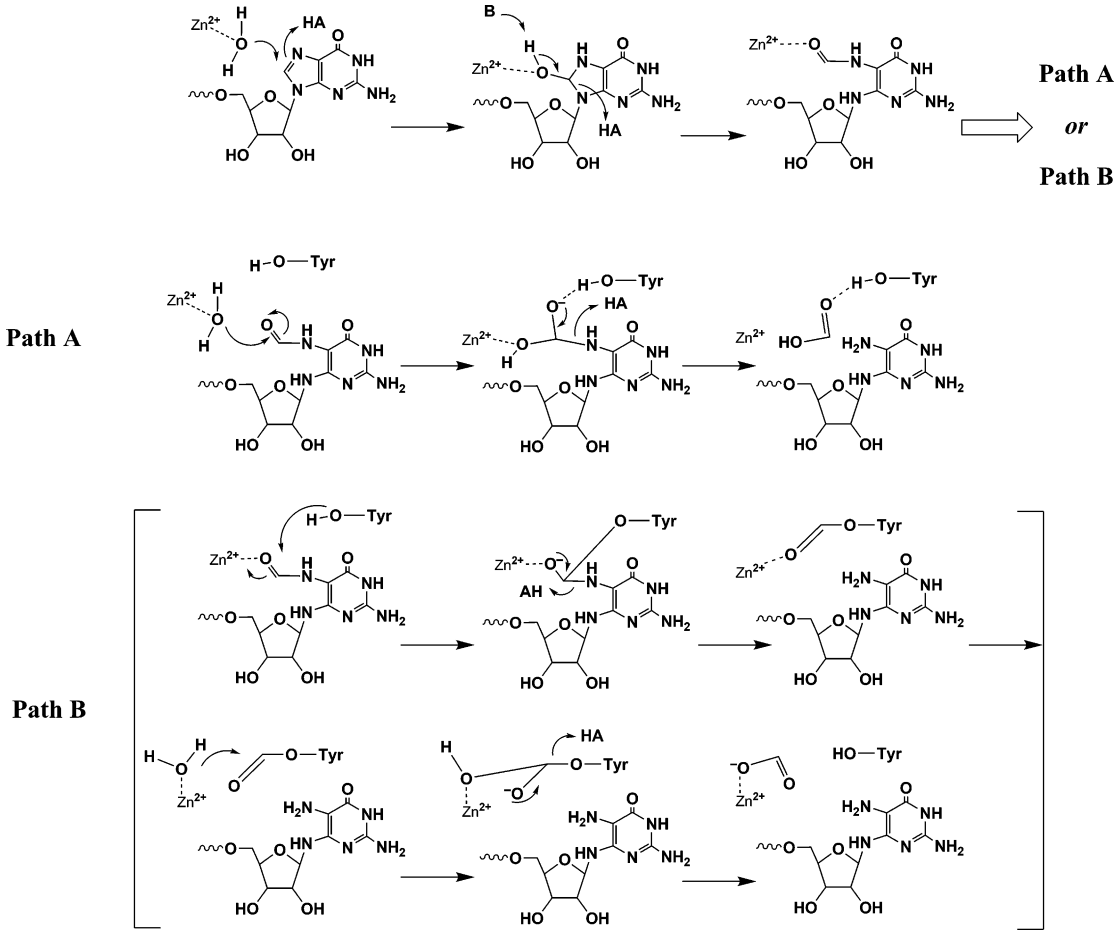

FIGURE 4: Proposed mechanism for the reaction catalyzed by GCH II.

(Tyr123Met), however, produces FAPy exclusively. In the X-ray crystal structure of *E. coli* GCH II, two of these residues, Arg83 and Asp127 (64 and 109 in *E. coli* numbering, respectively), form a salt bridge (Figure 1); Gly209 (*E. coli* residue 191) was in a disordered region and was not observed. The fourth, Tyr123, forms a hydrogen bond with a water molecule, which is a ligand to the zinc metal ion in the active site of the protein. This study shows that site-directed interchange of Tyr123 of SCO 2687 with a Met found in SCO 6655 leads to a complete change in the fate of GTP; the reversals in the fate of GTP by introduction of reciprocal mutations in SCO 6655 (Met → Tyr) (*11*) and SCO 1441 (Tyr → Met) further reinforce the key role that the residue at this position plays in dictating the observed product distribution. These observations suggest that the presence of a half-reaction that produces a stable product (FAPy) may permit GCH II to take advantage of the existing catalytic machinery to produce this compound exclusively.

