Article No. jmbi.1999.2669 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (1999) 288, 177–190

JMB

Interaction of Mitochondrial Presequences with DnaK and Mitochondrial hsp70

Xiao-Ping Zhang[1], Arne Elofsson[1], David Andreu[2] and Elzbieta Glaser[1]*

[1]*Department of Biochemistry Arrhenius Laboratories for Natural Sciences, Stockholm University, S-106 91 Stockholm, Sweden*

[2]*Department of Organic Chemistry, University of Barcelona, Spain*

**Corresponding author*

Mitochondrial heat shock protein 70 (mt-hsp70) functions as a molecular chaperone in mitochondrial biogenesis. The chaperone in co-operation with its co-proteins acts as a translocation motor pulling the mitochondrial precursor into the matrix. Mt-hsp70s are highly conserved when compared to the bacterial hsp70 homologue, DnaK. Here we have used DnaK as a model to study the interaction of mitochondrial presequences with mt-hsp70 applying a DnaK-binding algorithm, computer modeling and biochemical investigations. DnaK-binding motifs have been analysed on all available, statistically relevant mitochondrial presequences found in the OWL database by running the algorithm. A total of 87 % of mammalian, 97 % of plant, 71 % of yeast and 100 % of *Neurospora crassa* presequences had at least one DnaK binding site. Based on the prediction, five 13-mer presequence peptides have been synthesized and their inhibitory effect on the molecular chaperone (DnaK/DnaJ/GrpE) assisted refolding of luciferase has been analysed. The peptide with the highest predicted binding likelihood showed the strongest inhibitory effect, whereas the peptide with no predicted binding capacity showed no inhibitory effect. A 3D structure of the pea mt-hsp70 has been constructed using homology modeling. The binding affinities of the 13-mer presequence peptides and additional control peptides to DnaK and pea mt-hsp70 have been theoretically estimated by calculating the buried hydrophobic surface area of the peptides docked to DnaK and to the mt-hsp70 structural model. These results suggest that mitochondrial presequences interact with the mt-hsp70 during or after mitochondrial protein import.

Keywords: hsp70; DnaK; molecular chaperone; mitochondrial presequence; computer modeling

Introduction

Heat shock protein 70 (hsp70) functions as a molecular chaperone in a variety of biochemical processes including the prevention of aggregation of newly synthesized or unfolded polypeptides, facilitation of translocation of nascent chains across membranes, mediating the assembly or disassembly of multimeric protein complexes, and targeting proteins to lysosomes or proteasomes for degradation (Hartl, 1996; Rassow *et al.*, 1997). In yeast mitochondria, hsp70 in co-operation with mdj1 and Yge1p (homologues of bacterial DnaK and its co-chaperones DnaJ and GrpE, respectively) acts as an ATP-driven translocation-motor and pulls the emerging precursor into the mitochondrion (Glick, 1995; Horst *et al.*, 1996). Hsp70 is also involved in the folding and assembly of imported mitochondrial proteins. It is of great interest to investigate the interaction of mitochondrial precursors with mitochondrial hsp70 (mt-hsp70).

Some information on the interaction of hsp70 with mitochondrial preproteins is available. Yeast cytosolic hsp70 (Ssa1p) can bind to synthetic peptides with amino acid compositions typical of mitochondrial presequences (Endo *et al.*, 1996). The precursor of mitochondrial aspartate aminotransferase overexpressed in *Escherichia coli* has been shown to be complexed with DnaK *via* the presequence (Schmid *et al.*, 1992). The binding kinetics of a 21-residue presequence peptide of aspartate

Abbreviations used: mt, mitochondrial; hsp70, heat shock protein 70.

E-mail address of the corresponding author: e_glaser@biokemi.su.se

0022-2836/99/160177–14 $30.00/0

aminotransferase to DnaK has been extensively studied (Pierpaoli *et al.*, 1997; Schmid *et al.*, 1994). In a cell-free transcription-translation system, hsp70 binds to the newly synthesized mitochondrial precursor of aspartate aminotransferase (Lain *et al.*, 1994, 1995) and both the presequence and the N-terminal region of the mature part of the precursor are responsible for the binding (Lain *et al.*, 1995). Due to the difficulties in obtaining biochemical quantities of purified mitochondrial hsp70, most of the knowledge about the interaction of the hsp70 chaperone with its substrate has come from the studies on the bacterial hsp70 homologue, DnaK, cytosolic hsp70, or the endoplasmic reticulum hsp70, BiP. Previous studies have shown that DnaK, Hsc70 and BiP preferentially interact with hydrophobic amino acids (Ile, Leu, Val, Ala, Phe, Trp and Tyr) and positively charged amino acids (Arg and Lys) but not with acidic residues (de Crouy-Chanel *et al.*, 1996; Fourie *et al.*, 1994; Gething *et al.*, 1995; Hightower *et al.*, 1994; Richarme & Kohiyama, 1993). The minimal length of sequence which can bind to these chaperones should be seven amino acid residues long (Fourie *et al.*, 1994; Gething *et al.*, 1995). BiP has the specific binding motif Hy(W/X)-HyXHyXHy (Blond-Elguindi *et al.*, 1993), where Hy is a large hydrophobic or aromatic amino acid (most frequently Trp, Leu, or Phe), and X is any amino acid. In the most recent study, the DnaK binding motif was identified to be ++++HyHyHyHyHy++++ (Rüdiger *et al.*, 1997). The five-residue hydrophobic core is enriched particularly in L, but also in I, V, F and T. The two four-residue flanking regions are enriched in basic resides. Acid residues are excluded from the core and disfavoured in flanking regions.

The hsp70 family is a class of highly conserved proteins. All of these proteins show a high degree of identity with DnaK (58-60 %). Hsp70 proteins contain two functional domains, an N-terminal ATP-binding domain (~44 kDa) and a C-terminal substrate-binding domain (~27 kDa). The ATP-binding domain is highly conserved (>60 %). The substrate-binding domain displays less identity than the ATP-binding domain (~45 %; Hartl, 1996). According to structural information on DnaK (Zhu *et al.*, 1996), the substrate binding domain of DnaK contains two structurally separated subdomains. The N-terminal subdomain is folded into a compact β-sandwich, and the C-terminal subdomain is composed of five extended α-helices. The N-terminal β subdomain is responsible for substrate binding. Alignment of the substrate binding domain of four higher plant mitochondrial hsp70 proteins and DnaK shows that the β-sheet region has about

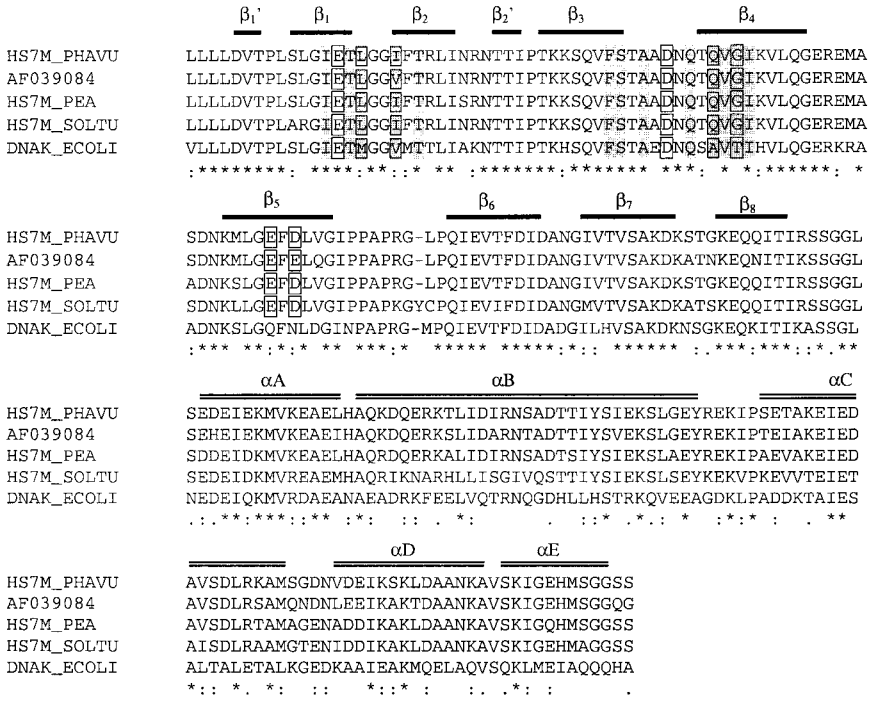

Figure 1. Alignment of the substrate-binding domain of DnaK (residues 389-607) with four mitochondrial hsp70s from kidney bean (HS7M_PHAVU), spinach (AF039084, Genebank access number), pea (HS7M_PEA) (residues 381-599) and potato (HS7M_SOLTU) (residues 381-600). The symbol * indicates the fully conserved residues and the : and . indicate strong and weak similarity of residues, respectively. Residues which directly interact with a model DnaK-binding peptide, NRLLLTG through hydrogen bonding or van der Waals forces, are shaded. The shaded residues in frame are not conserved. α(A-E) and β(1-8) stand for α-helical and β-strand regions, respectively. The residues in frame (with no shading) correspond to negative charges located on the side of the β-substrate-binding subdomain which faces the C-terminal of the substrate peptide. See Results and Discussion for details about surface charges.

