

# Divergent Hsc70 Binding Properties of Mitochondrial and Cytosolic Aspartate Aminotransferase

IMPLICATIONS FOR THEIR SEGREGATION TO DIFFERENT CELLULAR COMPARTMENTS\*

(Received for publication, September 4, 1998, and in revised form, October 16, 1998)

Antonio Artigues, Douglas L. Crawford, 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

**Cytosolic Hsc70 discriminates between the homologous mitochondrial and cytosolic isozymes of aspartate aminotransferase, binding exclusively the mitochondrial form. By screening a library of synthetic peptides spanning the sequence of the mitochondrial enzyme, we have identified binding sites in this polypeptide that interact with Hsc70. These potential binding sites are scattered over the entire sequence and map to secondary structure elements, particularly the  $\alpha$ -helix, that are partly exposed on the surface of the native protein. Several peptides corresponding to analogous positions in the cytosolic enzyme sequence do not bind to Hsc70. Phylogenetic analyses suggest that Hsc70 binding sequences have diverged as a consequence of biochemical specialization ensuring differential interaction of each isozyme with the cellular machinery in charge of protein folding and translocation.**

The Hsp70<sup>1</sup> family of molecular chaperones comprises a set of abundant proteins that assist a large variety of protein folding processes in different cellular compartments (1, 2). The ATP-regulated transient association of Hsp70 with unfolded polypeptides prevents either further folding or aggregation of the chains (3, 4). Hsp70 proteins display a certain degree of peptide sequence specificity, which extends beyond recognition of the unfolded states. Several studies using random synthetic peptide or peptide-display phage libraries have revealed the preference of Hsp70s for binding of short peptides (heptamers) enriched in hydrophobic residues (5–7). However, different members of the Hsp70 family differ to some extent in their peptide sequence specificities. For instance, the presence of basic residues favors binding of peptides to bacterial DnaK and Hsc70, the cytosolic heat shock cognate protein, but not to BiP, the endoplasmic reticulum Hsp70 (8, 9). Although high affinity binding sites for Hsp70 have been identified in the sequence of a few proteins (10–12), the mechanism of substrate recognition

by Hsp70, particularly in the context of full-length protein substrates, remains unclear.

In this report we analyze the molecular basis for the specific interaction of Hsc70 with only one of two homologous eukaryotic proteins, both of which are synthesized on cytoplasmic ribosomes but localize to different cellular compartments. The cytosolic (cAAT) and mitochondrial (mAAT) isozymes of rat liver aspartate aminotransferase have approximately 50% sequence identity and have nearly identical three-dimensional structures (13), but only mAAT interacts with Hsc70 either during synthesis in cell-free extracts (14) or refolding *in vitro* from acid-denatured states (15). Because Hsc70 appears to play a role in the translocation of proteins into mitochondria (16–18), the exclusive interaction of mAAT with Hsc70 may be important to ensure the segregation of the isozymes to their respective cellular locations following synthesis in the cytoplasm. As is true for most translocated mitochondrial proteins, rat liver mAAT is synthesized as a precursor (pmAAT) with a N-terminal 29-residue presequence peptide (19) that targets the protein for translocation into mitochondria. Mitochondrial presequences have been shown to bind to cytosolic Hsp70s (14, 20, 21). However, because mAAT lacking the presequence peptide also shows affinity for Hsc70 (14), additional recognition sites must exist within the mature portion of the protein. We used a collection of synthetic peptides spanning the complete sequence of mAAT, or selected regions of cAAT, to identify binding sites for Hsc70 in the sequence of these enzymes. Phylogenetic analyses of different AAT isozymes were also carried out to explore possible evolutionary features of the Hsc70 binding sites identified in mAAT.

## EXPERIMENTAL PROCEDURES

**Protein Purification**—Purification of pmAAT was carried out as described previously (22). Protein concentrations were estimated using the absorbance of the bound pyridoxal 5'-phosphate cofactor at 356 nm and a molar absorption coefficient of 8,500 M<sup>-1</sup> cm<sup>-1</sup> and  $M_r = 46,597$  for pmAAT. Hsc70 was purified from bovine brain following published procedures (23). Following the last ammonium sulfate precipitation, the protein was exhaustively dialyzed against 25 mM Tris-HCl, 20 mM NaCl, 10 mM  $\beta$ -mercaptoethanol and kept at 10 °C until use. Hsc70 protein concentration was measured from the calculated molar absorption coefficient at 280 nm of 47,800 M<sup>-1</sup> cm<sup>-1</sup> and  $M_r = 70,000$  (24).

**Unfolding and Refolding of pmAAT**—Acid unfolding of pmAAT was performed by titrating a stock solution of the enzyme in 2 mM Tris-HCl, pH 7.5 to pH 2.0, with diluted HCl (final protein concentration, 8  $\mu$ M), followed by incubation for 90 min at room temperature. Refolding of pmAAT was initiated by rapid dilution of the unfolded protein at 10 °C in refolding buffer (40 mM Hepes, 0.1 mM EDTA, 1 mM dithiothreitol, 10  $\mu$ M pyridoxal 5'-phosphate, pH 7.5) to 1.8  $\mu$ M final monomer concentration. Reactivation was followed by monitoring the appearance of transaminase activity over time.

