Conformation of Aspartate Aminotransferase Isozymes Folding under Different Conditions Probed by Limited Proteolysis*

(Received for publication, May 11, 1998, and in revised form, July 1, 1998)

Joseph R. Mattingly, Jr., Claudia Torella, Ana Iriarte, and Marino Martinez-Carrion‡

From the Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri, Kansas City, Missouri 64110-2499

The partially homologous mitochondrial (mAAT) and cytosolic (cAAT) aspartate aminotransferase have nearly identical three-dimensional structures but differ in their folding rates in cell-free extracts and in their affinity for binding to molecular chaperones. In its native state, each isozyme is protease-resistant. Using limited proteolysis as an index of their conformational states, we have characterized these proteins (a) during the early stages of spontaneous refolding; (b) as species trapped in stable complexes with the chaperonin GroEL; or (c) as newly translated polypeptides in cell-free extracts. Treatment of the refolding proteins with trypsin generates reproducible patterns of large proteolytic fragments that are consistent with the formation of defined folding domains soon after initiating refolding. Binding to GroEL affords considerable protection to both isozymes against proteolysis. The tryptic fragments are similar in size for both isozymes, suggesting a common distribution of compact and flexible regions in their folding intermediates. cAAT synthesized in cell-free extracts becomes protease-resistant almost instantaneously, whereas trypsin digestion of the mAAT translation product produces a pattern of fragments qualitatively akin to that observed with the protein refolding in buffer. Analysis of the large tryptic peptides obtained with the GroEL-bound proteins reveals that the cleavage sites are located in analogous regions of the N-terminal portion of each isozyme. These results suggest that (a) binding to GroEL does not cause unfolding of AAT, at least to an extent detectable by proteolysis; (b) the compact folding domains identified in AAT bound to GroEL (or in mAAT fresh translation product) are already present at the early stages of refolding of the proteins in buffer alone; and (c) the two isozymes seem to bind in a similar fashion to GroEL, with the more compact C-terminal portion completely protected and the more flexible N-terminal first 100 residues still partially accessible to proteolysis.

Proteins synthesized in the cytosol and destined for translocation into mitochondria must retain a flexible conformation in order to be efficiently imported (1, 2). Since mitochondrial protein translocation is thought to be largely a post-translational process (3), a mechanism to prevent the untimely folding of mitochondrial precursor proteins in the cytosol must exist. However, several cytosolic molecular chaperones appear to be involved in maintaining mitochondrial precursors in an import competent conformation. The cast of these chaperones has yet to be firmly established (4). In addition, little is known about the extent of cotranslational versus post-translational folding occurring for large oligomeric proteins synthesized in the cytosol of eukaryotic cells (5) but localizing to different cellular compartments.

One of the many open questions regarding the mechanism by which chaperones regulate protein folding relates to the structural features that determine the binding of an unfolded protein to a given chaperone and the structure of the substrate while bound to the chaperone. Molecular chaperones appear to protect intermediates along the folding pathway, dictated by the protein's primary structure, from undesirable reactions such as aggregation (6). It has been suggested that some chaperones can even unfold aggregated species (7, 8). Thus, the question arises as to the nature and structural features of folding intermediates recognized by chaperones and whether binding of these intermediates to chaperones affects the conformational properties of these partially folded species.

Naturally occurring isozymes such as the cytosolic (cAAT) and mitochondrial (mAAT) forms of aspartate aminotransferase, found within the same cells in eukaryotes, provide unique systems to probe some of the above questions. Much structural and functional information is available for these homodimeric enzymes. With 63% sequence similarity and nearly identical three-dimensional structures (9, 10), each of the two identical active sites is composed of residues from both subunits. Their biological fate within the cell appears to be encoded in their sequence. After synthesis, cAAT remains in the cytoplasm, while mAAT, synthesized as a precursor (pmAAT) with an N-terminal presequence (29 residues long for rat liver pmAAT; Ref. 11), is post-translationally transported into the mitochondrial matrix. Early hints of the intracellular selection show different folding rates measured for each isozyme after their in vitro synthesis in cell-free extracts. After translation in rabbit reticulocyte lysate (RRL), cAAT folds during the dead time of the experiment (<5 min), whereas pmAAT folding proceeds post-translationally at a very slow rate (t1/2 ~ 1 h) at 15 °C (12). The presequence was shown to play a minor role in this distinctive folding rate, since it slightly slowed down the folding of mAAT and had no effect on the folding rate of the cytosolic enzyme when it was fused to the N-terminal end of this protein (pcAAT) (12, 13). On the other hand, when these...
isozymes are allowed to refold in buffer from their GdnHCl- or acid-unfolded states, they followed similar folding paths, with the cysteine enzyme folding only about twice as fast as the mitochondrial form (14, 15). Moreover, the rate at which mAAT folds in RRL is substantially lower than the rate of refolding of the chemically unfolded protein. Thus, we hypothesized that the structural features that account for the slower post-translational folding of mAAT or pmAAT in RRL are likely to be protein determinants that certain RRL chaperones recognize (12, 13, 16). These distinguishing features seem to appear during folding of both homologous proteins, for the overall structures of the completely folded isozymes are remarkably similar (9, 10).

The identities of the RRL chaperones, which regulate the folding of large dimeric proteins as the AAT isozymes have not been fully determined (16), although hsp70 is clearly involved (17). On the other hand, analysis of *in vitro* folding reactions showed that at least two well characterized chaperones can discriminate between the mAAT and cAAT isozymes (13, 15). While purified hsp70 binds only to mAAT during refolding following denaturation at low pH (15), the *Escherichia coli* chaperonin GroEL, although it binds to both, shows a significantly higher affinity for refolding mAAT than for cAAT (13). The basis for this discrimination is unclear, in part because the features responsible for binding of substrates to these chaperones are not well understood but also because we do not know how dissimilar the conformations of refolding cAAT and mAAT might be and where the exposed segments of each protein reside in the folding intermediates.

In this work, we address some of these issues by comparing the conformations of mAAT and cAAT either at the early stages of spontaneous unassisted refolding, while in a stable complex with GroEL, or in a cytosolic-like environment as a freshly translated polypeptide in RRL. Given the considerable protease resistance of these two proteins in the folded state (18), we use limited proteolysis as an analytical tool. Protease digestion patterns reflect the amount of structure already present in partially folded species, since they depend not only on the steric accessibility of potential proteolytic sites in the folding intermediates but also on the flexibility of the surrounding segments of the polypeptide chain (19). The results presented here show that such an approach can be successfully employed for studies of the structural properties of folding intermediates of proteins under different environmental conditions and that the intermediates trapped by GroEL are the same as those appearing during unassisted refolding or, in the case of mAAT, during folding after synthesis in a cell-free extract.

