

Molecular characterization of a plant mitochondrial chaperone GrpE

Malla Padidam^{1,*}, Vijay S. Reddy², Roger N. Beachy¹ and Claude M. Fauquet¹

¹*International Laboratory for Tropical Agricultural Biotechnology (ILTAB/ORSTOM-TSRI), Division of Plant Biology-BCC206, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA (*author for correspondence);* ²*Department of Molecular Biology-MB31, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.*

Received 4 August 1998; accepted in revised form 22 October 1998

Key words: *Arabidopsis thaliana*, chaperone, GrpE structure, mitochondrial GrpE, mitochondrial transit peptide, *Nicotiana tabacum*

Abstract

Escherichia coli DnaK (Hsp70) cooperates with DnaJ and GrpE in its essential role as a molecular chaperone. Function of mitochondrial Hsp70 (mHsp70) in protein folding and organellar import in eukaryotes is critically dependent on GrpE. We cloned two genes from tobacco (*Nicotiana tabacum*) BY2 cells based on peptide sequences from a purified protein. The predicted amino acid sequences of both clones resembled that of GrpE from *E. coli* and its homologues from eukaryotes, and a cDNA clone from *Arabidopsis thaliana*. One gene (Type 1) encoded a deduced protein that was identical to the purified protein while the other (Type 2) encoded a deduced protein that has 80% sequence identity to Type 1. Both tobacco and *Arabidopsis thaliana* GrpE homologues bound to DnaK and ATP inhibited this binding. The tobacco GrpE homologue contained a typical N-terminal mitochondrial target presequence of 64 residues and the presequence directed the green fluorescent protein to tobacco mitochondria. The tobacco GrpE homologue also associated with mHsp70 when reintroduced into BY2 protoplasts, and this association was disrupted by ATP. A three-dimensional structure for the tobacco GrpE homologue was modeled based on the X-ray structure of *E. coli* GrpE complexed with DnaK. The modeled structure has the same overall structure as *E. coli* GrpE. We propose that the tobacco GrpE homologue interacts with mHsp70 in a manner analogous to *E. coli* GrpE with DnaK and designate it as tobacco mitochondrial GrpE (NtmGrpE).

Introduction

Heat shock proteins (Hsp) are ubiquitous and have multiple functions, including protein folding and translocation across membranes. Among the heat shock proteins, the Hsp70 family of proteins are best studied; these proteins act by binding and releasing extended peptide segments enriched in hydrophobic side chains in an ATP-dependent manner [3, 26, 30, 35]. Hsp70 homologues in eukaryotes are found in the cytoplasm, nucleus, ER/Golgi, plastids, and mitochondria [7, 10, 12, 25, 29]. The Hsp70 homologues found in mitochondria and chloroplasts are similar to

E. coli DnaK (Hsp70) in both sequence and function. The function of DnaK as molecular chaperone is enhanced more than 50-fold by association with the chaperone DnaJ and the nucleotide exchange factor GrpE [32, 35, 44]. ATP-bound DnaK binds and releases peptides in a rapid manner while the ADP-bound DnaK binds peptides in a stable manner. DnaJ stimulates the ATP hydrolysis activity of DnaK and peptide release is catalyzed by GrpE. Recent studies in yeast and mammals established the presence of GrpE and DnaJ homologues in mitochondria. In yeast and mammals, the function of mitochondrial Hsp70 (mHsp70), but not of the cytosolic Hsp70, critically depends on GrpE and DnaJ homologues [2, 6, 11, 23, 28]. They facilitate protein folding and import into mitochondria.

The nucleotide sequence data reported will appear in the GenBank Nucleotide Sequence Database under the accession numbers AF098635 and AP098636.

In plants, nuclear-encoded homologues of Hsp70 that occur in mitochondria and chloroplasts have been characterized [7, 20, 22, 29, 40, 46–48]; each has N-terminal transit peptide that is necessary for translocation to mitochondria and chloroplasts. They are constitutively expressed, but their synthesis is enhanced upon heat stress. DnaJ homologues have been shown to occur in *Arabidopsis* mitochondria [21], pea chloroplasts [41], and onion cytosol [4, 5]. Presence of a GrpE-like protein in pea chloroplasts was demonstrated by western blot analysis using antiserum raised against *E. coli* GrpE [41]. GrpE homologue has not been characterized from plant mitochondria. Characterizing plant mitochondrial GrpE homologue will guide efforts to understand the function of mHsp70 in protein folding and import. This paper reports molecular, biochemical, and structural characterization of mitochondrial GrpE from tobacco (*Nicotiana tabacum*).

