Proc. Natl. Acad. Sci. USA
Vol. 75, No. 11, pp. 5276–5280, November 1978
Biochemistry

Three-dimensional structures of aspartate carbamoyltransferase from *Escherichia coli* and of its complex with cytidine triphosphate

(allosteric behavior/enzyme complex/protein crystallography/x-ray diffraction/conformational changes)

HUGO L. MONACO, JAMES L. CRAWFORD, AND WILLIAM N. LIPSCOMB

Gibbs Chemical Laboratory, Harvard University, Cambridge, Massachusetts 02138

Contributed by William N. Lipscomb, July 10, 1978

ABSTRACT X-ray diffraction studies to nominal resolutions of 3.0 Å for unliganded aspartate carbamoyltransferase (EC 2.1.3.2) (*R*32 crystal symmetry) and of 2.8 Å for the complex of aspartate carbamoyltransferase with cytidine triphosphate (*P*321 crystal symmetry) have yielded traces of the polypeptide chains of the catalytic (C) and regulatory (R) chains in the hexameric C_6R_6 molecules. The independent molecular structures of the liganded and unliganded forms of the enzyme are very nearly identical. In the regulatory chain there is a CTP-binding domain that interacts with an adjacent regulatory subunit and a zinc-binding domain that interacts with the catalytic subunit. In the catalytic chain a polar domain shows interactions between adjacent pairs of C chains to form each trimer C_3 while an equatorial domain shows intramolecular C_3–C_3 interactions. The active site is at or near the interface between adjacent C chains within the trimers. Probably each active center involves amino acid residues from adjacent C chains.

Aspartate carbomyltransferase (EC 2.1.3.2) catalyzes the formation of carbamoyl-L-aspartate from L-aspartate and carbamoylphosphate (1, 2). In *Escherichia coli* this reaction is the first unique step in the biosynthesis of pyrimidines. The end-product of this pathway, cytidine triphosphate (CTP) inhibits this allosteric enzyme by interacting with a regulatory chain R, which is distinct from the catalytic chain C (3, 4). Also ATP, a purine, stimulates this enzyme so that a tendency toward balance in the pyrimidine and purine syntheses is achieved. The enzyme is hexameric (5, 6), and can be dissociated by use of mercurials into two catalytic trimers C_3 and three regulatory dimers R_2 (7). Molecular weights are 34,000 for C and 17,000 for R (5, 8, 9). Each regulatory unit contains one Zn^{2+}, which has been presumed to coordinate to the four sulfhydryl groups of this chain (10, 11), and which is necessary for reconstitution of the enzyme from the subunits C_3 and R_2. Other properties of this enzyme have been summarized recently (12).

Two crystallographic forms of molecular aspartate carbamoyltransferase have been studied at low resolution in this laboratory. Crystals of *R*32 symmetry of the unliganded form (containing neither allosteric effectors nor substrate analogues) show at 5.5-Å resolution that the molecular symmetry is D_3, that the two catalytic trimers are almost eclipsed when viewed along the molecular threefold axis, and that there is a large aqueous cavity in the center of the molecule (13). Crystals of symmetry *P*321 of the form grown in the presence of CTP showed that there are no gross structural differences when compared to the unliganded R32 form (14). The additional electron density associated with CTP was verified by a further difference-density study in which 5-iodo-CTP was bound to the enzyme. In this brief paper we summarize x-ray diffraction studies at 3.0-Å resolution for the unliganded *R*32 form and at 2.8-Å resolution for the CTP-*P*321 form. The major features of secondary structure are presented below.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