**EXPERIMENTAL PROCEDURES**

**Protein Construction and Purification**—To facilitate the purification of radiolabeled AAT isozymes, a tag consisting of six histidine residues was introduced at the C-terminal end of the proteins by cloning the corresponding cDNAs into pET23a (Novagen). The cDNA encoding the carboxyl terminus of mAAT in the plasmid pBSKS-6 (12) was altered to introduce the XhoI site, which was needed for subcloning into pET23a. This involved polymerase chain reaction mutagenesis (Strategene Quick Change kit) of the pBSKS–6 plasmid to insert the hexanucleotide XhoI recognition site immediately upstream of the stop codon. The NdeIXhoI fragment from the resulting plasmid was subcloned into pET23a to produce plasmid pET23a-1. The protein expressed from this plasmid contains the entire sequence of pmAAT plus an additional Leu-Glu-His8, at the carboxyl terminus. An XhoI site was similarly introduced upstream of the stop codon of the pmAAT cDNA contained in the plasmid pBSKS-4 (20) prior to subcloning into pET23a to produce pET23a-2. pmAAT mRNA produced by *in vitro* transcription under the T7 promoter is inefficiently translated in rabbit reticulocyte lysate. We therefore altered the stop codon of the pmAAT-HT cDNA in pET23a-2 to introduce a BamHI site. This allowed the NdeIBamHI fragment from the resultant plasmid (pET23a-3) to be cloned into pBluescript KS to produce pBSKS-11. pmAAT-HT mRNA was synthesized from this plasmid using T3 RNA polymerase and translated in rabbit reticulocyte lysate as described previously (20). Radiolabeled pmAAT-HT was expressed from the pET23a-3 plasmid.

Radiolabeled proteins were prepared by growing *E. coli* BL21 (DE3) pLysS+ containing the appropriate expression vector to late log phase in minimal medium with 1 mM IPTG and 0.4 g/l T7 RNA polymerase by adding isopropyl-1-thio-β-D-galactopyranoside to 0.5 mM and incubating for 30 min, inhibiting the endogenous *E. coli* RNA polymerase by adding rifampicin to 150 μg/ml, adding 1 μCi of Trans32S–Label (an amino acid hydrolysate from ICN primarily containing [35S]Met) per 100 ml of culture 15 min later, incubating for 1 h, and finally adding chemically synthesized amino acids to 0.2% and incubating for a further 1 h. The soluble form of either radiolabeled fusion protein could be purified in a single step by affinity chromatography using an immobilized Ni2+ matrix following the protocol supplied by the manufacturer (Qiagen). Fractions containing the purified protein were pooled, and the protein was concentrated by ultrafiltration subsequent to changing the buffer to refolding buffer (0.1 M HEPES, 0.1 mM EDTA, pH 7.5). The concentration of the histidine-tagged proteins was determined from the absorbance at 280 nm using calculated molar absorbivities of 55,610 M–1 cm–1 for the pmAAT-HT subunit and 66,810 M–1 cm–1 for the pmAAT-HT subunit. The calculated subunit molecular weights of pmAAT-HT and pcAAT-HT were 45,379 and 50,182, respectively, according to their amino acid sequences (11, 21). The specific radioactivity of the purified proteins is typically about 200,000 dpm/μg. The trypsin susceptibility and specific enzymatic activities of the modified proteins were indistinguishable from those of the unmodified proteins. The purified proteins were stored at 4 °C. Since we used the histidine-tagged proteins in most of the experiments described in this work, we will omit the “-HT” qualifier hereafter and refer to these proteins as pmAAT and pcAAT for simplicity and clarity of presentation. GroES and GroEL were purified as described previously (13), having been overexpressed using the pGroEL plasmid.

**In Vitro Refolding Reaction**—The native proteins were unfolded by incubating with GdnHCl for 30 min at room temperature (14). In the case of pmAAT, the protein concentration was 0.32 mg/ml, and the GdnHCl concentration was 4 M. In the case of pcAAT, the protein concentration was 0.54 mg/ml, and the GdnHCl concentration was 8 M. Both unfolding reactions contained refolding buffer plus 10 μM dithiotreitol. Refolding was initiated by diluting the denatured proteins either 40–fold (pmAAT-HT) or 60–fold (pcAAT-HT) in ice-cold refolding buffer. Complexes between the refolding transaminases and GroEL were formed by diluting solutions of the GdnHCl-denatured proteins into refolding buffer containing GroEL at a concentration such that the ratio of GroEL molecules to AAT molecules was 14 (14).  

**Trypsin Digestion Conditions and Quantitation of Trypsin Digestion Products**—Trypsin digestions of either the refolding protein alone or the complex between GroEL and the refolding protein were performed by adding a small aliquot ( ≤5% of the reaction volume) of a concentrated stock solution of trypsin in 1 mM HCl to give a final trypsin concentration of 10 μg/ml. This digestion was made within 1 min after refolding. *In vitro* translation reactions of pmAAT were diluted 10–fold into 20 mM HEPES, 150 mM NaCl, pH 8.3 prior to digesting with trypsin as described above. Samples were incubated on ice, and the progress of proteolysis was examined by removing an aliquot at various times and quantifying the digestion by adding a 2:1 (w/v) excess of soybean trypsin inhibitor over trypsin and 10 mM phenylmethylsulfonyl fluoride (from a stock solution in 0.2% ethanol). After 10 min on ice, samples were analyzed by SDS-PAGE on 12% acrylamide gels as described previously (12). Products of the trypsin digestions were detected by silver staining or by autoradiography using a Molecular Dynamics PhosphorImager. Molecular weights were estimated by comparing the migration of the unknown to that of 14C-labeled molecular weight markers purchased from Life Technologies, Inc.

The distribution of radiolabeled products in each of the gel lanes was determined using the Molecular Dynamics ImageQuant program as follows. The average pixel value across the width of each lane was calculated and plotted as a function of the distance from the top of the gel. The base line was set to reflect the background pixel intensity, and the peak maximum was manually defined to line up to the base line to mark the beginning and end of each band. Some bands were easily recognized in only some of the lanes, being either better resolved from adjacent peaks because of their relative intensity or being more intense and easily discerned from the background. The starting and ending migration distances defining these peaks were used to designate corresponding regions in adjacent lanes where the peaks may not be so
evident. The area under the line graph was calculated for each of the peaks. Under the exposure conditions used in these experiments, the pixel values are proportional to the actual amount of radiolabel in the gel.

Characterization of pcAAT Trypsinolysis Products—Denatured pcAAT was diluted 60-fold to 30 µg/ml into the appropriate volume of ice-cold refolding buffer (100 mM HEPES, pH 7.5, 0.1 mM EDTA) containing 1 mM dithiothreitol, 50 mM KCl, and 0.8 µM GroEL 14-mer (~1.2-fold molar excess over pcAAT monomer). Trypsin (10 µg/ml) was immediately added (~1 min after starting refolding), and the sample was incubated for 15 min on ice. After quenching proteolysis with a 2:1 (w/v) excess of trypsin inhibitor over trypsin, an aliquot was analyzed by SDS-PAGE, and the rest of the reaction was fractionated by HPLC on an RP-304 analytical column (4.6 × 250 mm, Bio-Rad) to separate the pcAAT fragments from GroEL. The chromatography was performed at a flow rate of 1 ml/min using a 95% acetonitrile/water gradient (20–45% over 25 min, followed by an increase from 45 to 55% over 20 min) containing 0.1% trifluoroacetic acid. Each fraction was analyzed by SDS-PAGE to determine the distribution of the proteolytic fragments observed in the electrophoretic pattern of the unfractionated reaction. Fractions collected manually from the reverse phase HPLC column were directly analyzed by electrospray mass spectrometry using the LCQ ion trap system (Finnigan MAT). Samples were injected directly into the ion source via a syringe injection at a flow rate of 5–10 µl/min, and data were acquired and elaborated using the Navigator software (Finnigan). All masses are reported as average values. The identification of the origin of the fragments within the known primary structure of cAAT (21) was based on their molecular mass determined by electrospray mass spectrometry and the presence of a histidine-tagged C-terminal end.