**Enzymatic Activities**—The pmAAT transaminase activity was measured at 37 °C using aspartate and  $\alpha$ -ketoglutarate as substrates, using a coupled assay with malate dehydrogenase as described previously (25). To measure the ATPase activity of Hsc70, the chaperone (0.5  $\mu$ M)

\* This work was supported by National Institutes of Health Grants HL-38412 and GM-38341 (to M. M.-C.) and National Science Foundation Grant IBN-9419781 (to D. L. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri, 5007 Rockhill Rd., Kansas City, MO 64110. Tel.: 816-235-5246; Fax: 816-235-5158; E-mail: mshoemaker@cctr.umkc.edu.

<sup>1</sup> The abbreviations used are: Hsp70, 70-kDa heat shock protein; AAT, aspartate aminotransferase; cAAT, cytosolic aspartate aminotransferase; mAAT, mitochondrial aspartate aminotransferase; pmAAT, precursor to mitochondrial aspartate aminotransferase; HPLC, high pressure liquid chromatography; Hsc70, heat shock cognate 70-kDa protein.

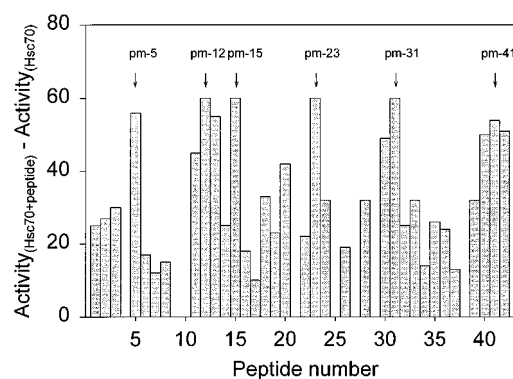
was incubated at 37 °C in 40 mM Hepes, 45 mM KCl, 120  $\mu$ M MgCl<sub>2</sub>, 60  $\mu$ M ATP, pH 7. At different incubation times, a 20- $\mu$ l aliquot was withdrawn and assayed for ATP content on a Turner TD 20-e luminometer, using the ATP bioluminescence assay kit from Sigma and following the manufacturer's instructions. The assay estimates ATP concentration from light emitted when firefly luciferase catalyzes the formation of adenylate-luciferin, which spontaneously decomposes (26). When ATP is the limiting reagent, the light emitted is proportional to the ATP present in the sample, and the concentration of ATP can be calculated by reference to an ATP concentration standard curve. The rate of spontaneous hydrolysis of ATP was estimated in samples incubated at 37 °C in the absence of Hsc70.

**Synthesis and Purification of Peptides**—The collection of synthetic peptides spanning the complete amino acid sequence of rat liver pmAAT (19) was a generous gift from Dr. B. M. Conti-Fine (University of Minnesota). The collection consists of 43 peptides (pm-1 to pm-43) starting with the N-terminal sequence of the pmAAT presequence corresponding to 14-residue sequences of the pmAAT sequence (except pm-43, which has only 10 residues). Each peptide contains a 4-residue overlap with the previous peptide in the pmAAT amino acid sequence. The peptides were synthesized as described previously (27). The purity of the peptides was assessed by reverse phase HPLC using a C18 column (Vydac 218TP, 250, 4.6 mm) and an acetonitrile/water gradient (5–70% over 30 min) containing 0.1% trifluoroacetic acid. The purity of the peptide preparations ranged from 50 to 95%. Most of the contaminating peptides represented truncated peptides randomly missing amino acids from incomplete coupling. Selected peptides were further purified by reverse phase HPLC on a Bio-Rad C18 HiPore RP-318 semipreparative column (250, 10 mm), using the same gradient as before. Major peaks were collected, and the full-length peptide peak was identified by mass spectrometry and amino acid composition analyses. The presequence peptide, MALLHSGRVLSGMAAAFHPGLAAAA-SARA (pre-p), and the 14-residue peptides from the rat liver cAAT sequences (28) corresponding to several Hsc70-binding regions in pmAAT, SFAQVQPAPPVLV (pc-5), GADFLGRWYIGTDN (pc-15), VLRVLSQMEKIVRI (pc-31), and VEYLVNEKHIYLM (pc-41), were synthesized in an Applied Biosystems 433A peptide synthesizer at the Molecular Core Facility and purified by reverse phase HPLC as described above.