STRUCTURE DETERMINATION

The unliganded *R*32 form has one RC group per asymmetric unit in a unit cell of dimensions a = 132 Å and c = 199 Å, while the CTP-P321 form has two RC units per crystallographic asymmetric unit in a unit cell of dimensions a = 122 Å and c = 142 Å. The x-ray diffraction data were collected by use of the oscillation method (15) from the native enzyme, and from heavy atom derivatives listed in Table 1. Conditions of preparation of these derivatives have been given previously (13, 14), except for the third CTP-*P*321 derivative, for which crystals were immersed in 5 mM $KAu(CN)_2$ and 40 mM K_2HPO_4 at pH 5.8 for two days. Heavy atom coordinates were found in difference electron density maps, starting from the phases obtained at low resolution, and were refined by alternate cycles of refinement of phases and least squares parameters (13, 14, 16). For the unliganded *R*32 form we used initially only those 5452 independent reflections that were observed in the native protein and in all four derivatives, and in later refinements phased 9518 independent reflections. In the *P*321 form we employed those 21,983 independent reflections present in the native crystals and at least two heavy atom derivatives. In all, we phased 74% of the theoretical number of reflections to 3.0 Å for *R*32, and 76% to 2.8 Å for *P*321 crystals.

In the CTP-*P*321 form there is a noncrystallographic twofold axis that was obeyed to observable accuracy for the C_3 trimer and the Zn domain (but not the CTP-binding domain) of the R_2 dimer. The phases of this electron density map were further improved by flattening the solvent density and averaging the molecular density, except in the CTP-binding domain, about this twofold axis in direct space, back transforming to reciprocal space, and combining the calculated phases with the observed structure factors for 24,238 independent reflections (17, 18).

From these final electron density maps a model was built in a Richards optical comparator (19) on a scale of 1.25 cm/Å.

DESCRIPTION OF THE MOLECULE

In Fig. 1 we show the gross molecular structure in D_3 symmetry. Three catalytic monomers are closely associated to form each trimer, and the two nearly eclipsed C_3 units are in close contact with each other in both the unliganded *R*32 and the CTP-*P*321 forms. Each catalytic monomer C is associated with a C monomer in the opposite C_3 subunit by interactions through a regulatory dimer. The two C units thus related through a given R_2 subunit are 120° apart about the threefold axis (Fig. 1).

In more detail, we show a C_3R_3 unit in Fig. 2 as the half of Fig. 1 closer to the reader. The polar domain (containing helices H1–H4 and a five-stranded parallel β sheet) of the C chain is closer to the threefold axis than is the equatorial domain (containing helices H6–H8 and a six-stranded parallel β sheet),

Abbreviations: C, catalytic chain of aspartate carbamoyltransferase; R, regulatory chain; MSA, 2-acetoxymercuri-4-nitrosalicylaldehyde.

Table 1. Heavy atom derivatives and refinement criteria

Derivative	No. of sites		Root mean square*					
			14.0–4.5 Å	4.5–3.8 Å	3.8–3.3 Å	3.3–3.0 Å	3.0–2.8 Å	All data
*R*32								
MSA	3	*Fc*	380	279	234	186	—	301
		E	165	142	122	92	—	140
$KAu(CN)_2$	3	*Fc*	150	117	97	78	—	122
		E	101	107	95	93	—	100
$K_2Pt(NO_2)_4$	9	*Fc*	280	185	142	119	—	218
		E	223	216	187	218	—	212
$IrBr_3$	7	*Fc*	285	166	115	88	—	212
		E	170	170	190	245	—	181
*P*321								
MSA	6	*Fc*	157	126	112	97	86	127
		E	83	76	66	62	58	72
$UO_2(NO_3)_2$	12	*Fc*	205	153	136	117	99	159
		E	103	105	97	79	67	96
$KAu(CN)_2$	5	*Fc*	80	64	60	55	49	66
		E	67	62	57	56	59	61
No. of reflections (unliganded *R*32)			3539	2120	2395	1464	—	9,518
No. of reflections (CTP-*P*321)			6711	4196	4889	3526	2661	21,983
Average figure of merit (unliganded *R*32)			0.738	0.647	0.604	0.528	—	0.652
Average figure of merit (CTP-*P*321)			0.668	0.629	0.602	0.562	0.523	0.611

MSA, 2-acetoxymercuri-4-nitrosalicylaldehyde.