68 % identity, whereas the different α-helices only have about 21 % identity (cf. Figure 1).

Here, we have modeled mt-hsp70 based on the DnaK structure and we have searched for DnaK binding sites on a collection of 147 mitochondrial presequences from plants, mammals, yeast and *Neurospora crassa* using the algorithm of Rüdiger *et al*. (1997). Based on the prediction results, five presequence peptides have been synthesized and the activity of these peptides as inhibitors of DnaK/DnaJ/GrpE assisted refolding of luciferase has been analysed. The binding specificity of these peptides to the mt-hsp70 model and to DnaK was analysed by calculating the buried hydrophobic surface area for the synthesized peptides docked to mt-hsp70 or DnaK. The calculated buried hydrophobic surface area on mt-hsp70 agreed well with the area on DnaK. Furthermore, the binding area correlated well with the experimentally determined affinity of the peptides to DnaK, indicating that the binding specificity of mt-hsp70 and DnaK are similar.

Results

Sequence conservation in the mt-hsp70 substrate-binding domain

In order to investigate the features of the substrate-binding domain of mt-hsp70, we aligned the four available plant mt-hsp70 sequences with the sequence of DnaK. The identity of these mt-hsp70s when compared to DnaK was 58-60 %. Based on the primary alignment, we identified a region from the mt-hsp70s (residues 381-599 from kidney bean, spinach and pea, residues 381-600 from potato) which corresponded to the substrate-binding domain of DnaK (residues 389-607; Figure 1). In the N-terminal subdomain of the substrate-binding domain of DnaK (Zhu *et al*., 1996), two sheets with four antiparallel β-strands in each form a β-sandwich structure which serves as a substrate-binding cleft. The C-terminal subdomain comprises five α-helices which appear to serve as a lid on the substrate cleft. Thirteen amino acid residues in the β-sandwich subdomain interact with a peptide NRLLLTG through hydrogen bonds and van der Waals forces (Figure 1). In mt-hsp70, nine of the 13 residues are highly conserved (marked with shading), the remaining four are not conserved (marked shading in frames). Met404 in the loop between β_1 and β_2, Ala435 and Thr437 within β_4 of DnaK are changed to Leu396, Gln427 and Gly429, respectively, in all mt-hsp70s. Val407 in β_2 of DnaK is sometimes changed to Ile399.

Structural modeling of mt-hsp70 substrate-binding domain

Based on the alignment and 3D structure of DnaK, a pea mt-hsp70 structural model was established using homology modeling (Figure 2). Comparison of the structure of the entire substrate-binding domain between DnaK and the pea mt-hsp70 model, shown in Figure 2(a), reveals a high degree of similarity. Also the location of the peptide NRLLLTG crystallized together with DnaK is similar to that seen when the peptide is docked to the model mt-hsp70. All the amino acid residues that interact with the substrate peptide NRLLLTG through hydrogen bonds and van der Waals forces are shown in Figure 2(b). There are seven hydrogen bonds (broken yellow lines) formed between the substrate peptide and five amino acid residues of DnaK (yellow script), Met404, Ser427, Ala429, Gln433 and Thr437. In the mt-hsp70, there are six hydrogen bonds (Table 1) formed between the peptide and amino acid residues corresponding to the above listed residues of the DnaK, Leu396, Ser419, Ala421, Gln425 and Gly429 (Figure 2(b), green script). Four hydrogen bonds are shown in Figure 2(b) (broken green lines). Two of the five residues forming hydrogen bonds are different between the DnaK and the mt-hsp70 model. Substitution of Met404 by Leu396 does not influence the formation of the hydrogen bond, but replacing Thr437 with Gly429 causes the loss of the hydrogen bond which existed between the O^{γ} of Thr437 and the peptide. Hydrogen bond lengths in DnaK and mt-hsp70 are comparable and are within the theoretical permissive length range (Table 1).

We investigated the surface charges of the substrate-binding β-subdomain of DnaK and mt-hsp70. Charges on the surface facing the N-terminal part of the NRLLLTG peptide were very similar between DnaK and mt-hsp70 (not shown), whereas the charges on the surface facing the C-terminal part of the peptide in DnaK and mt-hsp70, shown in Figure 2(c) and (d), respectively, showed some differences. In both structures, two negative charges located close to the substrate-binding cleft are conserved (Glu402, Asp431 in DnaK, the corresponding residues are Glu394 and Asp423 in mt-hsp70; see Figures 1 and 2). However, the charge of Asp423 in mt-hsp70 has moved to the outside of the cleft due to a slight variation of side-chain orientation, it is invisible in Figure 2(d). In DnaK, the third negative charge, Asp460, is distantly located from the substrate-binding cleft. The corresponding negative charge is missing in mt-hsp70. Two extra negative charges in mt-hsp70, Glu448 and Asp450 corresponding to Gln456 and Asn458 of DnaK, are located in close proximity to the substrate-binding cleft. These two negative charges are conserved in all known plant mt-hsp70s (Asp450 is substituted by Glu450 in spinach mt-hsp70 which maintains the charge; see Figure 1 residues in frame). Together with Glu394, these three negative charges form a strong negatively charged region just outside the substrate-binding cleft of mt-hsp70 on the C-terminal side of the peptide. This region of negative charges would provide an appropriate surface for interacting with the positively charged amphiphilic α-helix of mitochondrial presequences.

