Biosynthesis of Pteridines. Reaction Mechanism of GTP Cyclohydrolase I

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · MARCH 2003

Impact Factor: 4.33 · DOI: 10.1016/S0022-2836(02)01303-7 · Source: PubMed

doi:10.1016/S0022-2836(02)01303-7 J. Mol. Biol. (2003) 326, 503–516

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Biosynthesis of Pteridines. Reaction Mechanism of GTP Cyclohydrolase I

Jorge Rebelo[1,2], Günter Auerbach[1], Gerd Bader[1], Andreas Bracher[3] Herbert Nar[1], Cornelia Hösl[3], Nicholas Schramek[3], Johannes Kaiser[3] Adelbert Bacher[3], Robert Huber[1] and Markus Fischer[3]*

[1]*Abteilung Strukturforschung Max-Planck-Institut für Biochemie, Am Klopferspitz 18a 82152 Martinsried, Germany*

[2]*REQUIMTE/CQFB Departamento de Química FCT, Universidade Nova de Lisboa, 2829-516 Caparica Portugal*

[3]*Lehrstuhl für Organische Chemie und Biochemie Technische Universität München, Lichtenbergstraße 4 85747 Garching, Germany*

**Corresponding author*

GTP cyclohydrolase I catalyses the hydrolytic release of formate from GTP followed by cyclization to dihydroneopterin triphosphate. The enzymes from bacteria and animals are homodecamers containing one zinc ion per subunit. Replacement of Cys110, Cys181, His112 or His113 of the enzyme from *Escherichia coli* by serine affords catalytically inactive mutant proteins with reduced capacity to bind zinc. These mutant proteins are unable to convert GTP or the committed reaction intermediate, 2-amino-5-formylamino-6-(β-ribosylamino)-4(3H)-pyrimidinone 5′-triphosphate, to dihydroneopterin triphosphate.

The crystal structures of GTP complexes of the His113Ser, His112Ser and Cys181Ser mutant proteins determined at resolutions of 2.5 Å, 2.8 Å and 3.2 Å, respectively, revealed the conformation of substrate GTP in the active site cavity. The carboxylic group of the highly conserved residue Glu152 anchors the substrate GTP, by hydrogen bonding to N-3 and to the position 2 amino group. Several basic amino acid residues interact with the triphosphate moiety of the substrate. The structure of the His112Ser mutant in complex with an undefined mixture of nucleotides determined at a resolution of 2.1 Å afforded additional details of the peptide folding.

Comparison between the wild-type and mutant enzyme structures indicates that the catalytically active zinc ion is directly coordinated to Cys110, Cys181 and His113. Moreover, the zinc ion is complexed to a water molecule, which is in close hydrogen bond contact to His112. In close analogy to zinc proteases, the zinc-coordinated water molecule is suggested to attack C-8 of the substrate affording a zinc-bound 8*R* hydrate of GTP. Opening of the hydrated imidazole ring affords a formamide derivative, which remains coordinated to zinc. The subsequent hydrolysis of the formamide motif has an absolute requirement for zinc ion catalysis. The hydrolysis of the formamide bond shows close mechanistic similarity with peptide hydrolysis by zinc proteases.

Keywords: biosynthesis; folic acid; crystal structure; GTP cyclohydrolase I; tetrahydrobiopterin

Introduction

GTP cyclohydrolase I catalyses the formation of dihydroneopterin triphosphate (compound 5) from GTP (Figure 1). The enzyme product is the first committed intermediate in the biosynthetic pathways of tetrahydrofolate in plants and microorganisms and of tetrahydrobiopterin in animals.[1] The complex reaction sequence catalysed by GTP cyclohydrolase I is supposed to involve the hydrolytic opening of the imidazole ring of GTP under formation of a formamidopyrimidine type intermediate (compound 2),[2–6] the removal of the

Present addresses: G. Auerbach, Antisense Pharma GmbH, Josef-Engert-Straße 9, D-93053 Regensburg, Germany; A. Bracher, EMBL Grenoble, 6 rue Jules Horowitz, B.P. 181, 38042 Grenoble Cedex 9, France; H. Nar, Boehringer Ingelheim Pharma Deutschland, Structural Research, Birkendorferstraße 65, 88400 Biberach, Germany; C. Hösl, Institut für Pharmazie, Ludwig Maximilians Universität München, Butenandtstraße 7, 83177 München, Germany.

Abbreviations used: MPD, methyl pentane diol; superscripts A to D, mark the subunits A to D (e.g. ^{D}Arg65).

E-mail address of the corresponding author: markus.fischer@ch.tum.de

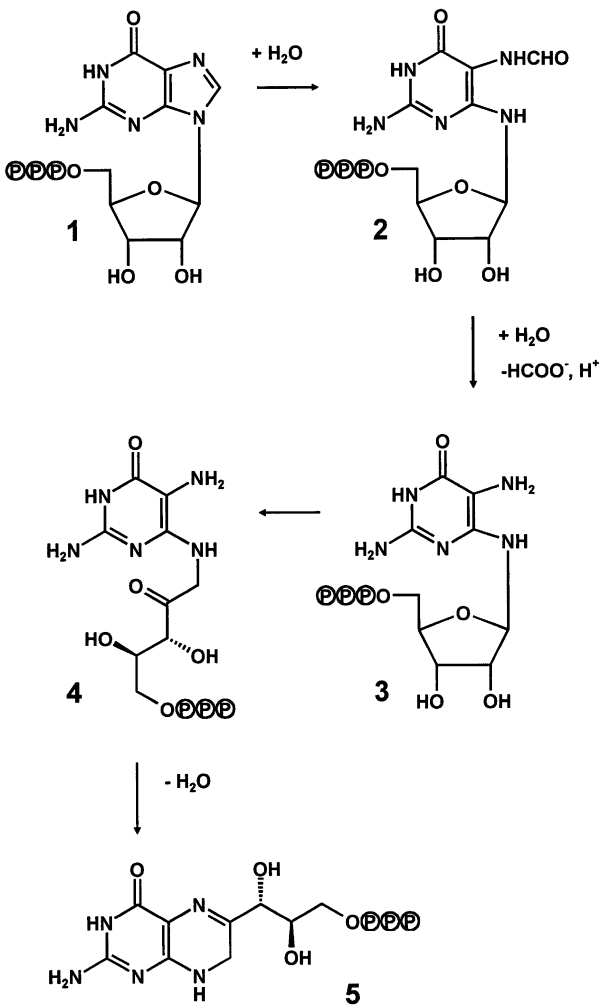

Figure 1. Hypothetical mechanism of the reaction catalysed by GTP cyclohydrolase I.[8]

purine carbon atom 8 of GTP as formate affording the triaminopyrimidine derivative (compound 3), an Amadori rearrangement of the ribose moiety, and finally closure of the dihydropyrazine ring by intramolecular condensation affording dihydroneopterin triphosphate (compound 5). By this reaction sequence involving the hydrolytic cleavage of two carbon–nitrogen bonds and one carbon–oxygen bond, carbon atoms 1′ and 2′ of the ribose side-chain of GTP become part of the pteridine ring system of the product.[7,8]

The crystal structure of wild-type *Escherichia coli* GTP cyclohydrolase I has been determined at a resolution of 2.6 Å.[9,10] The homodecameric protein has a mass of 247 kDa and the shape of a torus with dimensions of 65 Å × 100 Å. The D_5 symmetric structure comprises a β-barrel composed of 20 antiparallel strands surrounding a core of α-helical segments (Figure 2).

Recently, we have shown that the active sites of human and bacterial GTP cyclohydrolase I contain essential zinc ions that had escaped detection in earlier crystallographic studies.[11] The metal ion is supposed to act as a Lewis acid enabling the hydrolytic opening of the imidazole ring of GTP *via* activation of an attacking water molecule.

Earlier X-ray diffraction studies[9] on a complex of wild-type GTP cyclohydrolase I with the substrate analogue, 2′-deoxy-GTP, had shown that each of the ten active sites of GTP cyclohydrolase I is located at the interface of three adjacent subunits (subsequently labelled A, B, and D). However, the electron density had not been sufficiently clear in order to analyse the interaction between protein and bound ligand in detail.

The enzymes involved in the biosynthesis of tetrahydrofolate comprise important anti-infective drug targets.[12] The anti-microbial activity of sulfonamides and trimethoprim are based on inhibition of dihydropteroate synthase and dihydrofolate reductase, respectively.

The rapidly escalating resistance of human pathogens against all major drugs in current use necessitates the exploration of novel targets for anti-infective drugs.[13] A detailed analysis of potential target enzymes can serve as a basis for drug development. These considerations prompted the present study on the initial enzyme of the tetrahydrofolate pathway, GTP cyclohydrolase I, which suggests a detailed reaction mechanism on the basis of biochemical and crystallographic studies.