* rms $Fc = |\Sigma_h f_{hj}^2/n|^{1/2}$, in which f_{hj} is the heavy atom scattering amplitude for reflection *h* of derivative *j* and *n* is the number of reflections. The unit is electrons. The rms structure factor of the native protein is 1055 electrons for the unliganded *R*32 form and 552 electrons for the CTP-*P*321 form. rms $E = |\Sigma_h e_{hj}^2/n|^{1/2}$, in which e_{jh} is the lack of closure for reflection *h* of derivative *j* and *n* is the number of reflections. The unit is also electrons. The figure of merit is the mean value of the cosine of the phase angle error.

which is closer to the equatorial region containing the twofold axis. These domains are connected by helices H5 and H9, which is the COOH-terminal helix. Both the NH_2-terminus and the COOH terminus of the C chain are in the polar domain. These elements of secondary structure are well defined throughout the C chain, perhaps somewhat better in the polar domain, and have allowed us to fit the partial sequence (kindly supplied by

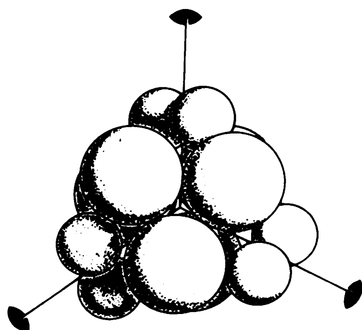

FIG. 1. Subunit assembly of two C_3 and three R_2 subunits in aspartate carbamoyltransferase. Each larger sphere represents a catalytic monomer C, and each smaller sphere represents a regulatory monomer R. The view is along the threefold axis, and the twofold axes are indicated by the usual symbol. This choice of enantiomer has been made on the basis of right-handed α-helices in the high-resolution structures.

W. Konigsberg) to the polar domain and to helices H5 and H9. The C–C interactions in the C_3 subunit are between helix H2 of one C chain and helix H3 and strand S3 of the adjacent C chain. Interactions between the two C_3 subunits occur between the loops of chain connecting S7 to H7 and S9 to S10.

In the regulatory chain, the fitting of the known sequence (20) is complete in the Zn domain (residues 102–152), but only partially complete in the CTP-binding domain (residues 1–96) because of the intrinsic disorder in the electron density maps. The CTP-binding domain includes helices H1′ and H2′ and a four-stranded β sheet. The CTP site, indicated in Fig. 2, is just to the left of this β sheet in the region of electron density that is present in the CTP-*P*321 form but absent in the unliganded *R*32 form. The approximately 4-Å distance of this "CTP density" from the four-stranded β sheet favors the previously preferred (14) assignment of the regulatory region. The part of this additional density associated with the pyrimidine ring has been tentatively assigned on the basis of the iodine position as determined at low resolution (14) from the 5-iodo-CTP derivative of this CTP-*P*321 crystal form. However, details of the binding of CTP are not yet clear because of some disorder in this region of the molecule. Strands S2′ and S3′ are connected, but we are unsure whether there are some 4 or 5 amino acids or as many as 15 amino acids in this region. The former choice would require that the extra 10 or 11 amino acids be added at the NH_2-terminus in a disordered region. Hence, we have to