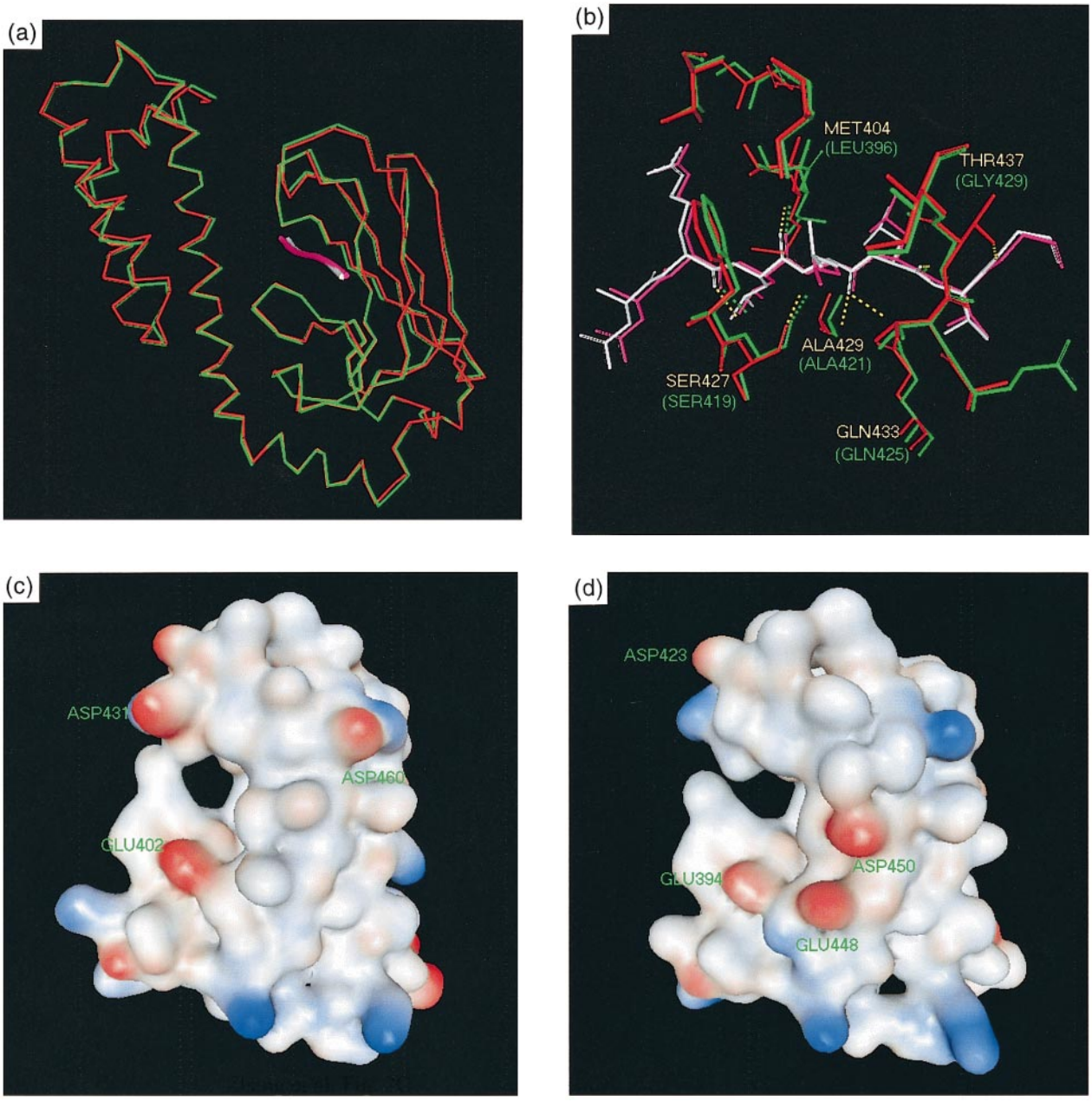

Figure 2. Mt-hsp70 model. (a) The mt-hsp70 (green) was constructed by homology modeling based on the structural information of DnaK (red). The location of peptide NRLLLTG in DnaK (white) and mt-hsp70 (pink) is shown. (b) Substrate-binding clefts of DnaK and the mt-hsp70 model. Hydrogen bonds (broken lines) in yellow or green to indicate the interaction between residues of DnaK (red sticks) and peptide NRLLLTG (white sticks) and between mt-hsp70 (green sticks) and the peptide (pink sticks) (two hydrogen bonds were not shown, another two hydrogen bonds were nearly perpendicular to the image). Amino acid residues of DnaK are indicated in yellow script and those of mt-hsp70 are in green script. In order to get a good view, hydrogen atoms are hidden. See Table 4 for further information. Surface charges of the substrate-binding pockets of DnaK and mt-hsp70 are shown in (c) and (d), respectively. Asp423 and Glu394 in mt-hsp70 correspond to the Asp431 and Glu402 in DnaK, which are conserved in DnaK and all known plant mt-hsp70. Glu448 and Asp450 in mt-hsp70 are two negative charges which are conserved in all known plant mt-hsp70 (Asp450 is substituted by Glu450 in spinach mt-hsp70 which maintains the charge).

DnaK binding sites on mitochondrial presequences

Overlapping 13-mer peptides of presequences were scored according to the DnaK-binding algorithm (Rüdiger *et al.*, 1997). We analysed every 13-mer peptide from a collection of 147 mitochondrial presequences (Schneider *et al.*, 1998) from plants, mammals, yeast and *N. crassa*. Our results showed that 87% of mammalian, 97% of plant, 71% of yeast and 100% of *N. crassa*, presequences have at least one DnaK binding site. Figure 3 shows the predicted results of 32 full-length plant mitochondrial presequences. Some presequences have sev-

Table 1. Hydrogen bonds between peptide NRLLLTG and DnaK/mt-hsp70

DnaK-NRLLLTG		Mt-hsp70-NRLLLTG	
Bond	Length (Å)	Bond	Length (Å)
Arg2 O—Ser427 N	2.97	Arg2 O—Ser419 N	3.06
Leu3 O—Met404 N	2.82	Leu3 O—Leu396 N	2.98
Leu4 N—Ser427 O	2.88	Leu4 N—Ser419 O	2.66
Leu4 O—Ala429 N	2.99	Leu4 O—Ala421 N	2.93[a]
Leu4 O—Gly433 $N^{\varepsilon 2}$	3.25	Leu4 O—Gln425 $N^{\varepsilon 2}$	3.55[a]
Leu5 O—Thr437 N	2.95	Leu5 O—Gly429 N	2.89
Gly7 N—Thr437 $O^{\gamma 1}$	2.49	—	-

[a] Hydrogen bonds which were not shown in Figure 2(b).

eral regions that contain a DnaK-binding motif. The DnaK-binding motif exists independently of whether the presequence is short (18 residues) or long (85 residues). In Table 2, we show the binding motifs with the lowest score in each presequence from yeast, *N. crassa* and mammals.

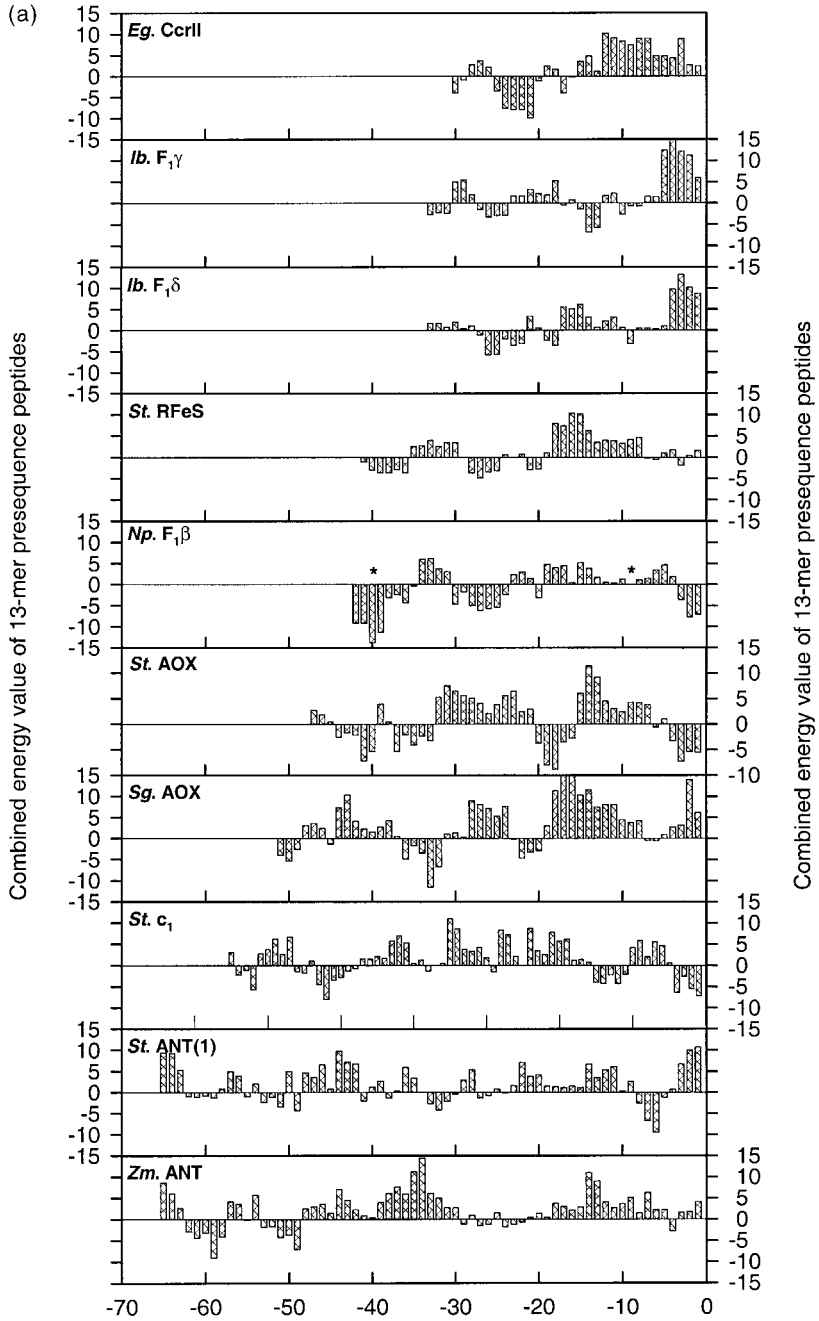

Figure 3(a) (*legend overleaf*)

Inhibition of DnaK-assisted refolding of the denatured luciferase by presequence peptides

Five peptides from plant mitochondrial presequences were synthesized according to the predicted DnaK binding motif (Table 3). The peptides 1-4 are binders and peptide 5 is a non-binder of DnaK according to the DnaK-binding algorithm (Rüdiger *et al.*, 1997). In competition experiments, the peptides were tested individually in the luciferase-refolding buffer before and during incubation of the unfolded luciferase with DnaK/DnaJ/GrpE. Peptides 1-4 inhibited the chaperone-assisted luciferase refolding process. Peptide 2, which has the highest predicted binding likelihood, showed the strongest inhibitory effect (>90 % at 50 μM), peptides 1, 3 and 4 with the intermediate predicted binding likelihood showed an intermediate inhibitory efficiency, whereas peptide 5 with no predicted binding capacity showed no inhibitory effect even at 100 μM (Figure 4(a)). Also, measurements of refolding after different incubation times confirmed the above conclusion (Figure 4(b)).

