Proc. Nat. Acad. Sci. USA
Vol. 70, No. 4, pp. 1117–1121, April 1973

Aspartate Transcarbamoylase from *Escherichia coli*: Electron Density at 5.5 Å Resolution

(enzyme activity/allosteric enzyme/x-ray diffraction/crystallography)

STEPHEN G. WARREN, BRIAN F. P. EDWARDS, DAVID R. EVANS, DON C. WILEY, AND WILLIAM N. LIPSCOMB

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

Contributed by William N. Lipscomb, February 5, 1973

ABSTRACT **The allosteric enzyme, aspartate transcarbamoylase (EC 2.1.3.2), has previously been shown in our x-ray diffraction studies to have D_3-32 symmetry. There are six catalytic (C) and six regulatory (R) chains in the molecular complex (R_6C_6). Our three-dimensional x-ray diffraction study of this enzyme (R32, a = 131 Å, c = 200 Å) at 5.5 Å resolution shows a spatial arrangement of the two catalytic trimers C_3 above and below an equatorial belt of three regulatory dimers R_2. The molecule is about 110 × 110 × 90 Å in largest dimensions, and is shown here to contain a large central aqueous cavity about 50 × 50 × 25 Å in size. Location of the single sulfhydryl of each catalytic chain, and correlation of its reactivity with enzymatic activity in the molecule, suggests that the nearby active sites are most probably accessible from the central cavity, but probably not directly from the external solution. The most obvious access to the central cavity consists of six channels, each about 15 Å in diameter, near the regulatory region. A component of the regulatory mechanism may be modulation of access of substrates through these channels.**

Regulation of enzyme activity is an important biochemical process within living cells. A number of mechanisms have been considered: (*a*) steric and conformational effects associated with the direct binding of small inhibitors at or near the active site, (*b*) the direct covering of the active site by another protein, (*c*) change of conformation at the active site upon binding of an effector in another region of the enzyme molecule, and (*d*) a conformational change transmitted from one protein molecule to another in an enzyme complex. We present here a 5.5 Å resolution structure of aspartate transcarbamoylase, features of which suggest a new possible component to general mechanisms of regulation: namely, the direct steric regulation of substrate access to the active site in a central cavity of the molecular complex.

Aspartate transcarbamoylase (ATCase) catalyzes the first committed step in pyrimidine biosynthesis (1, 2), the reaction of carbamoylphosphate with aspartate to yield carbamoylaspartate and phosphate. The activity of ATCase is modulated by allosteric effectors. Cytidine triphosphate, an end-product of the pyrimidine pathway, is a feedback inhibitor (3, 4), while ATP stimulates ATCase. These effectors thus provide one mechanism for achieving balance between pyrimidines and purines for biosynthesis of nucleic acids.

The catalytic and regulatory functions of ATCase (molecular weight 310,000, ref. 5) are associated with different polypeptide chains C and R, respectively (4, 5). The hexameric nature of ATCase (R_6C_6) was established by (*a*) the presence of elements of both 2-fold and 3-fold molecular symmetry in the x-ray study (6) and (*b*) the molecular weights of C and R, along with sequence analysis of the R chain (7). Dissociation of ATCase by reaction of the R-chain thiols with mercurials (5) has been interpreted (6, 7) to yield catalytic trimers C_3 and regulatory dimers R_2. The relative positions of the C_3 units above and below a wider equatorial belt of three regulatory dimers were established in the x-ray study (8), which also indicated the presence of a large solvent-filled region in the center of the ATCase molecule. We describe here a more detailed view of this and other features of the electron density of ATCase at 5.5-Å resolution from a greatly improved map.

Abbreviations: ATCase, aspartate transcarbamoylase (EC 2.1.3.2); C, catalytic chain; R, regulatory chain.

METHODS

The crystal form has space group R32, and hexagonal axes a = 131 Å and c = 200 Å. The molecule is situated at the intersection of a 3-fold and three 2-fold symmetry axes in the crystal. Hence, one-sixth of the molecule is in the crystallographic asymmetric unit.

