

The Binding of Inosine Monophosphate to *Escherichia coli* Carbamoyl Phosphate Synthetase*

(Received for publication, March 19, 1999)

James B. Thoden‡, Frank M. Raushel§, Gary Wesenberg‡, and Hazel M. Holden‡¶

From the ‡Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Madison, Wisconsin 53705 and the §Department of Chemistry, Texas A&M University, College Station, Texas 77843

Carbamoyl phosphate synthetase (CPS) from *Escherichia coli* catalyzes the formation of carbamoyl phosphate, which is subsequently employed in both the pyrimidine and arginine biosynthetic pathways. The reaction mechanism is known to proceed through at least three highly reactive intermediates: ammonia, carboxyphosphate, and carbamate. In keeping with the fact that the product of CPS is utilized in two competing metabolic pathways, the enzyme is highly regulated by a variety of effector molecules including potassium and ornithine, which function as activators, and UMP, which acts as an inhibitor. IMP is also known to bind to CPS but the actual effect of this ligand on the activity of the enzyme is dependent upon both temperature and assay conditions. Here we describe the three-dimensional architecture of CPS with bound IMP determined and refined to 2.1 Å resolution. The nucleotide is situated at the C-terminal portion of a five-stranded parallel β -sheet in the allosteric domain formed by Ser⁹³⁷ to Lys¹⁰⁷³. Those amino acid side chains responsible for anchoring the nucleotide to the polypeptide chain include Lys⁹⁵⁴, Thr⁹⁷⁴, Thr⁹⁷⁷, Lys⁹⁹³, Asn¹⁰¹⁵, and Thr¹⁰¹⁷. A series of hydrogen bonds connect the IMP-binding pocket to the active site of the large subunit known to function in the phosphorylation of the unstable intermediate, carbamate. This structural analysis reveals, for the first time, the detailed manner in which CPS accommodates nucleotide monophosphate effector molecules within the allosteric domain.

unstable precursor that initiates subsequent steps in the biosynthesis of both pyrimidine nucleotides and arginine. The end-product of the pyrimidine biosynthetic pathway, UMP, allosterically *inhibits* this enzyme. Conversely, ornithine, the co-substrate with carbamoyl phosphate in the next step of the arginine biosynthetic pathway, *enhances* the catalytic activity of CPS. Each of these effectors primarily, but not exclusively, modulates the activity of CPS by raising or lowering the dissociation constant for Mg²⁺ATP by approximately one order of magnitude (7, 8).

As isolated from *E. coli*, CPS is a large multidomain heterodimeric protein. The smaller of the two subunits has a molecular weight of ~42,000 and is catalytically responsible for the hydrolysis of glutamine and the subsequent translocation of the ammonia product to the large subunit (9). The larger subunit (molecular weight ~118,000) is composed of four major regions as indicated in Fig. 1 (1, 10). Two of these structural units, known as the carboxyphosphate (Met¹-Glu⁴⁰³) and carbamoyl phosphate (Asn⁵⁵⁴-Asn⁹³⁶) synthetic components are homologous to one another and contain all of the catalytic machinery necessary for the final assembly of carbamoyl phosphate. These two synthetase units are linked together by a third domain (Val¹⁴⁰⁴-Ala⁵⁵³) whose functional role in the structure and catalytic properties of CPS is still not well understood. However, this region of the protein appears to play a role in the oligomerization of the α,β -heterodimer into an $(\alpha,\beta)_4$ -tetrameric species (1). The remaining domain (Ser⁹³⁷-Lys¹⁰⁷³) has been shown by both x-ray crystallographic and biochemical methods to harbor the binding sites for such allosteric effectors as ornithine, UMP, and IMP.