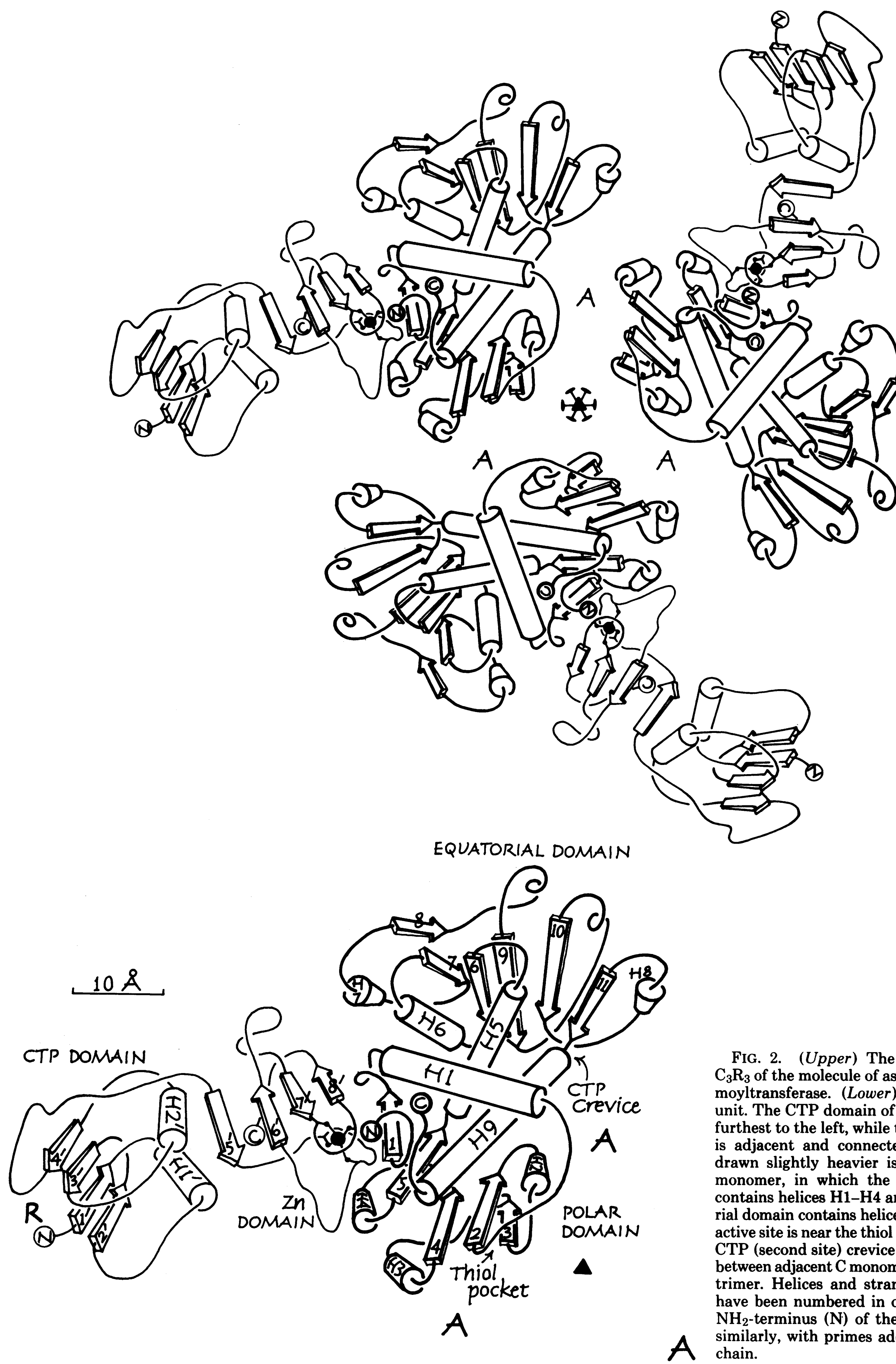

FIG. 2. (*Upper*) The "upper half" C_3R_3 of the molecule of aspartate carbamoyltransferase. (*Lower*) Labeled RC unit. The CTP domain of the R chain is furthest to the left, while the Zn domain is adjacent and connected. The chain drawn slightly heavier is the catalytic monomer, in which the polar domain contains helices H1–H4 and the equatorial domain contains helices H6–H8. The active site is near the thiol pocket and the CTP (second site) crevice at a boundary between adjacent C monomers within a C_3 trimer. Helices and strands of β sheet have been numbered in order from the NH_2-terminus (N) of the C chain, and similarly, with primes added, for the R chain.