Following our initial attempts to solve for the multiple heavy-atom positions in CTP-ATCase (6), we decided on a careful chemical program which would yield only one heavy atom on each RC unit of ATCase. This procedure was expected to yield a solvable difference Patterson function, from which these heavy-atom positions could be located. The phases of structure amplitudes from this derivative would then be employed to solve the more complicated multiple-site heavy-atom derivatives. Preparation of the 2-chloromercuri-4-nitrophenol derivative of the single relatively unreactive sulfhydryl group on each of the six catalytic sites has been described (9). This single-site derivative was prepared by reaction with ATCase in solution, and was subsequently crystallized. The three other derivatives were prepared by soaking crystals of ATCase in the following solutions: Au (8), 6.2 mM $KAu(CN)_2$, for 2–4 days, buffer of 0.05 M Tris–0.05 M imidazole (pH 6.35); Ir, saturated (10 mM) $IrBr_3$ for 5 days, buffer of Tris–imidazole (pH 6.35); Pt, 1 mM K_2PtCl_4, for 2–4 days, buffer of 0.1 M triethanolamine·HCl (pH 6.35).

Procedures for the collection of three dimensional x-ray diffraction data have been described (8). Multiple films were taken by screened precession photography using an Elliott rotating anode x-ray generator. Intensities were read by an Optronics P-1000 film scanner on-line to a PDP-11 computer (Digital Equipment Corp.). The value of $R = \sum_h |I_h - \bar{I}_h| / \sum_h \bar{I}_h$ for all symmetry-related reflections h was 0.053, where $\bar{I}_h$

TABLE 1. *Statistics of phasing**

Derivative	RMS F_C, elec.†	RMS E, elec.‡	R_C§	R_K¶
Hg	476	216	0.44	0.062
Au	351	219	0.46	0.062
Pt	688	336	0.53	0.095
Ir	451	299	0.53	0.081

* The figure of merit (mean value of the cosine of the phase angle error) is $\bar{M} = 0.77$ for this refinement in which scale factors were refined for each reciprocal lattice level (photograph) for the protein and derivative.

† RMS $F_C = [\sum_n f_{hj}^2/n]^{1/2}$, where f_{hj} is the heavy-atom scattering amplitude for reflection h of derivative j, and n is the number of reflections. The RMS structure factor of the native protein is 2870 elec.

‡ RMS E $= [\sum_h \epsilon_{hj}^2/n]^{1/2}$, where ϵ_{hj} is the lack of closure for reflection h of derivative j, and n is the number of reflections.

§ $R_C = \sum_h ||F_{pH} - F_p| - f_H|/\Sigma|F_{pH} - F_p|$ calculated for centric reflections only, where F_{pH} is the structure factor of the derivative, F_p for the protein, and f_H for the heavy atoms.

¶ $R_K = \sum_h ||F_{pH}| - |F_p + f_H||/\sum|F_{pH}|$, calculated for all reflections h.

is the average on a given film. For different films, i and j, the value of $R_{3D} = \sum_h |I_{hi} - I_{hj}|/\Sigma|I_{hi} + I_{hj}|$ was 0.065 for the

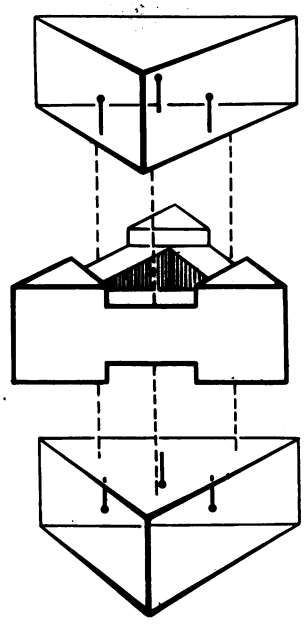

FIG. 1. Schematic view of the ATCase molecule expanded along the 3-fold axis. The central region is mostly regulatory, although some parts at the ends of *dotted lines* may belong to the catalytic trimers, most of which occur above and below this central region. The assembled molecule, 90 Å along 3 and 110 Å along 2 axes, has a central cavity (*shaded*) in overall dimensions about 25 Å along 3 and 50 Å along 2 axes. The Hg sites within C_3 units (*dots*) are presumed to be near the active site, which has direct access (*solid lines*) towards the central cavity. Six 15-Å openings (not shown) to the central cavity occur near the equatorial region.