Results

For the purpose of this study, single amino acid replacement mutant proteins carrying serine residues instead of Cys110, Cys181, His112 or His113 were purified from recombinant *E. coli* strains as described in Materials and Methods. The purified mutant proteins failed to produce dihydroneopterin 3′-triphosphate (compound 5) from GTP (Table 1).[9]

Recently, 2-amino-5-formylamino-(6β)-ribosylamino-4(3H)-pyrimidinone 5′-triphosphate (compound 2) has been identified as an early intermediate of the reaction catalysed by GTP cyclohydrolase I.[14] This intermediate was prepared by the catalytic action of the His179Ala mutant of GTP cyclohydrolase I, as described elsewhere,[14] and used as substrate for the mutants under study. As shown in Table 1, all mutants studied were unable to convert this reaction intermediate to dihydroneopterin 3′-triphosphate (compound 5).

Zinc was measured in the enzyme preparations by atomic absorption spectrometry. Whereas the wild-type protein contained about 0.9 zinc ion per subunit, the zinc concentration of the recombinant mutant proteins was below the level of detection (Table 1).

The His112Ser mutant was crystallized using the experimental conditions reported earlier for the wild-type protein. A complete data set of a single crystal was determined at a temperature of 95 K to a resolution of 2.1 Å. The structure was determined using model phases of the wild-type enzyme[9] and

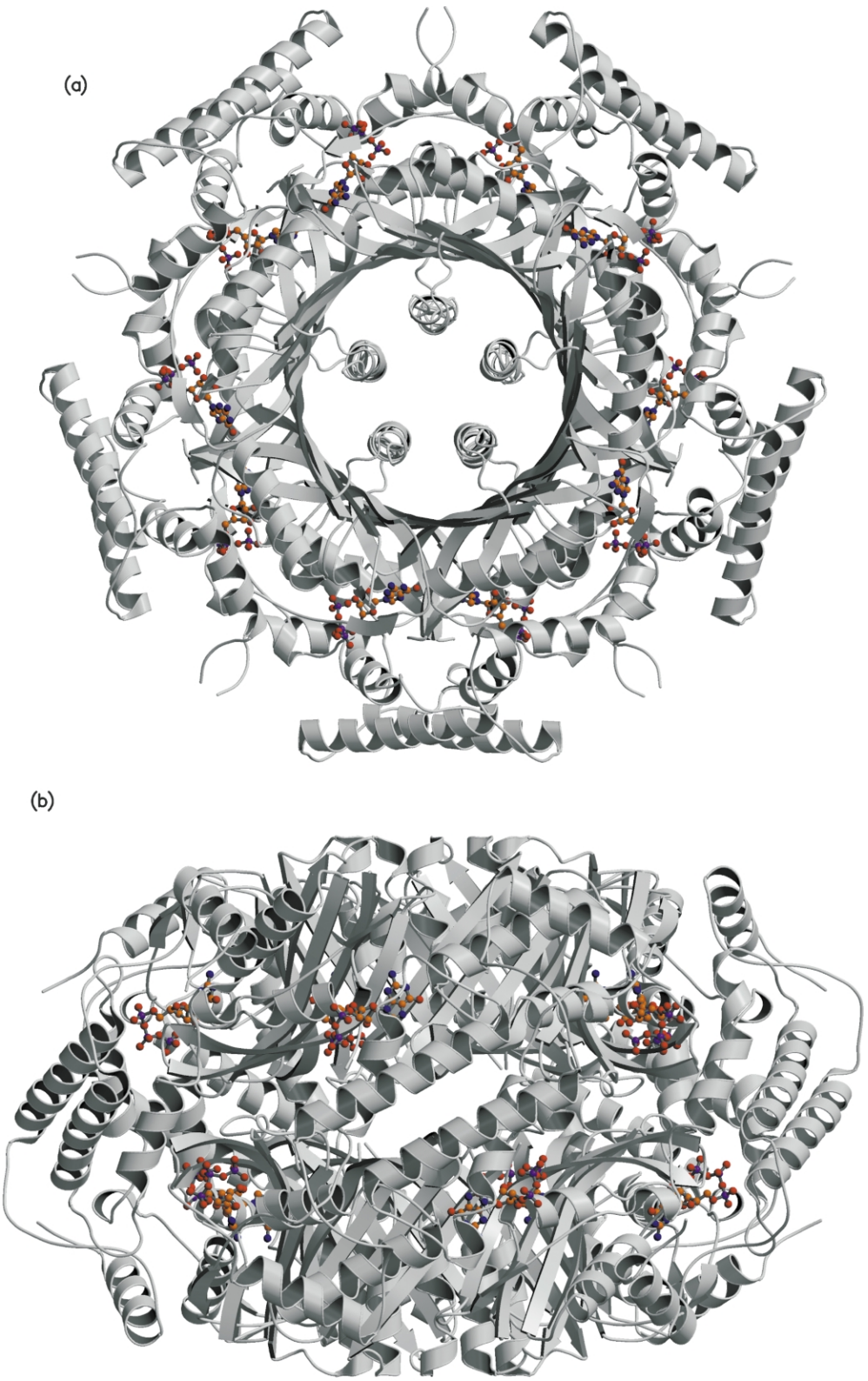

Figure 2. Ribbon representation of GTP cyclohydrolase I mutant His113Ser in complex with the natural substrate GTP. (a) The decamer is shown along the 5-fold symmetry axis and (b) after 90° rotation around the *x*-axis. The GTP molecules are drawn in ball-and-stick representation. The Figure was created using MOLSCRIPT[39] and Raster3D.[40]

refined to a resolution of 2.1 Å with an *R*-factor of 20.0%.

The inspection of the electron density map showed that the active site contained non-protein density that could be interpreted as a nucleoside triphosphate, although none had been added during purification or crystallization. Subsequent biochemical analysis revealed that the

Table 1. Properties of GTP cyclohydrolase I wild-type and mutants under study

Mutation	Zn content (mol/mol)	Activity (nmol mg^{-1} min^{-1}), Substrate	
		GTP	Compound 2
None	0.9	91	88
H112S	<0.2	<0.1	<0.1
H113S	<0.2	<0.1	<0.1
C110S	<0.2	<0.1	<0.1
C181S	<0.2	<0.1	<0.1

Zinc content and dihydroneopterin triphosphate formation from the natural substrate GTP and 2-amino-5-formylamino-6-ribofuranosylamino-4(3H)-pyrimidinone triphosphate (compound 2). The contents of bound zinc and GTP are referenced to the estimated number of GTP binding sites. Enzymatic activity is given as nmol dihydroneopterin triphosphate formed per minute with 1 mg of protein.

chromatographic purification of the Cys110Ser mutant, and the other mutants described here, affords the proteins in complex with a mixture of nucleotide triphosphates as well as some other low molecular mass phosphoric acid esters. In a typical preparation of the His112Ser mutant, the amounts of GTP and ATP present were estimated by HPLC to yield about 17% GTP and 23% ATP.

In light of this unexpected finding, all crystals were soaked with GTP prior to data collection in the following experiments. The Cys181Ser and His112Ser mutants could be crystallized under the experimental conditions used for the wild-type enzyme, albeit at an increased concentration of precipitant. The His113Ser mutant enzyme was crystallized using methyl pentane diol (MPD) as precipitant.

Complete data sets to resolutions of 2.5 Å, 2.8 Å and 3.2 Å were collected from single crystals of His113Ser, His112Ser and Cys181Ser mutant protein that had been soaked for 30 minutes with 10 mM GTP. The His113Ser and His112Ser mutant crystal diffraction sets were measured at a temperature of 95 K (Table 2) using synchrotron radiation. The diffraction data set of the Cys181Ser mutant was obtained at a temperature of 277 K with a Cu rotating anode X-ray source. Structures were determined using model phases of the wild-type enzyme[9] and refined to their limit resolutions (2.5 Å, 2.8 Å and 3.2 Å) with *R*-factors of 26.2%, 21.8% and 18.5%, respectively.