Binding capacity of presequence peptides to pea mt-hsp70

In order to show the affinity of the synthesized presequence peptides with mt-hsp70, we modeled heptapeptides derived from the synthesized peptides using ICM (Abagyan *et al.*, 1994). These short peptides were docked to the mt-hsp70 model and to DnaK by computer simulation. The buried hydrophobic surface area was calculated (Table 4). We also calculated the buried hydrophobic surface area for three heptapeptides for which the binding affinity to DnaK has been investigated with direct binding methods (Gragerov *et al.*, 1994). Two of these peptides, LRRASLG and LLTNRGL which did not bind to DnaK, showed very low values of the buried hydrophobic binding areas within DnaK. The peptide NRLLLTG, which was shown to bind to DnaK with a high degree of affinity, showed a high value of the buried hydrophobic surface area within DnaK. The five synthesized peptides shown in Table 3 were studied in DnaK/DnaJ/GrpE chaperone assisted luciferase-refolding experiments (cf. Figure 4). Peptide 5 with no inhibitory effect upon luciferase refolding experiments had a small buried hydrophobic surface area within DnaK, but peptides 2-4, which showed an inhibitory activity on refolding of luciferase, showed a large buried hydrophobic surface area value. A similar pattern of results was obtained with mt-hsp70. A greater binding area indicated a stronger interaction of the peptide with the mt-hsp70. These data indicate that the synthetic presequence peptides 1-4 have the potential to bind to

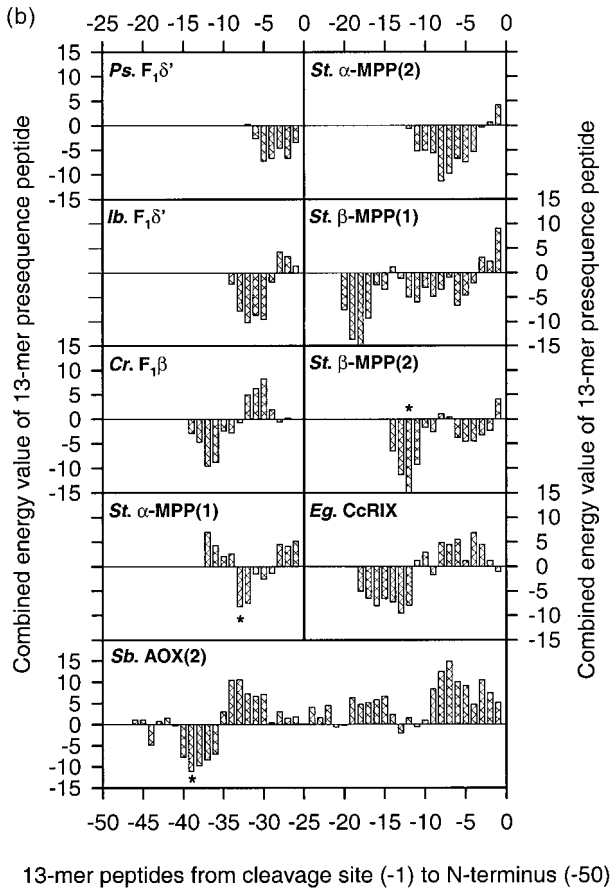

Figure 3(b) (*legend overleaf*)

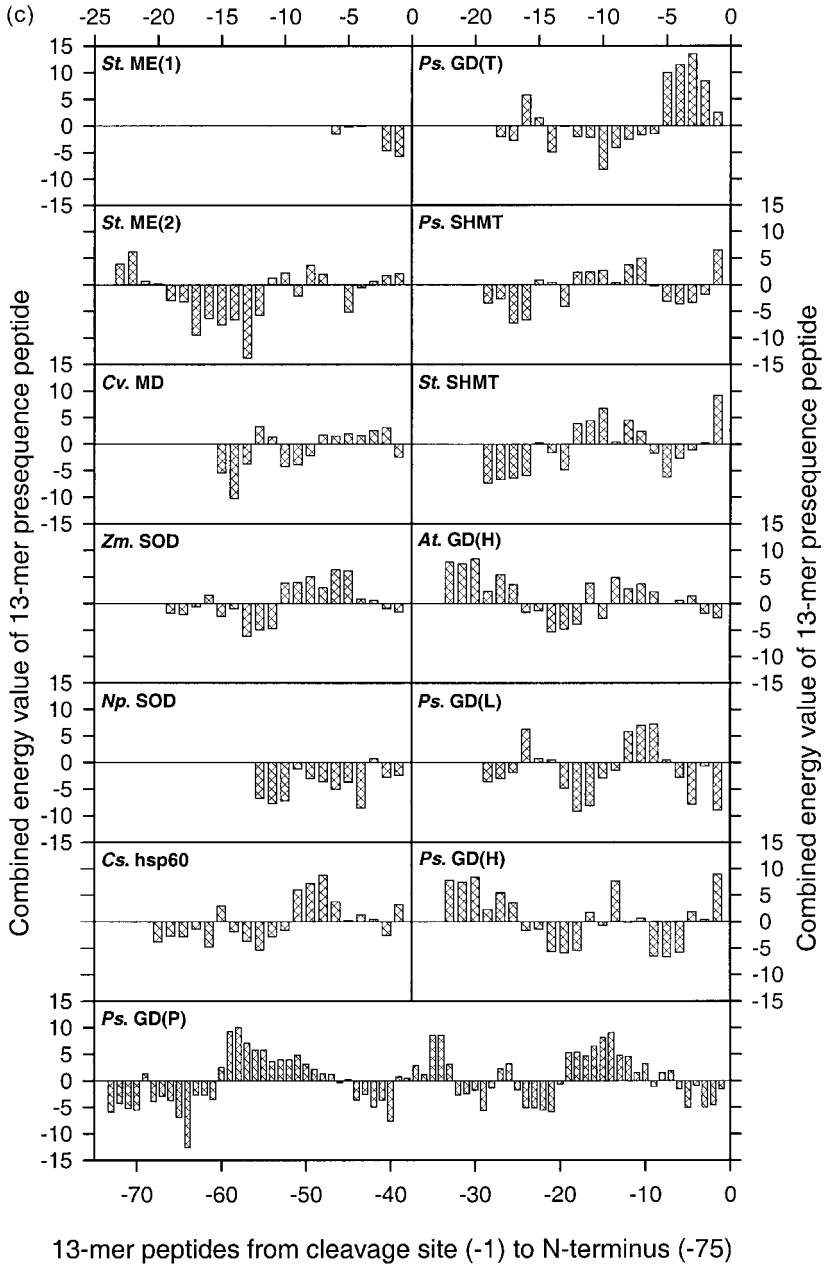

Figure 3. Prediction of the DnaK binding motif(s) on plant mitochondrial presequences. The prediction is performed by using the algorithm by Rüdiger *et al*. (1997). Overlapping 13-mer presequence peptides showing a combined energy of less than −5 are predicted to be DnaK binders. Presequences shorter than 13-residues are omitted. (a), (b) Presequences of mitochondrial inner membrane proteins; (c) presequences of matrix proteins. The asterisk * shows the 13-mer presequence peptides which have been used in the luciferase refolding studies. Abbreviation used: *At*, *Arabidopsis thaliana*; *Cr*, *Chlamydomonas reinhardtii*; *Cs*, *Cucurbita* sp.; *Cv*, *Citrullus vulgaris*; *Eg*, *Euglena gracilis*; *Ib*, *Ipomea batatas*; *Np*, *Nicotiana plumbaginifolia*; *Ps*, *Pisum sativum*; *Gm*, Glycine max; *Sg*, *Sauromatum guttatum*; *St*, *Solanum tuberosum*; ANT, adenine nucleotide translocator; AOX, alternative oxidase; c_1, cytochrome c_1; CcR, ubiquinol cytochrome *c* oxidoreductase complex (bc_1); $F_1\beta$, δ, γ, the β, δ, γ subunits of the ATP synthase; GD, glycine decarboxylase subunits; hsp, heat shock protein; MD, malate dehydrogenase; Me, malic enzyme; α-MPP, α-subunit of mitochondrial processing peptidase; SHMT, serine hydroxymethyltransferase; SOD, superoxide dismutase; RFeS, Rieske FeS protein.

mt-hsp70 and that the binding affinity of the peptides for both DnaK and mt-hsp70 seems to be very similar. Taken together, the presented findings imply that mitochondrial presequences could interact with mt-hsp70 during mitochondrial protein import.