**Phylogenetic Analyses**—Phylogenetic trees of mAAT proteins were obtained by parsimony analysis of AAT sequences after alignment with ClustalW and Macvector software from Oxford Molecular (version 6.0, 1997). Parsimony analyses of AAT protein sequences were performed with the use of PAUP (29). In all cases, the amino acid sequences associated with the leader sequence were excluded. For phylogenetic analysis, sequences identified as aspartate aminotransferase were obtained from GenBank (<http://www.ncbi.nlm.nih.gov>) and Swiss-Prot protein data base (<http://expasy.hcuge.ch>): *Bos taurus* cAAT (BovC), P33097; *B. taurus* mAAT (BovM), P12344; *Gallus gallus* cAAT (ChkC), P00504; *G. gallus* mAAT (ChkM), P00508; *Rattus norvegicus* cAAT (RatC), PIR JT0439; *R. norvegicus* mAAT (RatM), P00507; *Mus musculus* cAAT (MusC), P05201; *M. musculus* mAAT (MusM), P5202; *Homo sapiens* mAAT (HumM), P00505; *Sus scrofa* cAAT (PigC), P00503; *S. scrofa* mAAT (pigM), P00506; *Medicago sativa* cAAT (AlfC), P28011; *Arabidopsis thaliana* cAAT (ArbC), P46645; *A. thaliana* mAAT (ArbM), P46643; *A. thaliana* chloroplast precursor AAT (ArbP), P46644; *Oryza sativa* cAAT (RicC), P37833, PIR JC5124; *O. sativa* mAAT (RicM), PIR JC5125, D67643; *Panicum millaceum* cAAT (MilC), PIR S22378; *P. millaceum* mAAT (MilM), PIR S22379; *Caenorhabditis elegans* cAAT (CelC), Q22067; *C. elegans* mAAT (CelM), U39645; *Saccharomyces cerevisiae* cAAT (YstC), PIR S64854; *S. cerevisiae* mAAT (YstM), PIR S37933; *Escherichia coli* (Eco), P00509.

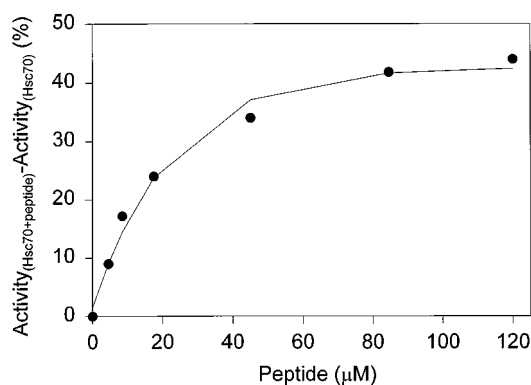
## RESULTS AND DISCUSSION

**Identification of Hsc70 Binding Sites**—The productive binding and release of substrates from Hsp70 proteins are coupled to the ATPase activity of the chaperone, which is controlled by the action of several cochaperones (3, 4). Consequently, in the absence of nucleotides and cochaperones, the binding of substrate to Hsc70 is basically irreversible and will impair the binding of a different substrate. In addition, the weak ATPase activity of Hsc70 is stimulated upon binding of substrate (30, 31). Thus, substrate binding to Hsc70 can be studied either by following competition with binding of another substrate (8) or by measuring the stimulation of Hsc70 ATPase activity. Because Hsc70 binds and stops refolding of acid-unfolded pmAAT



**FIG. 1. Binding of pmAAT peptides to Hsc70.** Hsc70 (1.8  $\mu$ M) was incubated for 16 h at 10 °C in refolding buffer containing 120  $\mu$ M of synthetic peptides corresponding to 14-residue overlapping segments of pmAAT. After addition of acid-unfolded pmAAT at a concentration of 1.8  $\mu$ M, the samples were incubated for an additional 120 min at 10 °C to allow complete refolding of the pmAAT that was not complexed with Hsc70. At the end of the incubation, the activity recovered ( $Activity_{Hsc70+peptide}$ ) was compared with that of a sample of pmAAT refolded in the presence of Hsc70 alone ( $Activity_{Hsc70}$ ). Data are expressed as percentage relative to the activity of a control native enzyme incubated under identical conditions. The maximum yield of pmAAT reactivation in the absence and presence of Hsc70 alone was  $75 \pm 5\%$  and  $20 \pm 2\%$ , respectively. The peptides showing the greatest competition effect are indicated on top.

(15), competition by a synthetic peptide should result in an increase in the fraction of pmAAT activity recovered. Thus, binding of the presequence peptide (pre-p) and 43 synthetic tetradecamers spanning the entire amino acid sequence of pmAAT was tested by comparing the yield of pmAAT reactivation in the presence of Hsc70 alone and Hsc70 that had been preincubated with a 70-fold molar excess of each peptide (32). These results are summarized in Fig. 1. Several peptides could not be tested either because they interfere with spontaneous refolding of pmAAT (N-terminal pm-4) or have low solubility in aqueous solutions. Seven peptides (pre-p, pm-5, pm-12, pm-15, pm-23, pm-31, and pm-41) were the most competitive. For these seven peptides, this competitive effect is concentration-dependent (data for the presequence is shown in Fig. 2), with  $k_{app}$  values in the 4–10  $\mu$ M range (Table I). Dissociation constants in the low micromolar range have been reported for the binding to different Hsp70 proteins of a number of small peptides (8–11, 33) as well as intact polypeptide ligands that remain stably unfolded under native conditions (34). As is true for most unfolded proteins, mAAT rapidly collapses into partially folded intermediate states upon transfer to native-like conditions. The affinity of intact mAAT (or pmAAT) for Hsc70 could not be determined because the initial interaction of unfolded mAAT with Hsc70 under native conditions results in the formation of insoluble aggregates of the chaperone-pmAAT complex (15). The differences in  $k_{app}$  may indicate that affinity is sequence-specific, but it is also possible that for each region of pmAAT the true binding sequence has been spliced into several fragments on sequential synthetic peptides, so that at least some of the peptides may not contain the complete high affinity binding sequence. In fact, for several regions of the pmAAT sequence, the peptides around that which shows maximum competition also bind to Hsc70 (Fig. 1). Stimulation of the ATPase activity correlated well with peptide binding data obtained from competition experiments. Those peptides showing maximal competition with pmAAT for binding to Hsc70 induced an ATPase stimulation ranging from 1.5- to 2-fold (Table I). According to these results, we can tentatively conclude that any of these potential Hsc70 binding sites within the pmAAT sequence could be responsible for the observed inter-



