Carbamoyl Phosphate Synthetase: Closure of the B-Domain as a Result of Nucleotide Binding[†,‡]

James B. Thoden,*[,§] Gary Wesenberg,[§] Frank M. Raushel,[||] and Hazel M. Holden*[,§]

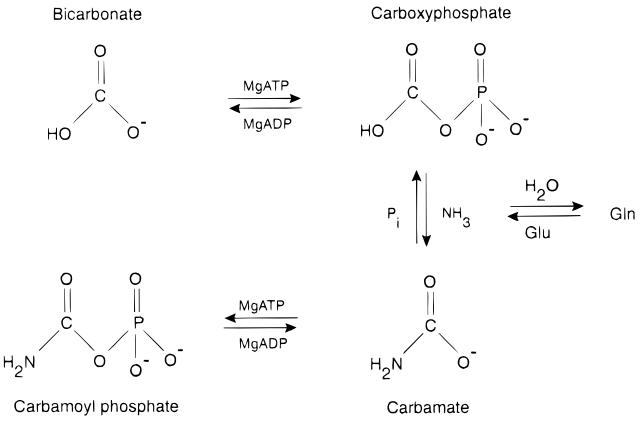
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin—Madison, 1710 University Avenue, Madison, Wisconsin, 53705, and Department of Chemistry, Texas A&M University, College Station, Texas, 77843

Received October 21, 1998; Revised Manuscript Received December 11, 1998

ABSTRACT: Carbamoyl phosphate synthetase (CPS) catalyzes the production of carbamoyl phosphate which is subsequently employed in the metabolic pathways responsible for the synthesis of pyrimidine nucleotides or arginine. The catalytic mechanism of the enzyme occurs through three highly reactive intermediates: carboxyphosphate, ammonia, and carbamate. As isolated from *Escherichia coli*, CPS is an α,β-heterodimeric protein with its three active sites separated by nearly 100 Å. In addition, there are separate binding sites for the allosteric regulators, ornithine, and UMP. Given the sizable distances between the three active sites and the allosteric-binding pockets, it has been postulated that domain movements play key roles for intramolecular communication. Here we describe the structure of CPS from *E. coli* where, indeed, such a domain movement has occurred in response to nucleotide binding. Specifically, the protein was crystallized in the presence of a nonhydrolyzable analogue, AMPPNP, and its structure determined to 2.1 Å resolution by X-ray crystallographic analysis. The B-domain of the carbamoyl phosphate synthetic component of the large subunit closes down over the active-site pocket such that some atoms move by more than 7 Å relative to that observed in the original structure. The trigger for this movement resides in the hydrogen-bonding interactions between two backbone amide groups (Gly 721 and Gly 722) and the β- and γ-phosphate groups of the nucleotide triphosphate. Gly 721 and Gly 722 are located in a Type III′ reverse turn, and this type of secondary structural motif is also observed in D-alanine:D-alanine ligase and glutathione synthetase, both of which belong to the "ATP-grasp" superfamily of proteins. Details concerning the geometries of the two active sites contained within the large subunit of CPS are described.

Carbamoyl phosphate synthetase, hereafter referred to as CPS, catalyzes the production of carbamoyl phosphate from two molecules of Mg^{2+}ATP, one molecule of bicarbonate, and either glutamine or ammonia as a nitrogen source (*1*). The catalytic mechanism for this remarkable transformation has been proposed to involve four discrete chemical steps that utilize carboxyphosphate, ammonia, and carbamate as unstable reaction intermediates (*2*). Experimental support for this catalytic mechanism has been obtained from partial reaction studies, kinetic investigations, and positional isotope exchange experiments (*2*−*5*). The chemical mechanism for this complex process is summarized below in Scheme 1. Note that carbamoyl phosphate is subsequently employed in the biosynthetic pathways leading to the formation of pyrimidine nucleotides or arginine.

Scheme 1



As isolated from *Escherichia coli*, CPS is an α,β-heterodimeric protein. The small subunit of molecular weight ∼42 000 functions in the hydrolysis of glutamine and the subsequent generation of the required ammonia (*6*). The larger subunit of molecular weight ∼118 000 contains the catalytic machinery necessary for the construction of carbamoyl phosphate from ammonia (*1*). In addition, this subunit provides the binding sites for the allosteric regulators UMP and ornithine. A primary structural analysis of the large subunit has revealed an apparent gene duplication during the evolution of CPS such that those residues contained within the region defined by Met 1 to Arg 400 are 40% identical to those lying within the sequence delineated by Ala 553 to Leu 933 (*7*). Each of these synthetase units has been shown

[†] This research was supported in part by grants from the NIH (GM55513 to H.M.H. and DK30343 to F.M.R.) and the NSF (BIR-9317398 shared instrumentation grant).

[‡] X-ray coordinates have been deposited in the Brookhaven Protein Data Bank (1BXR).

* To whom correspondence should be addressed.

[§] University of Wisconsin—Madison.

[||] Texas A&M University.

10.1021/bi982517h CCC: $18.00
Published on Web 02/06/1999

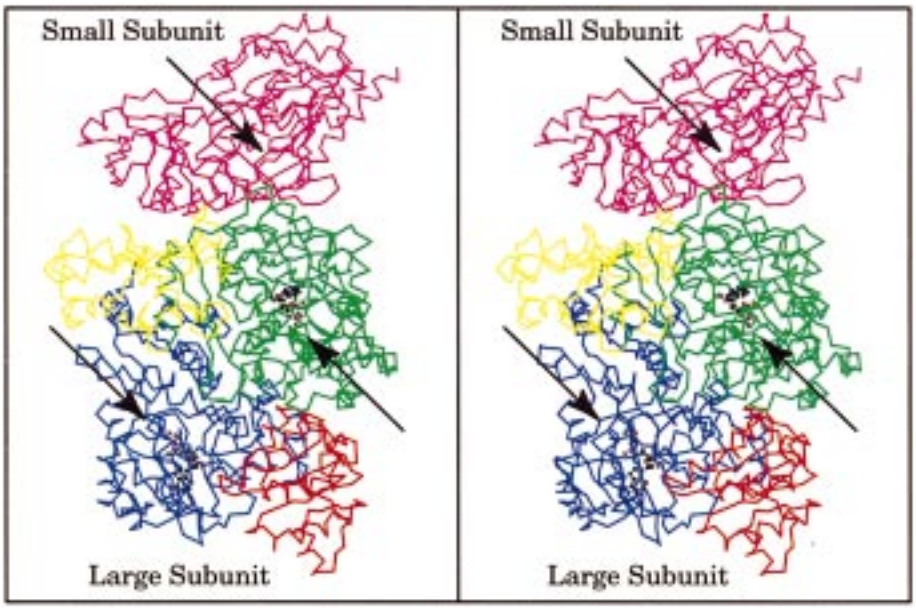


FIGURE 1: α-Carbon trace of the CPS α,β-heterodimer. The small subunit is displayed in magenta with the position of its active site marked by the black arrow. The polypeptide chain regions defined by Met 1−Glu 403, Val 404−Ala 553, Asn 554−Asn 936, and Ser 937−Lys 1073 in the large subunit are depicted in green, yellow, blue, and red, respectively. These regions correspond to the carboxyphosphate synthetic unit, the oligomerization domain, the carbamoyl phosphate synthetic unit, and the allosteric domain, respectively. The two ADP molecules bound in the large subunit are shown in ball-and-stick representations and are indicated by the black arrows. X-ray coordinates employed for this representation were from ref *10*.

to contain a binding site for one of the two molecules of Mg^{2+}ATP required for the overall synthesis of carbamoyl phosphate (*8*).