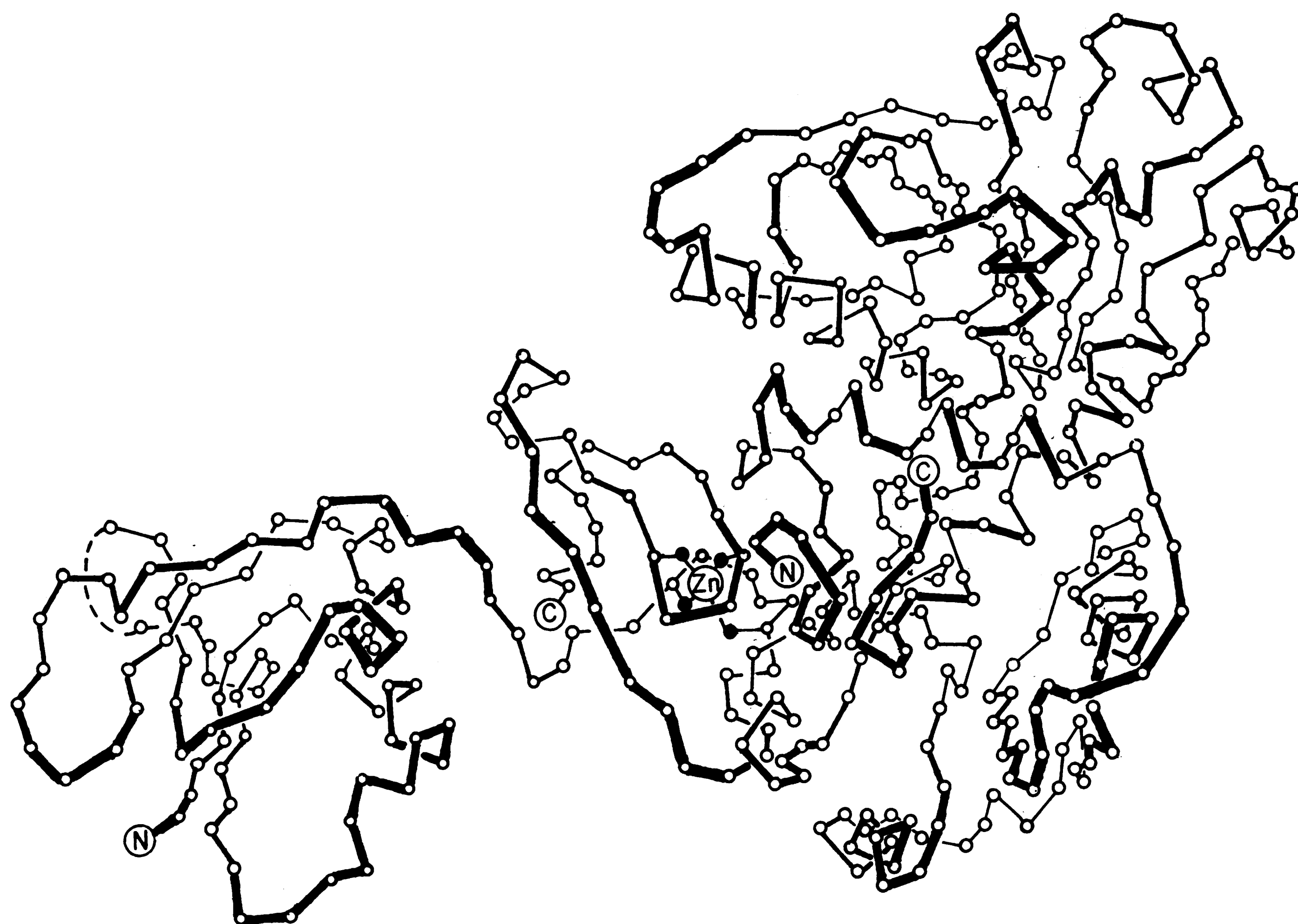

FIG. 3. Drawing of the R (left) and C (right) polypeptide chains in one RC unit of the hexameric C_6R_6 molecule. About 306 amino acids (±1% or so) are indicated in the catalytic chain in the electron density map.

leave a decision on this part of the trace for a later study. The four-stranded β sheet (S1′–S4′) of one regulatory chain is related to a similar sheet of the other regulatory chain of the dimer R_2 by a molecular twofold axis in such a way that an eight-stranded sheet is formed. Other structures that contribute to the stability of the dimer R_2 are helix H1′ and the chain that connects it to S2′, which are in contact with the chain that, in the other R unit, connects S2′ to S3′. Of course, a similar reverse interaction occurs, related by the molecular twofold axis.