Discussion

Mt-hsp70 has a predominant role in the mitochondrial protein import. Two models have been proposed for its function in the translocation process, the molecular ratchet model and the translocation motor model (Neupert, 1997; Pfanner & Meijer, 1995; Schatz & Dobberstein, 1996). Mt-hsp70 is proposed to bind to segments of the polypeptide as these emerge on the inner side of the inner membrane within the transport channel, with a net result of pulling precursors into mitochondria. Mitochondrial presequences contain information for targeting and processing. The mean length of known presequences from all organisms is 34 (±16) amino acid residues (Schneider *et al*., 1998). Considering the measured oscillation of the translocating protein with segments of about 40 residues (Ungermann *et al*., 1994), the presequence would be a primary target for mt-hsp70 on the condition that it contains consensus sequence for the mt-hsp70 binding. Presequences are rich in positively charged, hydrophobic and hydroxylated amino acids but have a very low content of negatively charged residues. The membrane potential required for the translocation of the presequence across mito-

Table 2. DnaK binding motifs predicted

Presequence	The motif with lowest score	The motif position/length of presequence	Score
A. *Yeast mitochondrial presequences*			
ATPA	RTAAIRSLSRTLI	−28 - −16/31	−8.98
ATPB	LPRLYTATSRAAF	−17 - −5/19	−5.73
ATPD	MGVRGLALRSVSK	−31 - −19/35	−8.35
ATPG	QAQVGILYKTNPV	−16 - −4/33	−12.60
COX4	RQSIRFFKPATRT	−21 - −9/25	−8.13
COX6	INRTLLRARPGAY	−29 - −17/40	−9.67
COX8	QQMIRTTAKRSSN	−24 - −12/27	−7.19
COXA	RAGGLSRITSVRF	−13 - −1/20	−6.05
CY1	TASTLLYADSLTA	−15 - −3/61/	−9.10
CYB2	KYKPLLKISKNCE	−78 - −66/80	−10.72
CYPC	IQQSRLFSNSASR	−14 - −2/20	−5.27
DLDH	MLRIRSLLNNKRA	−12 - −9/21	−10.82
HEMZ	RTQGSFLRRSQLT	−25 - −13/31	−9.05
HS77		-/23	-
IATP	RSALARSLQLQRG	−19 - −7/22	−8.29
MDHM		-/17	-
MPP1		-/13	-
MPP2	KFRNTRRLLSTIS	−13 - −1/20	−5.59
NDI1	SNKRLLTSTNTLV	−19 - −7/26	−8.51
ODP2	SSVLTRSLRLQLR	−15 - −3/28	−9.85
ODPX		-/30	-
RIM1	TQARFFHATTKKM	−13 - −1/17	−5.02
RM02		-/27	-
RM04		-/14	-
RM09	QGSIFSISKLHVR	−14 - −2/19	−5.72
RM13	RMSSLLKLHCIRP	−76 - −64/86	−10.59
RM27		-/16	-
RM31		-/12	-
RM32	QKQLGSIHRWLRE	−35 - −23/71	−7.40
RM36		-/33	-
RM37	ARSLGYRLISTSR	−19 - −7/21	−9.87
RM41	TKNMLYPLQKTLA	−38 - −26/45	−8.65
RPM2	QNVSLRSLKQQQS	−27 - −15/121	−8.17
RT28	GRNAILNLRISLC	−29 - −17/33	−9.17
SDH4		-/31	-
SODM	KGGLSLLSTTARR	−14 - −2/26	−11.47
SYH	SRSLNKVVTSIKS	−18 - −6/20	−7.12
SYV	KTFTFRLLNCHYR	−38 - −26/47	−10.22
UCR2		-/16	-
B. Neurospora crassa *mitochondrial presequences*			
ACP	RTAALTAARVARP	−44 - −32/46	−5.32
ATP9	ASTRVLASRLASQ	−34 - −22/35	−8.03
COX4	APALRRSIATTVV	−15 - −3/16	−5.89
COX5	TPTVSALVRNVAV	−24 - −12/27	−9.43
CY1	IAWYYHLYGFASA	−12 - −1/71	−6.40
MPP1	FRPARLVAQSSRC	−31 - −19/35	−6.58
MPP2	SRRLALNLAQGVK	−26 - −14/28	−10.02
NUAM	RSTLSRSAWRTGR	−31 - −19/33	−5.60
NUBM	TKASARTLSRAAA	−19 - −7/27	−5.55
NUEM	RRTPRIIVSNAFG	−18 - −6/26	−8.89
NUYM	TASAARMLRTSNA	−26 - −14/33	−6.21
ODP2	ASVARVALPSLTR	−15 - −3/28	−5.99
SYLM	ARPLGRLVPKLGA	−46 - −34/51	−10.37
SYYM	RTKALIRSGGSIA	−29 - −17/32	−7.79
UCRI	PARAVRALTTSTA	−15 - −3/32	−7.80
C. *Mammalian mitochondrial presequences*			
AATM_MOUSE		-/29	-
ACDL_RAT	MAARLLLRSLRVL	−20 - −8/30	−13.81
ACDM_RAT	RRGYKVLRSVSHF	−20 - −8/25	−10.26
ACON_PIG	MAPYSLLVTRLQK	−27 - −15/27	−11.88
ADX1_BOVIN	GRWRLLARPRAGA	−41 - −29/58	−12.03
ATP0_BOVIN	AAAVARALPRRAG	−37 - −25/43	−7.96
ATPB_HUMAN	PAQLLLRAAPTAV	−22 - −10/49	−12.23
ATPD_BOVIN	RPGLGRLVRQVRL	−14 - −2/22	−12.19
ATPG_BOVIN		-/25	-
ATPL_HUMAN	TRGLIRPVSASFL	−44 - −32/61	−9.03
ATPM_BOVIN	RRTSTVLSRSLSA	−54 - −42/67	−9.17
ATPO_BOVIN		-/23	-
ATPR_BOVIN	ILQRLFRLSSAVQ	−31 - −19/32	−8.99

BDH_RAT	TRHTLLFYPASFS	−20 - −8/47	−12.42
CAH5_HUMAN	TSAFSFLVEQMWA	−20 - −18/38	−5.11
CAH5_MOUSE	PRKPLAILRHVGL	−24 - −12/29	−11.21
CISY_PIG		−/27	-
COX4_BOVIN	ATRVFSLIGRRAI	−20 - −8/22	−12.74
COXA_HUMAN	PRGLLHSARTPGP	−23 - −11/41	−8.21
COXJ_BOVIN	MLRNLLALRQIAK	−23 - −11/23	−10.21
COXK_BOVIN	SQALVRSFSSTAR	−15 - −3/21	−8.17
COXM_BOVIN	AKNALSRLRVQSI	−23 - −11/32	−7.82
COXO_BOVIN	GQSIRRFTTSVVR	−14 - −2/16	−6.46
CP27_RAT	RMRLRWALLDTRV	−27 - −15/32	−7.19
CPM1_BOVIN	ARGLPLRSALVKA	−37 - −25/39	−6.12
CPN2_RAT		-/34	-
CPSM_RAT	GFGLANVTSKRQW	−22 - −10/38	−5.25
CPT2_RAT	MPRLLFRAWPRCP	−24 - −12/25	−11.13
D3D2_RAT	ARRVLLQAGSRLG	−23 - −11/28	−10.97
DHAM_HUMAN	AARFGPRLGRRLL	−12 - −1/17	6.74
DHAM_RAT	RAALSTARRGPRL	−16 - −4/19	−6.73
DLDH_HUMAN		-/35	-
FUMH_RAT	NRAFCLLARSRRF	−40 - −28/41	−6.35
GCSH_BOVIN	RAGAVRELRTGPA	−15 - −3/48	−5.1
GPDM_RAT	GGALATVLGLSQF	−28 - −16/42	−8.87
HEM6_HUMAN	SGTRATSLGRPEE	−31 - −19/36	−5.21
HUMOTC	LFNLRILLNNAAF	−31 - −19/32	−11.04
IATP_BOVIN	AATALAARTRQAV	−24 - −12/25	−5.94
IVD_HUMAN	MATATRLLGWRVA	−29 - −17/29	−9.06
JX0071		-/24	-
KCRS_HUMAN	TTGYLLNRQKVCA	−13 - −1/39	−10.62
MDHM_MOUSE		-/16	-
MPCP_BOVIN	YSSVVHLARANPF	−48 - −36/49	−7.13
MTDC_MOUSE	ALAVRLLRPTHGC	−27 - −15/35	−9.88
MUTA_HUMAN	KNQLFLLSPHYLR	−28 - −16/32	−12.13
NIAM_BOVIN	RAGVLGVRWLQKA	−24 - −12/29	−8.01
NNTM_BOVIN	KKNFLRTFHTHRI	−15 - −3/43	−7.91
NUAM_BOVIN	VRKALVGLSKSSK	−18 - −6/23	−10.61
NUBM_BOVIN	MLAARRLLGGSLP	−20 - −8/20	−6.44
NUCG_BOVIN	TPGLLSRLPVLPV	−16 - −4/48	−11.45
NUGM_BOVIN	AGRPSVLLLPVRR	−13 - −1/38	−8.69
NUHM_BOVIN	WGKHIRNLHKTAV	−15 - −3/32	−5.52
ODBB_BOVIN	AGWLLRLRAAGAD	−43 - −31/50	−12.17
ODP2_HUMAN	RNRLLLQLLGSPG	−17 - −5/54	−15.33
ODPA_HUMAN	MRKMLAAVSRVLS	−29 - −17/29	−6.69
P60_HUMAN	MLRLPTVFRQMRP	−26 - −14/26	−7.83
PCCA_RAT	SQQLLWTLKRAPV	−19 - −7/21	−13.18
PMIP_RAT	GTRYAYRLCGRRA	−28 - −16/33	−11.71
SODM_HUMAN	SRQLAPALGYLGS	−15 - −3/24	−5.84
SUCA_RAT	SRTFLLQQNGIRH	−13 - −1/27	−8.79
THIL_RAT	GVVRRPLLRGLLQ	−22 - −10/30	−10.74