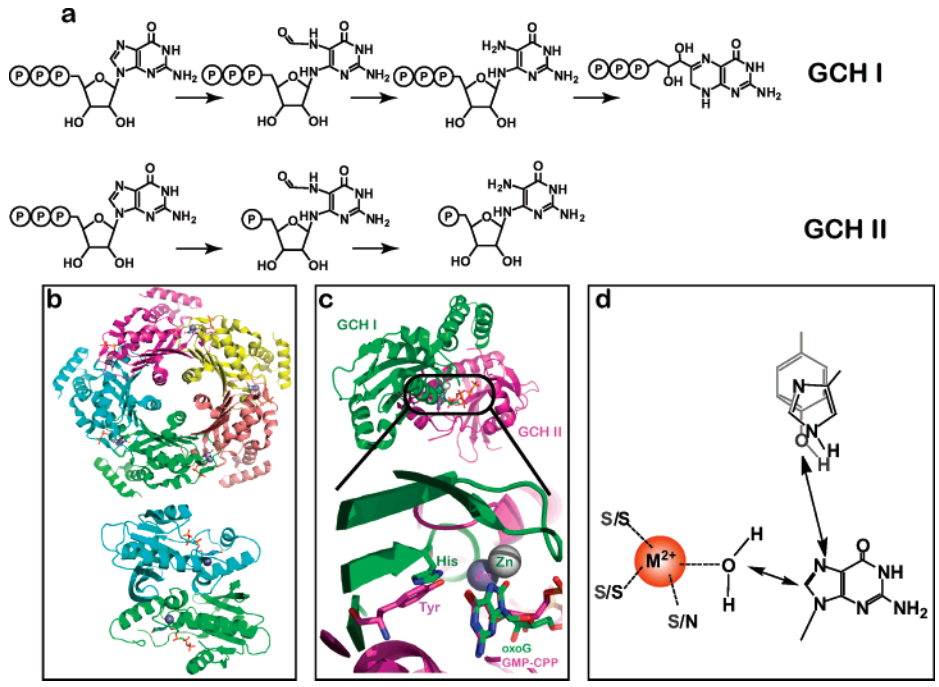

FIGURE 5: Comparison of the structure and function of GCH I and GCH II. Panel a shows the reactions that are catalyzed by GCH I and II. In each case, intermediates that have been postulated to be present or observed in the course of turnover are shown. Panel b compares the overall fold of the GCH I (PDB entry 1wur) and II (PDB entry 2bz0) proteins. The top of panel c shows the position of the monomers that one observes when the purine rings of 8-oxoGTP and GMP-CPP are aligned. The bottom of panel c shows a close-up view of the active site, showing the zinc metal ion in each protein and the catalytically essential His (GCH I) and Tyr (GCH II) residues. Panel d is a schematic representation of the structural features upon which GCH I and II have converged for hydrolysis of C-8 of GTP.

A mechanism for GCH II, consistent with the available data, is shown in Figure 4. We postulate that the zinc metal ion in the active site of GCH II serves as a Lewis acid to activate a water molecule for attack at C-8 of the substrate. In the X-ray crystal structure of the protein (*17*), a water molecule forms a hydrogen bond to the Tyr residue that we have identified as the key determinant of the fate of GTP. In the second half-reaction, it has been proposed (path A, Figure 4) that the FAPy formed in the first half-reaction is deformylated by a second water molecule, which presumably is also activated by the active site zinc. Although this mechanism is consistent with the available data, it does not suggest why the Tyr residue would be critical for the second half-reaction, but not the first. Our data are also consistent with an alternate pathway (path B, Figure 4) whereby the active site Tyr residue deformylates FAPy, producing APy, and that a water molecule activated by the zinc metal ion deformylates the formyl-tyrosine to regenerate the active site Tyr and release formate. The near-complete abolition of the production of APy in any protein lacking a Tyr residue at this position supports the notion that the active site Tyr may play a more intimate role in catalysis. However, one cannot rule out additional roles, for instance, ones where hydrogen bonding to the side chain hydroxyl lowers the energy required to achieve the transition state.