Materials and methods

Purification of a tobacco protein

Tobacco BY2 suspension cells were cultured in the dark at 27 °C in MS medium as described before [27]. Cells were collected on a sintered glass funnel one week after subculture and washed once with TBS (50 mM Tris-HCl pH 7.5, 150 mM NaCl). All subsequent operations were carried out at 4 °C. A 400 g portion (wet weight) of cells was mixed with 400 ml of TBS and lysed in a Waring blender (five 1-min pulses). The 37 000 × *g* supernatant was subjected to ammonium sulfate fractionation and proteins that precipitated at 50% and 60% saturation were pooled and subjected to size exclusion chromatography on Sephacryl S-200 (60 cm × 1.6 cm column, Amersham Pharmacia Biotech) equilibrated with TBS. Fractions containing a 27 kDa protein that cross-reacted with an affinity-purified polyclonal antibody directed against mammalian cell cycle inhibitor protein p21 (Santa Cruz Biotechnology) were identified by western blot analysis and pooled. The pooled fractions were then boiled for 10 min and centrifuged for 20 min at 37 000 × *g* to pellet precipitated proteins. The supernatant was again subjected to chromatography on Sephacryl S-200. Fractions containing the 27 kDa protein that cross-reacted with the anti-p21 antibody were pooled, concentrated by ultrafiltration, and bound to Mono Q anion-exchange resin (Amersham Pharmacia

Biotech) equilibrated with 20 mM Tris-HCl pH 8.5. Bound proteins were eluted from the column with a linear 0 to 1 M NaCl gradient. Fraction containing the 27 kDa protein that cross-reacted with the anti-p21 antibody was separated on a 13% SDS-PAGE and blotted onto PVDF membrane (Millipore). The 27 kDa protein band was cut out after staining the membrane with Ponceau-S.

N-terminal sequencing of protein and tryptic peptides

Protein that bound anti-p21 antibody on PVDF membrane was used for N-terminal sequence analysis as well as digestion with trypsin. Fragments resulting from trypsin digestion were separated on HPLC [18]. Limited N-terminal sequence of the protein and complete sequence of three tryptic peptides were determined using automated Edman degradation [24].

Cloning of the tobacco GrpE homologue

All RNA and DNA manipulations were carried out by standard methods [38]. The cDNA, synthesized from poly(A)⁺ RNA prepared from actively growing BY2 cells, was used for PCR amplification using degenerate oligonucleotide primers based on two tryptic peptide sequences of the purified 27 kDa tobacco protein. A partial cDNA (ca. 170 bp) was produced using the forward primer MP235 (5'-CAGAATTCTAYGAYCCNACNAAYGARCARTT-YGAYCC-3'; based on the peptide sequence PT-NEQFDP; *Eco*RI site is underlined) and reverse primer MP246 (5'-AAGAGCTCTTAGGCGTCCWSYTGRTICIGCYTCIGTYTCIACIGCIACNGT-3'; based on the peptide sequence IRPAEVLGV; *Sac*I site is underlined) and sequenced. A new specific reverse primer MP254 (5'-CAGGATCCGGTACTTGAATACTGCA-3', *Bam*HI site is underlined) was designed based on the partial cDNA sequence. Full-length clones were obtained by using a method for performing both 5' and 3' RACE (rapid amplification of cDNA ends) from the same template [42]. Briefly, double-stranded (ds) cDNA prepared from poly(A)⁺ RNA was ligated to the Marathon cDNA adapter with *Not*I site (Clontech). This ds cDNA was used in PCR (20 amplification cycles) to amplify 5' region of the cDNA using a primer (adapter primer AP1, 5'-CCATCCTAATACGACTCACTATAGGGC-3') that annealed to the adapter sequence at the 5' end of the cDNA and reverse primer MP254. Three 5' RACE clones were sequenced to design a new primer that annealed at the puta-

tive initiation codon (forward primer MP256; 5'-TTGTTCGACATGTCGGTGAGTAGAATCTCATCTCGG-3', *SalI* and *AflIII* sites are underlined). Subsequently, forward primer MP256 and adapter primer AP1 (that also anneals to the Marathon adapter at the 3' end of cDNA) were used in 3' RACE PCR to generate full-length fragments that contain sequences from the putative initiation codon to the polyadenylation site. Cloning and sequencing then identified two types of clones; those that encoded a deduced protein that was identical to the purified 27 kDa protein (Type 1 clones), and others that encoded a similar but not identical protein (Type 2 clones).

Arabidopsis cDNA clone AR192 (GenBank accession number D88745) that has sequence identity to tobacco Type 1 and Type 2 clones was cloned by PCR amplification of reverse-transcribed cDNA made from mRNA prepared from Columbia ecotype. Forward primer MP242 (5'-AAGGATCCCATATGTCTATGATGGATTTCGTTTG-3'; *BamHI* site is underlined) and reverse primer MP240 (5'-ATAGCGGCCGCTTAAGCATCAGACTCTTTCTTTTC-3'; *NotI* site is underlined) for PCR amplification were designed to amplify the putative coding region.

