

# In vivo formation of allosteric aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains: Implications for protein folding and assembly



PINGHUI ZHANG<sup>1</sup> AND H.K. SCHACHMAN

Department of Molecular and Cell Biology and Virus Laboratory, Wendell M. Stanley Hall,  
University of California at Berkeley, Berkeley, California

(RECEIVED February 7, 1996; ACCEPTED April 17, 1996)

## Abstract

Because the N- and C-terminal amino acids of the catalytic (c) polypeptide chains of *Escherichia coli* aspartate transcarbamoylase (ATCase) are in close proximity to each other, it has been possible to form in vivo five different active ATCase variants in which the terminal regions of the wild-type c chains are linked in a continuous polypeptide chain and new termini are introduced elsewhere in either of the two structural domains of the c chain. These circularly permuted (cp) chains were produced by constructing tandem *pyrB* genes, which encode the c chain of ATCase, followed by application of PCR. Chains expressed in this way assemble efficiently in vivo to form active, stable ATCase variants. Three such variants have been purified and shown to have the kinetic and physical properties characteristic of wild-type ATCase composed of two catalytic (C) trimers and three regulatory (R) dimers. The values of  $V_{max}$  for <sub>cp</sub>ATCase<sub>122</sub>, <sub>cp</sub>ATCase<sub>222</sub>, and <sub>cp</sub>ATCase<sub>281</sub> ranged from 16–21  $\mu\text{mol}$  carbamoyl aspartate per  $\mu\text{g}$  per h, compared with 15 for wild-type ATCase, and the values for  $K_{0.5}$  for the variants were 4–17 mM aspartate, whereas wild-type ATCase exhibited a value of 6 mM. Hill coefficients for the three variants varied from 1.8 to 2.1, compared with 1.4 for the wild-type enzyme. As observed with wild-type ATCase, ATP activated the variants containing the circularly permuted chains, as shown by the lowering of  $K_{0.5}$  for aspartate and a decrease in the Hill coefficient ( $n_H$ ). In contrast, CTP caused both an increase in  $K_{0.5}$  and  $n_H$  for the variants, just as observed with wild-type ATCase. Thus, the enzyme containing the permuted chains with widely diverse N- and C-termini exhibited the homotropic and heterotropic effects characteristic of wild-type ATCase. The decrease in the sedimentation coefficient of the variants caused by the binding of the bisubstrate ligand *N*-(phosphonacetyl)-L-aspartate (PALA) was also virtually identical to that obtained with wild-type ATCase, thereby indicating that these altered ATCase molecules undergo the analogous ligand-promoted allosteric transition from the taut (T) state to the relaxed (R) conformation. These ATCase molecules with new N- and C-termini widely dispersed throughout the c chains are valuable models for studying in vivo and in vitro folding of polypeptide chains.

**Keywords:** circular permutation; cooperativity; folding; protein engineering; stability

Reprint requests to: Howard K. Schachman, Department of Molecular and Cell Biology, University of California at Berkeley, 229 Stanley Hall #3206, Berkeley, California 94720-3206; e-mail: schach@garnet.berkeley.edu.

<sup>1</sup> Present address: AmCell Corp., Sunnyvale, California 94089.

**Abbreviations:** ATCase, aspartate transcarbamoylase; c, catalytic polypeptide chain; r, regulatory polypeptide chain; C, catalytic trimer; R, regulatory dimer; wt as subscript, wild-type; cp as subscript, circularly permuted; c and numbers following it in subscript designate the positions of the amino acid residues in the wild-type catalytic chain at which the circularly permuted catalytic chains start; PALA, *N*-(phosphonacetyl)-L-aspartate.

In earlier studies (Yang & Schachman, 1993a) of *Escherichia coli* aspartate transcarbamoylase (aspartate carbamoyltransferase, carbamoyl phosphate: L-aspartate carbamoyltransferase, EC 2.1.3.2), manipulation of the *pyrB* gene, which encodes the catalytic polypeptide chain of ATCase, was shown to yield constructs leading to efficient expression of active enzyme in which the c chains had new N- and C-termini and amino acid residues in the vicinity of the wild-type termini were linked covalently in a continuous polypeptide chain. These circularly permuted c chains could fold in vivo and associate to form active, stable

catalytic trimers which, with the addition of regulatory dimers, then could assemble into an ATCase-like holoenzyme ( $C_2R_3$ ) containing two C trimers and three R dimers. Analogous experiments leading to proteins containing circularly permuted polypeptide chains have been performed with phosphoribosyl anthranilate isomerase (Luger et al., 1989), dihydrofolate reductase (Buchwalder et al., 1992; Protasova et al., 1994), interleukin 1 $\beta$  (Horlick et al., 1992), T4 lysozyme (Zhang et al., 1993, 1994), ornithine decarboxylase (Li & Coffino, 1993), ribonuclease T1 (Mullins et al., 1994; Garrett et al., 1996),  $\beta$ -glucanase H (Hahn et al., 1994), interleukin 4 (Kreitman et al., 1994, 1995a, 1995b), glyceraldehyde 3-phosphate dehydrogenase (Vignais et al., 1995), and  $\alpha$ -spectrin SH3 domain (Viguera et al., 1995). Also, a combination of chemical linkage of the N- and C-termini in bovine pancreatic trypsin inhibitor followed by limited hydrolysis of a single peptide bond forming new termini yielded circularly permuted chains (Goldenberg & Creighton, 1983). In the case of concanavalin A, a posttranslational cleavage process and a transpeptidation reaction led to a final protein that was circularly permuted relative to its precursor and other lectins (Bowles et al., 1986).

Success in forming stable, active proteins composed of circularly permuted polypeptide chains requires that the covalent linking of residues at or near the original termini not place too much strain on the molecule and that the newly formed N- and C-termini be in appropriate locations to permit proper folding of the chains. As pointed out by Thornton and Sibanda (1983), many globular proteins, both those having single domains as well as others with several domains, have their N- and C-terminal residues in close proximity to each other in flexible regions of the chain at the surface of the protein. Closure of the original termini without conferring major instability of the folded chain, therefore, does not seem to constitute a formidable problem. However, little is known about restrictions in the introduction of new N- and C-termini. Where can they be located or, perhaps of greater importance, where can they not be located? Must they be in flexible loops between domains or can the new termini be within domains? Can the new termini be in the interior of the molecule rather than at the surface? In attempts to develop answers to some of these questions, we extended the previous work on ATCase by using a simpler method for circularly permuting the c chains of the enzyme. This led to the expression of five different active variants of ATCase in which the N- and C-termini were located in diverse regions of the molecule in either of the two domains of the folded c chains. Three such proteins have been purified and their kinetic and physical properties are described below. It is important to note that these variants, unlike the previously circularly permuted proteins (Yang & Schachman, 1993a) do possess the allosteric properties characteristic of wild-type ATCase. Thus, the circular permutation of the c chains with different locations of the termini did not perturb the delicate balance between the taut (T) and relaxed (R) states of ATCase that is so essential for its functioning as a regulatory enzyme. These results constitute convincing evidence that the conformations of the variant forms of ATCase containing circularly permuted chains are very similar to that of wild-type ATCase both in the unliganded and liganded form.

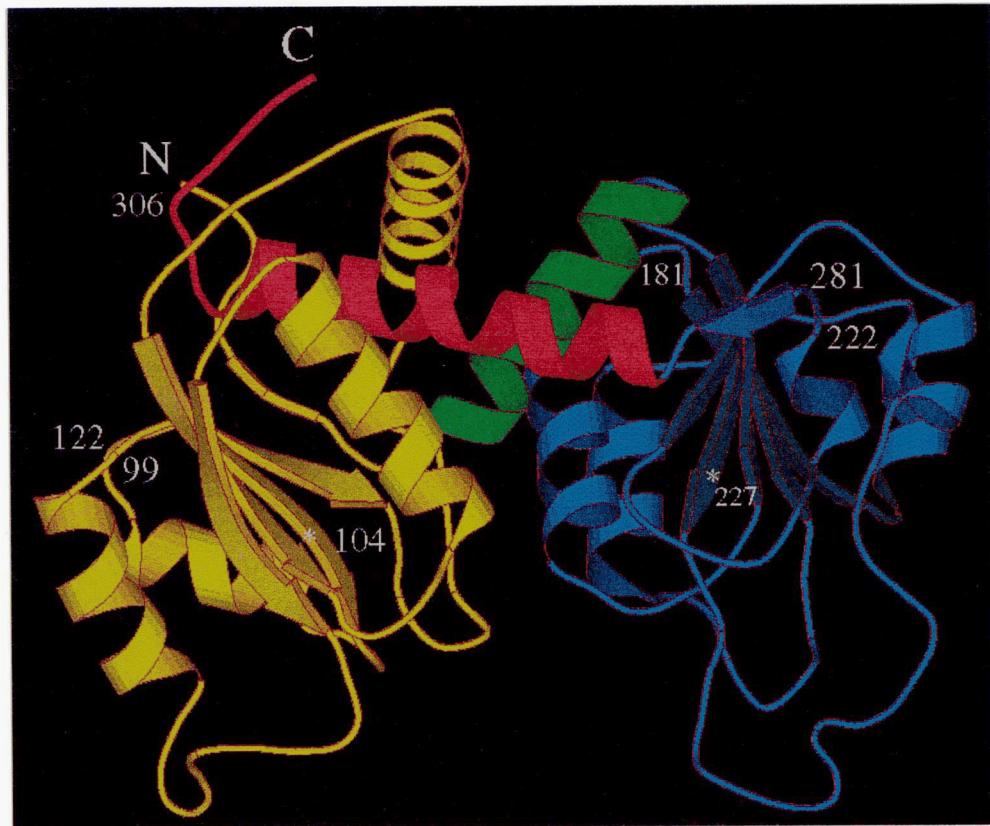
The availability of C trimers containing different circularly permuted polypeptide chains is of particular relevance to ongoing discussions about protein folding processes (Creighton, 1990; Dill, 1990; Kim & Baldwin, 1990; Jaenicke, 1991; Matthews,