RESULTS

Rationale—Both pmAAT and pcAAT can be unfolded in GdnHCl solutions and can refold unassisted when the denaturant is sufficiently diluted (14, 15). At 0 °C, the refolding process is slow (t1/2 ~ 140 min for pmAAT; t1/2 ~ 70 min for pcAAT), and this is referred to as in vitro refolding throughout this work. Previous studies had indicated that AAT, like many proteins, probably forms a compact, “molten globule”—like intermediate very early in the refolding reaction, which slowly converts into the final folded state (14, 15). We reasoned that it might be possible to assess the extent to which the putative “molten globule” intermediate(s) resembled the fully folded protein by a technique that could be applied to systems for which spectroscopic analysis is difficult, namely proteins synthesized in whole cell-free protein translation systems such as RRL. The technique chosen relies on the considerable protease resistance of the fully folded proteins (12, 18, 20, 22). Therefore, we hypothesized that if molten globule-like intermediates could be detected during the refolding of pmAAT or pcAAT, then their trypsin susceptibility should be greater than that of the fully folded protein but less than that expected for the fully unfolded protein. Thus, we compare the digestion pattern of the isoforms immediately after initiating refolding at 0 °C.

Trypsin Digestion of Refolding pmAAT—Trypsin digestion of the fully folded precursor protein is limited to the removal of the prosequence before cleavage away Arg-2 (negative numbering is used for the prosequence with the N terminus methionine as residue −29 and the C terminus alanine as residue −1) and the production of a single band corresponding to the mature protein with just an additional alanine at its amino terminus (22). On the other hand, trypsin treatment of the protein slowly refolding at 0 °C instead gives rise to an extremely reproducible pattern of fragments. SDS-PAGE analysis of the trypsin digestion (10 µg/ml) of pmAAT shows the presence of 16 bands, several strong and many weak ones (Fig. 1A), with the latter requiring PhosphorImager enhancement in order to be measured. The bands seem to cluster in four molecular weight ranges, which we designate A, B, C, and D. The bands within each apparent molecular weight range are numbered sequentially (Fig. 2, lane 1). A1 is intact pmAAT, and A2 is the species truncated within the prosequence and originally present before trypsin digestion. The species identified in Fig. 2 are listed in Table I along with their apparent molecular weights determined according to the mobility of molecular weight standards.

Five time-dependent behaviors are observed when the amount of the radiolabel in each band is plotted as a function of trypsin digestion time (see Fig. 3 for representative examples of four of these and Table I for a summary of the kinetic patterns of individual bands). The radiolabel in bands exhibiting type 1 behavior, A1 (intact pmAAT) and A2, shows a continuous de-
crease throughout the course of the reaction. The radiolabel in type 2 bands (C3 and D1) begins at a level that is significant relative to its maximum amount and subsequently decreases and then increases. (No example of this behavior is shown in Fig. 3). The amount of radiolabel in type 3 bands (A3, A4, A5, C2, and C4) also begins at a level that is significant relative to its maximum, but the intensity increases from its initial value and then decreases. The peak area at $t = 0$ depends on how the peaks are defined. For instance, the intensity of the A2 band in the $t = 0$ lane is so great that some spills over into the regions where the A3, A4, and possibly A5 bands migrate. The radiolabel in type 4 bands (A6, B1, B2, B3, and C1) behaves much the same way as the radiolabel in type 3 bands, increasing more or less rapidly to a maximum and then beginning to decrease, but the initial amount of radiolabel in these bands is insignificant relative to the maximum intensity achieved during the course of the digestion. The radiolabel in type 5 bands (D2 and D3) increases throughout the course of the trypsin reaction.

Thus appears that two populations of refolding pmAAT may exist. Type 3 and 4 bands exhibit the transient behavior expected of true reaction intermediates (Fig. 3). Fig. 4A shows an estimate of the maximum amount of radiolabel found in each of these species during the course of the reaction (expressed as a fraction of the initial amount of radiolabel present in the intact pmAAT band) as well as the time at which the radiolabel in each of the bands reaches a maximum. The amount of radiolabel in most type 3 and 4 bands reaches a maximum during the first 15 min of digestion (Fig. 4A, top). These species appear in a rough progression with the larger polypeptides (A3, A4, A5, B1, B2, and B3) accumulating earlier and the smaller fragments (C1 and C2) appearing later. The maximal accumulation values range from a low of about 2% (B3) to a high of about 7% (A3 and A5). It should be noted, however, that these amounts do not reflect the relative molar amounts of the various species, since each polypeptide may contain a different number of radiolabel as $[35S]$methionine residues (see Fig. 6 for distribution of methionine residues along the amino acid sequences). Nevertheless, assuming that the specific radioactivity of each polypeptide is proportional to its apparent molecular weight (Table I), one can estimate that the maximum fraction of pmAAT present in any of these intermediates ranges from

---

**Fig. 2. Identification and labeling of trypsin digestion products.** $^{35}$S-Labeled refolding pmAAT with or without GroEL (lanes 1 and 2), freshly synthesized pmAAT (lane 3), and refolding pcAAT with or without GroEL (lanes 4 and 5) were digested with trypsin for 20 min on ice essentially as described in the legend to Fig. 1. The contrast of these images was adjusted so as to make the minor components visible; consequently, the band intensities are not comparable even within a lane. The horizontal lines drawn to the right of each lane indicate the regions of the PhosphorImager scan that were considered to be separate bands. Each band is labeled sequentially according to the region in which it migrates.
about 2 to 11%. The type 5 bands (D2 and D3 in the top of Fig. 4A) reach a maximal accumulation at the end of the reaction (30 min), which suggests that they are relatively stable products of digestion. Assuming again that there is proportionality between specific radioactivity and molecular weight of the fragments, we can estimate that the molar fraction of total pmAAT

| TABLE I | Radiolabeled electrophoretic products of trypsin-digested AAT isozymes |
|---------|---------------------------------------------------------------------|
| Refolding pmAAT | pmAAT translation product |
| **No additions** | **+ GroEL** |
| **Peak** | **Kinetic pattern** | **M_r** | **Peak** | **Kinetic pattern** | **M_r** |
| A1 | 1 | A1 | 1 | 41.4 | A1 | 1 | A1 | 1 | 41.6 |
| A2 | 1 | A2 | 1 | 40.4 | A2 | 1 | A2 | 1 | 40.8 |
| A3 | 3 | A3 | 3 | 39.5 | A3 | 1 | A3 | 1 | 39.5 |
| A4 | 3 | A4 | 4 | 38.1 | A4 | 1 | A4 | 3 | 38.2 |
| A5 | 3 | A5 | 4 | 37.0 | A5 | 1 | A5 | 4 | 38.2 |
| A6 | 4 | A6 | 4 | 36.3 | A6 | 1 | A6 | 3 | 37.1 |
| B1 | 4 | B1 | 4 | ND | A7 | 1 | A7 | 3 | 36.4 |
| B2 | 4 | B2 | 5 | 33.8 | B1 | 3 | B1 | 5 | 35.3 |
| B3 | 4 | B3 | 5 | 30.2 | B2 | 3 | B2 | 5 | 33.2 |
| C1 | 4 | C1 | 2 | 28.2 | B3 | 3 | B3 | 5 | 32.3 |
| C2 | 3 | C2 | 2 | 26.1 | B4 | 1 | B4 | 5 | 31.5 |
| C3 | 2 | C3 | 2 | 24.7 | C1 | 3 | C1 | 5 | 27.5 |
| C4 | 3 | C4 | 2 | 22.9 | C2 | 3 | C2 | 5 | 26.8 |
| C5 | 3 | C5 | 4 | 21.3 | C3 | 3 | C3 | 5 | 25.3 |
| C6 | 3 | C6 | 4 | 20.5 | C4 | 0 | C4 | 5 | 24.6 |
| C7 | 4 | C7 | 1 | 23.2 | C5 | 4 | C5 | 5 | 23.7 |
| C8 | 3 | C8 | 1 | 22.9 | C6 | 4 | C6 | 5 | 22.8 |
| D1 | 2 | D1 | 0 | 20.5 | C7 | 1 | C7 | 5 | 21.3 |
| D2 | 5 | D2 | 0 | 17.0 | D1 | 5 | D1 | 5 | 20.6 |
| D3 | 5 | D3 | 4 | 16.1 | D2 | 5 | D2 | 5 | 18.8 |
| D4 | 4 | D4 | 3 | 15.3 | D3 | 5 | D3 | 5 | 18.6 |
| D5 | 3 | D5 | 3 | 15.6 | D4 | 5 | D4 | 5 | 17.4 |
| D6 | 2, 3 | D6 | 3 | 14.2 | D5 | 5 | D5 | 5 | 16.2 |
| D7 | 3 | D7 | 3 | 13.3 | D6 | 5 | D6 | 5 | 16.9 |
| D8 | 5 | D8 | 5 | ND | D7 | 5 | D7 | 5 | 15.9 |