The three-dimensional architecture of the enzyme from *E. coli*, with bound ADP, P_i, manganese, potassium, and ornithine, is now known from a high-resolution X-ray crystallographic analysis (*9*, *10*). This structural investigation revealed the relative location of the three independent active sites as indicated in Figure 1. Remarkably, the two active sites contained within the large subunit are separated by a linear distance of ∼35 Å and connected by an intramolecular tunnel. As can be seen in Figure 1, the small subunit forms molecular contacts with only the N-terminal half of the large subunit. Additionally, the active site identified in the small subunit for glutamine hydrolysis is ∼45 Å away from the binding site for the Mg^{2+}ATP that is required for the phosphorylation of bicarbonate. These observations have led to a proposal for the coupled synthesis of carbamoyl phosphate whereby the ammonia derived from the hydrolysis of glutamine migrates through a molecular tunnel from the small subunit to the large subunit. Bicarbonate is phosphorylated at the expense of the first Mg^{2+}ATP to produce carboxyphosphate and ADP in the first active site of the large subunit. The ammonia subsequently reacts with the carboxyphosphate intermediate in this first active site of the large subunit to produce carbamate. Finally, the unstable carbamate intermediate migrates from the site of its formation in the N-terminal half of the large subunit to the companion site in the C-terminal half of the protein. There, it is ultimately phosphorylated by the second molecule of Mg^{2+}ATP to form the product carbamoyl phosphate. Quite strikingly, the entire journey through the interior of CPS covers a distance of nearly 100 Å.

The three-dimensional motifs of the individual carboxyphosphate and carbamoyl phosphate synthetic units of the large subunit of CPS (Met 1 to Arg 400 and Ala 553 to Leu 933, respectively) are topologically similar but not identical to one another as can be seen in Figure 2a. This observation is in full accord with the different biochemical functions required of these two synthetase components. Each one of these two active sites must utilize Mg^{2+}ATP to phosphorylate either bicarbonate or carbamate, respectively. However, only the N-terminal half of the large subunit provides a binding site for the small subunit, a tunnel for the passage of ammonia and a region for the stabilization of the carboxyphosphate intermediate. Moreover, the allosteric effects exhibited by UMP and ornithine are predominantly transmitted through the C-terminal half of the large subunit (*11*).

In the original structure of CPS, the active site of the carboxyphosphate synthetic unit contained ADP, P_i, two manganese ions, and several potassiums while the carbamoyl phosphate component lacked the inorganic phosphate and one of the divalent metals. The most striking structural difference between these two halves of the large subunit was in the relative positioning of the so-called B-domains. In the carboxyphosphate synthetic unit, this B-domain was in a closed conformation as observed for the same motif in the related structure of D:alanine D:alanine ligase (*12*). The B-domain of the carbamoyl phosphate synthetic unit, however, was "more open" with several surface loops disordered as can be seen in Figure 2a. The openings and closings of these B-domains most likely influence both the catalytic activities of these two active sites in the large subunit and the movements of substrates and products within the enzyme. Furthermore, such conformational changes may transmit critical information for the coordinated control of the two active sites with one another. Presently, there is little known regarding the structural features that trigger the movements of the B-domains in CPS although such differences in the original structure may have been, in part, a function of the presence or absence of an inorganic phosphate in the carboxyphosphate and carbamoyl phosphate synthetic units,

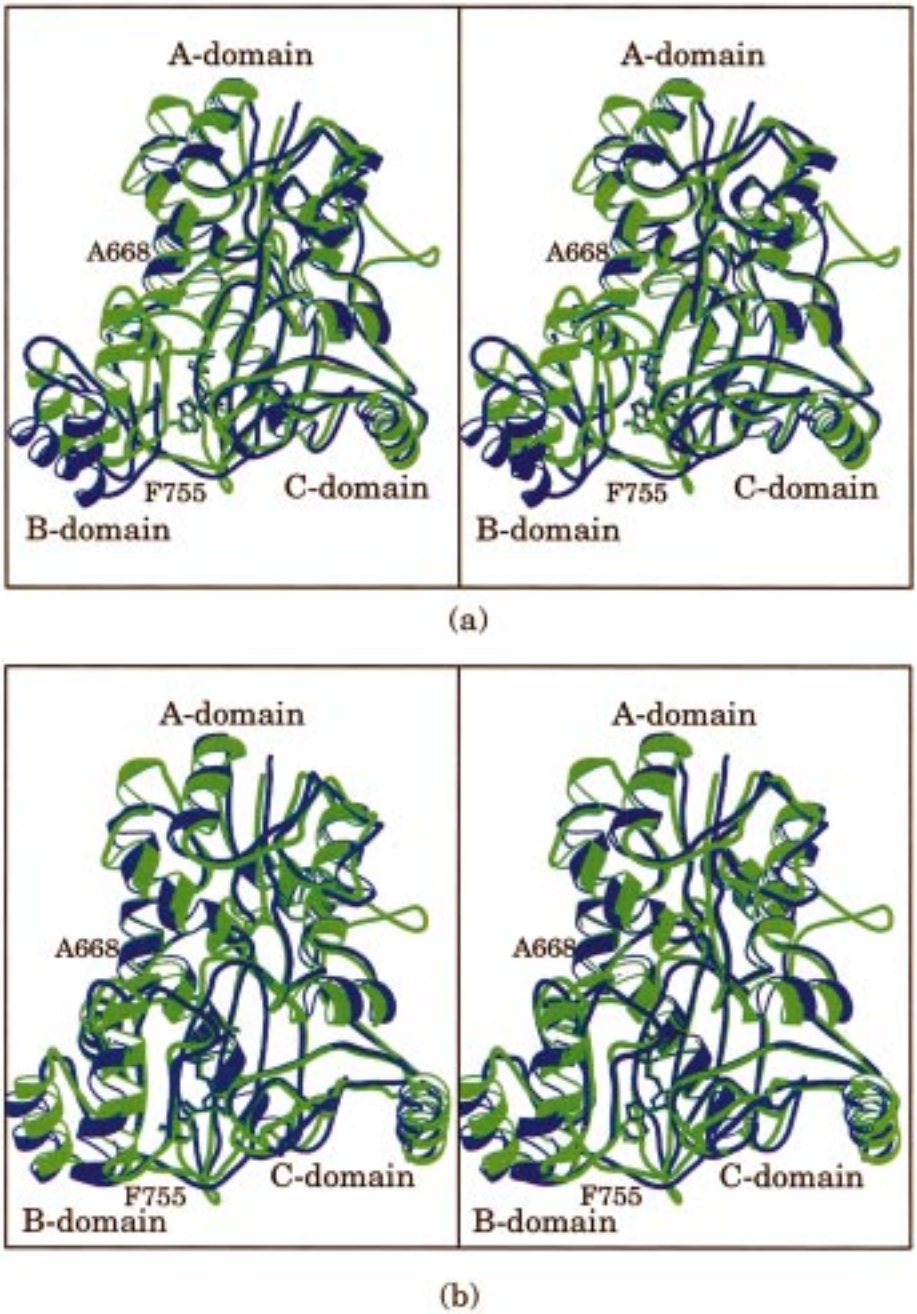


FIGURE 2: Superposition of the carboxyphosphate and carbamoyl phosphate synthetic units of the CPS large subunit. Shown in panel a is the carboxyphosphate synthetic component from Ser 9 to Glu 403 (depicted in green) superimposed onto the carbamoyl phosphate synthetic unit from Glu 560 to Asn 936 (displayed in blue). This superposition is based on the original structure (*10*) where the carboxyphosphate synthetic unit contained, in addition to ADP, a molecule of inorganic phosphate while the carbamoyl phosphate synthetic portion contained only ADP. Note that the B-domain in the carbamoyl phosphate synthetic unit is splayed away from the main body of the molecule. Shown in panel b are the same synthetic components superimposed upon one another, but in this case, both active sites contain AMPPNP. As can be seen, the two synthetase units are now virtually identical with respect to the disposition of the B-domains.

respectively. Here we report the structure of CPS cocrystallized in the presence of the nonhydrolyzable ATP analogue, AMPPNP[1] (5′-adenylylimidodiphosphate), first synthesized in Dr. Ralph Yount's laboratory (*13*). This structure mimics the Michaelis complex of ATP prior to the occurrence of any bond-making or bond-breaking steps. As expected, the B-domains of both synthetic units clamp down upon nucleotide binding. Details concerning the B-domain movements and the active-site geometries are presented.