1993; Baker & Agard, 1994; Viguera et al., 1995) and questions about the relationship between in vitro experiments with intact, disorganized polypeptide chains obtained by denaturation of proteins and in vivo studies where folding may be occurring as the polypeptide is being synthesized (Bergman & Kuehl, 1979; Peters & Davidson, 1982; Fedorov et al., 1992; Hahn et al., 1994).

### Rationale for construction of circularly permuted catalytic chains

Crystallographic studies of ATCase (Honzatko & Lipscomb, 1982; Kim et al., 1987; Ke et al., 1988; Lipscomb, 1994) have shown that the individual c chains are folded into an N-terminal domain comprising residues 1–134 and a C-terminal domain of residues 150–284 (Kinemages 1, 2, and 3). As seen in Figure 1 and Kinemage 3, these domains are linked covalently by an  $\alpha$ -helix composed of residues 135–149 and a second helix that crosses over from the C-terminal domain and threads through the N-terminal domain. This helix, containing residues 285–305, is followed by a short, flexible pentapeptide region terminating at residue 310, which is located about 14 Å from the N-terminal amino acid. Residue 306 is only about 5 Å from the N-terminal residue. Hence, the linkages used earlier (Yang & Schachman, 1993a) involved either the insertion of a flexible region of six amino acid residues to span the distance between residue 310 and 1, or the deletion of four residues from the C-terminus and the covalent linkage between residue 306 and 1. Both of these alternative designs were achieved by appropriate manipulation of regions of the *pyrB* gene (Yang & Schachman, 1993a, 1993b). For that research, only a single set of new N- and C-termini was introduced in the vicinity of residue 240, the site of a single cleavage by  $\alpha$ -chymotrypsin (Powers et al., 1993). The enzyme variants produced in this way were active and reasonably stable. However, they did not exhibit the allosteric properties characteristic of wild-type ATCase, and it seemed essential that constructs with different N- and C-termini be examined. The method designed for achieving that goal is described below, along with results for five different active ATCase variants containing circularly permuted c chains.

Since the demonstration by Peterson and Schachman (1991) that mutations in *pyrB* that introduced stop codons at positions within the helix comprising residues 285–305 resulted in pyrimidine auxotrophy, there has been much interest in the role of that helix in the folding and assembly of ATCase. In contrast, truncation of the polypeptide chain at positions 306 and beyond did not interfere with efficient expression in vivo of active, stable enzyme (Peterson & Schachman, 1991). In the light of these observations, the position selected for linking the original N- and C-terminal regions was 1–306. Moreover, it seemed desirable to start the polypeptide chain in the vicinity of the helical crossover from the C-terminal domain to the N-terminal domain. Hence, one of the variants was  $_{\text{cp}}\text{ATCase}_{\text{c281}}$ . With this genetic variant, the region of the polypeptide chain corresponding to the C-terminal helix was synthesized in vivo first, followed by the N-terminal domain, and finally the C-terminal domain. Clearly, efficient folding and assembly of stable, active ATCase from such a construct would be of interest in considerations of the in vivo formation of the enzyme in relation to the in vitro assembly from disorganized polypeptide chains.



**Fig. 1.** Ribbon diagram of the three-dimensional structure of a catalytic chain of ATCase<sub>wt</sub>. N-terminal domain (residues 1–134, in yellow) is on the left and C-terminal domain (residues 150–284, in blue) is on the right. The two domains are linked covalently by Helix 5 (in green) comprising residues 135–149. Helix 12 (shown in red), consisting of residues 285–305, crosses back from the C-terminal domain and interacts substantially with the N-terminal domain. Amino acid residues involved in the circular permutation experiments are labeled by their residue numbers. The approximate positions of residues 104 and 227 in the  $\beta$ -strands are indicated by \*. This diagram was generated with the program MOLSCRIPT (Kraulis, 1991) based on the crystal structure of the unliganded ATCase (Stevens et al., 1991) with the coordinates deposited in the Brookhaven Protein Data Bank (registry number 6AT1).

Because individual domains of multidomain proteins frequently act as independent folding units, it seemed worthwhile to examine variants in which the starting and stopping points of the polypeptide chain were within each of the domains of the c chains in ATCase. In effect, this required disrupting the chain connectivity in the segment encompassing residues 1–134 and 150–284. This was achieved, as illustrated schematically in Figure 2, with six constructs. In two of the circularly permuted chains,  $_{cp}c_{99}$  and  $_{cp}c_{122}$ , the N-terminal domain no longer was composed of a continuous polypeptide chain and part of the domain was synthesized early in the *in vivo* translation process and the remainder much later. In a similar way, the C-terminal domain was disrupted in terms of chain connectivity in the variants represented by  $_{cp}c_{181}$ ,  $_{cp}c_{222}$ , and  $_{cp}c_{227}$ . Taken together, the  $_{cp}pyrB$  variants designed to express ATCase containing the various  $_{cp}c$  chains serve as a test of the concept that folding involves a domain by domain process. Also, they provide evidence relevant to the question of whether folding *in vivo* is occurring contemporaneously with the elongation of the polypeptide chain.

Of the six constructs illustrated in Figure 2, five involve the placing of the new N- and C-termini in relatively flexible loops

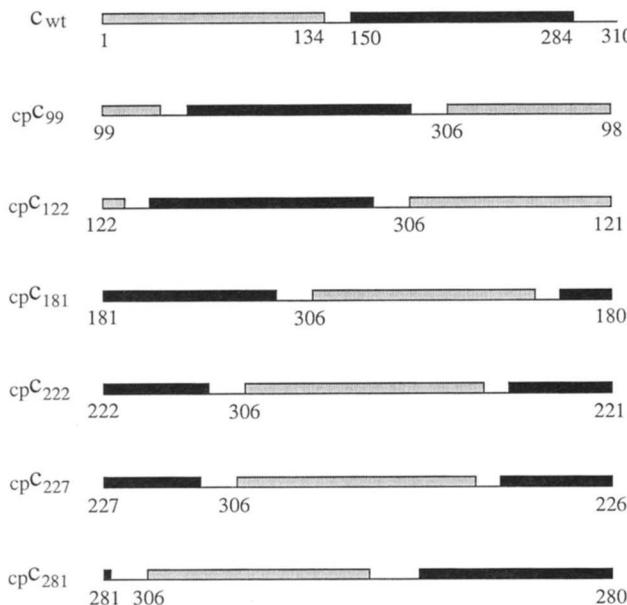
where the charged group would be in contact with solvent. One,  $_{cp}c_{227}$ , would have the N- and C-termini in a hydrophobic region if the tertiary structure of  $_{cp}ATCase_{c227}$  is similar to wild-type enzyme.

These differently permuted c chains not only vary considerably in terms of the temporal process of biosynthesis and chain connectivity, they also serve as different initial states for the folding and assembly process. If they yield final products that have similar or identical tertiary and quaternary structures, as indicated by the stability and activity of the holoenzymes as well as their allosteric properties, we would have evidence bearing on discussions about the kinetics and thermodynamics of folding of polypeptide chains.