*The peaks are identified as indicated in Fig. 2.

The kinetic patterns are described under “Results.”

*Estimated from mobility in SDS-PAGE relative to 14C-labeled standards. ND, not determined.

1*, a variant of type 1 behavior, the intensity of these species reaches a maximum within the first minute of digestion and then decreases rapidly.

**Fig. 3.** Time course of appearance of the various radiolabeled species observed during trypsin digestion of pmAAT. The amount of radiolabel associated with selected bands observed during the digestion of refolding pmAAT alone (panel A) or in a complex with GroEL (panel B) was determined as described under “Experimental Procedures.” Note that right axes differ in panels A and B. The amount of radiolabel in each band is expressed relative to the amount of radiolabel initially present in the intact pmAAT. The data were fitted to equations of the following forms using the implementation of the Marquadt algorithm found in the DeltaGraph program: type 1 kinetic pattern \( f(t) = A_1 e^{-k_1 t} + A_2 e^{-k_2 t} \), typified by band A1 (●) in panels A and B; type 3 kinetic pattern \( f(t) = A_3 e^{-k_3 t} + A_4 e^{-k_4 t} + A_5 e^{-k_5 t} \), typified by band B1 (●) in panel A; type 4 kinetic pattern \( f(t) = A_4 e^{-k_4 t} + A_5 e^{-k_5 t} + A_6 e^{-k_6 t} \), typified by band B2 (●) in panel A or A5 (●) and D3 (●) in panel B; type 5 kinetic pattern \( f(t) = A_6 e^{-k_6 t} + A_7 e^{-k_7 t} + A_8 e^{-k_8 t} \), typified by band D3 (●) in panel A or B2 (●) in panel B. The numbers in parenthesis identify the kinetic patterns as described under “Results.”

Trypsinolysis of Refolding AAT Isozymes
Trypsin Digestion of Refolding pmAAT Captured by GroEL—GroEL alone is rather stable to trypsin digestion under the conditions used for digesting the refolding pmAAT. Less than 10% of intact chaperone is lost in 30 min of digestion with 10 μg/ml trypsin at 0 °C (τ1/2 ~ 1.4 min). The presence of GroEL has only a minor effect on the number or the mobility of the trypsin digestion products originating from folding pmAAT. With the exception of two faint bands (m + GroEL-B3 and m + GroEL-D4), all the species observed during the digestion of pmAAT bound to GroEL are also produced during the digestion of folding pmAAT alone (Fig. 2, lanes 1 and 2). However, in the presence of GroEL, the relative amount of radiolabel present in the higher molecular weight products is much greater than that in the lower molecular weight species. Approximately 20–30% of the initial amount of radiolabel is lost rapidly (τ1/2 ~ 1.2 min) in the course of the trypsin digestion. After this early loss, the amount of radiolabel detected by SDS–PAGE essentially remains constant in the presence of GroEL, whereas another ~40% is slowly lost in the absence of GroEL (Fig. 5A). Thus, GroEL affords the trapped refolding pmAAT significant protection toward trypsin digestion.

Protection by GroEL does not arise from inhibition of trypsin, since the rate at which trypsin hydrolyzes the presequence peptide from native folded precursor is unaffected by the presence of GroEL (data not shown). Furthermore, most of the pmAAT radiolabel remains associated with GroEL after trypsinolysis of the pmAAT-GroEL complex, as evidenced by its co-sedimentation with GroEL on glycerol gradients (23), suggesting that the ability of GroEL to bind the resulting protein fragments remains unimpared under the trypsinolysis conditions. Finally, the addition of GroES and MgATP after trypsin digestion of the pmAAT-GroEL complex results in the release of all the bound polypeptides from GroEL (data not shown), as expected for a functional chaperonin. However, even the five largest and most abundant fragments (A4, A5, A6, B1, and B2) appear to be unable to refold properly, since they are recovered as insoluble aggregates following release from GroEL.

The total amount of radiolabel in the lanes corresponding to the various time points was calculated and is plotted as a function of reaction time in Fig. 5A. Approximately 30% of the radiolabel is lost very quickly (τ1/2 ~ 0.2 min); the remainder is lost at a rate that is over 100 times slower (τ1/2 ~ 27 min). Thus, in the 30 min over which the reaction is observed, approximately 70% of the radiolabel is lost from species retained on the gel. If the molar fraction of each species present after 30 min of digestion is calculated from the assumption of proportionality between specific radioactivity and apparent molecular weight and the resulting fractions are summed, about 75% of the pmAAT molecules initially present remain as truncated species, and over half of those are in D3. Thus, a large fraction of the potential trypsin digestion sites (see Fig. 6) in certain regions of the refolding protein remain protected, although not as effectively as in the fully folded protein; i.e., the refolding protein is in a conformation in which much of the protein spends a considerable amount of time in a native-like state in flux with less ordered conformations. Since the molecular chaperone GroEL has been shown to arrest the reactivation of refolding mAAT (13), we investigated whether it might stabilize putative “folding domains” toward trypsin digestion.