MATERIALS AND METHODS

Purification and Crystallization Procedures. Protein employed in this investigation was purified as described elsewhere (*14*). Crystals were grown by batch at 4 °C from 0.65 M tetraethylammonium chloride, 8% (w/v) poly(ethylene glycol) 8000, 100 mM KCl, 0.5 mM $MnCl_2$, 0.5 mM L-ornithine, 1.25 mM AMPPNP, and 25 mM HEPPS (pH 7.4) with a typical protein concentration of 3.5 mg/mL.

[1] AMPPNP is a competitive inhibitor versus ATP in the full forward reaction of CPS from *E. coli*. The K_i value is 0.4 mM versus ATP when the bicarbonate and glutamine concentrations are held at 1.0 mM and 0.15 mM, respectively (Dr. Grant E. Gibson, unpublished results).

Table 1: Intensity Statistics

	resolution range (Å)										
	overall	30.0−4.52	3.59	3.14	2.85	2.65	2.49	2.37	2.26	2.18	2.10
no. of measurements	1 657 168	297 176	313 021	281 952	136 922	125 905	116 674	106 167	95 317	84 230	72 734
no. of independent reflections	446 040	48 724	48 140	47 824	46 587	45 791	44 873	43 762	42 156	40 258	37 925
% completeness	93	99	100	99	97	96	94	92	88	84	80
avg I/avg σ (I)	29.5	46.2	49.1	40.6	24.3	17.1	11.8	8.1	5.8	4.2	3.1
R factor (%)[a]	5.2	4.1	4.9	6.7	7.1	8.6	11.1	14.4	18.0	22.6	26.5

[a] $R_{\text{factor}} = (\sum|I - \bar{I}|/\sum I) \times 100$.

Crystals generally appeared within 1 week and achieved maximum dimensions of 0.5 mm × 0.5 mm × 1.0 mm after 6−8 weeks. These crystals belonged to the space group $P2_12_12_1$ with unit cell dimensions of $a = 151.9$ Å, $b = 164.5$ Å, and $c = 332.6$ Å and one α,β-heterotetramer per asymmetric unit.

For X-ray data collection, the crystals were first transferred to a synthetic mother liquor containing 0.75 M tetraethylammonium chloride, 8% (w/v) poly(ethylene glycol) 8000, 100 mM KCl, 0.5 mM $MnCl_2$, 0.5 mM L-ornithine, 1.25 mM AMPPNP, and 25 mM HEPPS (pH 7.4). Following equilibration overnight, the crystals were transferred to a solution of 1.0 M tetraethylammonium chloride, 8% (w/v) poly(ethylene glycol) 8000, 125 mM KCl, 0.5 mM $MnCl_2$, 0.5 mM L-ornithine, 1.25 mM AMPPNP, and 25 mM HEPPS (pH 7.4) and allowed to equilibrate for 2−3 h. Finally, the crystals were transferred stepwise to a cryo-protectant solution containing 1.35 M tetraethylammonium chloride, 8% (w/v) poly(ethylene glycol) 8000, 250 mM KCl, 0.5 mM $MnCl_2$, 0.5 mM L-ornithine, 1.25 mM AMPPNP, 25 mM HEPPS (pH 7.4), and 10% ethylene glycol. The crystals were positioned on a loop of 20 μm surgical thread and flash-cooled to −150 °C in a stream of nitrogen gas.

X-ray Data Collection and Processing. The X-ray data employed for the structural analysis described here were collected using a MAR300 image plate system at the Stanford Synchrotron Radiation Laboratory, Beam Line 7-1. A high-resolution X-ray data scan consisting of 185 frames (1°/scan) was collected with the detector at its offset position and a crystal-to-detector distance of 420 mm. A low-resolution X-ray data scan consisting of 86 frames (1°/scan) was collected with the instrument at its centered position and again located at a crystal-to-detector distance of 420 mm. Only one crystal was required for a complete X-ray data set to 2.1 Å resolution.

The frames were processed with the software package DENZO and scaled with SCALEPACK (*15*). From the 271 frames, 1 657 168 reflections were integrated which reduced to 446 040 unique reflections after scaling. Relevant X-ray data collection statistics can be found in Table 1.

The structure of the CPS−AMPPNP complex was solved by molecular replacement (*16*) with the software package AMORE (*17*) and employing as a search model the complete tetrameric form of the CPS H353N site-directed mutant protein previously determined at 1.8 Å resolution (*18*). This mutant structure was employed as the search model since it was solved at higher resolution than the native structure. All nucleotides and effector molecules were removed from the coordinate set used in the molecular replacement trials. Following rigid-body refinement, the model was subjected to least-squares analysis at 2.1 Å resolution with the software package TNT (*19*) and omitting those amino acid residues between Leu 650 to Ile 775. Following this refinement cycle, the *R*-factor decreased to 35.0%. With over 5800 amino acid residues in the asymmetric unit, the goal of the model-building process was to lower the *R*-factor as much as possible using an "averaged" α,β-heterodimer before finally rebuilding the entire tetramer in the asymmetric unit. Consequently, to expedite the refinement and rebuilding processes, the electron densities corresponding to the four α,β-heterodimers in the asymmetric unit were averaged according to the algorithm of Bricogne (*20*) and one α,β-heterodimer was rebuilt into the averaged map, including residues Leu 650 to Ile 775. Following this rebuilding process, the entire tetramer was reconstructed from the "averaged" α,β-heterodimer and placed back into the unit cell. Additional cycles of least-squares refinement decreased the *R*-factor to 27.0%. The molecular averaging and solvent-flattening procedure was repeated with the new model, and the "averaged" α,β-heterodimer was adjusted according to the new electron density map. The two AMPPNP molecules and solvents conserved between all four α,β-heterodimers constituting the $(\alpha,\beta)_4$-heterotetramer were included at this stage. Again, the entire tetramer was reconstructed from the "averaged" α,β-heterodimer and placed back into the unit cell. Least-squares refinement reduced the *R*-factor to 22.0%. Finally, the model was adjusted within the unit cell, additional solvents were added, and the model was subjected to another round of least-squares refinement. The *R*-factor was reduced to 19.5% for all X-ray data between 30.0 and 2.1 Å. Relevant refinement statistics can be found in Table 2. The following side chains were modeled as multiple conformations: Asp 757, Met 772, Lys 881, and Arg 912 in α,β-heterodimer I, Arg 104, Asn 936, and Glu 1009 in

Table 2: Least-Squares Refinement Statistics for the CPS/AMPPNP Complex

resolution limits (Å)	30.0−2.1
R-factor (%)[a]	19.5%
no. of reflections used	446 040
no. of protein atoms	44 699
no. of solvent atoms	3242[b]
weighted root-mean-square deviations from ideality	
bond length (Å)	0.010
bond angle (deg)	2.12
planarity (trigonal) (Å)	0.007
planarity (other planes) (Å)	0.012
torsional Angle (deg)[c]	18.3

[a] R-factor $= \sum|F_o - F_c|/\sum|F_o|$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude. [b] In addition to these water molecules, there were eight AMPPNP molecules, 12 manganese ions, nineteen potassium ions, 14 chloride ions, four ornithines, and four tetraethylammonium ions included in the refinement. [c] The torsional angles were not restrained during the refinement.

α,β-heterodimer III, and Arg 490 and Asn 936 in α,β-heterodimer IV. The only residues adopting dihedral angles well outside of the allowed regions of the Ramachandran plot are Cys 269 and Ala 356, both of which are located in the active site of the small subunit. Indeed Cys 269 is the active site nucleophile and is situated in a "nucleophile" elbow as discussed previously (*10*). Perhaps the most remarkable technical aspect of the structure presented here is the lack of disordered regions in such a large protein. Indeed, each large subunit within the $(\alpha,\beta)_4$-heterotetramer begins at Met 1 and ends with the C-terminal Lys 1073 without a single break in the polypeptide chain backbone. The only disordered residues are located in the small subunits of the $(\alpha,\beta)_4$-heterotetramer. Specifically, the first N-terminal and the last few C-terminal residues are weak in each small subunit. Electron densities corresponding to the two AMPPNP molecules bound to the first α,β-heterodimer in the asymmetric unit are displayed in Figure 3.