## Results

### Construction of circularly permuted *pyrB* genes and plasmids that also contain wild-type *pyrI* gene

As shown in Figure 3 and described in Material and methods, a tandem arrangement of two *pyrB* genes was constructed by using a non-palindromic *TfiI* restriction site GAATC embedded



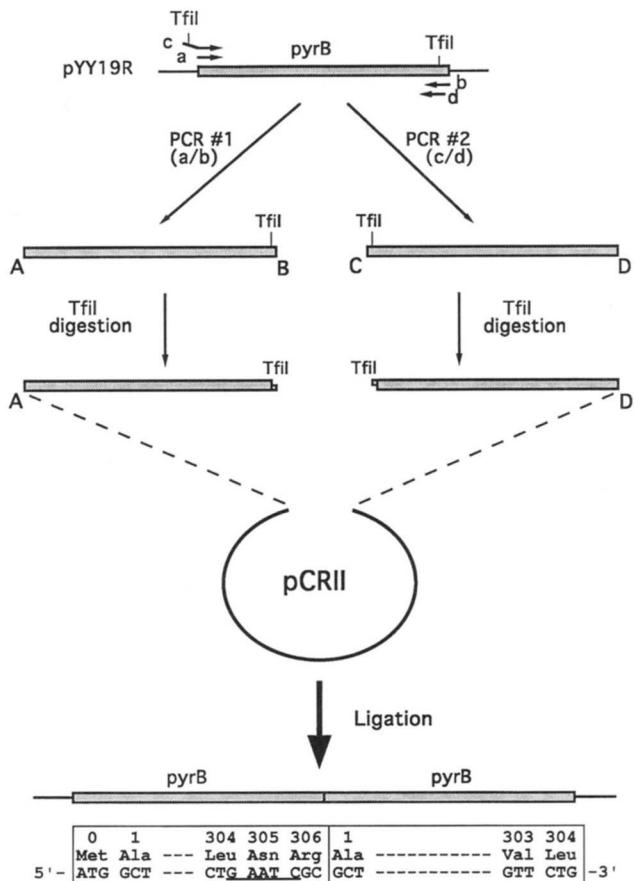
**Fig. 2.** Schematic diagram illustrating the amino acid sequence of the N-terminal domain (diagonal hatching) and C-terminal domain (solid black) in wild-type c chain ( $c_{wt}$ ) and circularly permuted c chains ( $cp_c$ ). All numbers refer to the positions in the wild-type sequence. The region between residues 134 and 150 is Helix 5, and that between 284 and 305 is Helix 12, which is followed by a flexible region of five residues terminating in the wild-type chain at residue 310. For the circularly permuted chains, the last four residues were removed through appropriate truncation of the wild-type gene.

in the codons for residues 304–306 and the TA cloning vector pCRII. With this tandem gene construct as a PCR template, various  $cp_pyrB$  genes were formed by incorporating an initiation codon ATG and a stop codon TAA into the forward and reverse primers, respectively. The start and end points of the  $cp_pyrB$  genes were controlled easily. This approach led to the expression of a series of  $cp_c$  chains containing Arg 306 linked covalently to Ala 1 and newly formed N- and C-termini distributed at various locations in the protein.

Because it was desirable to express the regulatory chains of ATCase along with the  $cp_c$  chains, the wild-type  $pyrI$  gene, including the preceding ribosome binding site, was amplified from plasmid PYY19R. After appropriate treatments, the  $cp_pyrB$  genes and wild-type  $pyrI$  genes were cloned into plasmid pT7-7 of the pT7 expression system (Tabor & Richardson, 1985). The resulting plasmids were then used to express ATCase molecules containing  $cp_c$  chains and wild-type r chains.

#### Formation of ATCase variants containing $cp_c$ chains

Five  $cp_pyrB$  genes were constructed to encode polypeptide chains with N- and C-termini in structural regions corresponding to surface loops and one  $cp_pyrB$  gene was designed to encode chains with termini that were buried in the wild-type enzyme. Nondenaturing PAGE patterns of the proteins in crude extracts (Fig. 4A) demonstrate that intact holoenzymes were formed from wild-type r chains and  $cp_c$  chains starting at residues Val 99, Val 122, Gly 181, Val 222, Met 227, and Pro 281. Somewhat lower mobilities were observed for these variants than for the wild-type enzyme; this decrease in mobility may be at-



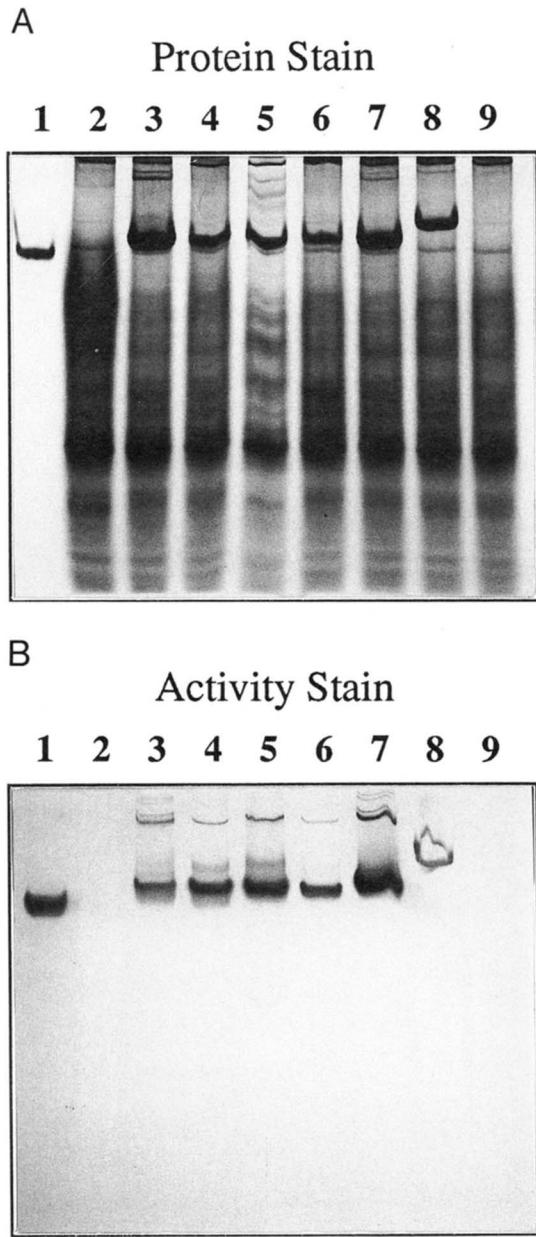
**Fig. 3.** Construction of tandem  $pyrB$  genes. Two separate PCR reactions (#1 and #2) with primer pairs **a/b** and **c/d**, respectively, were performed to amplify  $pyrB$  gene from pYY19R. The two products A-B and C-D were then subjected to Tfil restriction digestion. The resulting fragments A-Tfil and Tfil-D were isolated and ligated with TA cloning vector pCRII, yielding the recombinant plasmid carrying tandem  $pyrB$  genes shown at the bottom, along with some of the encoded amino acid residues. The Tfil restriction site is underlined.

tributable to the absence of the negatively charged Asp 307 in the  $cp_c$  chains. Staining of the native gels for enzyme activity (Bothwell, 1975) showed that all of the ATCase variants, with the exception of  $cp$ -ATCase<sub>c227</sub>, had activities comparable to that of wild-type ATCase (Fig. 4B).

An additional test of the in vivo assembly of ATCase variants containing  $cp_c$  chains was conducted with two  $cp_pyrB$  genes that encode chains with N- and C-termini within the interior of each of the two domains of the wild-type chains. As seen in Figure 4A, intact  $cp$ -ATCase<sub>c227</sub> was formed in good yield, but the enzyme activity (Fig. 4B) appeared to be substantially lower than that of the other variants. No intact enzyme could be detected by either the protein or the activity stain from the  $cp_pyrB$  gene encoding c chains starting at Met 104.

#### Catalytic and regulatory properties of ATCase variants containing $cp_c$ chains

Enzyme assays of the crude extracts of the five active ATCase variants containing  $cp_c$  chains with termini at surface loops demonstrated a sigmoidal dependence of activity on the con-



**Fig. 4.** In vivo formation of ATCase variants containing  $_{cp}c$  chains. Extracts of cells were subjected to electrophoresis on nondenaturing polyacrylamide gels that were then stained for protein with Coomassie blue stain (**A**) or for enzyme activity by an activity stain (**B**) according to the method of Bothwell (1975). The latter procedure involved 100 mM aspartate and 2 mM carbamoyl phosphate. Lane 1, ATCase<sub>wt</sub>; lane 2, crude extract of cells containing expression plasmid carrying *pyrI* gene only (negative control); lanes 3–8, cell extracts containing  $_{cp}ATCase_{c99}$ ,  $_{cp}ATCase_{c122}$ ,  $_{cp}ATCase_{c181}$ ,  $_{cp}ATCase_{c222}$ ,  $_{cp}ATCase_{c281}$ , and  $_{cp}ATCase_{c227}$ , respectively; lane 9, cell extracts of pGP1-2/H5533 harboring the recombinant plasmid encoding  $_{cp}ATCase_{c104}$ . Cell extracts were in 50 mM Tris-HCl, pH 7.5, containing 200 mM KCl, 0.2 mM EDTA, and 5 mM 2-mercaptoethanol.

centration of aspartate with Hill coefficients ( $n_H$ ) about 2, compared with 1.4 for ATCase<sub>wt</sub>. The concentration of aspartate corresponding to 0.5  $V_{max}$  varied from 4 to 17 mM.