Trypsinolysis of Refolding AAT Isozymes

Fig. 4. Maximum amount of radiolabel present in species generated by trypsin digestion of refolding pmAAT or pcAAT with or without GroEL. The maximum percentage of radiolabel found in each of the bands identified in Fig. 2 was determined from plots such as those shown in Fig. 3 and is expressed relative to the initial amount of radiolabel associated with the intact protein band at time 0. These values are plotted as solid bars on a time line indicating when this maximum amount was present in the reaction. The relative amount of radiolabel remaining in a band at the end of the 30-min digestion is also shown by the superimposed gray bar. Panel A summarizes the data for the digestion of pmAAT alone (upper time line) or plus GroEL (lower time line). Band D4 only appears in the digestion of pmAAT plus GroEL. Panel B summarizes the data for the digestion of pcAAT alone (upper time line) or plus GroEL (lower time line). The region of the gel containing bands C4 and C5 in the digestion of pcAAT alone was considered to contain only a single band, C3, in the digestion of pcAAT plus GroEL. The region of the gel containing bands C6 and D1 in the digestion of pcAAT alone was considered to contain a single band, C4, in the digestion of pcAAT plus GroEL.
Trypsin Digestion of pmAAT Synthesized in Rabbit Reticulocyte Lysate—Lanes 2–4 of Fig. 1B show an SDS-PAGE analysis of the trypsin digestion (10 μg/ml) of radiolabeled pmAAT (0.01 μg/ml) freshly synthesized in RRL. The total protein concentration in the reticulocyte lysate is approximately 100 mg/ml, with the vast majority of that protein being globin. No obvious degradation of RRL proteins is observed during the course of the trypsinolysis reaction (data not shown). The amount of radiolabel remaining in bands that are retained on an SDS-PAGE gel is presented as a function of reaction time in Fig. 5B. About 40% is lost in the first 4 min ($t_{1/2} \sim 0.5$ min), while another ~35% is lost in the subsequent 26 min ($t_{1/2} \sim 22$ min). As many as perhaps 30 species appear at one time or another in the course of the digestion of the translation product, significantly more than the 16 noted for refolding pmAAT. Nevertheless, the more prominent of the 16 species seen in the digestion of in vitro refolding pmAAT can be matched with species seen in the digestion of the translation product. These species are identified in Fig. 2B and are listed in Table I along with their estimated molecular weight and type of kinetic behavior. A similar digestion pattern was obtained for trypsinolysis of a postribosomal fraction of the RRL translation reaction (data not shown).

The molecular weights of the species produced in the digestions of the translation product or the refolding protein were determined by comparison with the mobility of standards, which had been prepared in sample matrices that attempted to replicate the matrix of their respective unknowns; i.e. the standards were diluted into either RRL or refolding buffer containing 0.1 M GdnHCl. This analysis gives quite different molecular weights for the prominent radioactive bands that migrate near the bromphenol blue dye front in samples from the two reactions (such as D5 in lane 2 and D7 in lane 3 in Fig. 2). However, if samples from the translation product digestion are diluted into refolding buffer containing 0.1 M GdnHCl prior to electrophoresis, most bands can be correlated (Table II). Since both CAAT and pcAAT synthesized in RRL acquire complete trypsin resistance within the dead time of the experiment (12), the tryptic pattern for these proteins would show exclusively a mature-like band (A2 in Fig. 2, lanes 4 and 5), and therefore it has not been included in Fig. 2.

Whereas the overall identity of the trypsin digestion products appears largely unaltered by the binding to GroEL, the kinetic behavior of some of these species is dramatically altered (see Table I and Fig. 3). This is particularly evident in the higher molecular weight intermediates. The bands exhibiting a type 1 behavior in the pmAAT–GroEL complex. By contrast, radiolabel in D1 remains virtually constant in the digestion of pmAAT–GroEL, defining a new kinetic type, which we designate type 0 (Table I). Changes in kinetic patterns are reflected in the times at which a maximum amount of radiolabel is found in these bands (Fig. 4A).

The maximum relative amounts of radiolabel present in the high molecular weight intermediates A3, A4, A5, A6, B1, and B2 increase anywhere from ~3.5- to almost 9-fold in the digestion of pmAAT bound to GroEL relative to the free pmAAT reaction, with B2 showing the greatest increase (Fig. 4A).
Trypsin Digestion of Refolding pcAAT—In vitro refolding of GdnHCl-unfolded pcAAT is only a little faster than that of its homologous pmAAT (13). By marked contrast, after synthesis in RRL at 30 °C, the folding of pcAAT is immediate (12), while that of pmAAT under the same conditions is so slow that it is apparently digested by endogenous proteases faster than it can fold (20). Since the interaction of in vitro refolding cAAT with GroEL is weaker than that of refolding mAAT (13), it seems that molecular chaperones amplify small differences in the intrinsic folding rates of the two isozymes and preferentially bind to mAAT. The lower affinity of GroEL for cAAT possibly reflects a more compact folding intermediate with fewer exposed hydrophobic residues or even fewer hydrogen bonds (24) for interaction of the protein with the chaperone. To test this hypothesis, we examined the trypsin susceptibility of refolding pcAAT under identical conditions as described for pmAAT alone.

If refolding pcAAT adopts a conformation that is more compact than that of refolding pmAAT, intermediate folding states of the former should be more trypsin-resistant than those of the latter. Such appears to be the case. Only 20% of the radiolabeled pcAAT is lost early ($t_{1/2} \approx 0.2 \text{ min}$; Fig. 5C). The remaining radiolabeled protein is lost at somewhat similar slower rates ($t_{1/2} \approx 37 \text{ min for pcAAT versus } t_{1/2} \approx 27 \text{ min for pmAAT}$). On the other hand, both the number and size of the partial digestion products arising from refolding pcAAT or pmAAT are remarkably similar (Fig. 1, A and C). The 18 species that can be detected during the digestion of refolding pcAAT are identified (Fig. 2, lane 4) and tabulated with their apparent molecular weights and kinetic patterns (Table I). Fig. 4B summarizes the time at which the maximum relative amount of radiolabel is found in each band as well as the actual maximum relative amounts. The maximal accumulation of the large pcAAT fragments (for instance, 26% for A2 and 12% for A4) is substantially greater than that observed for the corresponding fragments from pmAAT (about 7% for either A3 or A5).

Trypsin Digestion of Refolding pcAAT Captured by GroEL—Although GroEL does not bind refolding pcAAT tightly at 0 °C (13), it still affords the refolding protein a considerable degree of protection against trypsin digestion (Fig. 1C, lanes 6–9). The disappearance of radiolabel in species that can be detected by SDS-PAGE is biphasic (Fig. 5C). Following a rapid loss of about 20% of the initial protein, only 15% more is lost in a slower phase of the reaction ($t_{1/2} \approx 115 \text{ min}$) in the presence of GroEL. By contrast, in the absence of GroEL, an additional 40% is lost by the end of the reaction at 30 min. With few exceptions (band c + GroEL–D3, which is only found in the reaction of GroEL-bound protein, and bands C3, C4, C5, C6, and D1, which are present only in the digestion of pcAAT alone), all of the species observed during the digestion of pcAAT bound to GroEL are also produced during the digestion of pcAAT alone (Fig. 2C and Table I). The effect of GroEL on the kinetic patterns of these bands is qualitatively similar to its effect on the digestion of pmAAT (Fig. 4B). Many of the bands that reach a maximum in the first half of the reaction in the absence of GroEL (B1, B2, B3, and C1) continue accumulating until the end of the reaction in the presence of GroEL. Other bands (A2 and A4) reach a maximum at roughly the same time in both reactions or only a little later (A5) in the reaction containing GroEL. The major apparent low molecular weight end product of the refolding pcAAT digestion, D4, reaches a maximum about midway through the reaction containing GroEL but retains its type 5 kinetic behavior (intensity remains constant until the end).

Localization of Trypsin Cleavage Sites in GroEL-bound AAT—The reproducible pattern of tryptic fragments obtained with the GroEL-bound proteins suggests that only a limited number of peptide bonds are readily accessible or flexible enough to be cleaved by trypsin. Using a combination of reverse phase HPLC and electrospray mass spectrometry, we characterized the most prominent fragments and localized the cleavage sites within the amino acid sequences of the two proteins. For mAAT, the cleavage sites were identified as Arg$^{26}$ (fragment A4), Arg$^{54}$ (fragment A5), Lys$^{65}$ (fragment A6), and Arg$^{89}$ (fragment B2) (23). Although fragment B1 was not analyzed, it probably arises from cleavage after Lys$^{81}$, since a fragment of identical electrophoretic mobility produced by hydrolysis with elastase was found to originate by cleavage at Cys$^{80}$ (23).