TABLE 2. *Heavy-atom parameters**

Compound	Site no.	Occupancy (%)	Fractional coordinates x	y	z
2-Chloromercuri-4-nitrophenol (9)	1	95	0.089	0.100	0.104
$KAu(CN)_2$	1	64	0.082	0.133	0.097
6.2 mM†	2	20	0.275	0.126	0.028
	3	10	0.121	0.201	0.052
	4	13	0.728	0.251	0.155
	5	11	0.056	0.267	0.009
$IrBr_3$	1	79	0.008	0.153	0.071
10 mM†	2	38	0.040	0.130	0.074
	3	27	0.272	0.069	0.093
	4	18	0.124	0.069	0.125
	5	9	0.333	0.091	0.100
K_2PtCl_4	1	96	0.329	0.078	0.102
1 mM‡	2	37	0.0	0.0	0.157
	3	55	0.840	0.468	0.010
	4	31	0.333	0.006	0.112
	5	24	0.053	0.105	0.075
	6	16	0.905	0.294	0.163
	7	15	0.024	0.312	0.015

* The absolute hand of the entire heavy-atom constellation, and thus of the protein map, has not yet been determined.

† Buffer, 0.05 M imidazole–0.05 M Tris·HCl (pH 6.35).

‡ Buffer, 0.1 M triethanolamine·HCl (pH 6.35).

intensities of native ATCase crystals, and 0.078, 0.074, 0.096, and 0.082 for the intensities of the Hg, Au, Ir, and Pt derivatives, respectively.

Instead of proceeding with these five merged data sets for ATCase and four derivatives, refinement of phases and heavy atom parameters was done level-by-level (10), allowing the scale factor for each film (level) and the heavy-atom parameters to shift in the refinement. Merged three-dimensional sets of data were then prepared using these individual film scale factors. The final phases were calculated using these merged sets and the heavy-atom parameters without further refinements. In this way the statistics of phasing (Table 1) can be compared with published values. These phases were used to compute the electron-density map.

We also obtained phases after the "level-by-level" refinement by averaging the phases (weighted by figure of merit) for reflections determined on more than one film. Sections of the electron-density map of ATCase calculated using these phases were nearly identical to the corresponding sections of the map which we present here.

As an alternative to the level-by-level refinement we also obtained phases by carrying out the refinement on data sets prepared by the more usual procedure of merging films prior to refinement. The heavy-atom occupancies were similar, and the protein electron-density map was nearly identical, to results from level-by-level refinement, but the phasing statistics were somewhat worse (e.g., $\bar{M}$ is 0.74) as compared with the value of 0.77 shown in Table 1.

Occupancies and coordinates of heavy atoms are shown in Table 2. Because the resolution is only 5.5 Å in this study, we fixed the thermal parameter at 30 Å^2 for each heavy atom. Minor sites were dropped from the refinement if the occupancy fell below 9% or if a site already present in the phases

failed to appear in a subsequent difference electron-density map. The minor Au site in our previous study (8) disappeared in these refinements, while four other minor sites have appeared. Of course, we cannot be sure whether the smallest sites are real, due to conformational changes, or due to errors in phases or intensities. All of the heavy-atom sites which