**FIG. 2. Peptide concentration dependence of the competition between the presequence and pMAAT for binding to Hsc70.** Competition assays were performed as described in the legend to Fig. 1 except that Hsc70 was preincubated with different concentrations of the synthetic presequence peptide for 120 min at 10 °C in refolding buffer. The data were fitted to a single exponential equation (solid line), and the  $k_{app}$  value was estimated as the concentration of peptide at which 50% of the maximum reduction in pMAAT reactivation is achieved as reported in Table I.

TABLE I

*Hsc70 binding properties of synthetic peptides*

Binding of synthetic peptides to Hsc70 was tested by using the competition assay described in the legend to Fig. 1.

Peptide sequence <sup>a</sup>	Code	$k_{app}$ <sup>b</sup>	ATPase activity <sup>c</sup>	Stimulation <sup>d</sup>
		$\mu\text{M}$	$\text{nmol/min/mg}$	
None			0.92	1.00
<sup>-29</sup> <b>MALL</b> . . . <b>ASARA</b> <sup>-1e</sup>	pre-p	6.0	1.70	1.85
<b>MALL</b> HSGRVLSGMA <sup>-16</sup>	pm-1	1.8	ND <sup>f</sup>	ND
<b>SGMAA</b> AHFPLGAAA <sup>-6</sup>	pm-2	2.8	ND	ND
<b>LAAA</b> ASARASSWWT <sup>7</sup>	pm-3	3.7	ND	ND
<b>PDP</b> ILGVTEAFKRD <sup>27</sup>	pm-5	9.0	1.82	1.98
<b>PVL</b> VFKLIADFRDD	pc-5	NC <sup>g</sup>	0.75	
<b>ELAL</b> GENSEVLKSG <sup>98</sup>	pm-12	3.2	1.75	1.90
<b>GAS</b> FLQRFKFSRD <sup>130</sup>	pm-15	10.9	2.71	2.95
<b>GAD</b> FLGRWYIGTDN	pc-15	NC	0.98	1.07
<b>PRPE</b> QWKEMA <sup>213</sup> AVV	pm-23	4.1	1.63	1.77
<b>AKR</b> VESQLKILIRP <sup>293</sup>	pm-31	3.5	2.02	2.20
<b>VLR</b> VLSQMEKIVRI	pc-31	NC	0.89	0.97
<b>VER</b> LTKF <sup>383</sup> SVYMTK	pm-41	8.9	1.30	1.41
<b>VEY</b> LVNEKHIYLM	pc-41	NC	0.91	0.99
<b>FFD</b> MAYQGFASGDG <sup>233</sup>	pm-25	NC	ND	ND
<b>LAH</b> AHQVTK <sup>410</sup>	pm-43	NC	1.00	1.09

<sup>a</sup> Boldface is used to indicate mitochondrial peptides. The corresponding cytosolic peptides tested are included immediately below in italics. The numbers at the C terminus of each sequence identify the location of the peptide in the primary structure (19). The numbering of the residues is according to that of cAAT (412 residues), introducing gaps in the sequence of mAAT (402 residues) when appropriate to maximize homology. Negative numbering is used for the presequence peptide starting with the C terminus as residue -1.

<sup>b</sup>  $k_{app}$  values correspond to the concentration of peptide providing half of the maximum change in reactivation yields and were calculated after curve fitting of the experimental values to a single exponential equation, as shown in Fig. 2 for the presequence peptide.

<sup>c</sup> The ATPase activity of Hsc70 was measured by monitoring ATP hydrolysis using a bioluminescence assay as described under "Experimental Procedures."