The binding site for the positive allosteric effector, ornithine, was identified in the original x-ray structural analysis of CPS (1, 2). Ornithine was found to straddle the interface between the allosteric domain and the carbamoyl phosphate synthetic component. Specifically, the α -amino group of this compound was shown to lie within hydrogen bonding distance to O^γ of Tyr¹⁰⁴⁰ in the allosteric domain, whereas the δ -amino group was positioned within 3.0 Å from the carboxylate groups of Glu⁷⁸³ and Glu⁸⁹², and O of Asp⁷⁹¹, all of which originate from the carbamoyl phosphate synthetic component. In addition, the carboxylate group of ornithine was observed interacting with both the backbone amide nitrogen and O^γ of Thr¹⁰⁴². This same structure of CPS also provided the first hint for the locations of the nucleotide monophosphate binding sites. Indeed, within the allosteric domain, an inorganic phosphate was observed, hydrogen bonded, and ion-paired to Lys⁹⁵⁴, Thr⁹⁷⁴, Thr⁹⁷⁷, and Lys⁹⁹³. The hydroxyl group of Thr⁹⁷⁷ had previously been shown to be critical for the display of the allosteric properties of UMP (11), although somewhat later Lys⁹⁹³ was shown to be the residue modified upon photolabeling CPS with UMP (12). From the above mentioned data, a structural model was subsequently proposed by Thoden *et al.* (1) for the binding of UMP or

Carbamoyl phosphate synthetase (CPS)¹ is one of the few allosterically regulated enzymes whose three-dimensional structure is now known to high resolution (1, 2). A list of other regulated proteins, whose molecular architectures are known in quite extensive detail, includes phosphofructokinase (3), aspartate transcarbamoylase (4), glutamine synthetase (5), and glycogen phosphorylase (6), among others. The biochemical role of CPS is to catalyze the formation of carbamoyl phosphate, an

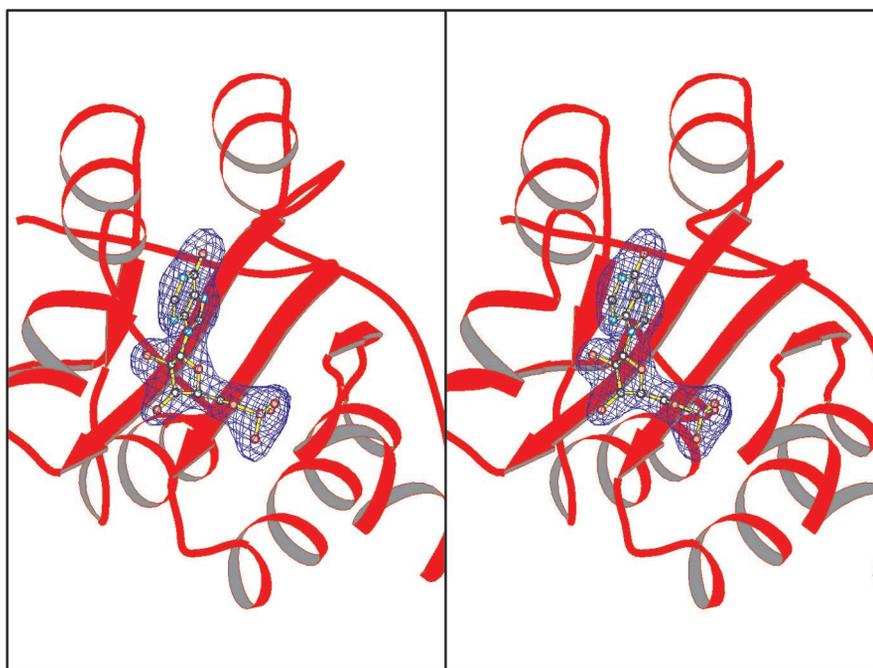
* This research was supported in part by Grants from the National Institutes of Health GM55513 (to H. M. H.) and DK30343 (to F. M. R.) and the National Science Foundation (BIR-9317398 shared instrumentation grant). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1CE8) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¶ To whom correspondence should be addressed: Dept. of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, 1710 University Ave., Madison, WI 53705. Tel.: 608-262-4988; Fax: 608-262-1319; E-mail: HOLDEN@ENZYM.WISC.EDU.

¹ The abbreviation used is: CPS, carbamoyl phosphate synthetase.

FIG. 2. Electron density corresponding to the IMP moiety. The electron density shown was calculated with coefficients of the form $(F_o - F_c)$ and contoured at 3σ . The portion of the polypeptide chain responsible for forming the IMP binding pocket in the allosteric domain is indicated by the ribbon representation.



collected. These frames were processed with DENZO and scaled with SCALEPACK (19). From 3,249,722 measurements, 1,855,323 reflections were integrated and reduced to 445,013 unique reflections after scaling. Scaling statistics are presented in Table I.