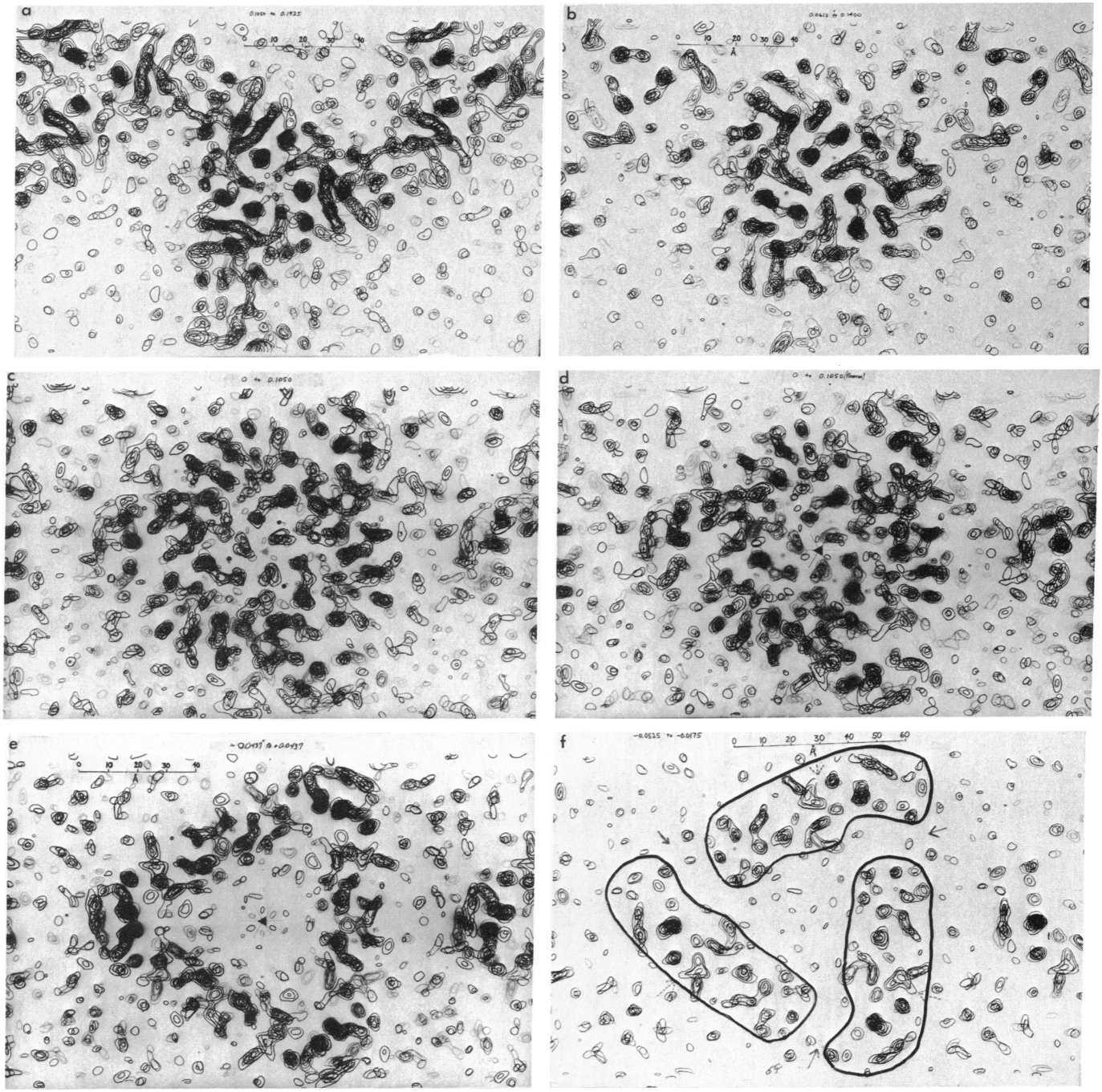

FIG. 2. Composites of the three-dimensional electron-density map of ATCase. (*a*) z = 0.1925–0.1050 (11 sections), the "top" of the C_3 units showing intermolecular contacts at the upper right and left, and prominent helical regions end-on near the center and at an inclined angle towards the outer parts of these units. These helical features are straight rods about 6 Å in diameter, and several are more than 20 Å in length. We now believe that access to the central cavity along the 3-fold axis is blocked or very greatly restricted by protein. (*b*) z = 0.1400–0.0612 (10 sections), near the central cross section of a C_3 unit. Densest features are helix regions, and the three faint *dots* 22 Å apart are Hg positions. (*c*) z = 0.1050–0 (13 sections), showing both catalytic (*near center*) and regulatory (*outer*) regions. The three *dots* are the Hg positions in section z = 0.1050. (*d*) z = 0–0.1050 (13 sections), showing the view from the molecular center (*solid triangle* at z = 0), looking toward the active sites (*faint dots*). (*e*) z = −0.0437–0.0437 (11 sections), spanning the center of the molecule give an excellent view of the central cavity. Most of the regulatory protein lies nearest the outer corners of this large central triangle of electron density. (*f*) z = −0.0525 to −0.0175 (5 sections), showing channels (*arrows*) about 15 Å in diameter leading from the outer solution to the central cavity and portions of the R and C chain shown enclosed in *solid lines*. *Dotted arrows* indicate the symmetry-related channels in the sections 0.0525–0.0175 not included here.