A similar approach was used to localize the tryptic cleavage sites within the amino acid sequence of cAAT. As found for mAAT, the cAAT fragments that could be unambiguously assigned arise from cleavages at the N-terminal region of the protein: Lys$^{59}$ (A4), Arg$^{89}$ (A5), Arg$^{99}$ (B1), and Arg$^{113}$ (B2) (Table III, Fig. 6). All of them contain an intact C-terminal end. This was established from the mass of the fragments estimated by electrospray mass spectrometry and by comparing the electrophoretic mobility of the tryptic fragments generated from wild type protein and from pcAAT having a histidine tag at the C-terminal end. All of the fragments obtained with pcAAT histidine-tagged migrated slightly more slowly than those observed with the untagged protein, as expected if the fragments extended to the C-terminal end of the polypeptide (data not shown).

DISCUSSION

For a peptide bond to be hydrolyzed by a protease, it must productively bind to the active site of the protease and be able to assume the geometry needed to fit into its active site. For isolated peptides, primary sequence is a most likely constraint, since such peptides have substantial conformational flexibility. However, when a peptide is part of a large, folded protein, additional features such as the local conformational flexibility of the substrate (peptide loop), as well as steric constraints imposed by the bulk of the protein must also be factored in. Thus, the greater proteolytic susceptibility observed for refolding AAT forms relative to that for the fully folded proteins is probably the net result of several factors including increased local conformational freedom and increased accessibility. Modeling studies have suggested that a loop consisting of a minimum of 10 residues could adopt a conformation resembling that of an optimum protease substrate (19). Therefore, local unfolding of a few residues of a substrate protein could greatly increase its protease susceptibility if those are in exposed regions of the conformation.

Since the unfolded AAT proteins are completely hydrolyzed by trypsin, the discrete intermediate species observed during the trypsin digestion of the refolding proteins immediately following initiation of the reaction suggest that the collapsed refolding intermediates contain folding domains, i.e. regions in
Trypsinolysis of Refolding AAT Isozymes

Identity of the main tryptic fragments determined by mass spectrometric analysis

| Fragment* | Observed mass | Calculated mass | AAT residue no.* | Cleavage site† |
|-----------|--------------|----------------|-----------------|---------------|
| GroEL-pmAAT |              |                |                 |               |
| A3        | 44,599.1     | 44,598.27      | -1 to 410        | Arg<sup>2</sup> |
| A4        | 41,805.1     | 41,804.07      | 27 to 410        | Arg<sup>24</sup> |
| A5        | 38,727.8     | 38,726.58      | 55 to 410        | Arg<sup>24</sup> |
| A6        | 37,825.3     | 37,829.54      | 64 to 410        | Arg<sup>24</sup> |
| B2        | 34,138.9     | 34,138.65      | 100 to 410       | Arg<sup>29</sup> |
| GroEL-pcAAT |              |                |                 |               |
| A2        | 46,498.1     | 46,495.72      | -1 to 412        | Arg<sup>2</sup> |
| A4        | 39,702.1     | 39,699.13      | 60 to 412        | Lys<sup>93</sup> |
| A5        | 37,292.3     | 37,294.41      | 81 to 412        | Arg<sup>20</sup> |
| B1        | 35,284.6     | 35,282.17      | 100 to 412       | Arg<sup>20</sup> |
| B2        | 34,026.6     | 34,029.4       | 114 to 412       | Arg<sup>113</sup> |

а Fragment labeling is as indicated in Fig. 2. These experiments were performed with untagged proteins.

* The numbering of the residues is according to that of pig cAAT (43) introducing gaps in the sequence of mAAT when appropriate to maximize sequence homology.

† Data taken from Ref. 23 for comparison.

An additional methionine residue was introduced at the junction between the presequence and cAAT sequence in the preparation of pcAAT (12).

Therefore, the mass of this fragment corresponds to that of cAAT plus a methionine residue in addition to the C-terminal alanine from the presequence.

which the polypeptide chain has lost the conformational flexibility or accessibility needed for protease susceptibility. The transient nature of these intermediates also suggests that these more stable conformations can interconvert to more flexible, protease-accessible states. Such a model is represented by the folding reactions shown in Scheme I, where \( U \) represents the entire range of unfolded conformations present in the de- naturant, and \( I \) represents an ensemble of molten globule-like intermediate states, with \( I_0 \) having a conformation in which most potential trypsin nick sites are highly reactive and \( I_1, I_2, \ldots, I_n \) representing conformations in which particular subsets of potential trypsin nick sites become unreactive. This ensemble of I-type intermediates, containing over 80% of the secondary structure of the native dimer, appears very rapidly (within the dead time of the manual mixing) (14, 15). This rapid step, also observed for the renaturation of many other denatured proteins (25), is followed by two slow isomerizations (14, 15) to generate the native conformation \( F \), which contains essentially no cleavage sites accessible to trypsin. Only the slow conversion to the native conformation is considered to be irreversible; the less ordered conformations are imagined to interconvert. The facility with which these interconversions occur is a measure of the flexibility of the nonnative protein. The intermediates represented in this simplified model for the refolding of AAT would correspond to local minima in a rough funnel-shaped energy landscape description of protein folding (26).

The existence of discrete folding domains in proteins has been well documented in other proteins, including another pyridoxal 5'-phosphate-dependent enzyme (27). In some cases, folding domains can be related to structural domains of fully folded protein (28). The folding domains implied by the trypsin-resistant fragments observed with refolding AAT likewise seem to be related to the structural organization of the folded protein. The first 29 N-terminal amino acids of pmAAT comprise the targeting presequence. Since this region is readily hydrolyzed by trypsin even in the folded protein (20, 22), it appears that it has few if any conformational or steric constraints imposed upon it by the remainder of the protein. Hydrolysis at the C-terminal end of the presequence (arginine), 2 residues upstream from its junction with the mature protein (Arg<sup>2</sup>), produces mature-like intermediates A3 (mAAT) and A2 (cAAT). Because these are the first intermediate species to reach a maximum, it seems reasonable to conclude that sites within the presequence are cleaved more rapidly than any other sites in the refolding protein and, therefore, that the mature portion of the refolding protein tends to be less accessible and flexible than the presequence, just as seen for the folded proteins.