The structure of the CPS-IMP complex described here was solved by the technique of molecular replacement (20) with the software package AMORE (21) and using, as a search model, the complete tetrameric form of CPS previously refined to 2.1 Å resolution (2). Following rigid body refinement, the model was subjected to least squares refinement at 2.1 Å resolution with the software package TNT (22). Because there were over 5,800 amino acid residues in the asymmetric unit, the model building and refinement processes were expedited by averaging the electron densities corresponding to the four α,β -heterodimers in the asymmetric unit as described previously (23). Alternate cycles of re-building using this averaged electron density map followed by least squares refinement of the "averaged" model expanded back into the unit cell reduced the R_{cryst} to 20.8% and the R_{free} to 26.5%. For calculation of R_{free} , 5% of the x-ray data were removed from the reflection file. At this point the entire tetrameric model was adjusted in the unit cell and an additional cycle of least squares refinement conducted leading to an R_{cryst} of 19.2% and an R_{free} of 25.7%. A final cycle of refinement was conducted with all measured x-ray data from 30 to 2.1 Å resulting in an R -factor of 19.3%. Relevant refinement statistics are presented in Table II. The CPS-IMP model described here includes eight ADP molecules, 12 manganese ions, 8 ornithines, 4 inorganic phosphates, 28 potassium ions, 12 chloride ions, 4,037 water molecules, 4 tetraethylammonium ions, and 4 IMP molecules. The following side chains were modeled as multiple conformations: Ser²⁹, Arg¹³⁰, Glu²⁷⁸, Asp⁴¹⁶, Glu⁶⁵⁵, Glu⁹¹⁰, Gln⁹⁶⁷, and Glu⁹⁸³ in the large subunit and Arg²¹⁵ in the small subunit of α,β -heterodimer I; Arg¹⁴⁵ in the large subunit and Met²⁸⁶ in the small subunit of α,β -heterodimer II; Arg⁸², Arg²²², Arg⁴⁹⁰, Arg⁵⁵⁹, Glu⁵⁹¹, Lys⁶³⁴, Gln⁶⁴⁵, Arg⁶⁵², Glu⁷⁷¹, Lys⁹⁴¹, Glu¹⁰⁰⁹, and Arg¹⁰²¹ in the large subunit and Gln⁷⁸ in the small subunit of α,β -heterodimer III and Lys³⁵ and Glu⁴⁷⁸ in the large subunit of α,β -heterodimer IV. In addition, Cys²⁶⁹ was modeled as sulfenic acid for all four small subunits contained within the asymmetric unit.

RESULTS AND DISCUSSION

The CPS crystals employed in this investigation contained a complete $(\alpha,\beta)_4$ -heterotetramer in the asymmetric unit. Because the four (α,β) -heterodimers are strikingly similar, however, the following discussion will refer only to (α,β) -heterodimer II as deposited in the Brookhaven Protein Data Bank. Shown in Fig. 2 is the electron density corresponding to the IMP ligand. The nucleotide fits the electron density best with the ribose in the C_2 -endo conformation. The purine ring adopts an anti-orientation and is tilted by approximately 156° from

the plane of the ribose ring. This is in sharp contrast to those conformations observed for the ADP moieties bound to CPS whereby the bases are tilted by approximately 128° and the ribose rings adopt C_3 -endo puckers. All the nucleotide monophosphates contained within the crystallographic asymmetric unit are very well ordered with the IMP bound to (α,β) -heterodimer II having an average temperature factor of 30.5 \AA^2 .