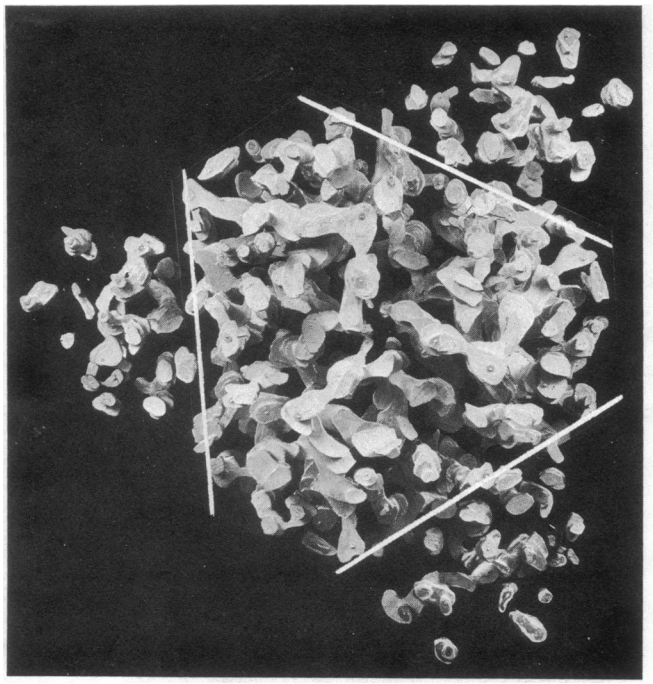

FIG. 3. A balsa wood model of ATCase, viewed down the 3-fold axis. Half of the molecule, i.e., $z > 0$, is shown. The C chain trimer is outlined in *white*.

satisfied these criteria are in regions of electron density of the protein, not in the extensive solvent regions in the crystal.

RESULTS AND DISCUSSION

The electron density was computed from all reflections between 5.5 Å and 15 Å, provided that they had been measured on the enzyme and at least three derivatives. Centroid phases (11) were used. Scaling was established by assuming nearly 100% occupancy for the Hg derivative (9), inasmuch as mercury analysis, spectrophotometric titration, and inactivation of the enzyme observed kinetically (9) indicate that the C-chain thiol has reacted quantitatively in this derivative.

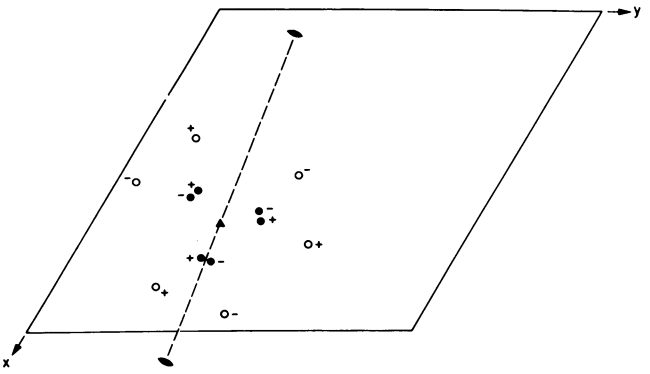

FIG. 4. Projection along the *c* axis of the heavy-atom positions within one ATCase molecule in its complex with CTP. The Hg positions on the six catalytic subunits are indicated by *solid dots*, and the Hg atoms replacing Zn are indicated by *open circles*. The plus and minus refer to *z* coordinates relative to the molecular center. Also, the position of one of the noncrystallographic (molecular) 2-fold axes is given as a *dashed line*. Mercury at the C-chain thiol is located ±21 Å in *z* from the molecular center. Hg at the Zn site is located ±11 Å in *z* from the molecular center.

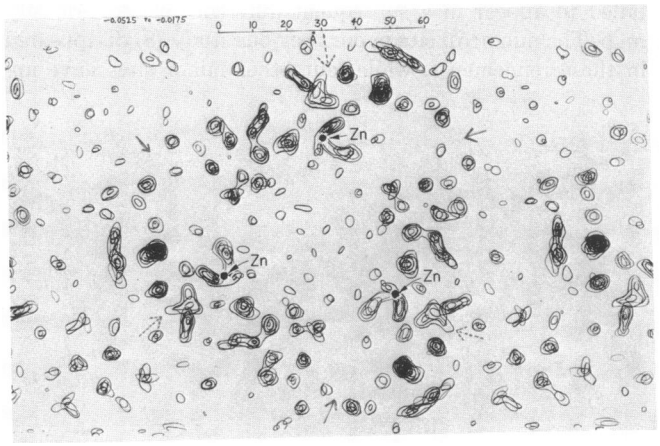

FIG. 5. $z = -0.0525$ to -0.0175. Five sections of electron density of R32 crystals. The tentative locations of the zinc ions are marked "Zn," at $z = -0.0525$. These positions are uncertain because they were located in the P321 crystals, in which the enzyme may have another conformation.