RESULTS

The crystals employed in this investigation contained a complete $(\alpha,\beta)_4$-heterotetramer in the asymmetric unit. A list of the ligands and ions contained within the X-ray coordinate data set is given in Table 2. Since the quaternary structure of the enzyme with bound AMPPNP is virtually identical to that observed in the original structural analysis of CPS (*9*), only the first α,β-heterodimer of the tetramer will be discussed here. From Figure 2b, it is immediately obvious that the local effect of AMPPNP binding to the carbamoyl phosphate synthetic component of CPS is the closure of the B-domain relative to the A- and C-motifs. Other than this movement, however, the molecular architecture of the complete CPS/AMPPNP α,β-heterodimer is remarkably similar to that previously observed for the enzyme with bound ADP. Indeed, the only significant differences in the polypeptide chain backbones between these two models reside in the region defined by Ala 668 to Phe 755. If these regions are removed from the calculation, the α-carbons for both models superimpose with a root-mean-square deviation of 0.43 Å.

A close-up view of the active site for the carboxyphosphate synthetic component is shown in Figure 4a, and a cartoon of potential electrostatic interactions between the nucleotide and the protein within 3.0 Å is displayed in Figure 4b. The adenine ribose of the AMPPNP adopts a $C_{3'}$-endo pucker, and both the 2′- and 3′- hydroxyl groups are anchored to the protein via the carboxylate group of Glu 215. The purine ring interacts only with polypeptide chain backbone atoms including O of Glu 208, NH of Leu 210, and NH of Gly 241. As observed in the original CPS model, the guanidinium groups of Arg 129 and Arg 169 and the backbone amide group of Gly 176 form electrostatic interactions with the α- and β-phosphate groups of AMPPNP. The γ-phosphate group of the nucleotide interacts with the guanidinium group Arg 306 and the backbone amide group of Gly 175. Indeed, these interactions were also seen in the original structure of CPS with ADP/P_i bound in the active site of the carboxyphosphate synthetic component and may, indeed, be the trigger for the closure of the B-domain as discussed below. In contrast to that observed in the original model of CPS with bound ADP/P_i, however, there is only one metal located in the active site. This manganese ion is ligated in an octahedral coordination sphere by $O^{\epsilon 1}$ and $O^{\epsilon 2}$ of Glu 299, $O^{\delta 1}$ of Asn 301, a water molecule, and two phosphoryl oxygens contributed by the β- and γ-phosphate moieties of the nucleotide. The observed metal:ligand bond distances range in length from 2.1 to 2.4 Å and the *B*-value for the metal ion is 24.3 $Å^2$. Gln 285, which served as a ligand to the second metal in the original CPS model, now lies within hydrogen-bonding distance to one of the α-phosphoryl oxygens of AMPPNP. Also note that Glu 299 functioned as the bridging ligand for the two manganese ions observed in the original CPS model (*10*). In addition to these major protein:nucleotide interactions, there are four well-ordered water molecules located within the active-site region.

Shown in Figure 5a is a close-up view of the active site for the carbamoyl phosphate synthetic component. The structural similarity between this active site and that described above for the carboxyphosphate synthetic unit is, indeed, striking. Again, the purine ring of the nucleotide is linked to the protein via backbone amide (Leu 756 and Gly 786) and carbonyl (His 754) groups, and the 2′- and 3′-hydroxyl groups of the ribose are bridged by the carboxylate group of Glu 761. The carboxylate group of Asp 753 also forms a hydrogen bond with the adenine ring as indicated in Figure 5b. In the carbamoyl phosphate synthetic unit, only two rather than three arginines are positioned within 3.0 Å of the phosphoryl oxygens of AMPPNP, namely Arg 715 and Arg 848. Arg 675, which is the structural counterpart of Arg 129 in the carboxyphosphate synthetic component, is located at 3.5 Å from one of the β-phosphoryl oxygens of the nucleotide. Both Gly 721 and Gly 722 interact with the γ- and β-phosphoryl moieties, respectively, through their backbone amide groups. Quite strikingly, in the carbamoyl phosphate synthetic component, the active site contains two manganese ions, both of which are octahedrally coordinated. The first metal, labeled Mn I in Figure 5a, is ligated by a β- and a γ-phosphoryl oxygen of the nucleotide, $O^{\epsilon 1}$ and $O^{\epsilon 2}$ of Glu 841, $O^{\delta 1}$ of Asn 843, and a water molecule. The metal:ligand bond distances range in length from 1.9 to 2.5 Å and the *B*-value for this manganese ion is 33.2 $Å^2$. The second metal is ligated by an α- and a γ-phosphoryl oxygen of the nucleotide, the bridging nitrogen of AMPPNP, $O^{\epsilon 1}$ of Gln 829, a water molecule, and $O^{\epsilon 2}$ of Glu 841 which acts as the bridging ligand between the two manganese ions. The *B*-value for this metal is 29.7 $Å^2$ and the metal:ligand bond distances range in length from 1.9 to 2.6 Å. There are three well-ordered water molecules located within the active site of the carbamoyl phosphate synthetic component as indicated in Figure 5.

The difference in metal content between the two active sites of the CPS large subunit described here is exactly opposite to that observed in the original model where the carboxyphosphate synthetic component contained two metal ions while the carbamoyl phosphate synthetic unit had only one (*10*). For all of the structural analyses of CPS accomplished to date, the crystals are always grown in the presence of 0.5 mM $MnCl_2$. The enzyme is catalytically competent with manganese (*21*), and the crystals simply grow more easily in the presence of this divalent metal ion as opposed to magnesium ions. Previous studies have, indeed, concluded that CPS requires a free metal ion, in addition to the divalent cations needed to complex the nucleotides, for full catalytic activity (*21*). It has been estimated that the