A superposition of the structure of His112Ser mutant in complex with the nucleoside triphosphate mixture at 2.1 Å resolution with the wild-type structure reveals no significant structural changes; r.m.s. deviations were less than 0.4 Å for all monomers. However, based on this first high-resolution structure of GTP cyclohydrolase I, the

Table 2. Crystallographic data

Data set	His112Ser/GTP	His113Ser/GTP	Cys181Ser/GTP	His112*Ser/GTP
A. *Data collection*				
Source[a]	DESY	DESY	RA	DESY
Space group	$C222_1$	$P4_32_12$	$C222_1$	$C222_1$
Observations (No.)	2,178,312	1,005,735	1,260,271	661,363
Unique reflections (No.)	105,682	100,128	70,145	238,566
R_{merge}(%)[b]	5.1	4.7	6.3	8.5
Resolution (Å)	2.8	2.55	3.2	2.1
Data completeness (%)	97.4	96.9	98.6	90.4
B. *Refinement*				
No. of atoms in refinement	26,509	17,827	26,475	27,375
Resolution limits (Å)	10.0–2.8	20.0–2.5	11–03.2	8.0–2.1
R-factor (%)[c]	21.8	26.2	18.5	20.0
R_{free}(%)[c,d]	27.2	29.3	22.8	24.6
r.m.s.d.[e] Bonds (Å)	0.009	0.013	0.008	0.008
r.m.s.d. Angles (degrees)	1.51	1.94	1.47	1.90
r.m.s.d. Monomers (NCS)[f] (Å)	–	–	–	0.25
B-factor (averaged, all monomers) ($Å^2$)	55.0	56.3	43.5	41.1
r.m.s.d., bonded *B*-factor ($Å^2$) (main-chain)	1.2	1.5	1.4	3.5
r.m.s.d., monomers *versus* mono A (wild-type) (Å)	–	–	–	0.358

[a] RA, rotation anode source; DESY, Synchrotron Hamburg, Germany.
[b] R_{merge}: $R_s = \Sigma(I-\langle I\rangle)/\Sigma(I)$, where I is the measured intensity and $\langle I\rangle$ is the averaged value; the summation is over all measurements.
[c] Observations above 3σ cut off.
[d] Free *R*-factor calculated by setting aside 10% of the reflections before refinement.
[e] Root-mean-square deviations from ideal values.
[f] Non-crystallographic symmetry.

secondary structure of the protein can be described in more detail. Two β-strands, β1A and β2B, and two α-helices, α5B and α6A, were additionally assigned (Figure 3).

The analysis of the main-chain atom temperature factors of the wild-type and mutant structures indicated that the most flexible region of the monomer is the loop between Gly18 and Asp31, connecting the α-helices α1 and α2. High main-chain *B*-factors were also observed at the chain termini. On the other hand, the active site is within a rather rigid region of the GTP cyclohydrolase I structure including the four-stranded antiparallel β-sheet.

Amino acid sequences have been reported for GTP cyclohydrolase I from a variety of prokaryotic and eukaryotic organisms.[15,16] The evolution of the C-terminal domain containing the regions relevant for oligomerization and enzyme catalysis appears to have proceeded in a relatively conservative way. With two exceptions, Arg139 and Ile132, every residue within a distance of 4 Å around the bound GTP is strictly conserved (Figure 3). Three adjacent subunits participate in each respective catalytic site. Specifically, ^{D}Arg65 (superscripts A to D mark the subunits A to D), ^{A}Cys110, ^{A}His112, ^{A}His113, ^{B}Tyr125 (contact distance 6 Å, *via* water molecule W2_A), ^{B}Ser-135, ^{B}Lys136, ^{B}Arg139, ^{A}Gln151, ^{A}Glu152, ^{A}His179, ^{A}Cys181, ^{A}Arg185 and ^{A}Thr196 are involved either in substrate binding and/or catalytic mechanism. Correlating the conserved sequence regions with the three-dimensional structure reveals that the region from ^{D}Ser58 to ^{D}Met69 narrows the entrance to the active site pocket. The loop between the β-strands β1 and β2, ^{A}Cys110 to ^{A}Val115 forms the active site loop and the conserved N-terminal region of helix 5, ^{A}Arg148 to ^{A}Thr155, provides the purine anchor Glu152. The sequence part ^{B}Val131 to ^{B}Arg139 contacts both the ribose and the triphosphate, ^{A}Ile175 to ^{A}His179 the nucleobase and ^{A}Cys181 to ^{A}Arg188 the triphosphate moiety (Figure 4).

Parts of the electron densities of the mutant proteins under study can be unequivocally attributed to bound substrate. The well-defined $F_{obs} - F_{calc}$ electron difference density maps clearly show the intact GTP molecule (Figure 5), thus confirming the complete catalytic inactivity of the mutants under study. In all structures, the GTP molecule is bound to ^{A}Glu152 *via* the ring N-1 nitrogen atom and the position 2 amino group of the 2-amino-4-oxopyrimidine moiety of the substrate, thus confirming the binding mode found for the 2′-deoxy-GTP complex.[5] Supporting this anchor, the amide hydrogen atom of the preceding peptide bond between ^{A}Val150 and ^{A}Gln151 forms a hydrogen bond to the 6-oxo group of the purine (Figure 4). Furthermore, a hydrogen bond is formed between the position 2 amino group of the guanine moiety and the carbonyl oxygen atom of residue ^{B}Ile132. Whereas the peptide bond of ^{A}Gln151 is involved in anchoring the guanidino moiety of the substrate, the side-chain of ^{A}Gln151 is hydrogen-bonded to both ^{A}His179 and ^{A}Thr196. The anchor residue, ^{A}Glu152, is located at the strictly conserved N-terminal region of helix α5 (Figure 3). It appears plausible that this hydrogen bond system remains in place throughout a substantial part of the reaction sequence, where it would initially contribute to the formation of the Michaelis complex and subsequently continue to serve as an anchor for the reaction intermediates (Figure 1).

While the guanidine moiety of GTP binds to the negatively charged part of the active site region, the triphosphate moiety is bound by a cluster of basic residues provided by the helices ^{D}α3, ^{B}α4 and ^{A}α5B (Figure 4). The charge compensation by these residues may be the reason why metal-assisted binding of the triphosphate group is not employed in GTP cyclohydrolase I. The strands ^{A}β2B and ^{A}β3, which are N-terminally adjacent to these α-helices (^{B}α4 and ^{A}α5B, respectively), form the widest part of the active site pocket by lining both the base and the ribose of the GTP molecule. The enzyme substrate interactions observed in the structure are summarized in Table 3.

In the difference electron density maps of the three complex mutant protein structures, the orientation of the hydroxyl groups in the GTP carbohydrate unit is clearly defined and thus the conformation of the ribose moiety can be unambiguously assigned to C-2′ *endo*. The 2′ and 3′ hydroxy groups of GTP form hydrogen bonds to the ^{B}Ser135 main-chain amino group and to ^{B}Ser135 O^{γ}, respectively. ^{B}Ser135 is, furthermore, hydrogen-bonded to the γ-phosphate group of GTP. Additionally, a nearby water molecule, W1_A (His113Ser), connects the 2′-hydroxyl group to the γ-phosphate.

The triphosphate chain of GTP was found in a bent conformation, which is almost identical to the conformation in the deoxy-GTP complex of the zinc-depleted wild-type enzyme.[10] Small GTP conformation differences were found in different mutants in the phosphate P^{β} to O-3 region, a solvent-exposed region where no fixed hydrogen bonds were assigned. The position of the γ-phosphate is fixed by interactions with the surrounding side-chains of residues ^{B}Ser135, Lys ^{B}136, Arg ^{B}139 and Arg ^{B}185 and virtually coincides with a bound inorganic phosphate in the wild-type structure.

The resolution of the His113Ser mutant structure allowed the location of the active site water molecules and the study of their interactions with substrate and enzyme. A water molecule, W2_A, extends the hydrogen bond network at the guanidino moiety, thus contributing to the tight anchoring of the substrate. This is in line with the strict conservation of the residues Tyr125 and Glu152, which interact with this water molecule. As suggested by the complex structure with the substrate analog, deoxy-GTP,[9] there is some free space on both sides of the imidazole portion of the nucleobase that might be occupied by solvent during catalysis. The water W3_A is held in position by hydrogen bonds to Gln149, Glu111,