The allosteric domain of CPS (Ser⁹³⁷ to Lys¹⁰⁷³) is characterized by a five-stranded parallel β -sheet flanked on either side by two and three α -helices, respectively. As can be seen in Fig. 2, the IMP ligand is situated at the C-terminal end of this sheet and, indeed, is completely contained within the allosteric domain as originally predicted (1). A close-up view of the IMP binding pocket is displayed in Fig. 3a and a cartoon of potential hydrogen bonds between the ligand and the protein is shown in Fig. 3b. As indicated by the gray spheres in Fig. 3a, there are three water molecules located within a 5 \AA sphere of the purine nucleotide. One side of the hypoxanthine ring is packed against a fairly hydrophobic pocket formed by Ile¹⁰⁰¹, Val¹⁰²⁸, and Ile¹⁰²⁹. The only direct electrostatic contact between the purine base and the protein occurs between O-6 of the hypoxanthine ring and N of Val⁹⁹⁴. A water molecule is located within 3.0 \AA to N-1 of the purine base. Both N⁸² of Asn¹⁰¹⁵ and O γ of Thr¹⁰¹⁷ serve to anchor the 2'-OH group of the ribose to the protein. The 3'-OH group of the ribose lies within hydrogen bonding distance to O of Thr¹⁰¹⁶. All three phosphoryl oxygens form direct hydrogen bonds with the protein via the side chain functional groups of Lys⁹⁵⁴, Thr⁹⁷⁴, Thr⁹⁷⁷, and Lys⁹⁹³ as indicated by the dashed lines in Fig. 3b. In addition, the backbone amide groups of Gly⁹⁷⁶ and Thr⁹⁷⁷ also lie within hydrogen bonding distance to two of the phosphoryl oxygens. Approximately 90% of the surface area for the nucleotide ligand is buried upon binding to CPS as calculated with the program GRASP and employing a search radius of 1.4 \AA (24).

The original structure of CPS was solved in the presence of both Mn^{2+} -ADP and ornithine to 2.8 Å resolution (1). In that model, two ornithine binding sites per α,β -heterodimer were identified. The first ornithine was shown to bridge the C-domain of the carbamoyl phosphate synthetic component to the allosteric domain. The second ornithine, along with an inorganic phosphate, was positioned at the C-terminal end of the

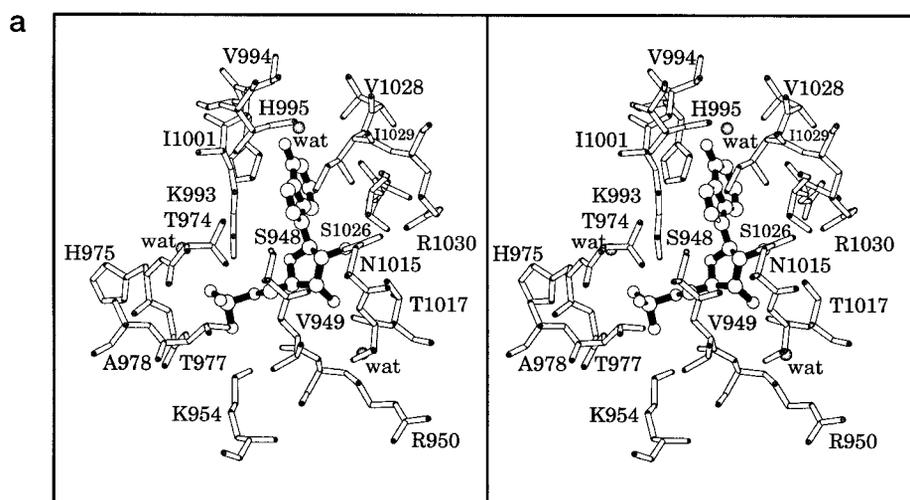


FIG. 3. The IMP binding pocket. Panel a, a close-up view of those amino acid residues located within approximately 5 Å of atoms of the IMP. Panel b, a cartoon of potential hydrogen bonds between the ligand and the protein, as indicated by the dashed lines.