Three variants,  $_{cp}ATCase_{c122}$ ,  $_{cp}ATCase_{c222}$ , and  $_{cp}ATCase_{c281}$ , were purified for further characterization. The similarity be-

tween  $_{cp}ATCase_{c122}$  and ATCase<sub>wt</sub> in terms of the sigmoidal dependence of enzyme activity on the concentration of aspartate is shown in Figure 5. This cooperativity is evident from the markedly curved Eadie plots for the two enzymes. Additional evidence demonstrating the similarity in the allosteric properties of the ATCase variants containing  $_{cp}c$  chains and the wild-type enzyme is illustrated in Figure 6, which shows the activation by ATP and the inhibition by CTP exhibited by  $_{cp}ATCase_{c222}$ .

The results for ATCase<sub>wt</sub> and the three ATCase variants containing  $_{cp}c_{122}$ ,  $_{cp}c_{222}$ , and  $_{cp}c_{281}$  chains are summarized in Table 1. Both the homotropic effects, as indicated by the values of  $n_H$ , and the heterotropic effects, revealed by the shifts in  $K_{0.5}$  and  $n_H$ , caused by the addition of ATP and CTP are very similar for the wild-type enzyme and the variants containing  $_{cp}c$  chains.

Because the three purified variants exhibited the homotropic and heterotropic effects characteristic of wild-type ATCase that have been interpreted earlier in terms of a ligand-promoted transition of the protein from a low-activity **T** state to a more active **R** conformation, it was of interest to determine whether a conformational change in the variants could be measured directly. Accordingly, measurements were made of the change in sedimentation coefficient caused by the binding of the bisubstrate analog PALA. As seen in Table 1, all of the variants showed a decrease in the sedimentation coefficient of about 2.5%, in excellent agreement with the value observed for the wild-type enzyme under similar conditions. Such a decrease in sedimentation coefficient corresponds to a ligand-promoted conformational change from a compact structure to a more swollen form (Gerhart & Schachman, 1968; Howlett & Schachman, 1977).

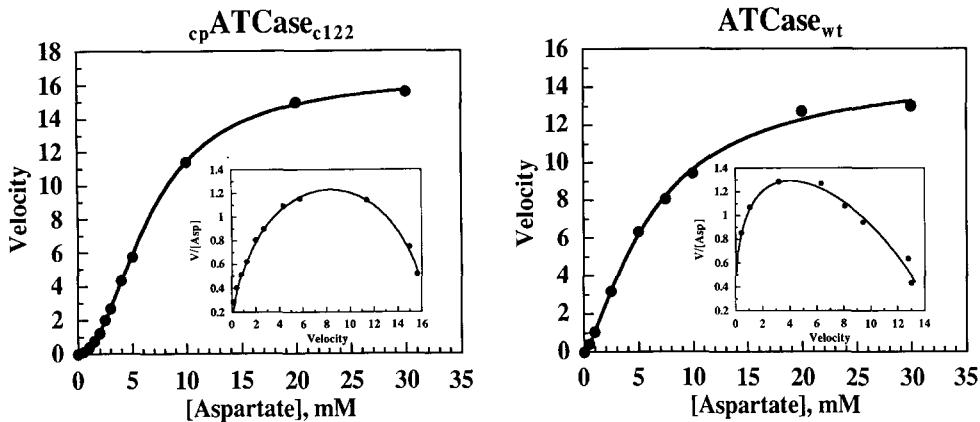
#### Thermal stability and solubility of ATCase variants

Preliminary studies by differential scanning microcalorimetry indicated a slight decrease in the stability of the ATCase variants containing  $_{cp}c$  chains (Table 1). Whereas ATCase<sub>wt</sub> exhibited two overlapping, relatively sharp transitions at 64 °C and 68 °C, a single endotherm with a  $T_m$  of 63 °C was observed for both  $_{cp}ATCase_{c122}$  and  $_{cp}ATCase_{c222}$ . In contrast,  $_{cp}ATCase_{c281}$  showed a minor transition at 49 °C and a major one at 54 °C.

All three purified ATCase variants were significantly less soluble than ATCase<sub>wt</sub> at neutral pH and low ionic strength. This decreased solubility proved a convenience in the purification process. It should be noted, however, that the variants were largely in soluble form in the crude cellular extract. Whether the decreased solubility is attributable to the net loss of one negatively charged residue, Asp 307, per chain as compared with wild-type chains has not as yet been determined.

#### Analysis of the sequences of the $_{cp}c$ chains

Complete nucleotide sequence determinations were performed on  $_{cp}pyrB_{122}$ ,  $_{cp}pyrB_{222}$ , and  $_{cp}pyrB_{281}$  and no secondary mutations were formed. With wild-type *pyrB*, the codon corresponding to amino acid residue 220 has been found to be either GCG or GTG, corresponding to Ala and Val, respectively. The DNA sequence analysis representing  $_{cp}c_{122}$  and  $_{cp}c_{281}$  chains showed that only Ala was at position 220, as expected from the starting plasmid pYY19R containing the wild-type *pyrB* gene. In contrast, the  $_{cp}pyrB$  gene encoding the  $_{cp}c_{222}$  chain contained



**Fig. 5.** Homotropic effect exhibited by  $_{cp}\text{ATCase}_{c122}$  and  $\text{ATCase}_{wt}$ . Assays were performed at 30 °C with saturating [ $^{14}\text{C}$ ]carbamoyl phosphate (5 mM) in 50 mM MOPS buffer, pH 7.0, containing 0.2 mM EDTA and 2 mM 2-mercaptoethanol. Enzyme activity is expressed as velocity in  $\mu\text{mol}$  of carbamoyl-L-aspartate formed per h per  $\mu\text{g}$  of C trimer as a function of the concentration of aspartate. Insets represent the Eadie plots of the same data.

the codon GTG encoding Val at position 220 instead of the expected GCG codon for Ala.

It was of interest to determine whether the  $_{cp}\text{c}$  chains retained the N-terminal Met residue encoded by the initiation codon ATG. After their isolation from the purified holoenzymes by reversed-phase HPLC, the  $_{cp}\text{c}$  chains were subjected to electron spray mass spectrometer analysis. The calculated molecular weights for  $_{cp}\text{c}_{122}$  and  $_{cp}\text{c}_{281}$  chains are 33,856.7 or 33,987.9 if the non-native N-terminal Met residue is included. Results of the mass spectrometric analyses of molecular weights and the relative abundance of the species (numbers in parentheses) were  $33,984.6 \pm 1.3$  (>90%) for  $_{cp}\text{c}_{122}$  and  $33,854.5 \pm 2.4$  (>90%) for  $_{cp}\text{c}_{281}$  chains. Thus, more than 90% of the  $_{cp}\text{c}_{122}$  and less than 10% of the  $_{cp}\text{c}_{281}$  chains contain the non-native N-terminal Met residue. For the  $_{cp}\text{c}_{222}$  chains that have Val at position 220 instead of Ala, the calculated molecular weight is 33,884.8 or 34,016.0 if the non-native N-terminal Met residue is included. Experimentally, the value was  $34,011.9 \pm 2.4$  (about 60%), indicating that about 60% of the  $_{cp}\text{c}_{222}$  chains contain the non-native N-terminal Met residue.

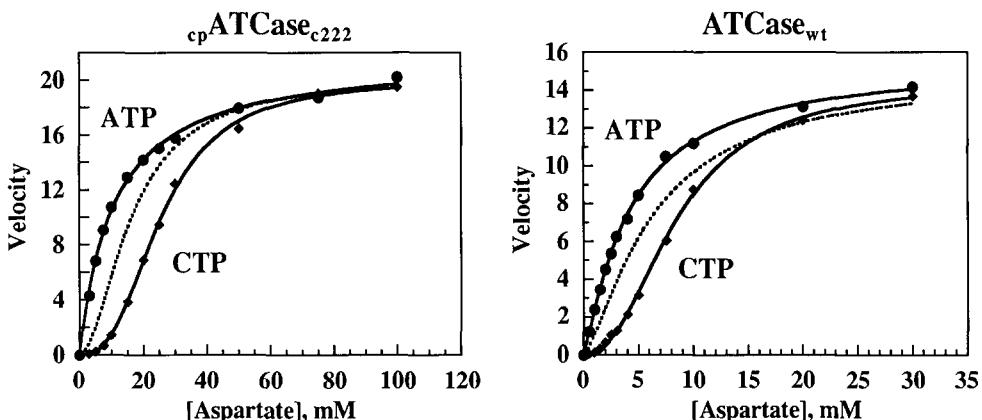
The molecular weight determinations for  $_{cp}\text{c}_{227}$  chains yielded  $33,721.0 \pm 3.1$  (about 50%) and  $33,852.6 \pm 3.7$  (about 50%), in comparison with the calculated values of 33,725.5 and 33,856.7 if the N-terminal residue is included. Thus, about 50% of the  $_{cp}\text{c}_{227}$  chains lost the native Met 227 residue.

Each of the four  $_{cp}\text{c}$  chains was subjected to four cycles of Edman degradation analysis to determine the sequence of amino acids at the N-termini of the  $_{cp}\text{c}$  chains. The data showed that the sequences at the N-terminus of each chain corresponded precisely with that expected from the DNA constructs and the molecular weight determinations from mass spectrometry.