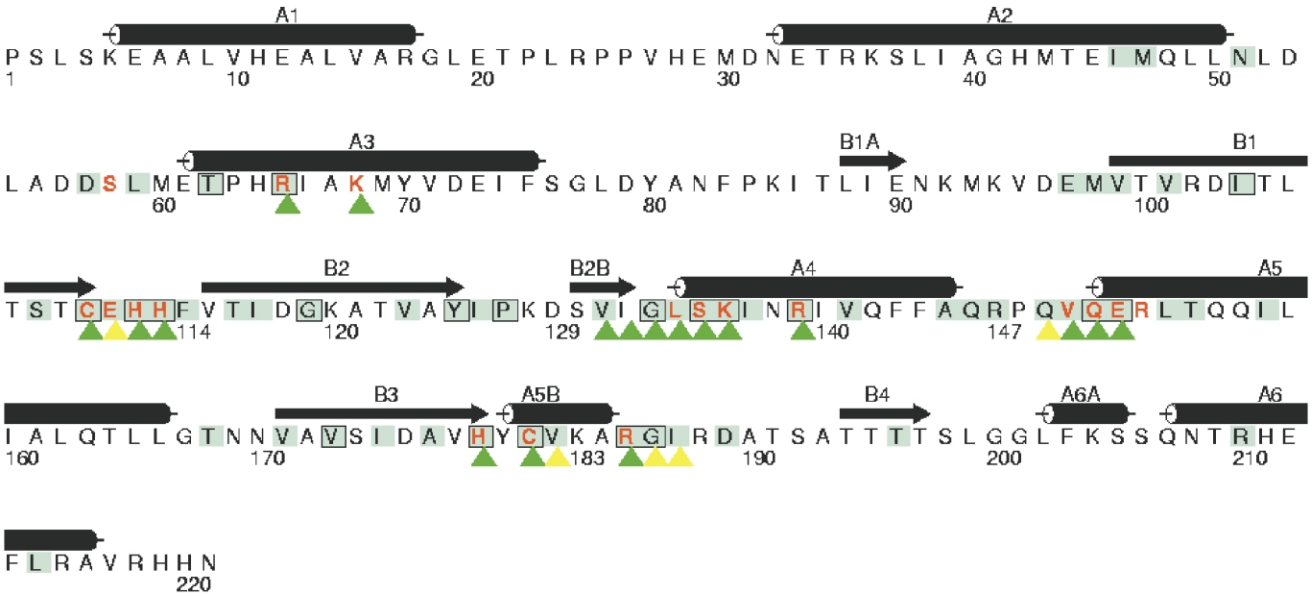

Figure 3. Amino acid sequence of *E. coli* GTP cyclohydrolase I. The secondary structure assignment is based on the wild-type crystal structure[10] and the (high-resolution) crystal structure of the His112Ser* mutant. Residues are shown in red if single-site mutation resulted in reduced catalytic activity.[9,14] *E. coli* GTP cyclohydrolase I was aligned with BLASTP against a non-redundant set of GTP cyclohydrolase I sequences.[41] Conserved residues are shown in blue, and strictly conserved amino acid residues are additionally boxed. Triangles below the amino acid sequence indicate residues closer than 6 Å to the substrate GTP. Green triangles designate residues closer than 4 Å to the substrate GTP. The Figure was created using ESPript.[42]

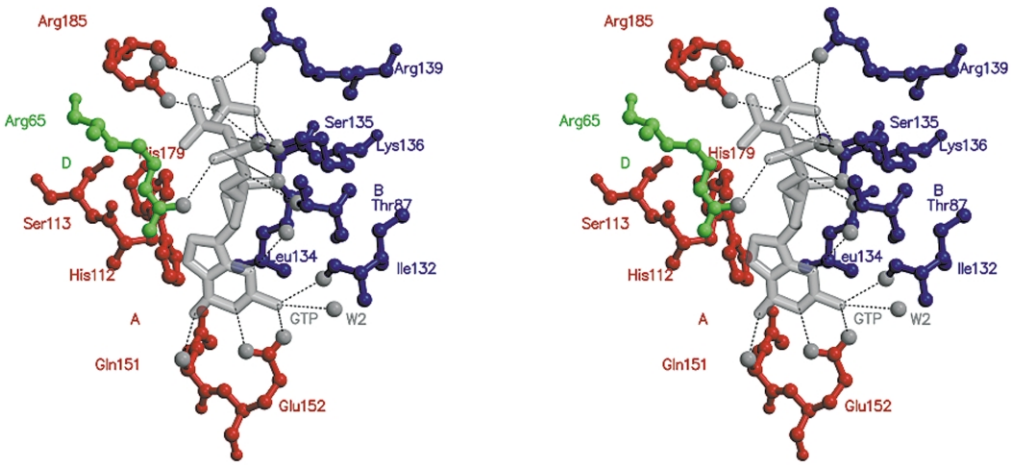

Figure 4. Stereo diagram from the active site of the *E. coli* GTP cyclohydrolase I His113Ser mutant in complex with the substrate GTP. The GTP molecule (shown as a transparent wire model representation) is embedded in a large hydrogen bond network (broken lines) within the active site. Amino acid residues are shown as ball-and-stick models coloured according to the subunit to which they belong: A, red; B, blue; and D, green. The Figure was created using MOLSCRIPT[39] and Raster3D.[40]

and His112 and lies in close proximity to the nucleotide base N-7. The water molecules W1_A, W4_A, and W5_A are hydrogen-bonded to the γ-phosphate. No water molecule was found around the exposed α-phosphate and β-phosphate.

Mutation analysis had indicated earlier that the amino acid residues Cys110 and Cys181 are both indispensable for catalytic activity. The active wild-type enzyme has both these amino acid residues bound to the essential zinc ion.[10,11] Replacement of either one of these by any of several amino acid residues yielded completely inactive protein. Moreover, no zinc was detected by atomic absorption spectrometry in both these mutants (Table 1). Finally, the inspection of the electron density maps for the present zinc-free Cys181Ser mutant model confirmed the complete absence of zinc in this mutant, as no residual electron density was observed on the $F_{obs} - F_{calc}$ density maps (Figure 5(a)).

Residues His112 and His113 have also been shown to be of major importance for the mechanism of this enzyme. The mutants His112Ser and His113Ser are devoid of catalytic activity (Table 1). The zinc-free crystal structure models of His112Ser and His113Ser, revealed important positive residual electron densities in the $F_{obs} - F_{calc}$ electron density maps. These densities occur on the region where the zinc metal ion is observed in the active wild-type form of the enzyme. Due to the fact that zinc was not detected by atomic absorption spectrometry (Table 1), a fully occupied disulphide group formed between Cys110 and Cys181 on the His113Ser mutant model was calculated but could not explain this residual electronic density. With both His/Ser mutants, the position and coordination of the residual electron density lead us to presume the presence of another metallic positive ion replacing zinc. As shown in Figure 5(b) for the mutant His113Ser, the detected $F_{obs} - F_{calc}$ electron density above 4.5σ cut-off level corresponding to the replaced metal is within the range of coordination of both Cys110 and Cys181 residues. For the mutant His112Ser, the weaker $F_{obs} - F_{calc}$ residual electron density detected on some of the active sites of the antropy unit (a.u.) (above 3.0σ) is shifted in the direction of His179 by about 3 Å to establish coordination to the imidazole side-chain *via* $N^{\delta 1}$.

In the structure of the wild-type enzyme,[11] the imidazole side-chain of ^{A}His112 forms an approximately 2.6 Å short hydrogen bond to the zinc-bound water, thus positioning it properly for nucleophilic attack on the substrate (Figure 6). A residue acting as a hydrogen bond acceptor for a metal-activated hydroxide is a common feature among zinc metalloenzymes.[17] In the Cys181Ser mutant the same orientation is observed for ^{A}His112. In the complex structure of deoxy-GTP and Zn-depleted GTP cyclohydrolase I, the imidazole ring plane of His112 was found rotated by about 90° in hydrogen bond distance to the furanosyl ring oxygen atom.[9] In the structure of the His112Ser mutant, ^{A}Ser112 forms a hydrogen bond with the side-chain from the adjacent residue ^{A}His113, also engaged in GTP binding by acting as hydrogen bond donor to the β-phosphate.

When the structures of the zinc bound wild-type enzyme and the His113Ser mutant are superimposed (Figure 6), a short distance (below van der Waals contact distance) is found between the zinc-bound water molecule and C-8 of GTP, suggesting a compressed state favouring nucleophilic attack of the oxygen atom on C-8. The product of this initial reaction could assume a more relaxed conformation.