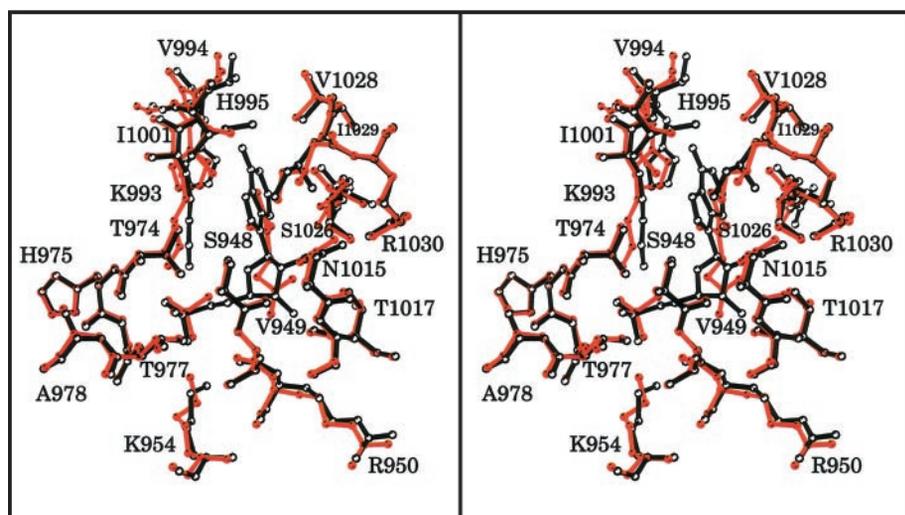
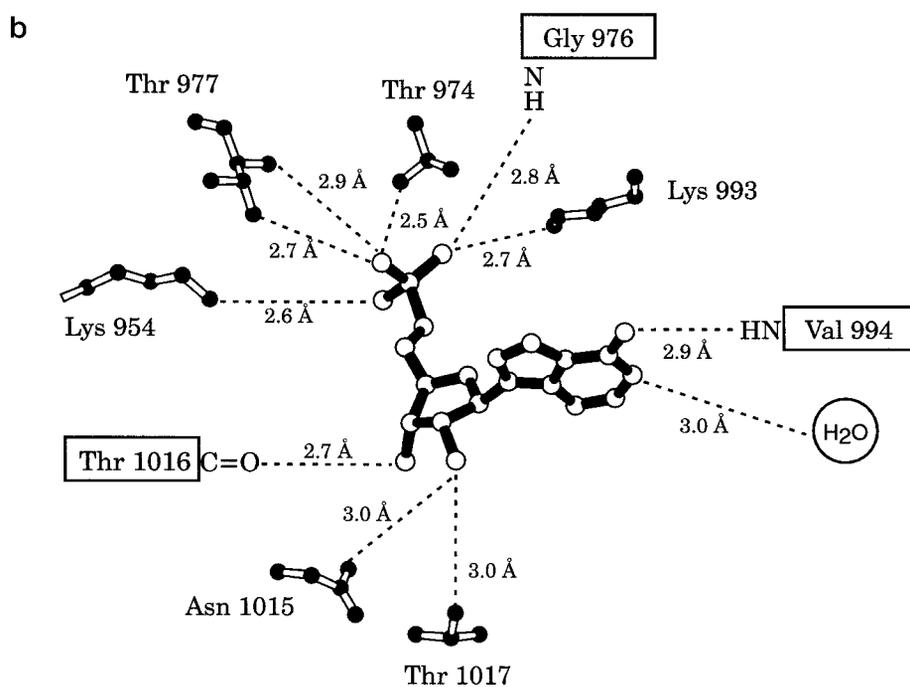
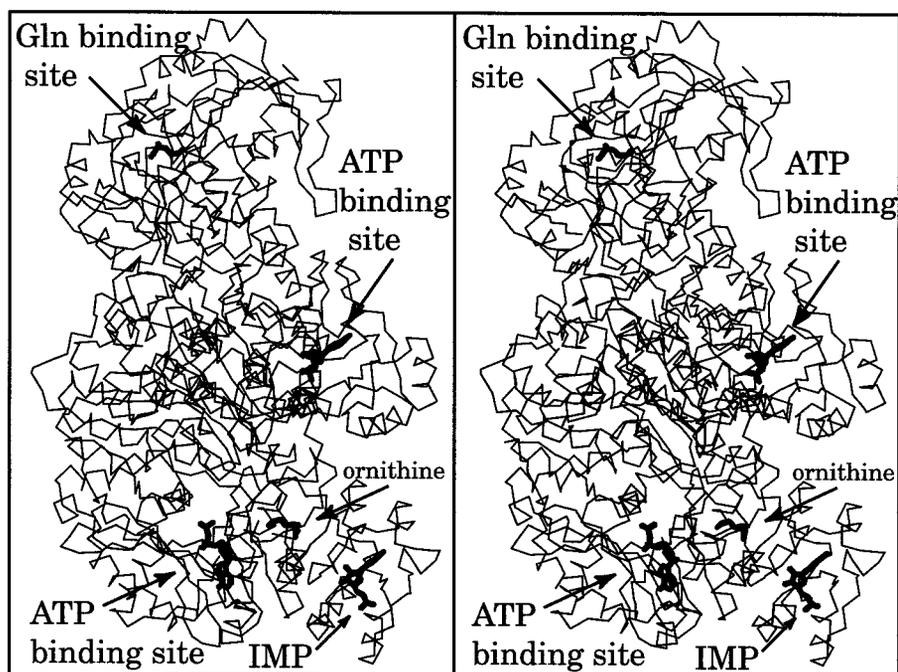