## Discussion

*Diverse ATCase molecules comprising  $_{cp}\text{c}$  chains exhibit the same allosteric transition from the T to the R conformation as observed with the wild-type enzyme*

Studies on a variety of proteins, such as bovine pancreatic trypsin inhibitor, phosphoribosyl anthranilate isomerase, dihy-



**Fig. 6.** Heterotropic effects exhibited by  $_{cp}\text{ATCase}_{c222}$  and  $\text{ATCase}_{wt}$ . Assays were performed at 30 °C with saturating [ $^{14}\text{C}$ ]carbamoyl phosphate (5 mM) in 50 mM MOPS buffer, pH 7.0, containing 0.2 mM EDTA and 2 mM 2-mercaptoethanol in the absence of nucleotides (dotted line), in the presence of 2 mM ATP (●), and in the presence of 0.5 mM CTP (◆). Enzyme activity is expressed as velocity in  $\mu\text{mol}$  of carbamoyl-L-aspartate formed per h per  $\mu\text{g}$  of C trimer as a function of the concentration of aspartate.

**Table 1.** Properties of ATCase variants

| Enzyme species               | $V_{max}^a$ |    |     | $K_{0.5}^a$ |    |     | $n_H^a$ |     |     | $\Delta s/s^b$<br>(%) | $T_m^c$<br>(°C) |
|------------------------------|-------------|----|-----|-------------|----|-----|---------|-----|-----|-----------------------|-----------------|
|                              | ATP         | —  | CTP | ATP         | —  | CTP | ATP     | —   | CTP |                       |                 |
| ATCase <sub>wt</sub>         | 16          | 15 | 14  | 4           | 6  | 9   | 1.2     | 1.4 | 2.2 | -2.6                  | 64, 68          |
| $c_p$ ATCase <sub>c122</sub> | 18          | 16 | 18  | 4           | 7  | 10  | 1.3     | 2.0 | 2.9 | -2.6                  | 63              |
| $c_p$ ATCase <sub>c222</sub> | 21          | 21 | 20  | 11          | 17 | 26  | 1.0     | 1.8 | 2.6 | -2.2                  | 63              |
| $c_p$ ATCase <sub>c281</sub> | 20          | 16 | 16  | 2           | 4  | 6   | 1.4     | 2.1 | 3.0 | -2.5                  | 49, 54          |

<sup>a</sup> Kinetics were performed at 30 °C in 50 mM MOPS buffer, pH 7.0, containing 0.2 mM EDTA and 2 mM 2-mercaptoethanol with saturating carbamoyl phosphate (5 mM) and varying concentrations of aspartate. Values of  $V_{max}$  are in units of mmole carbamoyl aspartate per mg of C trimer per h. The concentration of aspartate corresponding to 0.5  $V_{max}$  ( $K_{0.5}$ ) is expressed as mM. The concentrations of the allosteric effectors ATP and CTP were 2 mM and 0.5 mM, respectively. Hill coefficients,  $n_H$ , were evaluated from a fit of the assay data according to the Hill equation using the program Kaleidagraph (Synergy Software).

<sup>b</sup> Difference sedimentation velocity measurements of the enzymes in the absence and presence of PALA (250 μM) were made with a Beckman model XL-A analytical ultracentrifuge at 22 °C at concentrations of 1 mg/mL in 50 mM Tris, pH 8.0, containing 0.2 mM EDTA and 2 mM 2-mercaptoethanol.

<sup>c</sup> Values of  $T_m$  were measured using a Microcal MC-2 calorimeter. Samples contained proteins at 2–3 mg/mL in 40 mM potassium borate at pH 9.0 containing 0.2 mM EDTA and 5 mM 2-mercaptoethanol.

drofolate reductase, lysozyme, interleukin 1β, ornithine decarboxylase, ATCase, ribonuclease T1, β-glucanase, α-spectrin SH3 domain, interleukin 4, and glyceraldehyde 3-phosphate dehydrogenase, have now demonstrated that circularly permuted polypeptide chains can form stable, active molecules comprising widely different topologies. In all of the proteins examined thus far, the N- and C-termini of the wild-type molecules were in reasonably close proximity. Of these proteins, ATCase is the most complex with regard to both structure and function. Moreover, as shown above (Fig. 1; Kinemage 3), the newly formed termini in the circularly permuted chains can be located in diverse regions of each of the two structural domains of the c chains.

In a previous study, two  $c_p$ pyrB genes were constructed and expressed in vivo to produce  $c_p$ c chains that associated into active trimers (Yang & Schachman, 1993a). Both chains had an N-terminus at residue 235 of the wild-type sequence and ended at residue 242, with an overlap of eight residues at the C-terminus. In one protein, the regions near the wild-type N- and C-termini were linked from Arg 306 to Ala 1 because residue 306 is closer in space to residue 1 than is residue 310, the C-terminus of the wild-type chain. Instead of a deletion of four residues, the other  $c_p$ c chain had an insertion involving a six-amino acid linker from Leu 310 to Ala 1. The trimers containing these  $c_p$ c chains both associated with R dimers to form ATCase-like holoenzymes that were catalytically active, but devoid of homotropic and heterotropic effects. Sedimentation velocity experiments on these  $c_p$ ATCase<sub>235</sub> molecules indicated that they both were in the swollen, relaxed (R) conformation even in the absence of substrates, and therefore, in contrast to ATCase<sub>wt</sub>, no allosteric properties were observed (Yang & Schachman, 1993a). On the basis of these studies and related experiments (Newell & Schachman, 1990) on mutant forms of ATCase, Yang and Schachman (1993a) concluded that the loss of allosteric properties was not attributable inherently to the circular permutation of the c chains, but rather to the location of the new, flexible N- and C-termini in regions that destabilized the compact (T) conformation of the enzyme. This conclusion is supported by the evidence presented here (Figs. 5, 6; Table 1), demonstrating that different ATCase molecules con-

taining  $c_p$ c chains exhibit the homotropic and heterotropic effects characteristic of wild-type ATCase.

The values of  $V_{max}$ ,  $K_{0.5}$ , and  $n_H$  for  $c_p$ ATCase<sub>c122</sub>,  $c_p$ ATCase<sub>c222</sub>, and  $c_p$ ATCase<sub>c281</sub>, in the absence or presence of ATP or CTP, are remarkably similar to those determined for the wild-type enzyme under similar experimental conditions (Table 1). All of the data are consistent with a ligand-promoted transition of the molecules from the T conformation to the R state (Schachman, 1988). Direct evidence for this change in quaternary structure was first revealed for the wild-type enzyme by the 3% decrease in the sedimentation coefficient accompanying the ligand-promoted T to R transition (Gerhart & Schachman, 1968). As seen in Table 1, the addition of the bisubstrate analogue PALA leads to a decrease in sedimentation coefficient (Table 1) for  $c_p$ ATCase<sub>c122</sub>,  $c_p$ ATCase<sub>c222</sub>, and  $c_p$ ATCase<sub>c281</sub>, which is very similar to that observed for the wild-type enzyme under comparable conditions ( $\Delta s/s$  is about -2.5%). This sensitive test for the change in quaternary structure, coupled with the extensive kinetic data from enzyme assays, indicates that the enzymes containing the circularly permuted chains must have very similar structures to that of wild-type ATCase and that the ligand-promoted changes represent similar allosteric transitions.<sup>2</sup>

Detailed structural information for wild-type ATCase and some mutant forms is now available for both the T and R conformations (Lipscomb, 1994). One prominent difference between the two states is at the loop region comprising residues 230–245 of the c chain. In the T conformation, the side chains in this loop region of each c chain in one C trimer interact with the side chains of the c chains in the apposing C trimer (Kim et al., 1987). When the enzyme is in the R conformation, the two C trimers are separated from each other by about 12 Å along

<sup>2</sup>  $c_p$ ATCase<sub>c281</sub> has been crystallized and diffraction data have been obtained to about 2.8 Å resolution. The results at a preliminary stage of refinement indicate that there is no significant deviation of backbone structure from that of the wild-type enzyme in the T conformation except for the regions of the polypeptide chain at the linkage between the wild-type N- and C-terminal regions and the newly formed N- and C-termini (L. Gonzalez, P. Zhang, H.K. Schachman, & T. Alber, unpubl.).

the threefold symmetry axis, and the trimers are rotated slightly relative to one another. As a consequence, this interaction between the two C trimers is substantially disrupted (Ke et al., 1988). The importance of this loop region in determining whether the **T** or **R** conformation is dominant has been demonstrated with various mutant forms involving amino acid replacements in that segment of the polypeptide chain. When the wild-type enzyme is in the **T** conformation, Glu 239 of chains in one C trimer interact with Lys 164 and Tyr 165 in the apposing c chains in the other trimer. In contrast, Glu 239 interacts with Lys 164 of the same c chain when the enzyme is in the **R** conformation. An ATCase mutant in which Lys 164 was replaced by Glu was shown to be in the **R** conformation even in the absence of ligands (Newell & Schachman, 1990), suggesting that the conformational equilibrium between the **T** and **R** states is highly sensitive to changes in the loop region. It is not surprising, therefore, that cleavage of the polypeptide chain in this loop region resulted in enzymes that were in the **R** conformation even in the absence of any ligand and, as a result, were devoid of both homotropic and heterotropic properties (Powers et al., 1993; Yang & Schachman, 1993a, 1993b). As shown in this study, however, it is striking that cleavage of the c chains in loops on the surface of the holoenzyme, as contrasted to the internal region between the two C trimers, does not seem to have a significant impact on either the catalytic or regulatory properties of the enzyme.