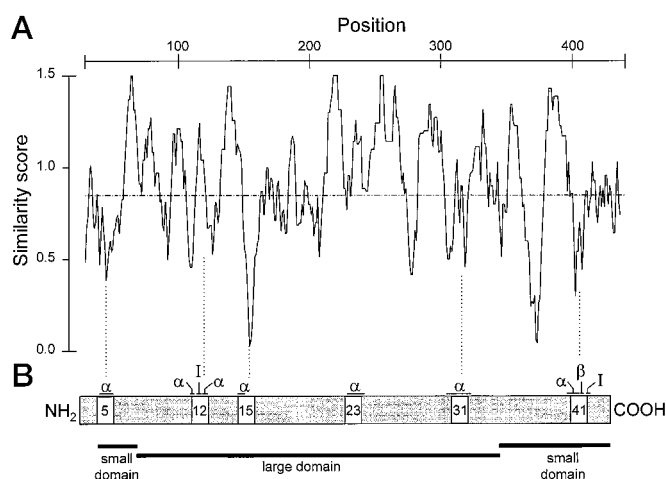
<sup>d</sup> Hsc70 ATPase activity in the presence of peptide relative to the basal activity measured in the absence of peptide.

<sup>e</sup> Only the N- and C-terminal end sequences are given for the intact 29-residue presequence peptide.

<sup>f</sup> ND, not determined; NC, does not compete with pMAAT binding to Hsc70.

action of the full-length protein.

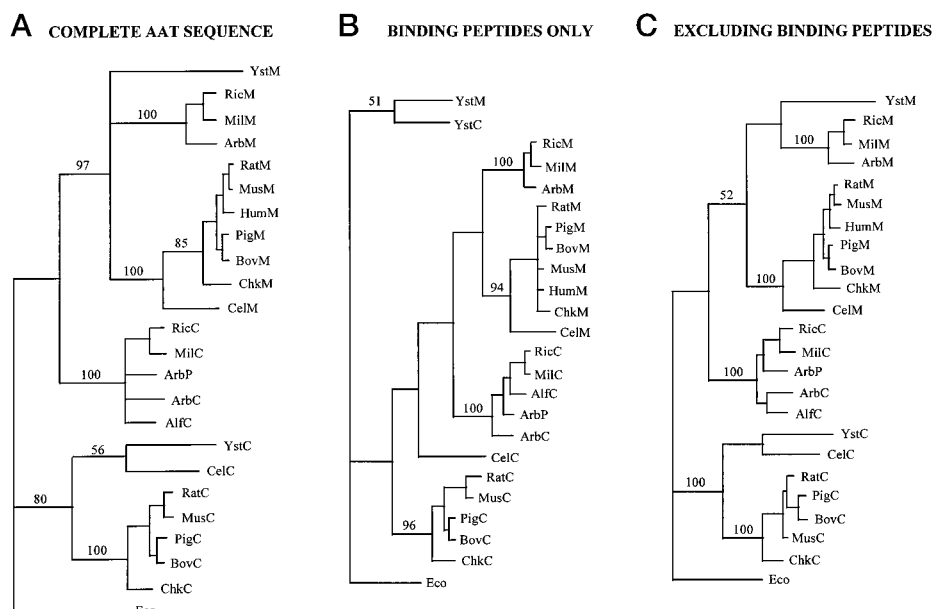
To test whether the inability of Hsc70 to recognize full-length cAAT as a substrate (14, 15) was because of the absence of potential recognition sites, the protein sequences in rat cAAT corresponding to four of the Hsc70-binding regions in pMAAT were identified by aligning the sequences of the two rat



**FIG. 3. Structural comparison between the AAT isozymes and location of potential binding sites for Hsc70 in mAAT.** A, similarity plot between rat liver mAAT and cAAT. The average sequence homology between cAAT and mAAT was calculated using the plot similarity program included in the Wisconsin Package of the Genetics Computer Group suite of programs (version 8.0, 1984) with a window size of seven residues, after the proteins were aligned inserting gaps where necessary to maximize homology. A score of 1.5 corresponds to a region of perfect homology. The dotted line represents the overall average similarity between the two proteins. B, the positions of potential binding sites for Hsc70 are indicated by white boxes in the bar representing the mAAT polypeptide. The peptides are identified by numbers as indicated in Fig. 1 and Table I. The native secondary structural elements predominating in the binding regions are marked above the bar as  $\alpha$  ( $\alpha$ -helix),  $\beta$  ( $\beta$ -sheet), and I (type I  $\beta$ -turn). The major domains in the mAAT structure (36) are indicated below the bar.

isozymes (19, 28). The four cAAT tetradecamer peptides (pc-5, pc-15, pc-31, and pc-41) were synthesized and used in competition assays as described before. None of the cytosolic peptides competed with unfolded pMAAT for binding to Hsc70 or stimulated the ATPase activity of the chaperone (Table I). It appears that cAAT may actually lack recognition sites for Hsc70, at least in regions analogous to those identified in mAAT.