FIG. 4. Superposition of the CPS allosteric binding pockets with bound IMP or glutamine/inorganic phosphate. The model described in this report is shown in black, whereas the structure of CPS complexed with glutamine and inorganic phosphate, as described by Thoden *et al.* (2), is displayed in red. Note the nearly exact correspondence between the inorganic phosphate and the phosphoryl moiety of the nucleotide.

β -sheet in the allosteric domain. At that time, it was postulated that the combination of ornithine and inorganic phosphate was most likely mimicking the mode of binding for such nucleotide

ligands as UMP and IMP. For the subsequent x-ray data collection and analysis of the CPS model to 2.1 Å resolution, the crystals were first soaked in a solution containing 25 mM glu-

FIG. 5. α -Carbon trace of the CPS α,β -heterodimer. The relative disposition of the three active sites and the allosteric effector binding pockets for CPS are indicated by the ball-and-stick representations. The active site for the small subunit is approximately 45 Å from the first Mn^{2+} ADP binding site of the large subunit which, in turn, is 35 Å from the second Mn^{2+} ADP binding motif. This second active site of the large subunit is positioned approximately 14 Å from the ornithine and 19 Å from the IMP ligands.



tamine in an attempt to label the active site of the small subunit. Although the glutamine did not bind in the active site in this study, it did, however, displace the ornithine from its binding pocket in the allosteric domain (2).

A superposition of this ligand binding site, as observed in the refined structure of CPS, onto the present CPS:IMP model is depicted in Fig. 4. The phosphoryl group of the nucleotide occupies a nearly identical position to that observed for the inorganic phosphate. This phosphate binding region is positioned at the N-terminal region of an α -helix formed by Thr.⁹⁷⁴ to Gly.⁹⁸². It can thus be speculated that both hydrogen bonds and electrostatic interactions with the positive end of the helix dipole serve to promote binding of phosphate moieties to this region of the allosteric domain. The only two side chains that adopt significantly different conformations between the two models are Ile.¹⁰⁰¹ and Ile.¹⁰²⁹ where their dihedral angles differ by approximately 120–130°. Other than these side chains, the two CPS complexes are remarkably similar, such that 5,744 of their backbone atoms coincide with a root mean square deviation of 0.31 Å. Recently, Bueso *et al.* (25) demonstrated that IMP bound to CPS can be photocrosslinked to the large subunit in very low yield. The only amino acid residue labeled after photoirradiation was shown to be His.⁹⁹⁵. Site-directed mutagenesis of His.⁹⁹⁵ to an alanine residue had relatively little effect on either the catalytic or regulatory properties of CPS. In the crystal structure of the CPS:IMP complex reported here, the imidazole ring of His.⁹⁹⁵ is approximately 4 Å from the hypoxanthine ring of the nucleotide monophosphate. The biochemical results of Bueso *et al.* (25) are thus entirely consistent with the fact that His.⁹⁹⁵ does not significantly contribute to the binding site for IMP.