Also in the regulatory chain, the Zn domain is dominated by the tetrahedral bonding of Zn^{2+} to cysteines 109, 114, 137, and 140 and the four-stranded antiparallel β sheet (S5′–S8′). The amino acid sequence can be reasonably well fitted to our electron density maps from Lys-56 to the COOH-terminal Asn-152. In particular, the Zn domain is a comparatively well-ordered region, and its contact with the polar domain of the C chain is well defined. Indeed, all of the substructures (e.g., loops ending at Cys residues) defined by the coordination of Zn^{2+} to the four Cys residues of the R chain participate in the interactions between R and C chains. It is therefore clear why mercurials that react with these Cys residues of the R chain dissociate the C_6R_6 molecule into smaller fragments. In particular, the regions of R chain from Ser-112 to Ser-121, and the Lys-138–Tyr-139 pair (in Cys . . . Cys loops) are in very close contact in the C chain with helix H4 and that region immediately before it in sequence. In addition, Glu-141 and Lys-142 of the R chain interact with another part of the C chain in the region before H4 in the amino acid sequence. A more detailed drawing of the R and C chains is shown in Fig. 3.

Chemical modification of Lys-84 (21, 22) destroys catalytic activity. We have placed this residue towards the NH_2-terminal direction at the beginning of helix H3. Also, chemical modification of the only Cys residue of the C chain suggests that it is near the active site: modification by small uncharged groups leaves the enzyme active (23) while modification by 2-chloromercuri-4-nitrophenol inactivates the enzyme (24). Also, oxidation of this C-chain thiol to sulfonate by potassium permanganate inactivates the enzyme (25). From the mercury derivatives (Table 1) we have located this thiol in a pocket between polar β sheet and helix H3 of one monomer and helix H2 of the adjacent monomer. This "thiol pocket" can accommodate an organomercurial such as MSA, and it opens most directly into the central cavity of the hexameric C_6R_6 molecule. Nearby, there is a crevice, in which a second 5-iodo-CTP molecule binds (14); this CTP crevice (Fig. 2) is between strands S6, S9, and S10 of the equatorial domain and helices H2, H5, and H9, all within the same C chain. The CTP crevice also opens into the central cavity. Helix H2 is at the interface between these two cavities, forming one boundary of the CTP crevice within the same C chain, and also of the thiol pocket of the adjacent C chain.

The portion of regulatory chain that is closest to these pockets is the part of the large Cys-114–Cys-137 loop that is nearer Cys-114. This region of R is in contact with helix H4 and in close proximity with helix H3. In the thiol pocket, adjacent to helix H3, reaction of this C-chain thiol with MSA also involves modification of Lys-84 (26). The nearby CTP crevice (Fig. 2) may also be a candidate for a binding region for substrates or inactivating reagents. If so, the active site appears to be shared

between adjacent C chains within a given catalytic trimer C_3, perhaps in a manner resembling that suggested (27) for glutathione reductase.

Our description here indicates that there is very little difference between the electron density of the unliganded molecule (*R*32) and the CTP-liganded molecule (*P*321), in spite of the different buffers (Tris/imidazole and citrate), their concentrations (0.05/0.05 and 0.04 M), and their pH values (6.35 and 5.8), respectively. In fact, the major difference is density ascribed to CTP in the *P*321 crystals as compared with no electron density in the corresponding regions of the *R*32 crystals. It is likely that this conformation is therefore of biochemical significance. However, conclusions with respect to activity are to be made very cautiously, because this is not the active conformation of aspartate carbamoyltransferase.

We have grown two crystal forms of the complex of aspartate carbamoyltransferase with its putative transition state analogue *N*-phosphonacetyl-L-aspartate. The crystal symmetry of one of them is *R*32 with unit cell dimensions $a = 123$ Å and $c = 709$ Å. The other crystalline form of this derivative of the enzyme has the space group *P*321 and unit cell dimensions of $a = 122$ Å and $c = 157$ Å. One can argue from these data that the symmetry of this enzyme when bound to molecules of this substrate analogue is D_3 from the former and at least C_3 from the latter crystals. If, as sedimentation ultracentrifuge experiments suggest (28), there is a large conformational change, it probably is at least an elongation of the molecule along the threefold axis, a change that is consistent with these unit cell measurements but that is not proved by them.

With this paper, we recognize Prof. E. Havinga on the occasion of his 70th birthday anniversary.

We thank Professor William Konigsberg for his preliminary results on the polypeptide sequence of the catalytic subunit. This work is based on earlier results by B. F. P. Edwards, G. Eisele, D. R. Evans, S. G. Warren, D. C. Wiley, and B. Wishner. Research support by Grant GM 06920 from the National Institutes of Health is gratefully acknowledged.