Although Tyr is absolutely required for formation of APy from GTP, several additional site-directed variants at this position of SCO 6655 also produce APy; most, however, produce only a trace quantity (<0.2%) relative to the wild type. Intriguingly, substitution of Met in SCO 6655 with His leads to a protein that catalyzes the production of both FAPy and APy with equal efficiency. We and others have noted (*11*, *17*) an interesting parallel between the identity of the residue near N-7 of GCH II and GCH I. Although these proteins are structurally distinct (*23*−*25*), each has been shown to require a zinc metal ion, which presumably activates a water molecule for excision of C-8 of GTP (*8*, *10*, *11*, *23*, *26*), in the course of conversion of GTP to dihydroneopterin triphosphate or APy (Figure 5a). Superposition of the substrate analogues that were cocrystallized in the active sites of these proteins (8-oxoGTP and GMP-CPP in GCH I and II, respectively) reveals a startling fact: despite the differences in the overall folds (Figure 5b), the two proteins have converged upon identical strategies for hydrolysis of C-8 of GTP. Each protein places a zinc metal ion in the active site near C-8 of the substrate analogue, within 3.1 and 4.1 Å of C-8 in GCH I and GCH II, respectively. In GCH II, the key Tyr residue that we have identified is found 4.5 Å from N-7 of the substrate. By contrast, in GCH I, a histidine residue occupies the same position and is ~4.1 Å from N-7 of the substrate and forms a hydrogen bond to a water molecule which in turn interacts with N-7 (Figure 5c). Is it possible that the His in GCH I plays a role identical to the Tyr residue in GCH II, perhaps involving covalent catalysis.

If GCH I and GCH II utilize identical strategies for excision of C-8, one would predict that mutation of the structurally analogous His in GCH I should lead to a protein that is analogously impaired in production of dihydroneopterin triphosphate but produces the triphosphate analogue of FAPy. In fact, site-directed mutagenesis studies with GCH I have shown that mutation of this residue in *E. coli* GCH I (H179A) leads to termination of the reaction upon formation of FAPy triphosphate, while mutation of H179 to D, F, K, N, Q, R, or S leads to formation of little or no deformylated product (*27*). These results mirror our findings with the Met 120 variants of SCO 6655. Therefore, remarkably, despite the lack of structural similarity between GCH I and GCH II

scaffolds, both have evolved means of hydrolyzing C-8 of the substrate, which utilize the Tyr or His in their respective active sites, for functionally identical purposes. Furthermore, a role for Tyr or His in covalent catalysis is plausible in both.

The studies reported here clearly point to a set of structural determinants that would be required for hydrolytic cleavage of C-8 from GTP as summarized in Figure 5d. Although the role of Tyr or His in GCH I and GCH II remains to be established unambiguously, the elements depicted in Figure 5d provide powerful criteria that should permit identification, *in silico*, of GCH II homologues that catalyze canonical (GTP → APy) or noncanonical reactions (GTP → FAPy). Indeed, a recent survey of bacterial genomes has revealed a large number of intragenomic GCH II homologues do not maintain the essential tyrosine residue and/or the zinc metal ion, suggesting the evolution of additional function(s) in the scaffold (S. D. Morrison and V. Bandarian, manuscript in preparation).

In addition to GCH I and GCH II, hydrolysis of the C-8−N-9 bond of GTP is a reaction catalyzed in Nature by two newly discovered GTP cyclohydrolases, GCH III and MptA, from *Methanocaldococcus jannaschii* (*28*, *29*). On the basis of sequence comparisons, neither protein has arisen from a GCH I- or GCH II-like ancestor. The reaction catalyzed by MptA is distinct from that of GCH I and GCH II in that the protein has been shown to require Fe^{2+} and to convert GTP to 7,8-dihydro-D-neopterin 2′,3′-cyclic phosphate. GCH III, however, catalyzes the conversion of GTP to FAPy. Intriguingly, the enzyme has been shown to require Mg^{2+} for activity and is activated by monovalent ions such as potassium and ammonium (*28*). We have recently shown that GCH III does not contain Zn^{2+} and determined the X-ray crystal structure of the protein (*34*). The structure reveals that the protein likely utilizes a two-metal ion catalytic mechanism, akin to that of adenylyl cyclase or DNA polymerase (*30*−*32*), suggesting that the paradigm for catalysis observed in GCH I and GCH II does not represent the only mechanistic route for hydrolysis of C-8 of GTP. However, metal ion-assisted activation of a water molecule to initiate the chemistry appears to be common to GCH I, GCH II, and GCH III.

In this paper, we present a case study identifying a specific mutation, in a set composed of three intragenomic homologues of GCH II, which appears to be responsible for the functional divergence of one member of the family. It is tempting to suggest that the fact that canonical function is readily "unmasked" by a single site-directed mutation supports the notion that the SCO 6655 activity arose from the canonical one guided by selective pressure; however, we have not established paralogy between these proteins. Nevertheless, the remarkable simplicity by which functional interchange can be achieved in GCH II highlights the potential utility of studies on intragenomic homologues as gateways to understanding functional divergence.