In CPS, the IMP ligand binds to the C-terminal portion of a parallel β -sheet. A search of the Brookhaven Protein Data Bank reveals that other enzymes have been solved in the presence of IMP, or analogs thereof, including adenylosuccinate synthetase (26). Detailed x-ray crystallographic analyses of this enzyme have demonstrated that the IMP ribose adopts the C_2 -endo pucker with the purine base oriented similarly to that observed in CPS (26). Additionally, the phosphoryl group of the nucleotide is positioned within approximately 5 Å of the N-terminal end of an α -helix formed by Gly.¹³² to Ala.¹⁴². The IMP

moiety is deeply buried within the protein such that only about 1% of its surface area is exposed to the solvent. Unlike that observed in CPS, the purine ring makes numerous hydrogen bonding contacts with the protein. Note that adenylosuccinate synthetase catalyzes the production of adenylosuccinate from IMP, aspartate, and GTP, and in this case, the nucleotide monophosphate is a true substrate. With regard to the role of IMP acting solely as an effector molecule, however, the only other observed structural example of IMP binding within an allosteric pocket is that of glycogen phosphorylase (27). In this enzyme, the ribose of the nucleotide adopts the C_3 -endo pucker. The ligand is positioned in a rather open region of the molecule with approximately 48% of its surface area exposed to the solvent, the hypoxanthine ring facing outwards and no direct protein contacts to the base. Other than the similarities in ribose puckers between CPS and adenylosuccinate synthetase, the tertiary motifs employed for IMP binding in these three enzymes are quite different.

The overall spatial relationships between the active sites of the small and large subunits of CPS and the IMP and ornithine effector binding pockets are depicted in Fig. 5. The ornithine and IMP binding sites are separated by approximately 12 Å, whereas the ornithine binding region and active site in the carbamoyl phosphate synthetic component are situated approximately 14 Å apart. Previously, it was demonstrated that allosteric ligands such as IMP and ornithine affect the carbamoyl phosphate-dependent ATP synthesis reaction more so than the other two partial reactions catalyzed by CPS (7). The manner in which the IMP allosteric effector site communicates to this Mg^{2+} ATP binding site is most likely via a complicated series of hydrogen bonds. As an example, there is a hydrogen bonding pathway leading directly from the 2'-hydroxyl group of the IMP ribose to one of the phosphoryl oxygens of the Mn^{2+} ADP. This hydrogen bonding network initiates via the interaction between N.⁶² of Asn.¹⁰¹⁵ and the 2'-hydroxyl group of IMP (Fig. 3b). The backbone amide group of Asn.¹⁰¹⁵ interacts with O of Asp.¹⁰⁴¹. In this region, O γ of Thr.¹⁰⁴² forms a hydrogen bond with the carboxylate group of ornithine. In turn, the δ -amino group of ornithine lies within hydrogen bonding distance to O of Asp.⁷⁹¹, whereas the backbone amide group of Asp.⁷⁹¹ interacts with O of His.⁷⁸⁸. Finally, the imidazole side

chain of His⁷⁸⁸ hydrogen bonds to a phosphoryl oxygen of the Mn²⁺ADP moiety. Examination of the CPS-IMP complex model described in this report demonstrates that additional hydrogen bonding networks may also exist between the IMP binding pocket and the active site of the carbamoyl phosphate synthetic component.

From previous biochemical studies (13, 14, 16), it is known that IMP and UMP compete for the same binding site on CPS. It can thus be speculated that the structure of CPS described here reveals not only the manner in which IMP is accommodated on the enzyme but also the mode of UMP binding to the allosteric domain, at least as far as the phosphoryl and sugar moieties are concerned. Recent detailed kinetic analyses of the behavior of CPS in the presence of IMP, UMP, or ornithine or combinations thereof have revealed a quite complicated set of catalytic behaviors (17). In addition, these studies have demonstrated that though both ornithine, an activator, and UMP, an inhibitor, can bind simultaneously to CPS, the activating effect of ornithine on the catalytic activity of CPS clearly prevails. What is not apparent from the present structural studies is the manner in which the enzyme discriminates between UMP and IMP. It would appear likely that the same molecular contacts are made to the ribose-phosphate moieties of UMP and IMP in the bound complexes. However, in the structure of IMP complexed to CPS, relatively few contacts are made to the hypoxanthine ring system. We anticipate that a significant molecular rearrangement occurs upon the binding of UMP to CPS that is not observed in the present model. These structural manifestations can only be elucidated via a combination of biochemical and x-ray crystallographic experiments and, indeed, these investigations are in progress.