The cleavage sites within the mature portion of the proteins corresponding to the most prominent tryptic fragments have been identified for both mAAT (23) and cAAT bound to GroEL. The instability and only transient accumulation of tryptic fragments arising from the proteins refolding in the absence of GroEL precluded their characterization. However, since the main digestion products comigrate in SDS-PAGE with the fragments originating from the GroEL-AAT complexes and also contain an intact C-terminal end, we can conclude that they must result from cleavage at nearly identical sites in the sequence. The qualitatively similar proteolytic pattern of free and GroEL-bound refolding proteins, particularly with regard to the most abundant fragments, A4, A5, A6, B1, and B2, suggests that the folding domains uncovered by the tryptic pattern are not determined by the binding to the chaperonin but rather represent an intrinsic feature of the folding intermediates. The distribution of trypsin-susceptible and -resistant regions is remarkably similar in the two isozymes (Fig. 6). Since the three-dimensional arrangement of the backbone in the native AAT forms is almost identical (10), it seems probable that the more compact domains detected in the refolding intermediates resemble those same segments in the folded proteins. In both proteins, the peptide bonds with greatest trypsin susceptibility are clustered in regions of the N-terminal portion of the polypeptides that are located at the bottom of the large domain in the native structures (Fig. 7). In mAAT, trypsin also cleaves in the N-terminal bridging peptide (Arg<sup>29</sup>), although the resulting fragment (A4) does not accumulate substantially in the absence of GroEL. This is the only region that is susceptible to
Trypsinolysis of Refolding AAT Isozymes

FIG. 7. Structural models of dimeric mAAT and cAAT. The diagrams were generated using the program VMD (Theoretical Biophysics Group, University of Illinois at Urbana-Champaign) and the x-ray structures of mAAT (pyridoxal form, Protein Data Bank entry code 7AAT (30)) and cAAT (maleate complex of pyridoxal form, Protein Data Bank entry code 2CST (10)) from chicken heart. The direction of view is along the 2-fold axis with the upper subunits drawn as tubes and the lower subunits in a dotted representation. In this view, the small domains (residues 1–48 and 326 to the C terminus) are at the upper right and lower left corners, with the large domains (residues 49–325) occupying the center of the model. The residues identified as tryptic cleavage sites in the mature protein (Table III) are represented as van der Waals spheres in the upper subunits.

Trypsin in the folded protein, although only at much higher concentrations of trypsin than those used here (18, 29). The central core of the large domain and the C-terminal component of the small domain are more resistant to proteolysis and therefore probably much more structured. Thus, the C-terminal portion appears to adopt a compact conformation somewhat faster than the N-terminal region. In the native structure, several of the trypsin-accessible sites identified (Lys59 and Arg113 in cAAT; Arg14 and Lys40 in mAAT) are close to the main subunit interface area around the 2-fold symmetry axis. Others (Arg26 in both proteins) are located in the vicinity of the region in the large domain that interacts with the N-terminal end from the other subunit (Fig. 7; Refs. 10 and 30). Therefore, it seems likely that the final organization of the N-terminal segment is achieved only after formation of the dimer.

Considering that the maximum amounts of the larger pcAAT intermediates detected are significantly greater than those of comparable pmAAT intermediates, the cytosolic protein while refolding seems to be somewhat more trypsin-resistant than pmAAT. This could indicate that the fraction of pcAAT present in less reactive species (those to the right of the spectrum in Scheme I) is greater and that these species may interconvert less readily with the more reactive species. The lower affinity of GroEL for refolding pcAAT relative to pmAAT (13) is also consistent with this interpretation if we suppose that GroEL binds preferentially to species to the left of the spectrum in Scheme I and that the decrease in protease sensitivity is correlated with decreased exposure of (hydrophobic) residues to solvent. Binding to GroEL confers substantial protection against proteolysis to both enzymes. This is evident from the retardation in the appearance of some of the large fragments (A5, B1, and B2) and their increased stability against further proteolysis (notice the greater accumulation of these peptides in the bottom panels of Fig. 4). This protection probably results from steric interference with the access of the protease to the proteins bound to GroEL.

The degree of protection appears to be more pronounced for mAAT than cAAT (Fig. 4 and 5). This may reflect differences between the two bound isozymes either in the flexibility or accessibility of the potential nick sites in the N-terminal region to the protease. However, GroEL forms a very stable complex with mAAT (arrests its refolding), but binding to cAAT is much weaker (only slows its refolding) (13). Thus, the GroEL-cAAT complex is in equilibrium with some free protein to which trypsin could have unrestricted access.

The precise location(s) where a given polypeptide binds to GroEL is as yet unknown. Yet, it is unlikely that the central cavity of resting GroEL could completely enclose a partially folded protein the size of either AAT isozyme (31), let alone accommodate both this polypeptide substrate and the added trypsin. Furthermore, recent studies of GroEL have indicated that residues near the opening of the cavity are involved in the interaction with polypeptide substrates (24, 32) that may bind at multiple sites in GroEL (33), while cryoelectron microscopy has shown that GroEL holds substrates between the apical domains that form the mouth of the cavity (34). Our data support a model for the binding of mAAT and cAAT to GroEL in which the more compact C-terminal portion of the molecule, which includes most of the small domain (residues 326–410 or 412 in mAAT or cAAT, respectively) and the core of the large domain (residues 115–325), is buried in the binding cavity of GroEL. On the other hand, several peptide segments from the more flexible N-terminal region (highlighted in Fig. 7) might be partially exposed at the mouth of the cavity.

GroEL has been shown to cause the unfolding of a number of monomeric proteins (8, 35–40) (β-lactamase, dihydrofolate reductase, cyclophilin, barnase, and human carbonic anhydrase II) or misfolded proteins (7); however, it has no effect on either
of the native (folded) dimeric AAT isozymes (data not shown). Unfolding can simply be the result of GroEL displacing the preexisting equilibrium between different folding states of the protein by binding to the less folded species. Yet, the intermediate species produced by trypsin digestion of AAT, whether the protein is alone during in vitro refolding or when it is bound to GroEL, are similar. This suggests that refolding AAT either does not become substantially more unfolded upon binding to GroEL or, if it does, the newly exposed trypsin sites remain relatively inaccessible and consequently beyond detection by limited proteolysis. Additional studies using more sensitive approaches such as H/D amide exchange will be required to clarify this issue. The compact structure we attribute to the trapped intermediate formed during the refolding of large proteins such as AAT is consistent with recent studies showing moderate to considerable amounts of native-like structure for smaller monomeric GroEL-bound proteins (41, 42), which seem to fit the central cavity of the chaperone. Furthermore, the apparent failure of GroEL to extensively unfold the AAT intermediate suggests that the intramolecular interactions in partially folded AAT are greater than the strength of the interactions between GroEL and partially unfolded AATS. Thus, the extent to which GroEL-bound polypeptides are unfolded may depend upon the particular protein substrates (36) and may be of even less importance in larger proteins (like the AATS), where intramolecular interactions are more likely to act as stabilizing forces rather than intermolecular contacts in the GroEL-folding protein complex.

The correspondence between the major intermediates generated in the digestion of pmAAT refolding in buffer or folding after synthesis in a cell-free extract (A3, A6, A7, B1, B4, and C4) suggests that both species share a basically common structural organization, despite the fact that folding of the nascent chain is probably vectorial during its elongation N-terminal first at the ribosome (5) whereas refolding in vitro involves the collapse of the complete chain upon its transfer from denaturing to native conditions. However, the pmAAT translation product is considerably more susceptible to proteolysis than the protein bound to GroEL. This result can be interpreted as due to either lower protection of the newly synthesized protein by the endogenous RRL chaperones responsible for slowing down its folding or to an intrinsically more flexible conformation of the folding species they recognize and stabilize. The identity of this set of chaperones is still unknown. In any case, the conformation of newly synthesized pmAAT seems to be determined primarily by its intrinsic folding properties and not by interactions with the chaperones present in the RRL, which presumably control its rate of folding in these cell-free extracts.