ACKNOWLEDGMENT

We acknowledge Jesus Hernandez for preparation of the four His_6-SCO 2687 variant overexpression plasmids, Joon S. Kim for initial studies on the His_6-SCO 2687 Y123M protein as well as the generation of the Met120Phe, -Gly, and -Ala SCO 6655 overexpression plasmids, Ivan Ogloblin for the purification of Met120Ala SCO 6655, and Jonathan Halevy for purification of Met120Gly and -His SCO 6655. V.B. acknowledges numerous insightful discussions with Dr. Matthew Cordes on functional and structural evolution of proteins.

SUPPORTING INFORMATION AVAILABLE

Table listing mutagenesis primers and figures showing spectral changes observed with Phe variants of SCO 6655 and 2687, as well as proteinase K digest of SCO 1441 and the Phe variant. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Ohno, S. (1970) *Evolution by gene duplication*, Springer-Verlag, New York.
2. Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D., and Kozarich, J. W. (1993) On the origin of enzymatic species, *Trends Biochem. Sci. 18*, 372−376.
3. Gevers, D., Vandepoele, K., Simillon, C., and Van de Peer, Y. (2004) Gene duplication and biased functional retention of paralogs in bacterial genomes, *Trends Microbiol. 12*, 148−154.
4. Gerlt, J. A., and Babbitt, P. C. (2000) Can sequence determine function, *Genome Biol. 1*, 1−10.
5. Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), *Nature 417*, 141−147.
6. Bacher, A., Eisenreich, W., Kis, K., Landenstein, R., Richter, G., Scheuring, J., and Weinkauf, S. (1993) in *Bioorganic Chemical Frontiers* (Dugas, H., and Schmidtchen, F. P., Eds.) Springer-Verlag, Berlin.
7. Bacher, A., Richter, G., Ritz, H., Eberhardt, S., Fischer, M., and Krieger, C. (1997) Biosynthesis of riboflavin: GTP cyclohydrolase II, deaminase and reductase, *Methods Enzymol. 280*, 382−389.
8. Kaiser, J., Schramek, N., Eberhardt, S., Puttmer, S., Schuster, M., and Bacher, A. (2002) Biosynthesis of vitamin B2. An essential zinc ion at the catalytic site of GTP cyclohydrolase II, *Eur. J. Biochem. 269*, 5264−5270.
9. Ritz, H., Schramek, N., Bracher, A., Herz, S., Eisenreich, W., Richter, G., and Bacher, A. (2001) Biosynthesis of riboflavin: Studies on the mechanism of GTP cyclohydrolase II, *J. Biol. Chem. 276*, 22273−22277.
10. Schramek, N., Bracher, A., and Bacher, A. (2001) Biosynthesis of riboflavin. Single turnover kinetic analysis of GTP cyclohydrolase II, *J. Biol. Chem. 276*, 44157−44162.
11. Spoonamore, J. E., Dahlgran, A. L., Jacobsen, N. E., and Bandarian, V. (2006) Evolution of New Function in the GTP Cyclohydrolase II Proteins of *Streptomyces coelicolor*, *Biochemistry 45*, 12144−12155.
12. Levenberg, B., and Kaczmarek, D. K. (1966) Enzymic release of carbon atom 8 from guanosine triphosphate, an early reaction in the conversion of purines to pteridines, *Biochim. Biophys. Acta 117*, 272−275.
13. Levenberg, B., and Linton, S. N. (1966) On the biosynthesis of toxoflavin, an azapteridine antibiotic produced by *Pseudomonas cocovenenans*, *J. Biol. Chem. 241*, 846−852.
14. Suzuki, F., Sawada, H., Azegami, K., and Tsuchiya, K. (2004) Molecular characterization of the *tox* operon involved in toxoflavin biosynthesis of *Burkholderia glumae*, *J. Gen. Plant Pathol. 70*, 97−107.
15. Foor, F., and Brown, G. M. (1975) Purification and properties of guanosine triphosphate cyclohydrolase II from *Escherichia coli*, *J. Biol. Chem. 250*, 3545−3551.