#### *Thermal stability of ATCase molecules containing <sub>cp</sub>c chains*

It is interesting that <sub>cp</sub>ATCase<sub>c122</sub> and <sub>cp</sub>ATCase<sub>c222</sub> both exhibited only slightly decreased  $T_m$  values compared with ATCase<sub>wt</sub> (Table 1). In contrast, <sub>cp</sub>ATCase<sub>c281</sub> is significantly less stable. This marked decrease in thermal stability may be attributable to a reduced domain-domain interaction in <sub>cp</sub>c<sub>281</sub> chains compared with that in wild-type c chains. In c<sub>wt</sub> chains, Helix 12 crosses over from the C-terminal domain and becomes partially buried in the N-terminal domain between Helix 1 and residues in other parts of the domain. Helix 12 has been shown to be important for the folding of the c chains and the assembly of the intact enzyme (Peterson & Schachman, 1991). In <sub>cp</sub>c<sub>281</sub> chains, Helix 12 is no longer covalently linked to the C-terminal domain; instead, there is a free N-terminus at residue 281 and its interaction with that domain would be only through the short  $\beta$  strand (residues 282–284) preceding the helix (assuming that  $\beta$  strand still exists in <sub>cp</sub>ATCase<sub>281</sub>). Thus, the cleavage of the chains at residue 281 could weaken the interaction between the domains, resulting in a decreased thermal stability of <sub>cp</sub>ATCase<sub>281</sub>. Further interpretation of these results must await structural determinations to evaluate slight movements in the vicinity of the linkage and the breaks in the chains.

#### *N-terminal methionine processing of the <sub>cp</sub>c chains*

The availability of different ATCase molecules comprising circularly permuted c chains with various N-termini provides an opportunity to examine the suggestion of Hirel et al. (1989) that the extent of N-terminal methionine excision from *E. coli* proteins is governed by the side chain of the amino acid following the N-terminal methionine. This influence of the next amino acid residue is attributable presumably to the effect on the binding

of the N-terminal peptide to the enzyme methionylaminopeptidase, which is responsible for the methionine processing. With Val, Gly, or Pro at the second position, the extent of processing, according to Hirel et al. (1989), would be 84%, 97%, and 88%, respectively. Thus, mass spectromic and Edman degradation analyses of the ATCase molecules containing <sub>cp</sub>c chains designed with N-termini at Val 99, Val 122, Gly 181, Val 222, and Pro 281 would be expected to show only small amounts of the non-native, N-terminal methionine residue. In the <sub>cp</sub>c chains for <sub>cp</sub>c<sub>281</sub>, the observed removal of the N-terminal methionine is greater than 90%, in good agreement with the value of 88% predicted by Hirel et al. (1989). Less than 10% methionine processing was observed with <sub>cp</sub>c<sub>122</sub>, in contrast to the estimated amount of 84%. This discrepancy can be attributed to Pro 123 because it has been found that a proline residue following the residue next to the N-terminal amino acid is inhibitory to methionine removal (Ben-Bassat et al., 1987; Hirel et al., 1989). For <sub>cp</sub>c<sub>222</sub>, the amount of methionine processing observed was about 40%, whereas the value estimated by Hirel et al. (1989) was 84%. As yet, there is no explanation for this discrepancy; it does seem likely that factors other than those discussed by Hirel et al. (1989) may affect the processing of N-terminal methionine in *E. coli* (Dalboge et al., 1990). It is of significance that the presence of non-native N-terminal methionine residues on some of the <sub>cp</sub>c chains does not appear to affect the behavior of the ATCase variant containing those <sub>cp</sub>c chains. Additional studies of this type are clearly warranted.

Considerations of N-terminal methionine processing are of particular interest with regard to <sub>cp</sub>ATCase<sub>c227</sub> because the N-terminus in those <sub>cp</sub>c chains is Met 227. Moreover, in the wild-type structure, Met 227 is located toward the end of  $\beta$  strand 9 and buried in the interior of the C-terminal domain. <sub>cp</sub>ATCase<sub>c227</sub> is formed in vivo with high efficiency and is low in activity. At present, we have no information regarding the degree to which the newly created N- and C-termini are exposed or buried in the interior. Nor is there convincing evidence as to whether N-terminal methionine processing occurs co-translationally or only after the polypeptide chain is folded into secondary and tertiary structures. Hence, speculation as to the expected extent of N-terminal methionine processing is extremely hazardous. It is relevant for further considerations to note that the observed extent of N-terminal processing in <sub>cp</sub>ATCase<sub>c227</sub> was only about 50%, compared with 90% according to Hirel et al. (1989) or Dalboge et al. (1990).

#### *Location of N- and C-termini and the role of domains in the folding of polypeptide chains*

Most of the N- and C-termini in the <sub>cp</sub>c chains described above were located in flexible loops separating elements of secondary structure. Such <sub>cp</sub>c chains folded and assembled in vivo with R dimers to give stable ATCase-like holoenzymes.<sup>3</sup> This was also observed with <sub>cp</sub>ATCase<sub>c227</sub>, which represented a striking exception in that the new termini were not in a flexible loop. In the wild-type enzyme, Met 227 is part of a  $\beta$  strand that is largely buried in the C-terminal domain. Creating N- and C-termini in

<sup>3</sup> Additional ATCase molecules containing circularly permuted c chains were formed with N-termini at positions 131 and 153 on either end of Helix 5 as well as at position 236 of the wild-type sequence (Y.R. Yang & H.K. Schachman, unpubl.).

what was a continuous polypeptide segment of a  $\beta$  strand clearly did not interfere with the folding of the chains and assembly of trimers and holoenzyme in good yield. Some local structural change must have occurred, however, because  $_{cp}\text{ATCase}_{c227}$  had a markedly decreased enzyme activity. Although neither Tyr 226 nor Met 227 have been implicated in the catalytic mechanism, it has been shown that amino acid replacements of the neighboring residues Arg 229 and Gln 231 cause changes in activity that have been attributed to alterations in the binding of substrates. With the availability of  $_{cp}\text{ATCase}_{c227}$  as a purified enzyme, it should be possible to determine whether there are changes in the local secondary and tertiary structures resulting from the insertion of an ion pair in a hydrophobic region of the molecule. Whereas a good yield of enzyme was obtained with  $_{cp}\text{c}_{227}$  chains, an equivalent effort with  $_{cp}\text{c}_{104}$  yielded no assembled enzyme. A  $_{cp}\text{pyrB}$  construct aimed at placing N- and C-termini at position 144 in the middle of an  $\alpha$ -helical region yielded no assembled ATCase. Although it is clear that N- and C-termini can be introduced throughout the polypeptide chains in regions corresponding to surface loops in the tertiary structure, it is not yet clear what restrictions limit the formation of ATCase with  $_{cp}\text{c}$  chains having termini in elements of secondary structure such as helices or  $\beta$  strands. Determining where N- and C-termini cannot be introduced into circularly permuted chains without interfering with the folding and assembly process will require many more constructs than those described here. In this regard, the development of a technique involving the random formation of a collection of circularly permuted chains followed by an *in vivo* selection of those chains competent to form stable proteins would be of great value in determining the restrictions for placing the new termini (R. Graf & H.K. Schachman, unpubl.).