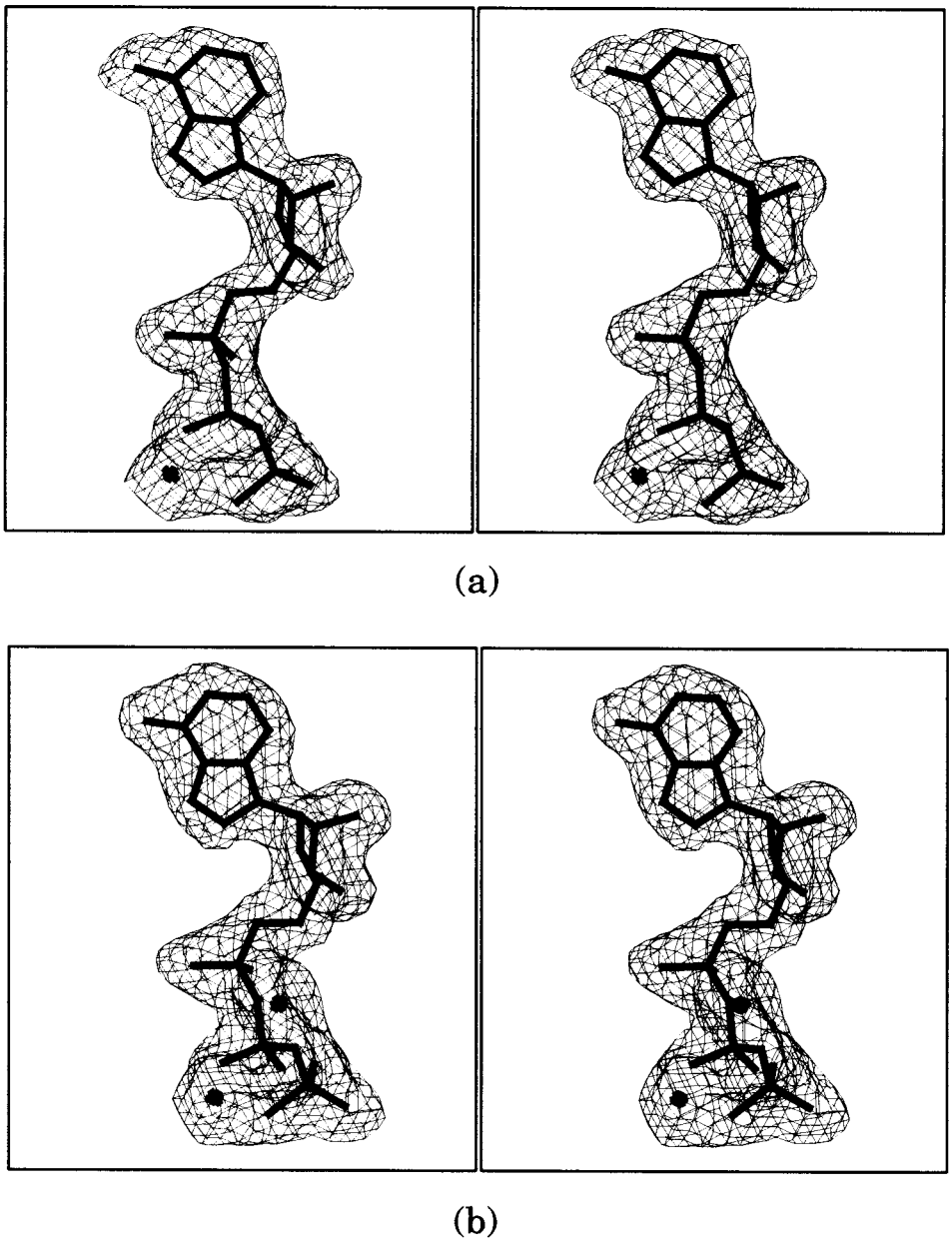


FIGURE 3: Electron densities for the two AMPPNP molecules bound to the CPS large subunit. The electron density maps shown were calculated with coefficients of the form $(F_o - F_c)$, where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude from the models lacking the coordinates for the metals and the nucleotides. The maps were contoured at 3σ. Shown in panels a and b are the observed electron densities for the nucleotides bound in the carboxyphosphate and carbamoyl phosphate synthetic components, respectively. The positions of the manganese ions are indicated by the asterisks.

dissociation constant for this manganese ion binding site is approximately 35 μM (*21*). The physiological significance of this difference in metal content between the original CPS model and the structure described here is presently not understood. What is absolutely clear, however, is that both active sites have the proper geometrical constraints to position two divalent metal ions given the opportunity and both Glu 299 and Glu 841 act as the bridging ligands when two metals are bound in the active sites. A structural analysis of CPS crystals grown in the presence of magnesium ions is presently underway.

A comparison of the active sites for the carboxyphosphate synthetase units of CPS with either bound ADP/P_i or AMPPNP is depicted in Figure 6a, and as is obvious, there are few structural changes between the two complexes. The inorganic phosphate observed in the original structure of the enzyme occupies nearly the same position as the γ-phosphate group of AMPPNP with its phosphorus atom displaced by approximately 1.2 Å relative to that observed for the γ-phosphorus of the nucleotide triphosphate. As such, the CPS−AMPPNP complex structure most likely mimics the prehydrolysis state.

As shown in the comparison depicted in Figure 6b, the structural changes accompanying AMPPNP binding to the active site of the carbamoyl phosphate synthetic component are substantial in the region defined by Leu 720 to Gly 722. Indeed, this region was completely disordered in the original electron density maps of CPS crystallized in the presence of ADP (*9*). Due to the hydrogen-bonding interactions formed between the backbone amide groups of Gly 721 and Gly 722 and the γ- and β-phosphoryl oxygens of the AMPPNP, the B-domain now clamps down thereby ordering

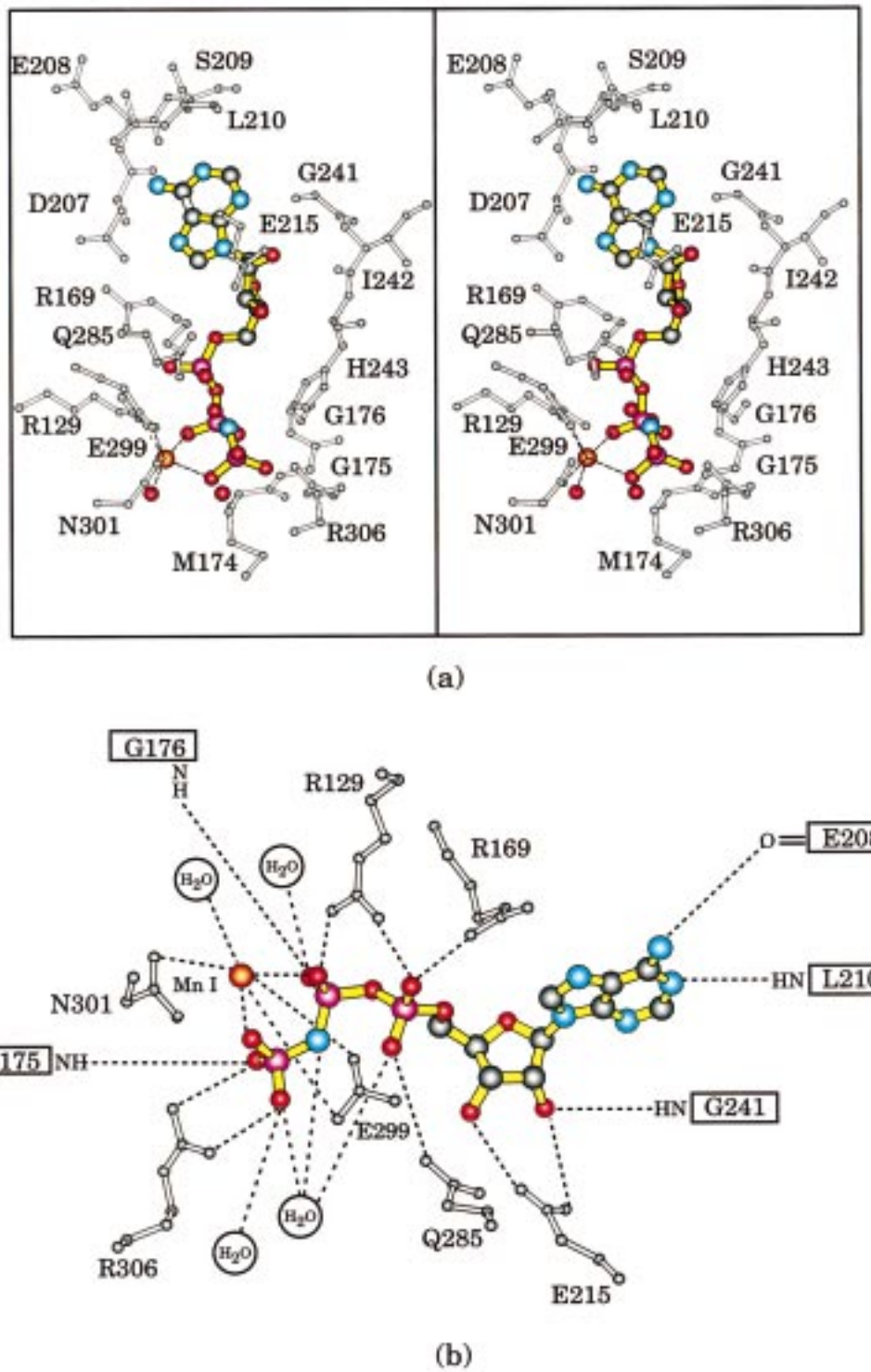


FIGURE 4: Closeup view of the active site for the carboxyphosphate synthetic component. (a) Those residues located within approximately 3.0 Å of the nucleotide are shown. Ordered water molecules are indicated by the red spheres. A cartoon representation of potential electrostatic interactions between AMPPNP and CPS is displayed in panel b. The dashed lines indicate distances equal to or less than 3.0 Å between atoms capable of participating in hydrogen-bonding interactions.

this region of polypeptide chain. Gly 721 and Gly 722 are situated in a Type III′ turn and are ideally suited to provide the necessary conformational flexibility that allows a "flap", delineated from Arg 715 to Met 725, to close over a nucleotide triphosphate species. Not surprisingly, a similar pair of glycine residues, namely Gly 175 and Gly 176, are located in a Type III′ turn in the carboxyphosphate synthetic component and they likewise interact with the γ- and β-phosphates of the AMPPNP, respectively. Indeed, previous studies have shown that replacement of these residues via site-directed mutagenesis results in the disruption of the phosphorylation events associated with the CPS reaction mechanism (*8*).