The potential binding sites for Hsc70 identified in mAAT are scattered over the entire sequence of the protein in regions of both the small and large domains that in the native protein are folded into secondary structure elements, primarily  $\alpha$ -helix (Fig. 3). A recent NMR solution structure of the DnaK substrate binding domain (35) shows the peptide binding groove occupied with the extended C-terminal tail of itself, which is  $\alpha$ -helical in native DnaK. Interestingly, the mAAT binding regions are for the most part exposed on the surface of the native protein (36). The local secondary structure of the binding sites rather than their degree of exposure on the native structure might determine the discrimination between properly folded and unfolded mAAT by Hsc70. This implies that, because Hsp70 proteins bind peptide substrates in an extended conformation (5–7), either the chaperone is able to unwind structured protein segments in the partially unfolded protein or only extensively unfolded conformations of mAAT are suitable substrates for Hsc70. Such an essentially unfolded conformation is expected for nascent chains on cytoplasmic ribosomes or imported polypeptides emerging into the matrix, presumably the pMAAT substrates for cytosolic and mitochondrial Hsp70, respectively, in the intact cell. Both Hsp70s are important components of the machinery responsible for translocation of nuclear-encoded proteins into mitochondria (16–18, 37–39). Cytosolic Hsc70 appears to be involved in maintaining translocated proteins in a partially unfolded, import-competent conformation until they engage the translocation machinery in the target membrane. Proteins remaining in the cytoplasm such as



**FIG. 4. Phylogenetic analyses of AAT.** Parsimony analyses of the complete AAT amino acid sequence (excluding the presequence peptides) (A), using just the six Hsc70 binding sequences shown in Table I (B), or excluding these six binding regions (C). The dendrograms were constructed from alignment of the AAT proteins as indicated under "Experimental Procedures." Bootstrap values greater than 50% ( $n = 500$ ) are listed only for branches leading to plant (*Ric*, *Mil*, *Arb*, and *Alf*), yeast (*Yst*) or vertebrate (*Rat*, *Pig*, *Bov*, *Mus*, *Hum*, and *Chk*) cytosolic (C) and mitochondrial (M) forms of AAT. They are not listed for branches within these nodes. See "Experimental Procedures" for an explanation of the abbreviations of the sources of the enzymes analyzed.

cAAT would not have this requirement and therefore could afford to fold quickly without the intervention of Hsc70. This proposal is supported by our previous finding that, after translocation in cell-free extracts, cAAT folds much faster than (p)mAAT (40). Whether Hsc70 displays similar selectivity toward other sets of proteins according to their cellular destinations is unknown. Many nascent chains have been found associated with Hsc70 in mammalian cells (41, 42), but the distribution between cytosolic residents and proteins destined for translocation among them was never investigated.

The 14-residue binding sequences shown in Table I display some of the features of the consensus motifs proposed for binding to BiP (large hydrophobic/aromatic residues) (7, 43) or DnaK (three to five hydrophobic residues flanked by basic residues) (12). Although no consensus sequence has been proposed for cytosolic Hsc70, this chaperone seems to favor peptides with hydrophobic residues flanked on either side by basic residues (8, 9, 33). All of the peptides in Table I have basic residues either following at least two hydrophobic residues (pm-5 and pm-12) or flanking small clusters of hydrophobic residues (pm-15, pm-23, pm-31). However, given the small number of binding peptides identified, their size about twice the apparent length of the Hsp70 binding motif (seven residues) (7), and the presence of some of the features described above in the non-binding cAAT peptides, it was not possible to develop a clear consensus sequence.

**Phylogenetic Analyses**—The failure of rat liver cAAT peptides to bind to Hsc70 and the mapping of the binding sites identified to regions in mAAT with the lowest similarity score between the two isozymes (Fig. 3) suggest that differences in primary structure are likely responsible for the selective recognition of mAAT by Hsc70. The generality of this concept can be explored by phylogenetic analysis of different AAT isozymes. The three cladograms shown in Fig. 4 represent the phylogenetic analyses using the (i) complete AAT amino acid sequence (excluding the presequence peptide), (ii) 84 amino acids from six Hsc70 binding peptides (pm-5, pm-12, pm-15, pm-23, pm-31, and pm-41; see Table I), and (iii) AAT sequence excluding these six peptides. All three analyses indicate that the duplication of mAAT and cAAT is ancient, as previously proposed (44), and that all plant and vertebrate mAATs share a common origin. Analyses with the complete AAT sequences or sequences that exclude the six binding peptides (Fig. 4, A and C) indicate that yeast mAAT shares a common origin with all

other mAAT proteins, and yeast cAAT shares a common origin with animal cAATs. Both observations are supported by large bootstrap values (97 and 80%, respectively). However, there is one striking difference among the three trees; when using the six Hsc70 binding peptides, yeast mAAT does not group with the other mitochondrial forms nor does the yeast cAAT group with the other cytosolic forms (compare B with A or C in Fig. 4). This observation suggests that yeast cytosolic Hsp70s may recognize different targeting sequences than Hsp70s from other eukaryotes. In fact, the affinity of Ssa1p, a yeast cytosolic Hsp70, for mitochondrial presequences appears to depend less on the specific primary sequence of a peptide than on the physical properties of the peptide such as its amphiphilic character (21).