16. Kobayashi, M., Ohara-Nemoto, Y., Kaneko, M., Hayakawa, H., Sekiguchi, M., and Yamamoto, K. (1998) Potential of *Escherichia coli* GTP cyclohydrolase II for hydrolyzing 8-oxo-dGTP, a mutagenic substrate for DNA synthesis, *J. Biol. Chem. 273*, 26394−26399.
17. Ren, J., Kotaka, M., Lockyer, M., Lamb, H. K., Hawkins, A. R., and Stammers, D. K. (2005) GTP cyclohydrolase II structure and mechanism, *J. Biol. Chem. 280*, 36912−36919.
18. R Foundation for Statistical Computing (2005) *R: A language and environment for statistical computing*, Vienna, Austria.
19. de Korte, D., Haverkort, W. A., Roos, D., and van Gennip, A. H. (1985) Anion-exchange high performance liquid chromatography method for the quantitation of nucleotides in human blood cells, *Clin. Chim. Acta 148*, 185−196.
20. Schaffer, A. A., Aravind, L., Madden, T. L., Shavirin, S., Spouge, J. L., Wolf, Y. I., Koonin, E. V., and Altschul, S. F. (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements, *Nucleic Acids Res. 29*, 2994−3005.
21. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
22. Thompson, J. D., Higgins, D. G., and 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, *Nucleic Acids Res. 22*, 4673−4680.
23. Auerbach, G., Herrmann, A., Bracher, A., Bader, G., Gutlich, M., Fischer, M., Neukamm, M., Garrido-Franco, M., Richardson, J., Nar, H., Huber, R., and Bacher, A. (2000) Zinc plays a key role in human and bacterial GTP cyclohydrolase I, *Proc. Natl. Acad. Sci. U.S.A. 97*, 13567−13572.
24. Rebelo, J., Auerbach, G., Bader, G., Bracher, A., Nar, H., Hosl, C., Schramek, N., Kaiser, J., Bacher, A., Huber, R., and Fischer, M. (2003) Biosynthesis of pteridines. Reaction mechanism of GTP cyclohydrolase I, *J. Mol. Biol. 326*, 503−516.
25. Tanaka, Y., Nakagawa, N., Kuramitsu, S., Yokoyama, S., and Masui, R. (2005) Novel reaction mechanism of GTP cyclohydrolase I. High-resolution X-ray crystallography of *Thermus thermophilus* HB8 enzyme complexed with a transition state analogue, the 8-oxoguanine derivative, *J. Biochem. (Tokyo) 138*, 263−275.
26. Kaiser, J., Schramek, N., Puttmer, S., Schuster, M., Eberhardt, S., Rebelo, J., Auerbach, G., Bader, G., Bracher, A., Nar, H., Hosl, C., Fischer, M., Huber, R., and Bacher, A. (2002) Essential zinc ions at the catalytic sites of GTP cyclohydrolases, in *Flavins and Flavoproteins* (Chapman, S., Perham, R., and Scrutton, N., Eds.) pp 241−246, Rudolf Weber, Berlin.
27. Bracher, A., Fischer, M., Eisenreich, W., Ritz, H., Schramek, N., Boyle, P., Gentili, P., Huber, R., Nar, H., Auerbach, G., and Bacher, A. (1999) Histidine 179 mutants of GTP cyclohydrolase I catalyze the formation of 2-amino-5-formylamino-6-ribofuranosylamino-4(*3H*)-pyrimidinone triphosphate, *J. Biol. Chem. 274*, 16727−16735.
28. Graham, D. E., Xu, H., and White, R. H. (2002) A member of a new class of GTP cyclohydrolases produces formylaminopyrimidine nucleotide monophosphates, *Biochemistry 41*, 15074−15084.
29. Grochowski, L. L., Xu, H., Leung, K., and White, R. H. (2007) Characterization of an Fe^{2+}-dependent archaeal-specific GTP cyclohydrolase, MptA, from *Methanocaldococcus jannaschii*, *Biochemistry 46*, 6658−6667.
30. Steitz, T. A., and Steitz, J. A. (1993) A general two-metal-ion mechanism for catalytic RNA, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6498−6502.
31. Hurley, J. H. (1998) The adenylyl and guanylyl cyclase superfamily, *Curr. Opin. Struct. Biol. 8*, 770−777.
32. Tesmer, J. J. G., Sunahara, R. K., Johnson, R. A., Gosselin, G., Gilman, A. G., and Sprang, S. R. (1999) Two-metal-ion catalysis in adenylyl cyclase, *Science 285*, 756−760.
33. DeLano, W. L. (2002) *The PyMol molecular graphics system*, DeLano Scientific, San Carlos, CA.
34. Morrison, S. D., Roberts, S. A., Zegeer, A. M., Montfort, W. A., and Bandarian, V. (2008) A new use for a familiar fold: The X-ray crystal structure of GTP-bound GTP cyclohydrolase III from *Methanocaldococcus jannaschii* reveals a two metal ion catalytic mechanism, *Biochemistry 47*, 230−242.

BI702263Z