In the original structure of CPS, the B-domain was in a closed conformation in the carboxyphosphate synthetic component, and this was most likely due to the presence of an inorganic phosphate plus ADP in its active site. From the present studies, it is absolutely clear that the trigger for

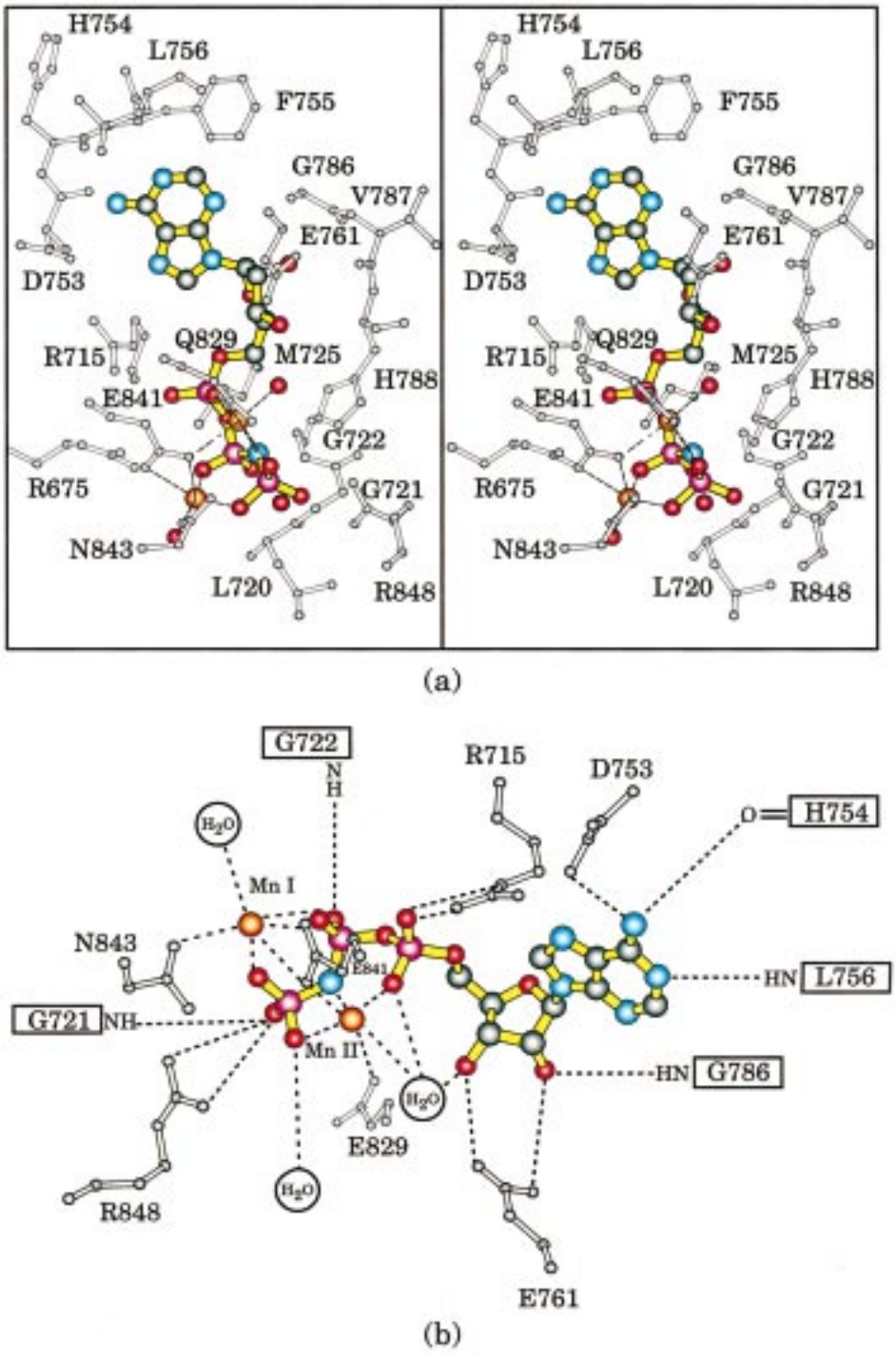
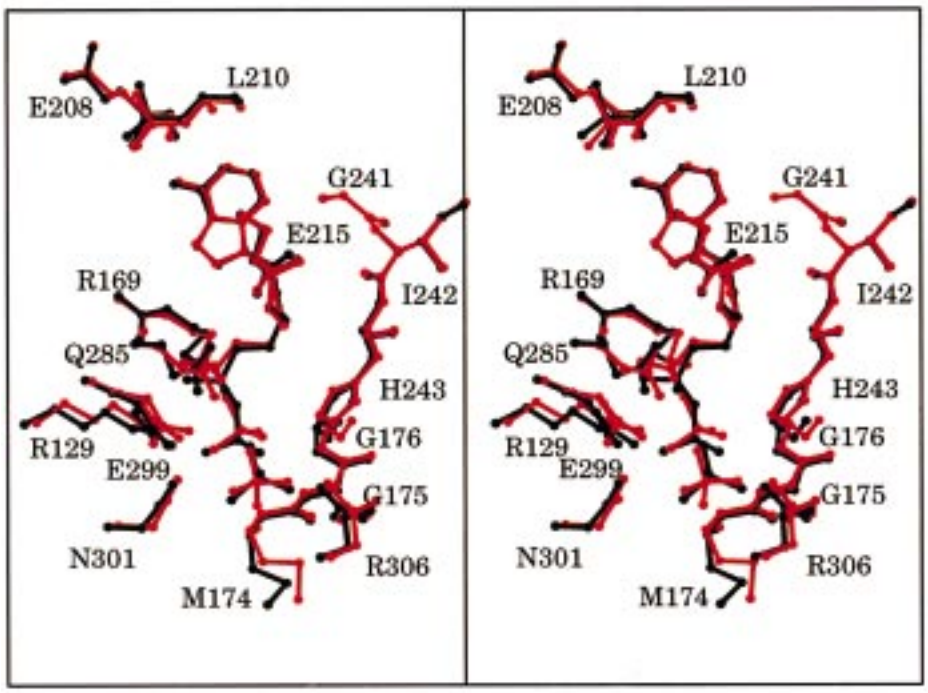