Contour levels were drawn at intervals of 0.07 eÅ^{-3} starting from the lowest contour at 0.14 eÅ^{-3} (F_{000} omitted). The highest density was about 0.5 eÅ^{-3}. Errors, as judged from the solvent region, appear in general to be limited to features less than 6 Å long at the lowest contour level.

The molecular boundary and central cavity are very clearly shown (Figs. 1–3). Regions previously assigned to the C_3 units are close to the 3-fold axis (Figs. 1 and 2*a–c*), while the major parts of the regulatory chains are near the $z = 0$ plane and mostly on the outer periphery of the equatorial region of the molecule (Figs. 1 and 2*e*). These basic structural features are in agreement with our earlier x-ray results (8), and with the later results from electron microscopy (12).

One important revision of our earlier interpretation (8) eliminates the funnel-like channel down the 3-fold axis towards the central cavity. In our improved map this channel has filled with electron density of the protein, blocking access from the top of the molecule to the central cavity. Reexamination of our earlier electron-density map, based upon only two heavy atom derivatives, shows that this new electron density was indeed present in the old map as very low contours, at about the same level as the noise.

Rod-like regions are attributed to α-helix, provided that they have a central density above 0.34 eÅ^{-3}, are at least 6 Å in diameter, and are at least 10 Å from any other rod. These rods are especially clear in the catalytic trimers, which are presently estimated to contain about 35% of helix. The longest rod is over 20 Å in length, parallel to the molecular 3-fold axis and only 10 Å from this axis. On the other hand, the helix content is estimated as only about 10% for the R units. At 5.5 Å resolution it is difficult to see short regions of helix, and we are uncertain about assignment of protein to C or to R in the inner parts of the equatorial region of the molecule. Hence, these percentages may be slightly low, and may need a small redistribution between types of subunits.

The molecular cavity (Fig. 2*e* and *f*) is about 25 Å in total extension along the 3-fold direction, and 50 Å in diameter in the plane of the 2-fold axes. The molecular center is halfway along these dimensions. Access to this central cavity is largest at six channels just above or below the equatorial plane, and in a region which certainly involves the R units and may also

be near the C units (Fig. 2*f*). Each channel is about 15 Å in diameter.

The Hg atoms on the catalytic subunits, believed to be near the active site (9), are 22 Å from each other, and at $z = \pm 21$ Å from the equatorial plane of the ATCase molecule (Figs. 1 and 2). The Hg sites are readily accessible from the internal cavity of the molecule, but appear not to be directly accessible from the outer surface of the enzyme molecule. This Hg derivative, prepared by reaction of ATCase with 2-chloromercuri-4-nitrophenol, is enzymatically inactive (9) and no longer binds succinate, an analogue of aspartate. Hence, it seems probable that the active site is also accessible from the central cavity. Location of this binding site in a difference electron density map is required in order to provide additional evidence that these results are not due to transmission of conformational effects associated with a more distant location of the active site external to the central cavity.

These features of the structure of ATCase stimulate speculation as to the purpose of the central cavity, its access channels, and the probable internal location of the active sites. It is tempting to suggest that some type of restriction of access of substrates to this central cavity might play a role in regulatory processes in ATCase, or in other regulatory complexes. For example, partial blocking of the channels may occur by part of the R (or C) chains in the complex of ATCase with CTP.

We therefore tentatively suggest a possible component of the regulatory mechanism in which the diffusion of substrates into the central cavity is the rate-limiting step. Kinetic schemes which involve restricted diffusion of both substrate and product through the access channels are qualitatively consistent with the observed kinetic effects of CTP. Restriction of access to the central cavity would not alter the maximum velocity, but would shift the position of the substrate saturation curve and explain the substrate dependence of maximal CTP inhibition. Furthermore, the location of the active sites in a central cavity of the molecule may explain the decrease in enzymatic activity relative to isolated C-subunit. The observation that substrate binding at the active site changes the conformation of the regulatory chain (13) may indicate that substrate, as well as CTP or ATP, can modulate access to the central cavity. Cooperative effects may arise in this scheme if substrate binding at an active site increases substrate access to the central cavity. More detailed aspects of this scheme should await confirmation of the location of the active site in the cavity and direct evidence of ligand-induced changes in the size of the access channels.