In many proteins, the polypeptide chain is segmented into discrete domains comprising about 100 amino acid residues, and there is only a single, small region of the chain connecting the domains covalently. For such proteins, the folding of the chains may occur stepwise as the nascent polypeptide chain is being synthesized (Chantrenne, 1961; Wetlaufer & Ristow, 1973; Peters & Davidson, 1982; Fedorov et al., 1992). A co-translational folding process is extremely plausible for large multidomain proteins whose individual domains are formed from a continuous stretch of amino acid residues, with only a single polypeptide segment linking the domains (Jacobson et al., 1994). Indeed, the individual domains of such proteins can form independently of each other. Folding of polypeptide chains in these proteins may proceed in a domain-by-domain fashion; each domain may be formed with the native tertiary structure as the polypeptide chain is being synthesized without requiring the completion of the entire chain. This type of co-translational folding of the polypeptide chain is not likely to be applicable for the two domains in a c chain such as that in ATCase (Fig. 1), because the N-terminal domain cannot achieve a stable folded structure until Helix 12, comprising residues 285–305, is synthesized. By truncating the gene encoding the c chains at various positions, Peterson and Schachman (1991) demonstrated that ATCase cannot be assembled *in vivo* if part or all of Helix 12 is not synthesized. The *in vivo* formation of  $_{cp}\text{ATCase}_{c281}$  demonstrates that it does not matter whether the region of the polypeptide chain corresponding to Helix 12 is attached covalently to the remainder of the chain at its N-terminus (i.e., residue 281 linked to 280) as in the wild-type enzyme or covalently at its C-terminus

(i.e., residue 306 linked to residue 1). Nor does it matter for the assembly of the enzyme whether that polypeptide segment is synthesized at the end of the translational process as with ATCase<sub>wt</sub> or at the initiation of translation. The findings for all six variants of ATCase containing  $_{cp}\text{c}$  chains, which are summarized in Figure 2, show that chain connectivity throughout each of the two domains is not a requirement for efficient folding of the chains and assembly of the holoenzyme. For  $_{cp}\text{c}_{99}$  and  $_{cp}\text{c}_{122}$  chains, the continuity of the polypeptide chain representing the N-terminal domain is disrupted in different positions. Similarly, the connectivity in the region of the chain representing the C-terminal domain is disrupted in  $_{cp}\text{c}_{181}$ ,  $_{cp}\text{c}_{222}$ ,  $_{cp}\text{c}_{227}$ , and  $_{cp}\text{c}_{281}$  chains. Apparently the different segments, on the basis of their amino acid sequences, have sufficient propensity to form secondary structural elements that then associate to form domains despite the absence of a continuous, intact polypeptide chain.

As yet there has been no success in isolating soluble C trimers from the various ATCase species containing the  $_{cp}\text{c}$  chains described above. This is in marked contrast with the trimers containing  $_{cp}\text{c}$  chains described by Yang and Schachman (1993a). When this solubility problem is overcome, physical chemical studies of the folding of a variety of circularly permuted polypeptide chains and their assembly into active trimers should yield valuable insights about the folding process and the relationship between the *in vivo* process and that involved in the reactivation of denatured proteins.

## Materials and methods

### *Materials and general procedures*

Oligonucleotides were either purchased from Operon, Inc. or synthesized on an Applied Biosystems 380B synthesizer and purified by thin layer chromatography. Sequences of the oligonucleotides are shown in the Electronic Appendix. Taq DNA polymerase was from United States Biochemical Corp. and Vent DNA polymerase was from New England BioLabs. All PCR reactions were performed following the standard protocol (Sambrook et al., 1989). DNA sequencing was performed with Sequenase from United States Biochemical Corp.

TA Cloning kit including the pCRII vector and the *E. coli* strain INVαF' was purchased from Invitrogen. SeaPlaque low melting agarose was from FMC Corp. Restriction enzymes and T4 DNA ligase were from Pharmacia, United States Biochemical Corp., or New England BioLabs.

After restriction digestion, DNA fragments were separated on low-melting agarose gel. Gel slices containing the appropriate fragments were used directly in subsequent ligation reactions.

### *Construction of tandem pyrB genes, recombinant plasmids for expression of ATCase variants, and $_{cp}\text{pyrB}$ genes*

As illustrated in Figure 3, *pyrB* from pYY19R, which contains the *pyrBI* operon, was amplified by the PCR technique, using Taq DNA polymerase. Two primer pairs, **a/b** and **c/d**, were used in two separate reactions. The resulting 0.9-kb PCR products, A-B and C-D, were digested with *T*fil, yielding A-*T*fil and *T*fil-D. These two fragments, together with the 3.9-kb linear vector pCRII, were then subjected to a single-step ligation followed

by transformation into a recombination negative *E. coli* strain, INV $\alpha$ F'. The *lacZ* gene carried by pCRII and the genotype of the INV $\alpha$ F' strain allowed blue/white screening of the recombinant plasmids. Among the white colonies screened, more than 50% were found to contain the recombinant plasmids carrying tandem *pyrB* genes.

Wild-type *pyrI* gene, including the ribosome binding site preceding it, was amplified from pYY19R. *Bam*H I and *Cla* I recognition sequences were incorporated into the forward and reverse PCR primers, respectively. The resulting 0.5-kb PCR fragment was then digested with *Bam*H I and *Cla* I and inserted into pT7-7 at *Bam*H I and *Cla* I sites.

Circularly permuted *pyrB* genes were constructed from the tandem *pyrB* genes using PCR. The initiation codon ATG and the stop codon TAA were added in the forward and reverse PCR primers, respectively. These primers were also designed to include the *Nde* I recognition sequence CATATG, which contains the initiation codon and *Bam*H I site GGATCC, respectively. Circularly permuted *pyrB* genes flanked by the *Nde* I and *Bam*H I sites were then cloned directionally into the above plasmid carrying the wild-type *pyrI* gene.

#### Expression and purification of ATCase variants

The two-plasmid system described by Tabor and Richardson (1985) was used to express the ATCase variants. Recombinant plasmids containing  $_{cp}$ *pyrB* and wild-type *pyrI* genes were introduced into *E. coli* HS533 cells harboring pGP1-2. Cells carrying the two plasmids were grown in LB medium containing ampicillin and kanamycin (50 mg/L each) at 30 °C to OD<sub>590</sub> = 0.4–1.0. The temperature was then shifted to 42 °C for 30 min to induce protein synthesis, followed by continued incubation at 37 °C for 2–3 h.

Purification of  $_{cp}$ ATCase<sub>c122</sub>,  $_{cp}$ ATCase<sub>c222</sub>, and  $_{cp}$ ATCase<sub>c281</sub> was achieved by the following procedure. After lysis of the cells, debris was removed by centrifugation at 30,000 × g for 30 min, and the extracts in 100 mM Tris·HCl, pH 8.5, containing 0.2 mM EDTA and 5 mM 2-mercaptoethanol, were chromatographed on a Q-Sepharose fast-flow column. Gradient elutions were performed with the same Tris buffer containing 0.1–0.4 M KCl. The crude preparations were then dialyzed against 50 mM MOPS, pH 7.0, containing 0.2 mM EDTA and 5 mM 2-mercaptoethanol. All three ATCase variants appeared as white, shining precipitates after dialysis, which were then redissolved in 100 mM Tris, pH 8.5, containing 0.2 mM EDTA and 5 mM 2-mercaptoethanol. The enzymes were further purified by chromatography on a Sephadryl S-300 column when necessary.

#### Other procedures

Differential scanning microcalorimetry was performed using a Microcal MC-2 calorimeter as described by Peterson and Schachman (1991). Samples contained protein at 2–3 mg/mL in 40 mM potassium borate buffer, pH 9.0, containing 0.2 mM EDTA and 5 mM 2-mercaptoethanol.

Nondenaturing PAGE was performed by the method of Jovin et al. (1964) using 7% polyacrylamide gels. The gels were stained for total protein using Commassie brilliant blue G250 or, for ATCase activity by the method of Bothwell (1975), using 2 mM carbamoyl phosphate and 100 mM aspartate.

The procedure of Davies et al. (1970) was used to measure the activities of the enzymes with <sup>14</sup>C-labeled carbamoyl phosphate (New England Nuclear). Assays were performed at 30 °C in 50 mM MOPS, pH 7.0, containing 0.2 mM EDTA and 2 mM 2-mercaptoethanol with saturating carbamoyl phosphate (5 mM) and variable concentration of aspartate. When examining the effects of nucleotides (2 mM ATP or 0.5 mM CTP), MgOAc was added to the MOPS buffer to 3 mM. Assay data were fitted to Hill equation, using the program Kaleidograph (Synergy Software).

A Beckman model XL-A analytical ultracentrifuge equipped with absorption optics was used to measure the change in sedimentation coefficient ( $\Delta s/s$ ) of the enzyme caused by the addition of PALA. Protein concentrations were 1 mg/mL in 50 mM Tris, pH 8.0, containing 0.2 mM EDTA and 2 mM 2-mercaptoethanol. The concentration of PALA was 250 μM and the calculations of  $\Delta s/s$  were performed as described by Heath (1994).

Molecular masses of the peptide chains were determined by David King, using a Hewlett Packard 5989 A electrospray ionization mass spectrometer. N-terminal sequence analysis was performed by Charlene Zhou. Samples were chromatographed on a Vydac C18 HPLC column prior to the analyses.

#### Acknowledgments

This research was supported by National Institute of General Medical Sciences research grant GM 12159 and by National Science Foundation grant DMB 9020605. We are indebted to Ying R. Yang for her help and valuable suggestions during the course of this work. We thank Dr. David King of the Howard Hughes Medical Institute, University of California at Berkeley, for determining the molecular weights of the polypeptide chains by mass spectrometry and Charlene Zhou for performing the N-terminal sequence analysis. We thank Yu-Chin Fang and Alex Santoso for their help in certain phases of this work.