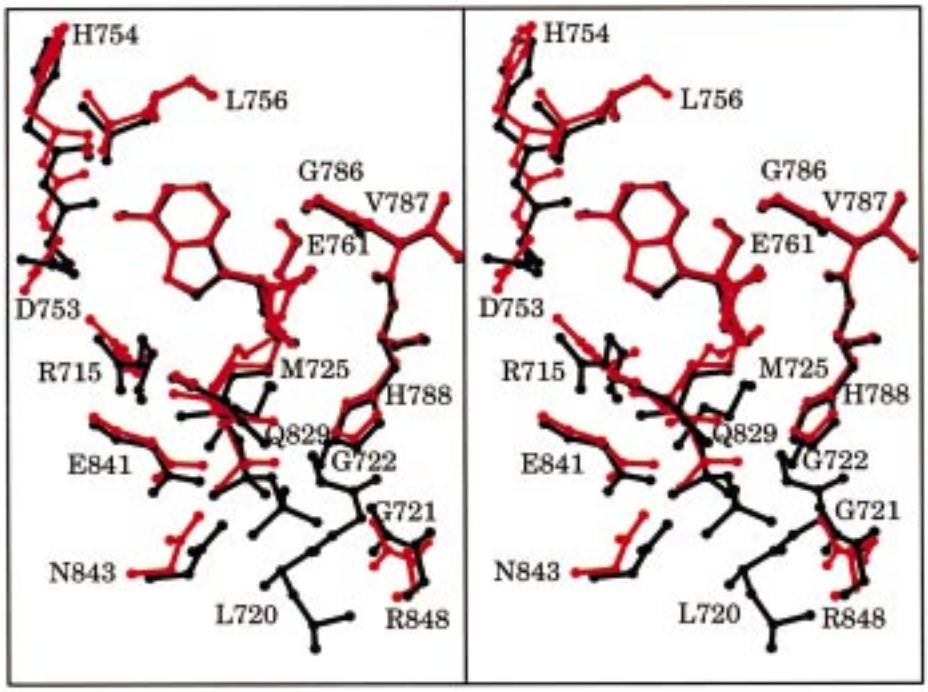
FIGURE 5: Closeup view of the active site for the carbamoyl phosphate synthetic component. (a) Those residues located within approximately 3.0 Å of the nucleotide are shown. Ordered water molecules are indicated by the red spheres. A cartoon representation of potential electrostatic interactions between AMPPNP and CPS is displayed in panel b. The dashed lines indicate distances equal to or less than 3.0 Å between atoms capable of participating in hydrogen-bonding interactions.

closure of the B-domain is occupation of the phosphate-binding site with either P_i or the γ-phosphate group of a nucleotide. The mode of AMPPNP binding exhibited by CPS is substantially different from that observed for enzymes containing the so-called P-loops as exemplified by myosin subfragment-1 (*22*). In these enzymes, the β- and γ-phosphate groups of the nucleotide are positioned by amino acid side chains residing in a loop connecting a strand of β-sheet to an α-helix. In both active sites of the CPS large subunit, however, the β- and γ-phosphate groups lie near the Type III′ turns described above and these secondary structural motifs serve to connect two strands of antiparallel β-pleated sheet.

One of the many intriguing aspects of the original structure of CPS was the presence of an inorganic phosphate group in the carboxyphosphate synthetic unit but not in the carbamoyl phosphate synthetic component (*9*). While the original crystallization conditions did not include inorganic phosphate, there was clearly such a molecule present in the active site of the carboxyphosphate synthetic component and most likely resulted from either a contaminant in the ADP samples employed in the crystallization medium or the



(a)



(b)

FIGURE 6: Differences in the active site geometries when ADP versus AMPPNP is bound. Shown in panel a is the carboxyphosphate synthetase active site with bound ADP/P_i (depicted in red) superimposed onto the carboxyphosphate synthetase active site with bound AMPPNP (displayed in black). Shown in panel b is the carbamoyl phosphate synthetase active site with bound ADP (depicted in red) superimposed onto the carbamoyl phosphate synthetase active site with bound AMPPNP (displayed in black). For the sake of clarity, the manganese ions were removed from the figure.

breakdown of ADP. Regardless of the origin of the inorganic phosphate, it was curious as to why it was only observed binding in the first half of the CPS large subunit and not in the second half. At the time, it was speculated that this was, indeed, the region of protein responsible for stabilizing the carboxyphosphate intermediate. Now that the structure of the CPS−AMPPNP complex is known and the loop defined by Ala 668 to Phe 755 is ordered, it is instructive to compare the protein environments responsible for binding the γ-phosphates of the AMPPNP molecules in the two halves of the CPS large subunit. From Figure 6, it is clear that the regions immediately surrounding the γ-phosphates of the AMPPNP molecules are similar in both synthetase portions. The only obvious difference is the replacement of Met 174 for Leu

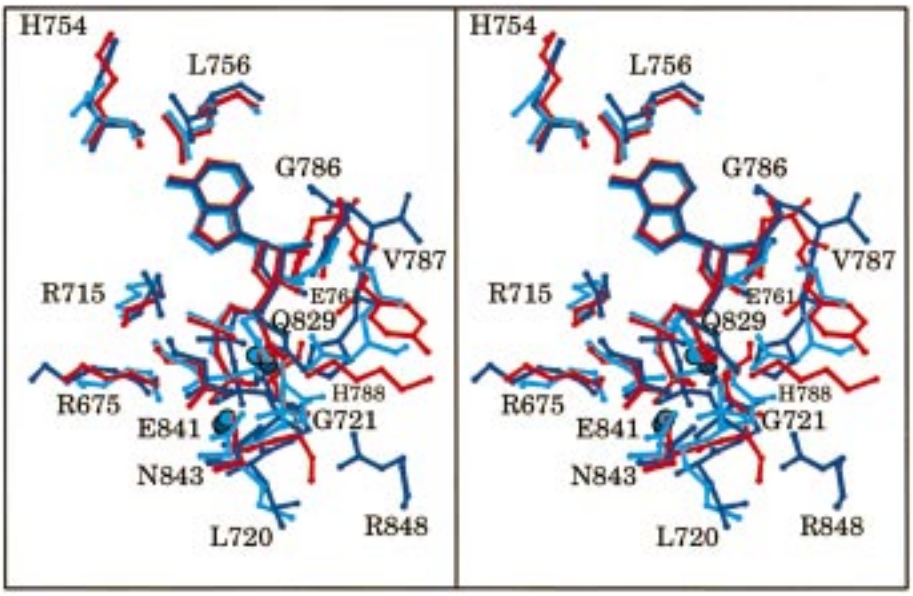


FIGURE 7: Comparison of the protein environments surrounding the nucleotides in the CPS carbamoyl phosphate synthetic component, D-alanine:D-alanine ligase (PDB file name 1IOV), and glutathione synthetase (PDB file name 1GSA). CPS, D-alanine:D-alanine ligase, and glutathione synthetase are displayed in blue, red, and cyan, respectively. Note that the glutathione synthetase model contains bound ADP and a sulfate ion. The labeled residues correspond to those found in CPS.

720 near the γ-phosphate-binding pocket. Additionally, searches of the regions within 5 Å of the γ-phosphate groups reveal virtually identical numbers of charged and/or polar side chains with the exception of Thr 173 of the carboxyphosphate synthetic component versus Val 713 in the carbamoyl synthetic unit. This small change in hydrophobicity is not enough to argue why an inorganic phosphate was observed in the carboxyphosphate synthetic component but not in the carbamoyl phosphate unit in the original CPS model. There are clearly other subtle factors regarding inorganic phosphate binding that cannot be distinguished at this time. It is possible that the active site of the carboxyphosphate synthetic component has a slightly higher affinity for inorganic phosphate, and had the original crystals been grown in the presence of millimolar quantities of inorganic phosphate, the active site of the carbamoyl phosphate synthetic component would have also bound such molecules. Crystallization experiments with millimolar quantities of phosphate ions are presently underway. In addition, carboxyphosphate mimics are being synthesized in order to address the manner in which CPS stabilizes the carboxyphosphate intermediate which has a half-life of approximately 70 ms (*23*).