DnaK binding motifs predicted on yeast (A), *Neurospora crassa* (B) and mammalian (C) mitochondrial presequences. Only the DnaK-binding motif with the lowest score in each presequence is shown. Presequences shorter than 13 amino acid residues were omitted. Presequences which lack a 13-mer residue segment scoring lower than −5 are predicted to have no DnaK-binding motif, represented by -.

chondrial membranes is negative inside the mitochondrion supporting the pull-in force of the positively charged presequence. If mt-hsp70 pulls precursors into mitochondria as suggested (Pfanner *et al.*, 1997), it most likely has to interact first with mitochondrial presequences.

DnaK is an obvious model system for investigating the interaction of mitochondrial presequences with mt-hsp70 as this bacterial chaperone is very highly conserved in comparison to mt-hsp70s and its binding specificity has been well characterized (Rüdiger *et al.*, 1997; Zhe *et al.*, 1996). The mitochondrion is believed to have evolved from bacteria following endosymbiosis (Margulis, 1975).

All hsp70 proteins have two functional domains, an ATP-binding domain and a sub-

Table 3. Synthesized presequence peptides used in this work

Peptide No.	Amino acid residues	Presequence	Prediction
1	NH_2-SRRLLASLLRQSA-$CONH_2$	$pF_1\beta$	−13.82
2	NH_2-TRHLLNLTRRRSR-$CONH_2$	st. β-MPP(1)	−16.20
3	NH_2-TSSRLRALKVRGT-$CONH_2$	st. α-MPP(1)	−11.31
4	NH_2-STVRRALLNGRNG-$CONH_2$	AOX(2)	−11.13
5	NH_2-ISRSLGNSIPKSA-$CONH_2$	$pF_1\beta$	1.46

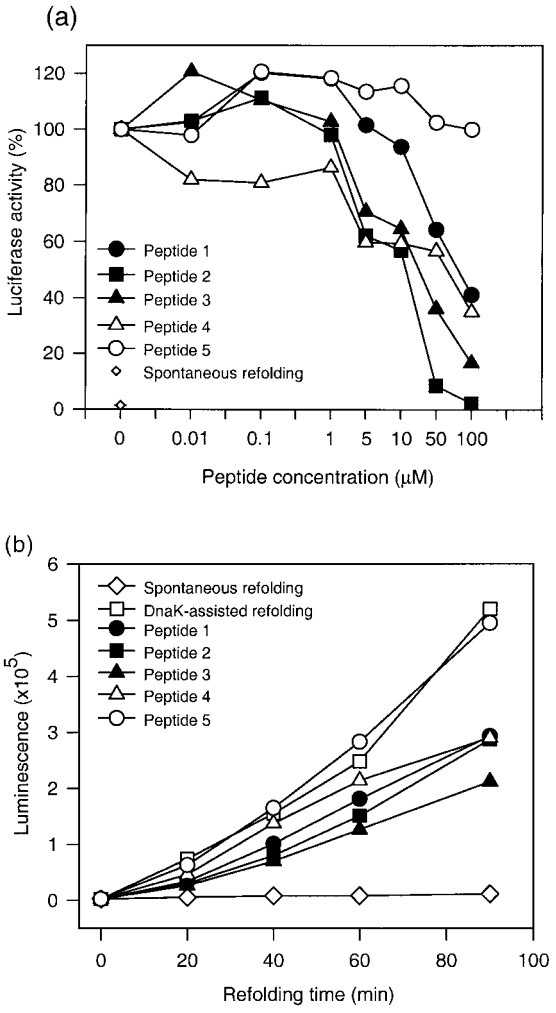

Figure 4. Inhibition of DnaK-assisted refolding of unfolded firefly luciferase by synthetic plant mitochondrial presequence peptides. (a) Molecular chaperones DnaK/DnaJ/GrpE were preincubated with various concentrations of the synthetic peptides, as indicated in the Figure, at room temperature for ten minutes. The denatured luciferase and ATP were then added, thoroughly mixed and incubated at room temperature for 60 minutes. Luciferase activity was detected with a luminometer as described in Materials and Methods. 100% represents the chaperone-assisted luciferase refolding in the absence of any peptide. (b) Time course of the inhibition of DnaK-assisted refolding of unfolded firefly luciferase by synthetic plant mitochondrial presequence peptides. The concentration of peptides was 10 μM. After adding luciferase, samples were withdrawn at various incubation times, and luciferase activity was analysed.

strate-binding domain. The former binds and hydrolyses ATP; this process supplies energy for the action of hsp70. The ATP-binding domain is highly conserved amongst all hsp70s, whereas the substrate-binding domain is less conserved (Hartl, 1996). However, structural analysis of the substrate-binding domain showed that the β-subdomain, which is responsible for peptide binding, is very highly conserved as are the amino acid residues which directly contact a substrate peptide through hydrogen bonding and van der Waals forces (Figure 1). In pea mt-hsp70, in the β-subdomain of the substrate-binding domain, out of 13 residues involved in the interaction with the substrate peptide only four residues are not conserved in comparison to DnaK. The alterations of Met404 and Val407 in DnaK to Leu396 and Ile399 in pea mt-hsp70, respectively, only slightly affected the hydrophobicity of the substrate-binding cleft. These residues of the chaperones interact with residues at the (−1, 1) and (−2) positions of the NRLLLTG substrate peptide. The third change of DnaK Ala429 to mt-hsp70 Gln425, conserved in all mt-hsp70s, makes the substrate-binding cleft more polar (cf. Figures 1 and 2). This increased polarity may be crucial for mt-hsp70 and it may ensure a higher binding affinity between highly hydrophobic and positively charged presequences and mt-hsp70. Changing Thr437 in DnaK to Gly429, which has a smaller side-chain, in the mt-hsp70 removes one of the hydrogen bonds present in the interaction of DnaK with the peptide. This change allows the substrate-binding cleft to accept a peptide with larger side-chain amino residues. The similarity of the interaction of DnaK and mt-hsp70 with the substrate peptide is, however, best evidenced by the agreement of the calculated hydrogen bond lengths formed to the substrate (cf. Table 1). Despite four out of 13 residues not being conserved, the hydrogen bonding pattern between the substrate peptide and the mt-hsp70 was very similar to the pattern between the peptide and DnaK. Furthermore, analysis of surface charges on the β-subdomain of the substrate binding domain of mt-hsp70 revealed an area of high negative charge on one side of the substrate binding cleft (Figure 2(d)). This was due to the grouping of the three negatively charged residues, conserved in all known plant mt-hsp70s, resulting in a potential site of interaction with the positively charged amphiphilic α-helix of mitochondrial presequences. This may reflect a more specific presequence binding function of plant mt-hsp70s in comparison to DnaK.