#### References

- Baker D, Agard DA. 1994. Kinetics versus thermodynamics in protein folding. *Biochemistry* 33:7505–7509.
- Ben-Bassat A, Bauer K, Chang SY, Myambo K, Boosman A, Chang S. 1987. Processing of the initiation methionine from proteins: Properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *J Bacteriol* 169:51–75.
- Bergman LW, Kuehl WM. 1979. Formation of an intrachain disulfide bond on nascent immunoglobulin light chains. *J Biol Chem* 254:8869–8876.
- Bothwell MA. 1975. Kinetics and thermodynamics of subunit interactions in aspartate transcarbamylase [thesis]. University of California, Berkeley.
- Bowles DJ, Marcus SE, Pappin DJC, Findlay JB, Eliopoulos E, Maycox PR, Burgess J. 1986. Posttranslational processing of concanavalin A precursors in jackbean cotyledons. *J Cell Biol* 102:1284–1297.
- Buchwalder A, Szadkowski H, Kirschner K. 1992. A fully active variant of dihydrofolate reductase with a circularly permuted sequence. *Biochemistry* 31:1621–1630.
- Chantrenne H. 1961. *The biosynthesis of proteins*. Oxford/New York: Pergamon Press.
- Creighton TE. 1990. Protein folding. *Biochem J* 270:1–16.
- Dalbøge H, Bayne S, Pedersen J. 1990. In vivo processing of N-terminal methionine in *E. coli*. *Febs Lett* 266:1–3.
- Davies GE, Vanaman TC, Stark GR. 1970. Aspartate transcarbamylase. Stereospecific restrictions on the binding site for L-aspartate. *J Biol Chem* 245:1175–1179.
- Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* 29:7133–7155.
- Fedorov AN, Friguet B, Djavadi-Ohanian L, Alakhov YB, Goldberg ME. 1992. Folding on the ribosome of *Escherichia coli* tryptophan synthase β subunit nascent chains probed with a conformation-dependent monoclonal antibody. *J Mol Biol* 228:351–358.
- Garrett JB, Mullins LS, Raushel FM. 1996. Are turns required for the folding of ribonuclease T1? *Protein Sci* 5:204–211.
- Gerhart JC, Schachman HK. 1968. Allosteric interactions in aspartate trans-

- carbamylase. II. Evidence for different conformational states of the protein in the presence and absence of specific ligands. *Biochemistry* 7: 538–552.
- Goldenberg DP, Creighton TE. 1983. Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J Mol Biol* 165:407–413.
- Hahn M, Piotrzk K, Boriss R, Heinemann U. 1994. Native-like in vivo folding of a circularly permuted jellyroll protein shown by crystal structure analysis. *Proc Natl Acad Sci USA* 91:10417–10421.
- Heath RH. 1994. The global conformation determines the functional properties of the ligand sites of aspartate transcarbamoylase [thesis]. Berkeley, California: University of California at Berkeley.
- Hirel PH, Schmitter MJ, Dessen P, Fayat G, Blanquet S. 1989. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci USA* 86:8247–8251.
- Honzatko RB, Lipscomb WN. 1982. Interactions of phosphate ligands with *Escherichia coli* aspartate carbamoyltransferase in the crystalline state. *J Mol Biol* 160:265–286.
- Horlick RA, George HJ, Cooke GM, Tritch RJ, Newton RC, Dwivedi A, Lischwe M, Salemme FR, Weber PC, Horuk R. 1992. Permuteins of interleukin 1 $\beta$ —A simplified approach for the construction of permuted proteins having new termini. *Protein Eng* 5:427–431.
- Howlett GJ, Schachman HK. 1977. Allosteric regulation of aspartate transcarbamoylase. Changes in the sedimentation coefficient promoted by the bisubstrate analogue *N*-(phosphonacetyl)-L-aspartate. *Biochemistry* 16: 5077–5083.
- Jacobson RH, Zhang XJ, DuBose RF, Matthews BW. 1994. Three-dimensional structure of  $\beta$ -galactosidase from *E. coli*. *Nature* 369:761–766.
- Jaenicke R. 1991. Protein folding: Local structures, domains, subunits, and assemblies. *Biochemistry* 30:3147–3161.
- Jovin T, Chrambach A, Naughton MA. 1964. An apparatus for preparative temperature-regulated polyacrylamide gel electrophoresis. *Anal Biochem* 9:351–369.
- Ke H, Lipscomb WN, Cho Y, Honzatko RB. 1988. Complex of *N*-phosphonacetyl-L-aspartate with aspartate carbamoyltransferase. X-ray refinement, analysis of conformational changes and catalytic and allosteric mechanisms. *J Mol Biol* 204:725–747.
- Kim KH, Pan Z, Honzatko RB, Ke HM, Lipscomb WN. 1987. Structural asymmetry in the CTP-ligated form of aspartate carbamoyltransferase from *Escherichia coli*. *J Mol Biol* 196:853–875.
- Kim PS, Baldwin RL. 1990. Intermediates in the folding reactions of small proteins. *Annu Rev Biochem* 59:631–660.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Kreitman RJ, Puri RK, McPhie P, Pastan I. 1995a. Circularly permuted interleukin 4 retains proliferative and binding activity. *Cytokine* 7:311–318.
- Kreitman RJ, Puri RK, Pastan I. 1994. A circularly permuted recombinant interleukin 4 toxin with increased activity. *Proc Natl Acad Sci USA* 91:6889–6893.
- Kreitman RJ, Puri RK, Pastan I. 1995b. Increased antitumor activity of a circularly permuted interleukin 4-toxin in mice interleukin 4 receptor-bearing human carcinoma. *Cancer Res* 55:3357–3363.
- Li X, Coffino P. 1993. Degradation of ornithine decarboxylase: Exposure of the C-terminal target by a polyamine-inducible inhibitory protein. *Mol Cell Biol* 13:2377–2383.
- Lipscomb WN. 1994. Aspartate transcarbamoylase from *Escherichia coli*: Activity and regulation. *Adv Enzymol* 68:67–151.
- Luger K, Hommel U, Herold M, Hofsteenge J, Kirschner K. 1989. Correct folding of circularly permuted variants of a  $\beta\alpha$  barrel enzyme in vivo. *Science* 243:206–210.
- Matthews CR. 1993. Pathways of protein folding. *Annu Rev Biochem* 62:653–683.
- Mullins LS, Wesseling K, Kuo JM, Garrett JB, Raushel FM. 1994. Transposition of protein sequences: Circular permutation of ribonuclease T1. *J Am Chem Soc* 116:5529–5533.
- Newell JO, Schachman HK. 1990. Amino acid substitutions which stabilize aspartate transcarbamoylase in the R state disrupt both homotropic and heterotropic effects. *Biophys Chem* 37:183–196.
- Peters T Jr, Davidson LK. 1982. The biosynthesis of rat serum albumin. In vivo studies on the formation of the disulfide bonds. *J Biol Chem* 257: 8847–8853.
- Peterson CB, Schachman HK. 1991. Role of a carboxyl-terminal helix in the assembly, interchain interactions, and stability of aspartate transcarbamoylase. *Proc Natl Acad Sci USA* 88:458–462.
- Powers VM, Yang YR, Fogli MJ, Schachman HK. 1993. Reconstitution of active catalytic trimer of aspartate transcarbamoylase from proteolytically cleaved polypeptide chains. *Protein Sci* 2:1001–1012.
- Protasova NY, Kireeva ML, Murzina NV, Murzin AG, Uversky VN, Gryaznova OI, Gudkov AT. 1994. Circularly permuted dihydrofolate reductase of *E. coli* has functional activity and a destabilized tertiary structure. *Protein Eng* 7:1373–1377.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Schachman HK. 1988. Can a simple model account for the allosteric transition of aspartate transcarbamoylase? *J Biol Chem* 263:18583–18586.
- Stevens RC, Chook YM, Cho CY, Lipscomb WN, Kantrowitz ER. 1991. *Escherichia coli* aspartate carbamoyltransferase: The probing of crystal structure analysis via site-specific mutagenesis. *Protein Eng* 4:391–408.
- Tabor S, Richardson CC. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* 82:1074–1078.
- Thornton JM, Sibanda BL. 1983. Amino and carboxy-terminal regions in globular proteins. *J Mol Biol* 167:443–460.
- Vignais ML, Corbier C, Mulliert G, Branlant C, Branlant G. 1995. Circular permutation within the coenzyme binding domain of the tetrameric glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *Protein Sci* 4:994–1000.
- Viguera AR, Blanco FJ, Serrano L. 1995. The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. *J Mol Biol* 247:670–681.
- Wetlauffer DB, Ristow S. 1973. Acquisition of three-dimensional structure of proteins. *Annu Rev Biochem* 42:135–157.
- Yang YR, Schachman HK. 1993a. Aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains. *Proc Natl Acad Sci USA* 90:11980–11984.
- Yang YR, Schachman HK. 1993b. In vivo formation of active aspartate transcarbamoylase from complementing fragments of the catalytic polypeptide chains. *Protein Sci* 2:1013–1023.
- Zhang T, Bertelsen E, Alber T. 1994. Entropic effects of disulphide bonds on protein stability. *Struct Biol* 1:434–438.
- Zhang T, Bertelsen E, Benvegnù D, Alber T. 1993. Circular permutation of T4 lysozyme. *Biochemistry* 32:12311–12318.