The model fitting was partially done with the aid of the GRIP-75 molecular graphics system (29), built by the University of North Carolina Department of Computer Science. Major system contributions were made by E. G. Britton, J. S. Lipscomb, M. E. Pique, W. Siddall, J. E. McQueen, Jr., J. Hermans, W. V. Wright, and F. P. Brooks, Jr. (principal investigator), and the system development was supported by the National Institutes of Health, the National Science Foundation, the Atomic Energy Commission, and the IBM Corporation.

1. Jones, M. E., Spector, L. & Lipman, F. (1955) *J. Am. Chem. Soc.* **77,** 819–820.
2. Reichard, P. & Hanshoff, G. (1956) *Acta Chem. Scand.* **10,** 548–566.
3. Yates, R. A. & Pardee, A. B. (1956) *J. Biol. Chem.* **221,** 757–770.
4. Gerhart, J. C. & Pardee, A. B. (1962) *J. Biol. Chem.* **237,** 891–896.
5. Weber, K. (1968) *Nature (London)* **218,** 1116–1119.
6. Wiley, D. C. & Lipscomb, W. N. (1968) *Nature (London)* **218,** 1119–1121.
7. Gerhart, J. C. & Schachman, H. K. (1965) *Biochemistry* **4,** 1054–1062.
8. Meighen, E. A., Pigiet, V. & Schachman, H. K. (1970) *Proc. Natl. Acad. Sci. USA* **65,** 234–241.
9. Rosenbusch, J. P. & Weber, K. (1971) *J. Biol. Chem.* **246,** 1644–1657.
10. Rosenbusch, J. P. & Weber, K. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 1019–1023.
11. Nellbach, M. E., Pigiet, V. P., Gerhart, J. C. & Schachman, H. K. (1972) *Biochemistry* **11,** 315–327.
12. Jacobson, G. R. & Stark, G. R. (1973) in *The Enzymes,* ed. Boyer, P. D. (Academic, New York), Vol. 9, pp. 225–308.
13. Warren, S. G., Edwards, B. F. P., Evans, D. R , Wiley, D. C. & Lipscomb, W. N. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 1117–1121.
14. Edwards, B. F. P., Evans, D. R., Warren, S. G., Monaco, H. L., Landfear, S. M., Eisele, G., Crawford, J. L., Wiley, D. C. & Lipscomb, W. N. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 4437–4441.
15. Crawford, J. L. (1977) Dissertation (Harvard University, Cambridge, MA).
16. Dickerson, R. E., Kopka, M. L., Varnum, J. C. & Weinzierl, J. E. (1967) *Acta Crystallogr.* **23,** 511–522.
17. Fletterich, R. J. & Steitz, T. A. (1976) *Acta Crystallogr.* **A32,** 125–131.
18. Bricogne, G. (1976) *Acta Crystallogr.* **A32,** 832–847.
19. Richards, F. M. (1968) *J. Mol. Biol.* **37,** 225–230.
20. Weber, K. (1968) *Nature (London)* **218,** 1116–1119.
21. Greenwell, P., Jewett, S. L. & Stark, G. R. (1973) *J. Biol. Chem.* **248,** 5994–6001.
22. Kempe, T. D. & Stark, G. R. (1975) *J. Biol. Chem.* **250,** 6861–6869.
23. Vanaman, T. C. & Stark, G. R. (1970) *J. Biol. Chem.* **245,** 3565–3573.
24. Evans, D. R., McMurray, C. H. & Lipscomb, W. N. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 3638–3642.
25. Benisek, W. F. (1971) *J. Biol. Chem.* **246,** 3151–3959.
26. Landfear, S. M. (1978) Dissertation (Harvard University, Cambridge, MA).
27. Schulz, G. E., Schirmer, R. H., Sachsenheimer, W. & Pai, E. F. (1978) *Nature (London)* **273,** 120–124.
28. Gerhart, J. C. & Schachman, H. K. (1968) *Biochemistry* **7,** 538–552.
29. Britton, E. G., Lipscomb, J. S. & Pique, M. E. (1978) *Computer Graphics* **12** (3), 222–227.