Finally, given the similarity of the trypsinolysis patterns under the various conditions employed, it is also unlikely that a significant conformational rearrangement occurs in any AAT isozyme while bound to a chaperone such as GroEL; i.e. in the absence of ATP and GroES, GroEL traps similar folding intermediate(s) of both isozymes but binds the mitochondrial form with much higher affinity (13). The structural basis for this discrimination by GroEL between the two isozymes is still obscure. Differences in affinity for GroEL have been observed with other pairs of homologous isozymes such as the mitochondrial and cytosolic forms of malate dehydrogenase (44). In this case, tighter binding of the mitochondrial enzyme could be correlated with a higher overall hydrophobicity of this enzyme. However, no such correlation exists for the AAT enzymes (13). Other comparative studies have been reported using structurally related proteins from different species. For example, murine dihydrofolate reductase, but not the E. coli protein, forms a stable complex with GroEL (45). A short loop unique to the sequence of the mouse enzyme was identified as having an effect on binding (46), but whether it represents a specific recognition site for GroEL remains unclear. We can speculate that perhaps the few segments of the polypeptide sequences that are highly disimilar between pcAAT and pmAAT (12) might determine the strength of their interaction with GroEL. Alternatively, the higher protease resistance of refolding cAAT suggests that the stability of intermediates in the folding pathway per se could contribute to the relative affinity of each isozyme for binding to GroEL. Thus, a comparative analysis of the structural properties of these two homologous enzymes trapped by GroEL could shed light into the molecular mechanisms underlying recognition of putative substrates by GroEL.

Acknowledgments—We thank Dr. Juan I. Moreno for the preparation of one of the cDNAs for pmAAT-HT, Ellen Henderson for technical assistance, and John Bolin for the purification of GroEL.

REFERENCES

1. Eilers, M., and Schatz, G. (1986) Nature 322, 228–232
2. Chen, W. J., and Douglas, M. G. (1987) J. Biol. Chem. 262, 15605–15609
3. Schatz, G. (1986) J. Biol. Chem. 261, 15331–15337
4. Mihara, K., and Omura, T. (1996) Experientia 52, 1063–1068
5. Fedorov, A., and Baldwin, T. O. (1997) J. Biol. Chem. 272, 32715–32718
6. Ellis, R. J., and Harl, P. U. (1996) ASSEB J. 10, 19–26
7. Ranson, N. A., Dunster, N. J., Burston, S. G., and Clarke, A. R. (1995) Mol. Biol. 250, 581–586
8. Fenton, W. A., and Horwich, A. L. (1997) Protein Sci. 6, 743–760
9. Jansonius, J. N., and Vincent, M. G. (1987) in Biological Macromolecules and Assemblies (Jurnak, F., and McPherson, A. Eds) Vol. 3, pp. 187–285, John Wiley & Sons, New York
10. Malashkevich, V. N., Starokupovt, B. V., Borisov, V. V., Dauter, Z., Wilson, K. S. and Torchinsky, Y. M. (1995) J. Mol. Biol. 247, 111–124
11. Mattingly, J. R., Jr., Rodriguez-Remolar, F. J., Gordon, J., Iriarte, A., and Martinez-Carrion, M. (1987) Biochim. Biophys. Res. Commun. 149, 859–865
12. Mattingly, J. R., Jr., Iriarte, A., and Martinez-Carrion, M. (1993) J. Biol. Chem. 268, 36230–36237
13. Mattingly, J. R., Jr., Iriarte, A., and Martinez-Carrion, M. (1995) J. Biol. Chem. 270, 1138–1148
14. Reyes, A. M., Iriarte, A., and Martinez-Carrion, M. (1993) J. Biol. Chem. 268, 22281–22289
15. Artigues, A., Iriarte, A., and Martinez-Carrion, M. (1997) J. Biol. Chem. 272, 16852–16861
16. Lain, B., Iriarte, A., and Martinez-Carrion, M. (1994) J. Biol. Chem. 269, 15588–15596
17. Lain, B., Iriarte, A., Mattingly, J. R., Jr., Moreno, J. I., and Martinez-Carrion, M. (1995) J. Biol. Chem. 270, 24722–24729
18. Iriarte, A., Hubert, E., Kraft, R., and Martinez-Carrion, M. (1984) in Chemical and Biological Aspects of Vitamin B6 Catalysis (Evangelopoulos, A. E., Ed) Part B, pp. 269–276, Alan R. Liss, Inc., New York
19. Hubbard, S. J., Eisenmenger, F., and Thornton, J. M. (1994) Protein Sci. 3, 757–768
20. Mattingly, J. R., Jr., Youssef, J., Iriarte, A., and Martinez-Carrion, M. (1993) J. Biol. Chem. 268, 3925–3932
21. Paves-Peux, M., Ferry, N., Stup, J., Hanoue, J., and Barouki, R. (1988) J. Biol. Chem. 263, 17459–17466
22. Martinez-Carrion, M., Ahler, F., Iriarte, A., Mattingly, J. R., Jr., Youssef, J., and Wu, T. (1990) Ann. N. Y. Acad. Sci. 585, 346–356
23. Torella, C., Mattingly, J. R., Jr., Artigues, A., Iriarte, A., and Martinez-Carrion, M. (1998) J. Biol. Chem. 273, 3915–3925
24. Bux, A. M., Zahn, R., and Persat, A. R. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 3571–3575
25. Miranker, A. D., and Dobson, C. M. (1996) Curr. Opin. Struct. Biol. 6, 31–42
26. Dill, K. A., and Chan, H. S. (1997) Nat. Struct. Biol. 4, 10–19
27. Cui, K., and Schirch, V. (1997) J. Biol. Chem. 271, 2987–2994
28. Parker, M. J., Spencer, J., Jackson, G. S., Burston, S. G., Hossza, L. L., Craven, C. J., Waltho, J. P., and Clarke, A. R. (1996) Biochemistry 35, 15740–15752
29. Sandstrom, E., and Christensen, P. (1980) J. Biol. Chem. 255, 10284–10290
30. McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1995) J. Mol. Biol. 225, 495–517
31. Braig, K., Otwinowski, Z., Hegde, R., Boivert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1995) Nature 371, 578–586
32. Fenton, W. A., Kashi, Y., Partk, K., and Horwich, A. L. (1994) Nature 371, 614–619
33. Brazzi, B. T., Cleland, J. L., McDowell, R. S., Horwich, A. L., and Sigler, P. B. (1994) Nature 371, 578–586
34.分辨率，但不与 E. coli 蛋白质形成稳定复合物。
Trypsinolysis of Refolding AAT Isozymes

642–645

39. Persson, M., Carlsson, U., and Bergenhjem, N. C. (1996) Biochim. Biophys. Acta 1298, 191–198
40. Walter, S., Lorimer, G. H., and Schmid, F. X. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 9425–9430
41. Gervasoni, P., Staudenmann, W., James, P., Gehrig, P., and Pluckthun, A. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 12189–12194
42. Goldberg, M. S., Zhang, J., Sondek, S., Matthews, C. R., Fox, R. O., and Horwich, A. L. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 1080–1085
43. Doonan, S., Doonan, H. J., Hanford, R., Vernon, C. A., Walker, J. M., Airoldi, L., Bossa, F., Barra, D., Carloni, M., Fasella, P., and Riva, F. (1975) Biochem. J. 149, 497–506
44. Staniforth, R. A., Cortés, A., Burston, S. G., Atkinson, T., and Clarke, A. R. (1994) FEBS Lett. 344, 129–135
45. Clark, A. C., Hughes, E., and Frieden, C. (1996) Biochemistry 35, 5893–5901
46. Clark, A. C., and Frieden, C. (1997) J. Mol. Biol. 268, 512–525