Since rate of catalysis depends on the concentration of substrates within the central cavity, another possible mechanism of regulation could be the alteration of the steady-state concentration of substrates in the cavity. The alteration could be imagined to result from ligand-induced changes in the internal environment of the cavity. Whether regulation actually does involve restriction of access to the central cavity or change in environment in the central cavity in addition to conformational changes involving the active site will be more apparent when biochemical and x-ray studies of ATCase and of its complexes with effectors have been carried to higher resolution.

As a general phenomenon central cavities in enzyme complexes may be of importance where regulated diffusion of substrates and intermediates may be advantageous. This situation would be particularly desirable in multienzyme complexes if intermediates were not convalently (14) bound. Central cavities may arise simply as a consequence of the packing, of nearly globular subunits around positions of high symmetry, like the D_3-32 central position of ATCase.

Appendix: The crystal of ATCase with CTP

Crystals of ATCase grown in the presence of CTP have space group P321 with $a = 122$ Å and $c = 143$ Å[6]. The two molecules per unit cell are situated on 3-fold axes, so that one-third of the molecule is in the crystallographic asymmetric unit. We have now been successful in preparing a mercury derivative on each catalytic chain, by immersing a crystal in 90 μM 2-chloromercuri-4-salicylaldehyde in the presence of 0.14 mM 2-mercaptoethanol (to mask the Hg) and 0.5 mM CTP. The relationship of this constellation of Hg atoms to the Hg atoms of the crystals of ATCase itself has now been established (Fig. 5). In fact, these Hg positions within the ATCase molecule and within the CTP-ATCase are the same (assuming the same absolute configuration) within about 1-2 Å, the limit of our present resolution.

We have also solved for the Hg positions in another derivative of CTP-ATCase in which Hg replaces Zn, and have shown that this Zn site (Fig. 4) is 24 Å away from the nearest C-chain thiol group. The location of the Zn ion has been transferred to our electron-density map (Fig. 5), and it appears to be near the R-C interface, as has been suggested (15, 16). However, it must be kept in mind that the zinc position was determined in the CTP-ATCase crystals, whereas the map is of unliganded ATCase, and the enzyme may have different conformations in the two crystal forms.

These two heavy-atom derivatives will be used to obtain approximate phases in order to solve the multiple-site derivatives obtained from $UO_2(NO_3)_2$, $K_2Pt(NO_2)_4$, and $KAu(CN)_2$.

We thank the National Institutes of Health (Grant GM 06920) for support, and the Guggenheim Foundation for a fellowship to W.N.L. We thank Prof. F. H. Westheimer for stimulating discussions. This study was presented at the Third Enzyme Conference, University of California at Los Angeles, Dec. 27–29, 1972.

1. Jones, M. E., Spector, L. & Lippman, F. (1955) *J. Amer. 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. Gerhart, J. C. & Schachman, H. K. (1965) *Biochemistry* **4**, 1054–1062.
6. Wiley, D. C. & Lipscomb, W. N. (1968) *Nature* **218**, 1119–1121.
7. Weber, K. (1968) *Nature* **218**, 1116–1119.
8. Wiley, D. C., Evans, D. R., Warren, S. G., McMurray, C. H., Edwards, B. F. P., Franks, W. A. & Lipscomb, W. N. (1971) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 285–290.
9. Evans, D. R., McMurray, C. H. & Lipscomb, W. N. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 3638–3642.
10. Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. (1967) *Nature* **214**, 652–656.
11. Blow, D. M. & Crick, F. H. C. (1959) *Acta Crystallogr.* **12**, 794–802.
12. Richards, K. E. & Williams, R. C. (1972) *Biochemistry* **11**, 3393–3395.
13. Gerhart, J. C. (1970) *Curr. Top. Cell. Regul.* **2**, 275–325.
14. Reed, L. J. & Cox, D. J. (1966) *Annu. Rev. Biochem.* **35**, 57–84.
15. Rosenbusch, J. P. & Weber, K. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1019–1023.
16. Nelbach, M. C., Pigiet, V. P., Gerhart, J. C. & Schachman, H. K. (1972) *Biochemistry* **11**, 315–327.