A few reports have been presented of a direct interaction between the hsp70 family of proteins and individual mitochondrial presequences but none on the interaction of the mt-hsp70 with the presequences. The presequence of mitochondrial aspartate aminotransferase has been shown to bind to DnaK (Pierpaoli *et al.*, 1997; Schmid *et al.*, 1992) and to hsp70 in a cell-free transcription-translation system (Lain *et al.*, 1994, 1995). Cytosolic hsp70 (Ssa1p) can bind to synthetic peptides with amino acid compositions which form an amphiphilic α-helix, typical of mitochondrial presequences (Endo *et al.*, 1996). *In vitro* import experiments using folded proteins containing a longer presequence suggest indirectly that

Table 4. Buried hydrophobic surface area of presequence heptapeptides with DnaK and mt-hsp70

Peptides	DnaK Binding motif	DnaK binding area	Affinity	Mt-hsp70 Binding motif	Mt-hsp70 Binding area
LRRASLG	LRRASLG	497	–[b]	LRRASLG	499
LLTNRGL	LLTNRGL	528	–[b]	LLTNRGL	533
NRLLLTG	NRLLLTG	663	+[b]	NRLLLTG	666
Peptide1	RLLASLL	657	+[a]	RRLLASL	657
Peptide2	TRHLLNL	679	+[a]	TRHLLNL	705
Peptide3	LRALKVR	655	+[a]	SRLRALK	648
Peptide4	TVRRALL	676	+[a]	TVRRALL	671
Peptide5	NSIPKSA	540	-[a]	NSIPKSA	569

–, Peptides with low affinity to DnaK. +, Peptides with high affinity to DnaK.
[a] Affinity of peptides to DnaK deduced from luciferase refolding experiment.
[b] Peptide affinity to DnaK came from biochemical studies (Gragerov *et al.*, 1994; Zhu *et al.*, 1996).

mt-hsp70 can interact with presequences (Matouschek *et al.*, 1997).

In our study, we have analysed all available, statistically relevant mitochondrial presequences from mammals, plants, *N. crassa* and yeast. By using computer algorithms for the prediction of DnaK-binding sites (Rüdiger *et al.*, 1997). We found a potential DnaK-binding motif in almost all mitochondrial presequences (cf. Table 1, Figure 3). Most of the DnaK binding motifs are located at the N-terminal region or in the middle of the presequence. On the basis of these results, we chemically synthesised four peptides corresponding to the predicted strong DnaK-binding motifs and one peptide with no likelihood of binding. We have simulated docking of the peptides to the substrate-binding cleft of both chaperones and have calculated the buried hydrophobic surface binding area for interaction of these peptides with the chaperones (cf. Table 4). The data obtained showed a good correlation between the predicted likelihood of binding and a value of the buried hydrophobic surface binding area for the peptides that were synthesized on the basis of the prediction. This correlation was also seen for the three control peptides for which binding affinity to DnaK has been determined from direct biochemical binding studies. Also the effect of the synthetic presequence peptides on DnaK-assisted luciferase refolding confirmed the prediction. The peptide with no likelihood of DnaK-binding had no inhibitory effect upon luciferase refolding and had a low calculated buried hydrophobic surface area within both chaperones, whereas the peptide with the highest predicted likelihood of DnaK-binding, showed a high degree of inhibitory activity on refolding of luciferase and a high buried hydrophobic surface area value.

In summary, we have shown that mt-hsp70 is very similar to DnaK in terms of both structure and the predicted substrate binding specificity. Therefore, we suggest that DnaK is a suitable model system for studying interaction between the mitochondrial presequences and mt-hsp70. We have further shown that DnaK can bind mitochondrial presequences and on the basis of the modeling data we have extrapolated these results to the mt-hsp70. The fact that the predicted DnaK-binding motifs are present in almost all mitochondrial presequences of all phylogenetic groups strongly supports the idea that mt-hsp70 binds the presequence of the incoming mitochondrial precursors to the matrix. The fact that DnaK-binding motifs were found in such a broad range of presequences suggests that the interaction of the mitochondrial presequences with mt-hsp70 might be a general phenomenon with consequences for the mechanism for mitochondrial protein import.

Material and Methods

Chemicals

ATP, purified firefly luciferase was purchased from Sigma; luciferase assay reagent from Promega; DTT from Scientific Imaging Systems, NY, USA.

Molecular chaperones DnaK, DnaJ and GrpE

The overexpressed DnaK, DnaJ and GrpE were kindly provided by Dr Hans-Joachim Schönfeld (Schönfeld *et al.*, 1995).

Mitochondrial presequences

Mitochondrial presequence structures were collected from on-line OWL database (Schneider *et al.*, 1998). Only the presequences with cleavage sites determined by N-terminal sequencing were used. These presequences are statistically relevant. The numbers of presequences collected were: plants, 32; mammals, 61; yeast, 39; and *Neurospora crassa*, 15.

Synthesis of presequence peptides

Five peptides (Table 3) corresponding to four plant mitochondrial presequences, $pF_1\beta$ from *Nicotiana plumbaginifolia*, α-MPP(2) and β-MPP(1) from potato, and alternative oxidase (2) from soybean were synthesized by Boc/benzyl solid phase synthetic methods (Merrifield, 1986) on *p*-methylbenzhydrylamine resin, to provide the C-terminal carboxamides after HF deprotection and cleavage. The synthetic products were purified to homogeneity by reverse-phase chromatography and

gave amino acid analyses and MALDI-TOF mass spectra consistent with the expected structures.

hsp70 sequences and alignment

Sequences of DnaK, and mt-hsp70s from pea, potato and kidney bean were obtained from Swissprot and spinach mt-hsp70 was obtained from Genebank. Alignment was performed using Clustal W (1.7) (Thompson *et al.*, 1994) using default parameters.

Establishment of the structural model of pea mt-hsp70

The structure of the substrate-binding domain of DnaK was downloaded from PDB (1DKX, Brookhaven National Laboratory) which was elucidated by Zhu *et al.* (1996), comprising residues 389 to 607 of DnaK. Through alignment of pea mt-hsp70 precursor with DnaK, we picked out the region from 381 to 599 of mature protein for kidney bean (HS7M_PHAVU), spinach (AF039084, Genebank access number) and pea (HS7M_PEA), and the region from 381 to 600 for potato (HS7M_SOLTU) mature protein which correspond to the DnaK substrate-binding domain. The model of the substrate binding domain of pea mt-hsp70 was created using the standard method for homology modeling in ICM (Cardozo *et al.*, 1995). The complete model and the substrate-binding pocket together with a binding peptide (NRLLLTG) are shown in Figure 2.

The quality of the model was checked with WHATIF (Vriend, 1990), Errat (Colovos & Yeates, 1993), the verify3D server (Luthy *et al.*, 1992) and PROCHECK (Laskowski *et al.*, 1993). The quality of the model seemed good according to all methods except that bad contacts were found by Whatif for residues 450 and higher and that Errat reported two suspicious regions at residues 381 to 407 and around 460. The only residue involved in the binding that did not seem to be in a very good environment seemed to be Leu396.

Visualisation of molecular structure

All images of the molecular structures in Figure 2 were created using WeblabViewer Pro (Molecular Simulations Inc.). Hydrogen bonds in Table 1 were calculated using PdbViewer (Guex & Peitsch, 1997).

Prediction of DnaK-binding motifs on mitochondrial presequences

The prediction was performed according to the DnaK-binding algorithm by Rüdiger *et al.* (1997). Presequences were cut into short peptides of 13-amino acid residues before running the algorithm. The adjacent peptides were overlapped by 12 amino acid residues. Every peptide was scored according to the algorithm. The whole process was very similar to a scanning window moving along the presequence one amino acid residue per step.

Inhibition of refolding of denatured luciferase by synthesized peptides

Luciferase unfolding and refolding was performed as described by Buchberger *et al.* (1994). Luciferase activity was assayed with a luminometer constructed by Bo Höjer (Department of Biochemistry, Stockholm University), which is equipped with a sensitive photomultiplier tube (Model R268, Hamamatsu TV Co., Tokyo, Japan; see Möller *et al.*, 1994 for details). The assay was based on the protocol supplied by the manufacturer, Promega. A 66 μM luciferase stock solution was made by dissolving purified luciferase in 1 M glycylglycine (pH 7.4). The stock enzyme solution was diluted ten times with unfolding buffer (25 mM Hepes-KOH (pH 7.6), 50 mM KCl, 5 mM $MgCl_2$, 5 mM β-mercaptoethanol, 6 M guanidinium-HCl) and incubated at room temperature for two hours. The denatured luciferase was suspended at a final concentration of 55 nM in refolding medium containing 25 mM Hepes-KOH (pH 7.6), 5 mM Mg-acetate, 50 mM KCl, 1 mM DTT, 1 mM ATP and 0.5 μM DnaK/DnaJ/GrpE each. The enzyme activity was assayed by recording the luminescence after adding 1 μl of refolded enzyme to 50 μl of luciferase assay reagent.