Acknowledgment—We thank Dr. W. W. Cleland for helpful discussions. 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.

REFERENCES

1. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) *Biochemistry* **36**, 6305–6316
2. Thoden, J. B., Raushel, F. M., Benning, M. M., Rayment, I., and Holden, H. M. (1999) *Acta Crystallogr. Sec. D* **55**, 8–24
3. Schirmer, T., and Evans, P. R. (1990) *Nature* **343**, 140–145, and references therein
4. Kosman, R. P., Gouaux, J. E., and Lipscomb, W. N. (1993) *Proteins Struct. Funct. Genet.* **15**, 147–176, and references therein
5. Liaw, S. H., Jun, G., and Eisenberg, D. (1994) *Biochemistry* **33**, 11184–11188, and references therein
6. Gregoriou, M., Noble, M. E., Watson, K. A., Garman, E. F., Krulle, T. M., de la Fuente, C., Fleet, G. W., Oikonomakos, N. G., and Johnson, L. N. (1998) *Protein Sci.* **7**, 915–927, and references therein
7. Braxton, B. L., Mullins, L. S., Raushel, F. M., and Reinhart, G. D. (1992) *Biochemistry* **31**, 2309–2316
8. Braxton, B. L., Mullins, L. S., Raushel, F. M., and Reinhart, G. D. (1996) *Biochemistry* **35**, 11918–11924
9. Matthews, S. L., and Anderson, P. M. (1972) *Biochemistry* **11**, 1176–1183
10. Trotta, P. P., Burt, M. E., Haschemeyer, R. H., and Meister, A. (1971) *Proc. Natl. Acad. Sci. U. S. A.* **68**, 2599–2603
11. Czerwinski, R. M., Mareya, S. M., and Raushel, F. M. (1995) *Biochemistry* **34**, 13920–13927
12. Cervera, J., Bendala, E., Britton, H. G., Bueso, J., Nassif, Z., Lusty, C. J., and Rubio, V. (1996) *Biochemistry* **35**, 7247–7255
13. Boettcher, B., and Meister, A. (1981) *J. Biol. Chem.* **256**, 5977–5980
14. Boettcher, B., and Meister, A. (1982) *J. Biol. Chem.* **257**, 13971–13976
15. Kasprzak, A. A., and Villafranca, J. J. (1988) *Biochemistry* **27**, 8050–8056
16. Anderson, P. M. (1977) *Biochemistry* **16**, 587–593
17. Braxton, B. L., Mullins, L. S., Raushel, F. M., and Reinhart, G. D. (1999) *Biochemistry* **38**, 1394–1401
18. Miran, S. G., Chang, S. H., and Raushel, F. M. (1991) *Biochemistry* **30**, 7901–7907
19. Otwinowski, Z. (1993) in *Proceedings of the CCP4 Study Weekend: Data Collection and Processing, Oscillation Data Reduction Program* (Sawyer, L., Isaacs, N., and Bailey, S., eds) pp. 56–62, SERC Daresbury Laboratory, United Kingdom
20. Rossmann, M. G. (1972) *The Molecular Replacement Method*, Gordon and Breach Science Publishers, Inc., New York
21. Navaza, J. (1994) *Acta Crystallogr. Sec. A* **50**, 157–163
22. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sec. A* **43**, 489–501
23. Thoden, J. B., Miran, S. G., Phillips, J. C., Howard, A. J., Raushel, F. M., and Holden, H. M. (1998) *Biochemistry* **37**, 8825–8831
24. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11**, 281–296
25. Bueso, J., Cervera, J., Fresquet, V., Marina, A., Lusty, C. J., and Rubio, V. (1999) *Biochemistry* **38**, 3910–3917
26. Poland, B. W., Fromm, H. J., and Honzatko, R. B. (1996) *J. Mol. Biol.* **264**, 1013–1027
27. Oikonomakos, N. G., Zographos, S. E., Tsitsanou, K. E., Johnson, L. N., and Acharya K. R. (1996) *Protein Sci.* **5**, 2416–2428