The variation in the topologies of the phylogenetic trees suggests that the six binding sequences are evolving differently than the rest of the AAT protein. This supposition is supported by the analysis of amino acid variation among vertebrate AATs. Specifically, the six binding peptides in vertebrate mAATs have significantly less variation than the same 84 amino acids in cAAT (12% for mitochondrial and 34% for cytosolic; chi-square  $p < 0.01$ ). This is not because of more constraints on the mitochondrial form as a whole because in the protein excluding the six binding peptides, both mitochondrial and cytosolic isoforms have approximately 19% variation (Table II). The previous observation of a greater conservation in mAAT (45) is thus supported only for the regions of the protein that bind Hsc70. The unique features of the Hsc70 binding regions *versus* the rest of the protein are also reflected in the amount of fixed differences between the two vertebrate isoforms. (A fixed difference is a variable site where all proteins of one isozyme type share a common amino acid that is different from the homologous site in the other isozyme.) There are more fixed differences between cAAT and mAAT in the binding peptides (51% or 43/84) than in remaining portions of the protein (39% or 134/325). That is, among all vertebrate mAAT and cAAT proteins, these two isoforms are more likely to have different amino acid residues in the six Hsc70 binding regions than in any other section of the protein.

The evolutionary analyses also shed light on the sequence features that may determine or at least contribute to peptide recognition by Hsc70. Comparison of the regions in the mAAT proteins that bind to Hsc70 with the non-binding regions reveals that non-binding regions have fewer consecutive hydro-

TABLE II  
Variation within and between mAAT and cAAT proteins

The number and percentage of variable sites for the six Hsc70 binding sequences and the remaining non-binding regions were calculated for the vertebrate cAAT (5 sequences) and mAAT (6 sequences) proteins included in Fig. 4. Notice that the cAAT sequences that are homologous to the Hsc70 binding regions from mAAT have a significantly greater amount of variable sites (chi-square,  $p < 0.01$ ). Number of overall variable sites is not a simple summation because some variable sites within one isoform are also variable in the other isoform.

	Number of variable sites		
	mAAT	cAAT	Overall
Binding sequences (84 residues)	10 (12%)	29 (34%)	34 (40%)
Non-binding sequences (325 residues)	63 (19%)	61 (19%)	107 (33%)

phobic residues. However, comparing analogous peptides from vertebrate cAATs and mAATs indicates that binding to Hsc70 is not determined solely by the number or distribution of hydrophobic residues in the sequence. For example, peptides pm-5, pm-31, and pm-41 have fewer total or consecutive hydrophobic residues than their corresponding peptides from cAAT; yet the latter do not bind Hsc70. Additionally, the number of basic residues or their localization relative to hydrophobic residues does not appear to represent binding determinants either. There is, however, one striking pattern, (F/A)KR, detected in the binding peptides pm-5 and pm-31 that always occurs in the corresponding homologous peptides among vertebrate mAATs but is never found in any other vertebrate cAAT sequences or in the non-binding regions of mAAT. This short sequence might therefore represent an important element of at least one of the recognition motifs for Hsc70.

**Concluding Remarks**—Hsc70 interacts specifically with several mAAT peptides distributed throughout the complete protein sequence. The phylogenetic data indicate that the six binding peptides analyzed are evolving differently than the rest of the protein. These 84 amino acids (i) produce a phylogenetic topology that differs from the complete sequences, (ii) are either more (cAAT) or less (mAAT) variable, and (iii) have more fixed differences between the two isozymes. Together these molecular and evolutionary analyses indicate that sequence divergence observed between the mitochondrial and cytosolic isozymes might have evolved for the differential interaction of each enzyme form with the cellular machinery in charge of protein folding and sorting, and this may promote efficient translocation into mitochondria for mAAT and rapid folding in the cytosol for cAAT.