Another structurally characterized enzyme with a zinc cofactor bound by two cysteine

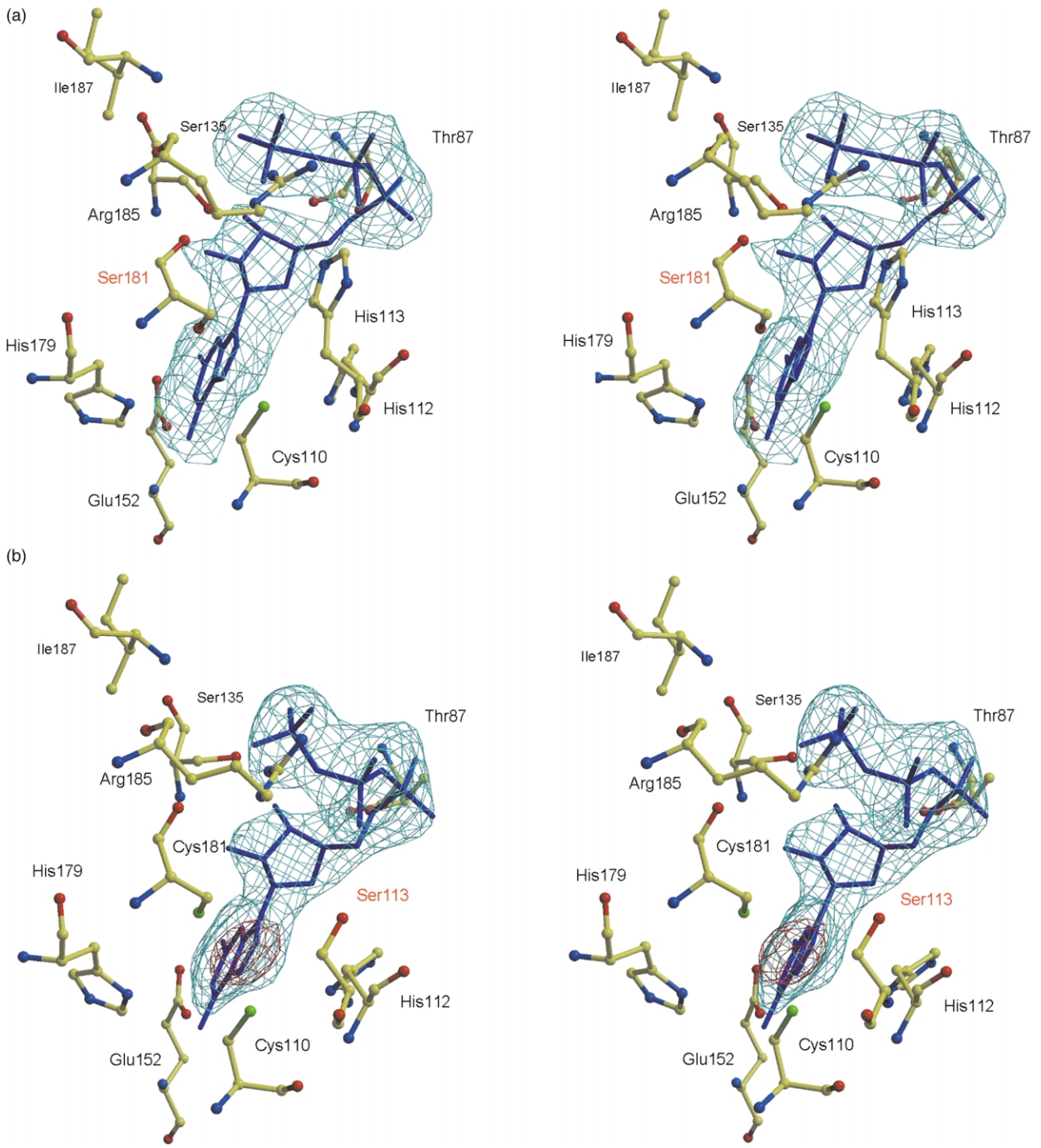

Figure 5. Stereo diagrams of the $F_{obs} - F_{calc}$ electron density maps immediately after rigid-body refinement. Wired molecular model of GTP (dark blue) is superimposed on the GTP assignable electron density maps. Protein amino acid residues are represented using ball-and-stick models. (a) Mutant Cys181Ser with a 2.0σ contour level $F_{obs} - F_{calc}$ electron density maps, coloured light blue. (b) Mutant His113Ser model with mutation. The $F_{obs} - F_{calc}$ electron density map, contoured at the 2.5σ level is coloured as light blue. Additionally red coloured $F_{obs} - F_{calc}$ electron density map contoured at the 4.5σ level shows the presence of an undetermined metal ligand occupying the wild-type zinc position. The Figure was elaborated using MOLSCRIPT[39] and Raster3D.[40]

residues, one histidine residue and one water molecule is cytidine deaminase, which catalyses the conversion of cytidine to uracil.[18–20] It has been demonstrated that cytidine deaminase binds transition state analogues formed *in situ* by hydration of 5-fluorozebularine and zebularine with very high affinity. In these complexes, the zinc-bound water molecule is replaced by the substrate analogue hydrate, a state resembling the transition state of the GTP imidazole ring opening step in the GTP cyclohydrolase I reaction.

Table 3. Hydrogen bonds

Guanine base (^{A}GTP)	Protein
N1	^{A}Glu152 O$^{\varepsilon 1}$
N2	^{A}Glu152 O$^{\varepsilon 2}$
	^{B}Ile132 O
O6	^{A}Gln151N
N3	^{B}Leu134N
Ribose (^{A}GTP)	
O2′	^{B}Ser135N
O3′	^{B}Ser135 O$^{\gamma}$
O5′	^{B}Thr87 O$^{\gamma 1}$
Triphosphate (^{A}GTP)	
O1A	^{B}Lys136 Nz
O2A	^{D}Arg65 NH1
O1G	^{B}Arg139 NH1
	^{A}Arg185 NH1
O2G	^{B}Arg139 NH1
	^{B}Ser135 O$^{\gamma}$
	^{B}Lys136 Nz
O3G	^{A}Arg185 NH2
	^{B}Ser135 O$^{\gamma}$

Hydrogen bonds established between the active site residues and the bound GTP substrate molecule, as observed on the crystallographic structure of the GTP cyclohydrolase I His113Ser mutant. Superscripts mark the molecular subunits.

Discussion

GTP cyclohydrolase I catalyses a complex reaction sequence. The formamide intermediate, compound 2, has been confirmed by spectroscopic analysis.[5] It has also been shown that the carbohydrate rearrangement involves the introduction of a proton from solvent into the pro-7*R* position of dihydroneopterin triphosphate.[21]

Whereas the enzyme reaction had been known for more than three decades, the presence of an essential zinc ion at the active site of the enzyme had escaped detection. Moreover, auto-oxidation of the zinc-depleted *E. coli* enzyme had resulted in the formation of an artifactual disulphide bond between cysteine residues 110 and 181, which were no longer engaged in zinc chelation.[10] Disulphide bond formation in the absence of zinc was also reported for cytidine deaminase, whose metal coordination resembles GTP cyclohydrolase I.[22] The presence of the metal cation at the active site was ultimately discovered during the crystallographic analysis of the human protein which had been purified without the use of chelating agents.[11]

We have now shown that the replacement of Cys110, Cys181, His112 or His113 by serine affords mutant proteins with reduced metal binding capability. Each of the mutants was unable to convert GTP to dihydroneopterin 3′-triphosphate, even after addition of $ZnCl_2$. Notably, each of these mutant proteins contained nucleotide triphosphate GTP that was not removed during purification. These findings leave no doubt that all four amino acid residues are required for the chelation of active zinc with significant affinity. However, N$^{\varepsilon}$ of His112 is 4.5 Å apart from the zinc ion and is unlikely to interact directly with the metal ion. The loss of activity in this mutant may result from a lack of fixation of the zinc-bound water molecule by its short hydrogen bond to His112 and in consequence lead to the loss of the reactive ligand of the zinc. On the other side, the hydrogen bond from the His113 imidazole ring to the mutated His112Ser effectively prevents zinc coordination (Figure 6(b)). On the His113Ser mutant, the absence of His113 imidazole ring allows an alternate side-chain conformation for residue His112 (Figure 6(a)).

The mutants under study are also unable to catalyse the release of formate from GTP and from the formamide type intermediate, compound 2. It follows that a correctly positioned zinc ion is not only required for the opening of the imidazole ring but also for the subsequent hydrolysis of the formamide motif. This reaction is highly reminiscent of the hydrolysis of peptide bonds by zinc proteases.[23,24]

The data suggest that zinc acts as a Lewis acid activating water molecules (i) for the hydration of the imidazole ring of GTP, which precedes the ring opening, and (ii) for the hydration of the formyl group of compound 2. The residue His112 appears to be required to properly orientate the zinc-activated water for nucleophilic attack on the substrate and to accept the proton released in this step (Figure 7). As corroborated by the orientation of His112 in the complex structure of zinc-depleted GTP cyclohydrolase I with deoxy-GTP, the resulting imidazolium group of His112 could rotate and transfer the proton to the ribose ring oxygen of the GTP hydrate. This would facilitate the opening of the imidazole ring in the substrate by intermediary formation of a Schiff base as proposed.[9] In analogy to the mechanism of zinc proteases, the coordination number of zinc could then be increased to five through complexation of an additional water molecule which would attack the zinc-complexed formyl group of the intermediate 9 conducive to a tetrahedral transition state (intermediate 10) for the abstraction of formate.[25–27] The imidazole side-chain of the adjacent residue His179 could support the cleavage of the formamide bond by acid catalysis (Figure 7). This role is supported by the reversible conversion of GTP to 2-amino-5-formylamino-6-(ribosylamino)-4(3H)-pyrimidinone 5′-triphosphate catalysed by His179 mutants of GTP cyclohydrolase I.