DISCUSSION

The two synthetase components of the CPS large subunit belong to a superfamily of proteins containing a characteristic "ATP-grasp" fold (*24*, *25*). Other members of this family with known three-dimensional structures include, for example, biotin carboxylase, D-alanine:D-alanine ligase, glutathione synthetase, succinyl-CoA synthetase, and pyruvate phosphate dikinase (*12*, *26*−*29*). From initial comparisons, it is known that the A-, B-, and C-domains of the two synthetase components of CPS are topologically most similar to those observed in biotin carboxylase (*9*). The structure of biotin carboxylase was solved in the absence of nucleotides, so it is not possible to conduct a detailed comparison of the active sites for these two enzymes at this time. It is instructive, however, to compare the immediate surroundings of the nucleotides bound in CPS (carbamoyl phosphate synthetic unit), D-alanine:D-alanine ligase (*30*), and glutathione synthetase (*31*) as shown in Figure 7. For the most part, the chemical characteristics of the binding pockets in these enzymes are similar but certainly not identical. For example, Leu 756 in CPS, which is positioned near the purine ring of the nucleotide, is absolutely conserved in the ligase (Leu 183) and the synthetase (Leu 201). The positive charges of Arg 675 and Arg 715 in CPS that interact with the negatively charged phosphoryl oxygens of the nucleotide are retained as Lys 97 and Lys 144 in the ligase and Lys 125 and Lys 160 in the synthetase. In both CPS and the ligase, Glu 761 and Glu 187, respectively, serve to anchor the 2′- and 3′-hydroxyl groups of the nucleotide ribose to the protein. In glutathione synthetase, this residue is Asp 208 which only hydrogen bonds to the 3′-hydroxyl group of the ribose. As can be seen in Figure 7, all three enzymes have an asparagine residue at position 843 in CPS which serves as a ligand to one of the metals. Gln 829, which functions as a ligand to the second metal in CPS, is replaced by Asp 257 and Asp 273 in the ligase and the synthetase, respectively. Those residues where the chemical nature is not conserved between these enzymes include Leu 720, His 754, Val 787, and His 788 in CPS which correspond to Gly 149, Lys 181, Tyr 210, and Lys 215 in D-alanine:D-alanine ligase and Met 165, Asn 199, Gly 234, and Asn 235 in glutathione synthetase. The metals bind in nearly identical positions in these three enzymes with Glu 841 in CPS, Glu 270 in the ligase, and Glu 281 in the synthetase serving as the bridging ligands. Perhaps the most striking similarity between these three enzymes is the Type III′ reverse turn delineated by Leu 720 to Gly 722 in CPS. The second and third residues of this secondary structural element form interactions, via backbone amide groups, with phosphoryl oxygens of the nucleotides. Indeed, Gly 721 and Gly 722 of the CPS carbamoyl phosphate synthetic unit correspond to Ser 150 and Ser 151 in D-alanine:D-alanine ligase and Gly 166 and Glu 167 in

glutathione synthetase. This Type III′ reverse turn is most likely a characteristic feature of those enzymes belonging to the ATP-grasp protein family. Indeed, the backbone amide hydrogen-bonding interactions between the second and third residues of the Type III′ turn with the γ- and β-phosphate groups of the nucleotide triphosphate are critical for closure of the B-domains.

In the CPS α,β-heterodimer, the active site for the small subunit is positioned 45 Å from the first active site of the large subunit. Likewise, the two active sites contained within the CPS large subunit are separated by approximately 35 Å. In light of these distances, it has been postulated that the three active sites of CPS communicate via domain movements. The structure described here, indeed, represents the first example of such a domain movement in CPS caused by the binding of a nucleotide triphosphate. This movement, triggered by the hydrogen-bonding interactions between the β- and γ-phosphates of AMPPNP and the backbone amide groups of Gly 721 and Gly 722, is accomplished by numerous small changes in the dihedral angles of the polypeptide chain backbone. These changes are propagated throughout the B-domain with some atoms moving by more than 7.0 Å and the α-carbons for the two domains (with or without bound nucleotide triphosphate) superimposing with a root-mean-square deviation of approximately 4.1 Å. Nevertheless, the overall secondary structural features of the B-motifs observed in the CPS−AMPPNP and the CPS−ADP complexes are maintained. The model presented here is one "snapshot" of the many anticipated conformational changes that occur in CPS during the reaction cycle. Undoubtedly additional domain movements occur upon the binding or release of substrates and/or effector molecules and these may play pivotal roles in the proper functioning of this most remarkable enzyme.

ACKNOWLEDGMENT

The high resolution X-ray data set employed in this structural analysis was collected at the Stanford Synchrotron Radiation Laboratory which is funded by the Department of Energy, Office of Basic Energy Sciences. We thank Dr. Cary B. Bauer for assistance during the X-ray data collection, Dr. Grant E. Gibson for measuring the inhibitory properties of AMPPNP, and Dr. W. W. Cleland for critically reading the manuscript.

REFERENCES

1. Trotta, P. P., Burt, M. E., Haschemeyer, R. H., and Meister, A. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2599−2603.
2. Anderson, P. M., and Meister, A. (1966) *Biochemistry 5*, 3164−3169.
3. Raushel, F. M., Anderson, P. M., and Villafranca, J. J. (1978) *Biochemistry 17*, 5587−5591.
4. Raushel, F. M., and Villafranca, J. J. (1979) *Biochemistry 18*, 3424−3429.
5. Raushel, F. M., Mullins, L. S., and Gibson, G. E. (1998) *Biochemistry 37*, 10272−10278.
6. Matthews, S. L., and Anderson, P. M. (1972) *Biochemistry 11*, 1176−1183.
7. Nyunoya, H., and Lusty, C. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4629−4633.
8. Post, L. E., Post, D. J., and Raushel, F. M. (1990) *J. Biol. Chem. 265*, 7742−7747.
9. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) *Biochemistry 36*, 6305−6316.
10. Thoden, J. B., Raushel, F. M., Benning, M. M., Rayment, I., and Holden, H. M. (1998) *Acta Crystallogr., Sect. D 55*, 8−24.
11. Czerwinski, R. M., Mareya, S. M., and Raushel, F. M. (1995) *Biochemistry 34*, 13920−13927.
12. Fan, C., Moews, P. C., Walsh, C. T., and Knox, J. R. (1994) *Science 266*, 439−443.
13. Yount, R. G., Babcock, D., Ballantyne, W., and Ojala, D. (1971) *Biochemistry 10*, 2484−2489.
14. Miran, S. G., Chang, S. H., and Raushel, F. M. (1991) *Biochemistry 30*, 7901−7907.
15. Otwinowski, Z. (1993) *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"*, pp 56−62, complied by L. Sawyer, N. Isaacs, and S. Bailey, SERC Daresbury Laboratory, England.
16. Rossmann, M. G. (1972) *The Molecular Replacement Method*, Gordon and Breach, New York.
17. Navaza, J. (1994) *Acta Crystallogr., Sect. A 50*, 157−163.
18. Thoden, J. B., Miran, S. G., Phillips, J. C., Howard, A. J., Raushel, F. M., and Holden, H. M. (1998) *Biochemistry 37*, 8825−8831.
19. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr., Sect. A. 43*, 489−501.
20. Bricogne, G. (1976) *Acta Crystallogr., Sect. A 32*, 832−847.
21. Raushel, F. M. Rawding, C. J., Anderson, P. M., and Villafranca, J. J. (1979) *Biochemistry 18*, 5562−5566.
22. Gulick, A. M., Bauer, C. B., Thoden, J. B., and Rayment, I. (1997) *Biochemistry 36*, 11619−11628.
23. Sauers, C. K., Jencks, W. P., and Groh, S. (1975) *J. Am. Chem. Soc. 97*, 5546.
24. Murzin, A. G. (1996) *Curr. Opin. Struct. Biol. 6*, 386-394.
25. Galperin, M. Y., and Koonin, E. V. (1997) *Protein Sci. 6*, 2639−2643.
26. Waldrop, G. L., Rayment, I., and Holden, H. M. (1994) *Biochemistry 33*, 10249−10256.
27. Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J., and Katsube Y. (1993) *J. Mol. Biol. 229*, 1083−1100.
28. Wolodko, W. T., Fraser, M. E., James, M. N. G., and Bridger, W. A. (1994) *J. Biol. Chem. 269*, 10883−10890.
29. Herzberg, O., Chen C. C., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., and Dunaway-Mariano D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2652−2657.
30. Fan, C., Park, I.-S., Walsh, C. T., and Knox, J. R. (1997) *Biochemistry 36*, 2531−2538.
31. Hara, T., Kato, H., Katsube, Y., and Oda, J. (1996) *Biochemistry 35*, 11967−11974.

BI982517H