Calculation of the hydrophobic binding area of the presequence heptapeptides with DnaK and mt-hsp70

Each sequence of the synthesized 13-mer peptide was cut into pieces of seven amino acid residues. The conformation of the heptapeptides was modeled after the NRLLLTG peptide crystallized with DnaK (Zhu *et al.*, 1996) using ICM (Abagyan *et al.*, 1994) according to the script supplied by the program. The heptapeptides were docked into the binding pocket of both mt-hsp70 and DnaK. The complete system was minimized in torsional space using default parameters from ICM. The buried hydrophobic surface area was then calculated for each heptapeptide in the modeling program. Three peptides for which binding affinity to DnaK had been tested with biochemical methods (Gragerov *et al.*, 1994; Zhu *et al.*, 1996) were used as control peptides. Of these three, two were negative controls, which did not bind DnaK and one of them was a positive control, which is a good binder to DnaK.

Acknowledgements

We thank Dr Stefan Rüdiger and Professor Bernd Bukav for kindly supplying the computer algorithm program for prediction of DnaK binding motif on natural protein. Molecular chaperones DnaK/DnaJ/GrpE were a very kind gift from Dr Hans-Joachim Schönfeld. We thank Dr Patrick Dessi for stimulating discussions and critical reading of the manuscript. This work was supported by grants from the Swedish Natural Science Research Council and The Swedish Strategic Foundation to E.G. and A.E. and from the Dirección General de Investigatión Cientifica y Técnica and Fondo de Investigaciones Sanitarias to D.A.

References

Abagyan, R. A., Totrov, M. M. & Kuznetsov, D. N. (1994). ICM - a new method for protein modeling and design. Applications to docking and structure prediction from the distorted native conformation. *J. Comp. Chem.* **15**, 488-506.

Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F. & Gething, M. J.

(1993). Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell,* **75**, 717-728.

Buchberger, A., Valencia, A., McMacken, R., Sander, C. & Bukau, B. (1994). The chaperone function of DnaK requires the coupling of ATPase activity with substrate binding through residue E171. *EMBO J.* **13**, 1687-1695.

Cardozo, T., Totrov, M. & Abagyan, R. (1995). Homology modeling by the ICM method. *Proteins: Struct. Funct. Genet.* **23**, 403-411.

Colovos, C. & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511-1519.

de Crouy-Chanel, A., Kohiyama, M. & Richarme, G. (1996). Specificity of DnaK for arginine/lysine and effect of DnaJ on the amino acid specificity of DnaK. *J. Biol. Chem.* **271**, 15486-15490.

Endo, T., Mitsui, S., Nakai, M. & Rosie, D. (1996). Binding of mitochondrial presequences to yeast cytosolic heat shock protein 70 depends on the amphiphilicity of the presequence. *J. Biol. Chem.* **271**, 4161-4167.

Fourie, A. M., Sambrook, J. F. & Gething, M. J. H. (1994). Common and divergent peptide binding specificities of hsp70 molecular chaperones. *J. Biol. Chem.* **269**, 30470-30478.

Gething, M. J., Blondelguindi, S., Buchner, J., Fourie, A., Knarr, G., Modrow, S., Nanu, L., Segal, M. & Sambrook, J. (1995). Binding sites for Hsp70 molecular chaperones in natural proteins. *Cold Spring Harbor Symp. Quant. Biol.* **60**, 417-428.

Glick, B. S. (1995). Can Hsp70 proteins act as force-generating motors? *Cell,* **80**, 11-14.

Gragerov, A., Zeng, L., Zhao, X., Burkholder, W. & Gottesman, M. E. (1994). Specificity of DnaK-peptide binding. *J. Mol. Biol.* **235**, 848-854.

Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis,* **18**, 2714-2723.

Hartl, F.-U. (1996). Molecular chaperones in cellular protein folding. *Nature,* **381**, 571-580.

Hightower, L. E., Sadis, S. E. & Takenaka, I. M. (1994). Interactions of vertebrate hsc70 and hsp70 with unfolded protein and peptides. In *The Biology of Heat Shock Proteins and Molecular Chaperones* (Morimoto, R. I., Tissières, A. & Georgopoulos, C., eds), pp. 179-207, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Horst, M., Oppliger, W., Feifel, B., Schatz, G. & Glick, B. S. (1996). The mitochondrial protein import motor: dissociation of mitochondrial hsp70 from its membrane anchor requires ATP binding rather than ATP hydrolysis. *Protein Sci.* **5**, 759-767.

Lain, B., Iriarte, A. & Martinez-Carrion, M. (1994). Dependence of the folding and import of the precursor to mitochondrial aspartate aminotransferase on the nature of the cell-free translation system. *J. Biol. Chem.* **269**, 15588-15596.

Lain, B., Iriarte, A., Mattingly, J. R., Jr, Moreno, J. I. & Martinez-Carrion, M. (1995). Structural features of the precursor to mitochondrial aspartate aminotransferase responsible for binding to hsp70. *J. Biol. Chem.* **270**, 24732-24739.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

Luthy, R., Bowie, J. U. & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature,* **356**, 83-85.

Margulis, L. (1975). Symbiotic theory of the origin of eukaryotic organelles; criteria for proof. *Symp. Soc. Exp. Biol.* **29**, 21-38.

Matouschek, A., Azem, A., Ratliff, K., Glick, B. S., Schmid, K. & Schatz, G. (1997). Active unfolding of precursor proteins dueing mitochondrial protein import. *EMBO J.* **16**, 6727-6236.

Merrifield, B. (1986). Solid phase synthesis. *Science,* **232**, 341-347.

Möller, A., Gustafsson, K. & Jansson, J. K. (1994). Specific monitoring by PCR amplification and bioluminescence of flirefly luciferase gene-tagged bacteria added to environmental samples. *FEMS Microbiol. Ecol.* **15**, 193-206.

Neupert, W. (1997). Protein import into mitochondria. *Annu. Rev. Biochem.* **66**, 863-917.

Pfanner, N. & Meijer, M. (1995). Protein sorting. Pulling in the proteins. *Curr. Biol.* **5**, 132-135.

Pfanner, N., Craig, E. A. & Honlinger, A. (1997). Mitochondrial preprotein translocase. *Annu. Rev. Cell. Dev. Biol.* **13**, 25-51.

Pierpaoli, E. V., Sandmeier, E., Baici, A., Schonfeld, H. J., Gisler, S. & Christen, P. (1997). The power stroke of the DnaK/DnaJ/GrpE molecular chaperone system. *J. Mol. Biol.* **269**, 757-768.

Rassow, J., von Ahsen, O., Bomer, U. & Pfanner, N. (1997). Molecular chaperones: towards a charcterization of the heat-shock protein 70 family. *Trends Cell Biol.* **7**, 129-133.

Richarme, G. & Kohiyama, M. (1993). Specificity of *Escherichia coli* chaperone DnaK (70-KDa heat shock protein) for hydrophobic amino acids. *J. Biol. Chem.* **268**, 24074-24077.

Rüdiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**, 1501-1507.

Schatz, G. & Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science,* **271**, 1519-1526.

Schmid, D., Jaussi, R. & Christen, P. (1992). Precursor of mitochondrial aspartate aminotransferase synthesized in *Escherichia coli* is complexed with heat-shock protein DnaK. *Eur. J. Biochem.* **208**, 699-704.

Schmid, D., Baici, A., Gehring, H. & Christen, P. (1994). Kinetics of molecular chaperone action. *Science,* **263**, 971-973.

Schneider, G., Sjöling, S., Wallin, E., Wrede, P., Glaser, E. & von Heijne, G. (1998). Feature-extraction from endopeptidase cleavage sites in mitochondrial targeting peptides. *Proteins: Struct. Funct. Genet.* **30**, 49-60.

Schönfeld, H. J., Schmidt, D., Schroder, H. & Bukau, B. (1995). The DnaK chaperone system of *Escherichia coli*: quaternary structures and interactions of the DnaK and GrpE components. *J. Biol. Chem.* **270**, 2183-2189.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.

Ungermann, C., Neupert, W. & Cyr, D. M. (1994). The role of Hsp70 in conferring unidirectionality on

protein translocation into mitochondria. *Science,* **266**, 1250-1253.

Vriend, G. (1990). WHATIF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52-56.

Zhu, X. T., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E. & Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science,* **272**, 1606-1614.

Edited by A. R. Fersht

(Received 1 October 1998; received in revised form 2 March 1999; accepted 5 March 1999)