Single turnover kinetic studies showed that the velocity of the ring opening reaction exceeds that of product formation by approximately one order of magnitude.[4,6,28] More recent studies suggest that the hydrolysis of the formamide bond is even more rapid than the ring opening step.[6,29] Thus, the hydrolytic reactions involving zinc ion catalysis are rapid by comparison with the isomerization of the carbohydrate moiety.

Comparison of the zinc-containing and zinc-free wild-type enzymes and the zinc-deficient mutants indicates that the topology of the active site is relatively rigid and apparently independent of a

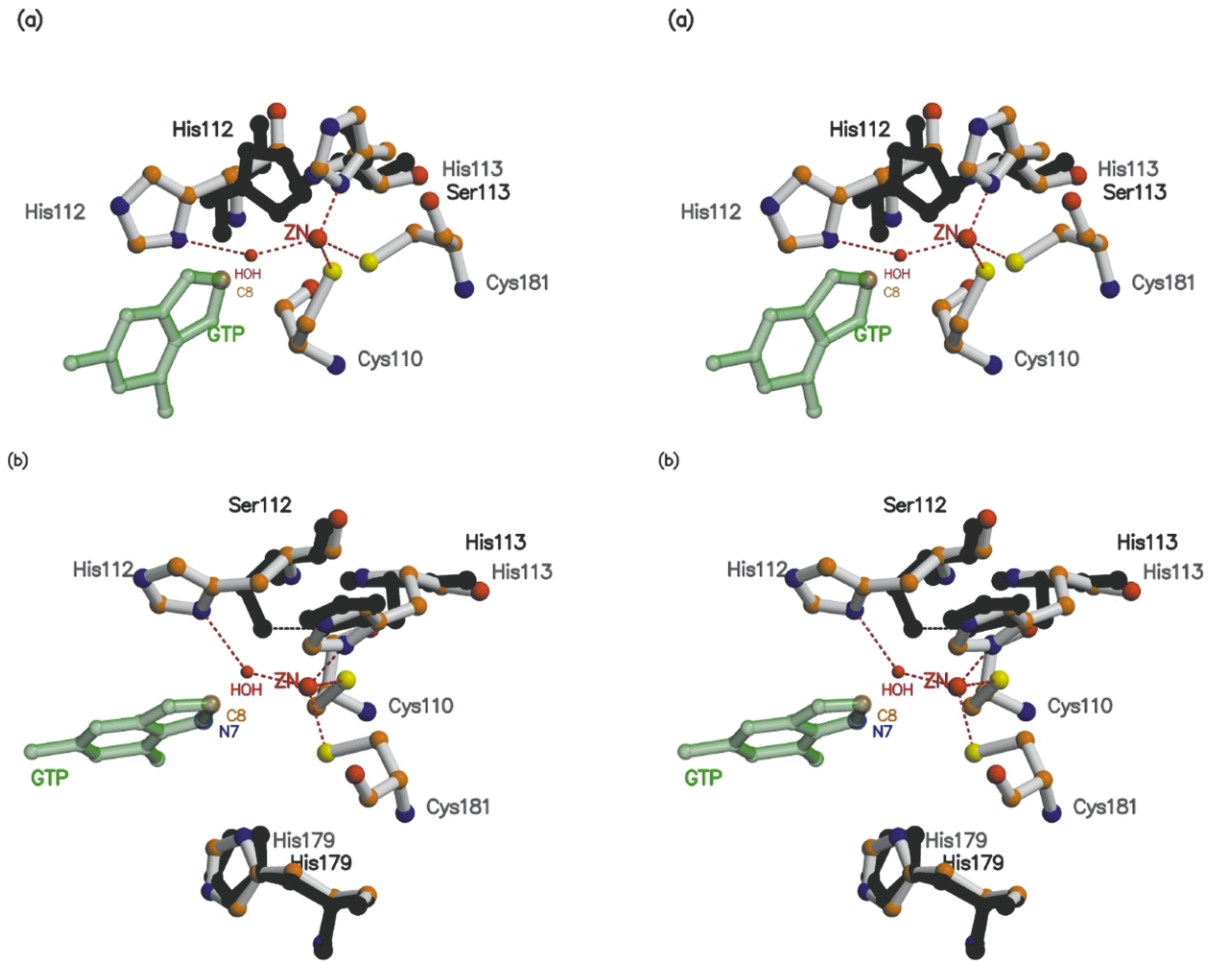

Figure 6. Orientation of the GTP substrate related to the zinc cofactor on the active site of GTP cyclohydrolase I. The peptide backbones from the wild-type GTP cyclohydrolase I and from the GTP complex of (a) His113ser and (b) His112Ser mutants were aligned. The resulting locations of the His112, His113 side-chains, of wild-type zinc and the mechanistically important zinc-bound water molecule are superimposed on the binding pocket. The wild-type residues are printed in grey. The zinc ion, the catalytic active water molecule and all the bonds are coloured red. Mutated structures and the hydrogen bonds within are coloured in black. The Figure was created using MOLSCRIPT[39] and Raster3D.[40]

possible thiol oxidation or metallic coordination. This view is also supported by the crystallographic temperature factors, which are low for the active site area. The structural comparison also suggests that the conformation of the substrate GTP is substantially constrained by the presence of the zinc ion. A putative deformation of the guanine moiety could further assist the water addition mediated by the zinc ion. The substrate GTP is most probably fixed on both ends during the enzymatic reaction. It is likely to be flexible only on the phosphate P^{α}, ribose and imidazole ring regions.

Materials and Methods

Purification of mutant GTP cyclohydrolase I proteins

Previously reported recombinant *E. coli* strains for hyper-expression of mutants of GTP cyclohydrolase I were grown and harvested as described.[9] No supplement of zinc was added to the growth solution. Cells suspended in 25 mM Tris–hydrochloride (pH 8.0), containing 20 mM KCl and 3 mM sodium azide were lysed by ultrasonic treatment. The suspension was centrifuged, and the supernatant was heated to 60 °C for ten minutes. After centrifugation, the supernatant was dialysed against 25 mM Tris–hydrochloride (pH 8.2), containing 20 mM potassium chloride, 3 mM sodium azide, and 24 mM GTP. The solution was applied to a column of Sepharose Q (Pharmacia; 3.8 cm × 11 cm), which was developed with a gradient of 100 mM to 600 mM potassium chloride. Fractions were combined and applied to a Superdex-200 column (Pharmacia), which was developed with 25 mM Tris–hydrochloride (pH 8.2), containing 100 mM KCl, and 3 mM sodium azide.

Enzyme assays

GTP cyclohydrolase I activity was determined as described.[14,21]