**Acknowledgment**—We thank B. M. Conti-Fine for providing the collection of pmAAT synthetic peptides.

#### REFERENCES

- Gething, M. J., and Sambrook, J. F. (1992) *Nature* **355**, 33–45
- Hartl, F. U. (1996) *Nature* **381**, 571–579
- McKay, D. B. (1993) *Adv. Protein Chem.* **44**, 67–98
- Bukau, B., and Horwich, A. L. (1998) *Cell* **92**, 351–366
- Zhu, X., Zhao, X., Burkholder, W. B., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrikson, W. A. (1996) *Science* **272**, 1606–1614
- Landry, S. J., Jordan, R., McMacken, R., and Giersach, L. M. (1992) *Nature* **355**, 455–457
- Flynn, G. C., Pohl, M. T., Flocco, M. T., and Rothman, J. E. (1991) *Nature* **353**, 726–730
- Fourie, A. M., Sambrook, J. F., and Gething, M.-J. H. (1994) *J. Biol. Chem.* **269**, 30470–30477
- Gragerov, A., and Gottesman, M. E. (1994) *J. Mol. Biol.* **241**, 133–135
- Gething, M.-J. H., Blond-Elguindi, S., Buchner, J., Fourie, A., Knarr, G., Modrow, S., Nanu, L., Segal, M., and Sambrook, J. (1995) *Cold Spring Harbor Symp. Quant. Biol.* **60**, 417–428
- Fourie, A. M., Hupp, T. R., Lane, D. P., Sang, B.-C., Barbosa, M. S., Sambrook, J. F., and Gething, M.-J. H. (1997) *J. Biol. Chem.* **272**, 19471–19479
- Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997) *EMBO J.* **16**, 1501–1507
- Jansonius, J. N., and Vincent, M. G. (1987) in *Biological Macromolecules and Assemblies* (Jurnak, F., and McPherson, A., eds) Vol. 3, pp. 87–285, John Wiley & Sons, Inc., New York
- Lain, B., Iriarte, A., Mattingly, J. R., Jr., Moreno, J. I., and Martinez-Carrion, M. (1995) *J. Biol. Chem.* **270**, 24732–24739
- Artigues, A., Iriarte, A., and Martinez-Carrion, M. (1997) *J. Biol. Chem.* **272**, 16852–16861
- Deshais, R. J., Koch, B. D., Werne-Washburne, M., Craig, A., and Scheman, R. (1988) *Nature* **332**, 800–805
- Sheffield, W. P., Shore, G. C., and Randall, S. K. (1990) *J. Biol. Chem.* **265**, 11069–11076
- Terada, K., Ohtsuka, K., Imamoto, N., Yoneda, Y., and Mori, M. (1995) *Mol. Cell. Biol.* **15**, 3708–3713
- Mattingly, J. R., Jr., Rodriguez-Bercoval, F. J., Gordon, J., Iriarte, A., and Martinez-Carrion, M. (1987) *Biochem. Biophys. Res. Commun.* **149**, 859–865
- Schmid, D., Baici, A., Gehring, H., and Christen, P. (1994) *Science* **263**, 971–973
- Endo, T., Mitsui, S., Nakai, M., and Roise, D. (1996) *J. Biol. Chem.* **271**, 4161–4167
- Altieri, F., Mattingly, J. R., Jr., Rodriguez-Bercoval, F. J., Youseff, J., Iriarte, A., Wu, T., and Martinez-Carrion, M. (1989) *J. Biol. Chem.* **264**, 4782–4786
- Welch, W. J., and Feramisco, J. R. (1985) *Mol. Cell. Biol.* **5**, 1229–1237
- Palleros, D. R., Welch, W. J., and Fink, A. L. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5719–5723
- Martinez-Carrion, M., Turano, C., Chiacone, E., Bossa, F., Giartossio, A., Riva, F., and Fasella, P. (1967) *J. Biol. Chem.* **242**, 2397–2409
- Leach, F. R., and Webster, J. J. (1986) *Methods Enzymol.* **133**, 51–70
- Houghten, R. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 5431–5435
- Pavé-Preux, M., Aggerbeck, M., Veyssier, C., Bousquet-Lemerrier, B., Hanoune, J., and Baruki, R. (1998) *J. Biol. Chem.* **263**, 17459–17466
- Swofford, D. L. (1989) in *PAUP: Phylogenetic Analysis Using Parsimony*, version 3.0b, Illinois Natural History Survey, Champaign, IL
- Ha, J.-H., and McKay, D. B. (1994) *Biochemistry* **33**, 14625–14635
- Gao, B., Greene, L., and Eisenberg, E. (1994) *Biochemistry* **33**, 2048–2054
- Artigues, A., Iriarte, A., and Martinez-Carrion, M. (1997) in *Techniques in Protein Chemistry* (Martshak, D. R., ed) Vol. VIII, pp. 481–492, Academic Press, New York
- Takenaka, I. M., Leung, S. M., McAndrew, S. J., Brown, J. B., and Hightower, L. E. (1995) *J. Biol. Chem.* **270**, 19839–19844
- Palleros, D. R., Shi, L., Reid, K. L., and Fink, A. L. (1994) *J. Biol. Chem.* **269**, 13107–13114
- Wang, H., Kurochkin, A. V., Dang, Y., Hu, W., Flynn, G. C., and Zuidewerg, E. R. (1998) *Biochemistry* **37**, 7929–7940
- McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1992) *J. Mol. Biol.* **225**, 495–517
- Neupert, W. (1997) *Annu. Rev. Biochem.* **66**, 863–917
- Glick, B. S. (1995) *Cell* **80**, 11–14
- Schatz, G. (1996) *J. Biol. Chem.* **271**, 31763–31766
- Mattingly, J. R., Jr., Iriarte, A., and Martinez-Carrion, M. (1993) *J. Biol. Chem.* **268**, 26320–26327
- Beckman, R. P., Mizzen, L. A., and Welch, W. J. (1990) *Science* **248**, 850–854
- Nagata, H., Hansen, W. J., Freeman, B., and Welch, W. J. (1998) *Biochemistry* **37**, 6924–6938
- Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., and Gething, M.-J. H. (1993) *Cell* **75**, 717–728
- Mehta, P. K., Hale, T. I., and Christen, P. (1993) *Eur. J. Biochem.* **214**, 549–561
- Sonderegger, P., Gehring, H., and Christen, P. (1977) *J. Biol. Chem.* **252**, 609–612