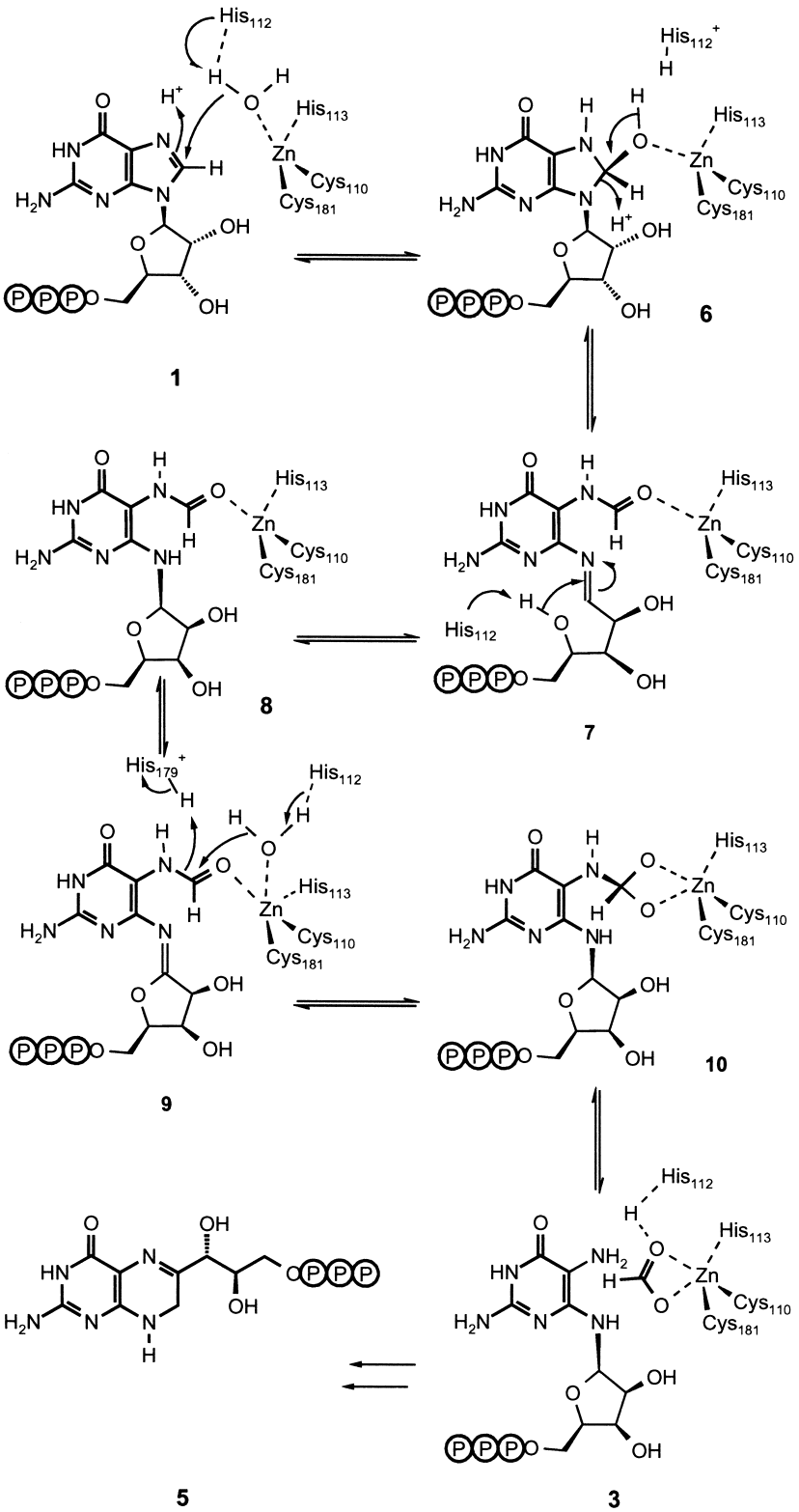

Figure 7. Detail from the hypothetical reaction mechanism for GTP cyclohydrolase I. This mechanism comprises the following steps. Nucleophilic attack by a zinc complexed water molecule to the position C-8 from the imidazole ring of the nucleotide base (transition from state 1 to state 6). Hydrolytic opening of the imidazole ring (transition from state 6 to state 8) with transient formation of a Schiff base (state 7). Hydration from the formyl group of the 2-amino-5-formylamino-6-(ribosylamino)-4(3H)-pyrimidone 5′-triphosphate (transition from state 8 to state 9). Formamide hydrated tetrahedral state supported by a penta-coordinated zinc (state 10). Release of formate (transition from state 10 to state 3). Amadori rearrangement of the ribose ring and subsequent closure of the dihydropyrazine ring by intramolecular condensation (transition from state 3 to state 5).

Determination of zinc

To a solution of enzyme (3 mg/ml) in 50 mM Tris–hydrochloride (pH 8.0), hydrochloric acid was added to a final acid concentration of 1 M. This mixture was heated to 96 °C for two hours. Zinc was then determined by atomic absorption spectrometry using a Unicam 919 spectrometer (Unicam, Cambridge, UK).

Determination of GTP cyclohydrolase I concentration

GTP cyclohydrolase I concentrations were measured by radial immunodiffusion using plates containing 0.56% (w/v) agarose, 100 mM potassium phosphate (pH 7.0), 3 mM sodium azide, and 0.25% (v/v) GTP cyclohydrolase I anti-serum from rabbit.[30] Wild-type GTP cyclohydrolase I was used as standard.

Crystallization and crystallographic data collection

Crystallization experiments were performed at room temperature employing the vapour diffusion technique. Crystals of GTP cyclohydrolase I mutant proteins were obtained under the following conditions. (a) Mutant His112Ser*/unsoaked crystals: 0.1 M sodium Mes (pH 6.0), containing 0.2 M sodium acetate and 3 mM sodium azide. (b) Mutants His112Ser and Cys181Ser: 0.1 M Mops (pH 7.0), with 10% (w/v) polyethylene glycol 6000 and 0.1 M ammonium sulphate. (c) Mutant His113Ser: 0.1 M Tris–hydrochloride (pH 8.5), containing 0.2 M ammonium dihydrogen phosphate and 50% (v/v) MPD. The orthorhombic crystals corresponding to the two first crystallization conditions (conditions (a) and (b)) belong to the space group $C222_1$ with 15 monomers in the asymmetric unit and lattice constant $a = 315.9$ Å, $b = 220.6$ Å, $c = 131.4$ Å for mutant His112Ser* (under cryo conditions), $a = 224.46$ Å, $b = 313.26$ Å, $c = 130.18$ Å for mutant His112Ser (under cryo conditions) and $a = 223.14$ Å, $b = 317.8$ Å, $c = 132.1$ Å for mutant Cys181Ser at room temperature. The crystals of mutant His113Ser obtained with MPD as precipitant (condition (c)) belong to the space group $P4_32_12$ with ten monomers in the asymmetric unit and lattice constants $a = b = 123.9$ Å, $c = 388.6$ Å. The cell angles are $\alpha = \beta = \gamma = 90°$ for all cases.

Crystals were soaked for 30 minutes in harvesting buffer supplemented with GTP to a final concentration of 10 mM GTP. After soaking, the crystals were immediately frozen in a nitrogen stream using glycerol or MPD (mutant His113Ser) as cryo-protectant.

The data set of the complex of Cys181Ser with GTP was collected to a limit resolution of 3.2 Å from a single crystal at room temperature on a Mar Research imaging plate system installed on a Rigaku rotating anode generator ($\lambda = 1.54$ Å; 0.5° frames). The other data sets were collected from single frozen crystals at the BW6 beam-line of DESY (Hamburg) ($\lambda = 1.05$ Å; 0.3° frames) at a temperature of 95 K using a CCD detector.

Data processing and refinement

Indexation and integration of the crystal diffraction images from the mutants His112Ser* and GTP-Cys181Ser was made using the MOSFLM package,[31] and from the mutants GTP-His112Ser and GTP-His113Ser using the DENZO/SCALEPACK[32] package. Further processing, was carried by the CCP4 suite of programs.[33] A total of 661,363 measured reflection intensities for the His112Ser* merged on to 238,566 unique reflections with R_{merge} values on intensities of 8.5%. This set represents 90.5% of the possible data to a resolution of 2.1 Å. For the His113Ser complex, 100,128 unique merged reflections (R_{merge} value of 4.7%) were obtained from a total of 1,005,735 collected measurements, representing 96.9% completeness of the data. The His112Ser complex yield 110,748 unique reflections from 2,178,312 intensity reflections measured, characterized by R_{merge} value of 5.1% and completeness of 97.4% for the total data to 2.55 Å resolution. Finally, the Cys181Ser complex 1,260,271 intensity measurements resulted on 70,145 unique reflections complete to 98.6% with overall R_{merge} value of 6.3% for all the data to a resolution of 3.2 Å (Table 2).

The model of the wild-type enzyme[9,10] was used for refinement of the mutant complex structures after molecular replacement with Molrep,[34] followed by rigid-body refinement with CNS.[35] The electron density of His112Ser* was modified applying non-crystallographic symmetry averaging with the program MAIN.[36] Model building was done with O[37] and refinement with CNS.[35,38] The 15 monomers of His112Ser*, His112Ser and Cys181Ser and the ten monomers of His113Ser per asymmetric unit were strongly constrained by non-crystallographic symmetry during initial refinement and released at the medium to final stages. All models refined to satisfying overall quality, with *R*-factors of $R_{His112Ser*} = 20.0\%$, $R_{His113Ser} = 26.2\%$, $R_{His112Ser} = 21.8\%$ and $R_{Cys181Ser} = 18.5\%$ obtained at 3σ cut-off level. Other important refinement information is also presented in Table 2.

Ribbon representations were drawn with the programs MOLSCRIPT[39] and rendered with Raster3D.[40]

Protein Data Bank accession numbers

Coordinates of the complex structures His112Ser*, His112Ser/GTP, His113Ser/GTP and Cys181Ser/GTP have been deposited in the RCSB Brookhaven Protein Data Bank with accession codes 1A8R, 1N3R, 1N3S and 1N3T, respectively.

Acknowledgements

This work was supported in part by grants from the Deutsche Forschungsgemeinschaft, the European Community (EC grant ERB CHRX CT93-0243 and ERB FMRX CT98-0204), the Fonds der Chemischen Industrie, the Hans Fischer Gesellschaft and PhD grant PRAXISXXI/BD/21493/99 (to J.R.) from the Fundação para a Ciência e Tecnologia (Portugal). We thank Dr H. D. Bartunik and his group for expert assistance during crystallographic data collection at the BW6 beamline at DESY (Hamburg). We also thank A. Werner for expert secretarial assistance and Dr K. Maskos for helpful discussions.

References

1. Bacher, A., Rieder, C., Eichinger, C., Arigoni, D., Fuchs, G. & Eisenreich, W. (1999). Elucidation of novel biosynthetic pathways and metabolite flux patterns by retrobiosynthetic NMR analysis. *FEMS Microbiol. Rev.* **22**, 567–598.
2. Shiota, T., Baugh, C. M. & Myrick, J. (1969). The assignment of structure to the formamidopyrimidine nucleoside triphosphate precursor of pteridines. *Biochim. Biophys. Acta*, **192**, 205–210.
3. Shiota, T., Palumbo, M. P. & Tsai, L. (1967). A chemically prepared formamidopyrimidine derivative of guanosine triphosphate as a possible intermediate in pteridine biosynthesis. *J. Biol. Chem.* **242**, 1961–1969.
4. Schramek, N., Bracher, A. & Bacher, A. (2001). Ring opening is not rate-limiting in the GTP cyclohydrolase I reaction. *J. Biol. Chem.* **276**, 2622–2626.
5. Bracher, A., Eisenreich, W., Schrameck, N., Ritz, H., Götze, E., Herrmann, A. *et al.* (1998). Biosynthesis of Pteridines. NMR studies on the reaction mechanism of GTP cyclohydrolase, pyruvolyltetrahydropterin

synthase and sepiapterin reductase. *J. Biol. Chem. Mol. Biol. Inc.* **273**, 28132–28141.
6. Schramek, N., Bracher, A. & Bacher, A. (2001). Biosynthesis of riboflavin. Single turnover kinetic analysis of GTP cyclohydrolase I. *J. Biol. Chem.* **276**, 44157–44162.
7. Weygand, F., Simon, H., Dahms, G., Waldschmidt, M., Schliep, H. J. & Wacker, H. (1961). Über die Biogenese des Leucopterins. *Angew. Chem.* **73**, 402–407.
8. Burg, A. W. & Brown, G. M. (1968). The biosynthesis of folic acid. VIII. Purification and properties of the enzyme that catalyzes the production of formate from carbon atom 8 of guanosine triphosphate. *J. Biol. Chem.* **243**, 2349–2358.
9. Nar, H., Huber, R., Auerbach, G., Fischer, M., Hösl, C., Ritz, H. *et al.* (1995). Active site topology and reaction mechanism of GTP cyclohydrolase I. *Proc. Natl Acad. Sci. USA*, **92**, 12120–12125.
10. Nar, H., Huber, R., Meining, W., Schmid, C., Weinkauf, S. & Bacher, A. (1995). Atomic structure of GTP cyclohydrolase I. *Structure*, **3**, 459–466.
11. Auerbach, G., Herrmann, A., Bracher, A., Bader, G., Gütlich, M., Fischer, M. *et al.* (2000). Zinc plays a key role in GTP cyclohydrolase I. *Proc. Natl Acad. Sci. USA*, **97**, 13567–13572.
12. Brown, G. M. (1962). The biosynthesis of folic acid II. Inhibition by sulfamides. *J. Biol. Chem.* **237**, 536–540.
13. Tenover, F. C. (2001). Development and spread of bacterial resistance to antimicrobial agents. An overview. *Clin. Infect. Dis.* **33**, S108–S115.
14. Bracher, A., Fischer, M., Eisenreich, W., Ritz, H., Schramek, N., Boyle, P. *et al.* (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.
15. Katzenmeier, G., Schmid, C., Kellermann, J., Lottspeich, F. & Bacher, A. (1991). Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I from *Escherichia coli*. *Biol. Chem. Hoppe-Seyler*, **372**, 991–997.
16. Maier, J., Witter, K., Gütlich, M., Ziegler, I., Werner, T. & Ninnemann, H. (1995). Homology cloning of GTP-cyclohydrolase I from various unrelated eukaryotes by reverse-transcription polymerase chain reaction using a general set of degenerate primers. *Biochem. Biophys. Res. Commun.* **212**, 705–711.
17. Christianson, D. W. & Cox, J. D. (1999). Catalysis by metal-activated hydroxide in zinc and manganese metalloenzymes. *Annu. Rev. Biochem.* **68**, 33–57.
18. Betts, L., Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr (1994). Cytidine deaminase. The 2.3 Å crystal structure of an enzyme: transition-state analog complex. *J. Mol. Biol.* **235**, 635–656.
19. Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr (1995). Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase. *Biochemistry*, **34**, 4516–4523.
20. Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr (1996). Cytidine deaminase complexed to 3-deazacytidine: a valence buffer in zinc enzyme catalysis. *Biochemistry*, **35**, 1335–1341.
21. Bracher, A., Eisenreich, W., Schramek, N., Ritz, H., Götze, E., Herrmann, A. *et al.* (1998). Biosynthesis of pteridines. NMR studies on the reaction mechanisms of GTP cyclohydrolase I, pyruvoyltetrahydropterin synthase, and sepiapterin reductase. *J. Biol. Chem.* **273**, 28132–28141.
22. Mejlhede, N. & Neuhard, J. (2000). The role of zinc in *Bacillus subtilis* cytidine deaminase. *Biochemistry*, **39**, 7984–7989.
23. Christianson, N. & Lipscomb, W. N. (1988). Structural aspects of zinc proteases mechanisms. *Mol. Struct. Energ.* **9**, 1–25.
24. Lipscomb, W. N. & Sträter, N. (1996). Recent advances in zinc enzymology. *Chem. Rev.* **96**, 2375–2433.
25. Hangauer, D. G., Monzingo, A. F. & Matthews, B. W. (1984). An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxymethyl dipeptides. *Biochemistry*, **23**, 5730–5741.
26. Christianson, D. W., David, P. R. & Lipscomb, W. N. (1987). Mechanism of carboxypeptidase A: hydration of a ketonic substrate analogue. *Proc. Natl Acad. Sci. USA*, **84**, 1512–1515.
27. Holden, H. M. & Matthews, B. W. (1988). The binding of L-valyl-L-tryptophan to crystalline thermolysin illustrates the mode of interaction of a product of peptide hydrolysis. *J. Biol. Chem.* **263**, 3256–3260.
28. Bracher, A., Schramek, N. & Bacher, A. (2001). Biosynthesis of pteridines. Stopped flow kinetic analysis of GTP cyclohydrolase I. *Biochemistry*, **40**, 7896–7902.
29. Schramek, N., Bracher, A., Fischer, M., Auerbach, G., Nar, H., Huber, R. & Bacher, A. (2002). Reaction mechanism of GTP cyclohydrolase I: single turnover experiments using a kinetically competent reaction intermediate. *J. Mol. Biol.* **316**, 829–837.
30. Schödon, G., Redweik, U. & Curtius, H. C. (1989). Purification of GTP cyclohydrolase I from human liver and production of specific monoclonal antibodies. *Eur. J. Biochem.* **178**, 627–634.
31. Leslie, A. G. W. (1991). Macromolecular data processing. In *Crystallographic Computing* (Moras, V. D., Podjarny, A. D. & Thierry, J. C., eds), pp. 27–38, Oxford University Press, Oxford.
32. Otwinowski, Z. & Minor, W. (1993). *DENZO: Film Processing Program For Macro-molecular Crystallography*, Yale University Press, New Haven, CT.
33. Bailey, S. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
34. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022–1025.
35. Brünger, A. T. (1992). *X-PLOR, Version 1.3: A System for Crystallography and NMR*, vol. 3, Yale University Press, New Haven, CT.
36. Turk, D., (1992). *Development and usage of a macromolecular graphics program*. PhD thesis, Technical University Munich, Germany
37. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. (1990). O: a macromolecule modelling environment. In *Crystallography and Modelling Methods in Molecular Design* (Bugg, C. & Ealick, S., eds), pp. 189–199, Springer Verlag, New York.
38. Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A*, **47**, 392–400.
39. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 945–949.
40. Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869–873.

41. Holm, L. & Sander, C. (1998). Removing near-neighbor redundancy from large protein sequence collections. *Bioinformatics*, **14**, 423–429.

42. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.

Edited by A. R. Fersht

(Received 2 August 2002; received in revised form 4 November 2002; accepted 11 November 2002)