Expression and purification of proteins in E. coli

Expression vector pET32a (Novagen) was used to express proteins in *E. coli*. The pET32a vector was designed for high-level expression of peptide sequences fused to 109 amino acid long thioredoxin protein, 6 His residues, and S-Tag sequence. Full-length Type 1 (coding for amino acids 1–299) and Type 2 (coding for amino acids 1–304) clones were digested with *AflIII* and *NotI* and the insert was cloned into *NcoI* and *NotI* sites on pET32a. These plasmids encode recombinant proteins with predicted molecular mass of 51 kDa. The sequence coding for amino acids 65–299 of Type 1 protein was cloned after PCR amplification with forward primer MP282 (5'-CAGGATCCTCTGCATCACCCCAACCT-3'; *BamHI* site is underlined) and reverse primer MP270 (5'-ATAGCGGCCGCTTAAGCTTCAGTGTTTTGATCGCCTG-3'; *NotI* site is underlined). The resulting clone encodes a recombinant protein of 43 kDa. *Arabidopsis* cDNA clone AR192 (sequence coding for amino acids 1–273) was cleaved with *BamHI* and *NotI* and ligated into pET32a that was also cleaved with *BamHI* and *NotI*, resulting in a plasmid that encodes a recombinant protein of 48 kDa. A control expression plasmid was constructed by cloning gene 5 from *E.*

coli phage M13 (encoding the single-stranded DNA binding protein of 9.7 kDa, [37]) into pET32a at the *BamHI* and *SacI* sites by PCR amplification. This plasmid encodes a fusion protein of 27 kDa.

E. coli AD494 strains containing the above expression plasmids were grown in LB medium at 37 °C until the OD₆₀₀ reached 0.5–0.7. Protein expression was induced by adding 1 mM IPTG and incubating at 25 °C for 4 h. Induced cells were pelleted and resuspended in S-Tag purification buffer (20 mM Tris pH 7.5, 150 mM NaCl, 0.1% Triton X-100) with or without 5 mM ATP and lysed by sonication. The soluble fusion proteins in the lysates were affinity-purified on S-protein agarose according to the manufacturer's suggestions (Novagen). Briefly, lysates were bound to S-protein agarose at room temperature for 30 min and unbound proteins were removed by six washes with purification buffer with or without 5 mM ATP. The bound proteins were eluted with 0.5 M acetic acid and pH of eluted fractions was neutralized by the addition of Tris buffer.

The purity and identity of the proteins were determined by Coomassie blue staining following SDS-PAGE (13%) and by western blotting with S-protein that binds to S-Tag (Novagen), with antisera directed against *E. coli* GrpE (gift of B. Bukau, Freiburg) or tomato mHsp68 (gift of D. Neumann, Halle).

Mitochondrial localization

The DNA sequence coding for amino acids 1–64 from Type 1 clone was amplified by PCR using forward primer MP256 and reverse primer MP293 (5'-TTCCATGGATGAGATTCCAAACCATTGC-3'; *NcoI* site is underlined). The amplified fragment was digested with *AflIII* and *NcoI* and cloned into plasmid pCATGFP at the *NcoI* site. pCATGFP is a plant expression vector containing a modified, brightly fluorescent GFP gene under the control of duplicated cauliflower mosaic virus 35S promoter (C. Reichel, personal communication). Cloning at the *NcoI* site present at the initiation codon of GFP resulted in N-terminal translational fusion.

Plasmid DNA was introduced into protoplasts prepared from BY2 cells as described previously [31]. Electroporation sample contained 1×10^6 protoplasts and 40 µg of plasmid DNA. After electroporation, protoplasts were cultured for 24 h at 27 °C in the dark and observed under a Nikon fluorescence microscope. Transfected protoplasts were also stained with mitochondrial vital dye Rhodamine 6G chloride [9] (Mole-



Figure 1. Comparison of the amino acid sequences of the tobacco proteins with GrpE proteins. The amino acid sequences deduced from the cloned tobacco genes Type 1 (N.t.1) and Type 2 (N.t.2) were aligned with the *Arabidopsis thaliana* cDNA clone AR192 (A.t., GenBank accession number D88745) and GrpE proteins from *E. coli* (E.c., X07863), *Saccharomyces cerevisiae* (S.c., D26059), and *Rattus norvegicus* (R.n., U62940). Also shown are conserved residues in 20 GrpE homologue sequences (C20; [28]). The putative cleavage site of N-terminal mitochondrial transit peptide of Type 1 tobacco protein is marked by an arrow. Regions of the deduced amino acid sequence that are identical to peptide sequences obtained from purified tobacco protein are underlined. The asterisks above the sequence indicate identical amino acids between the Type 1 tobacco protein and the *E. coli* protein. Five invariant residues in all GrpE homologues are boxed.

cular Probes) by incubating in 100 nM Rhodamine 6G chloride solution for 5–10 min before observing. Fluorescence due to GFP and Rhodamine 6G chloride were recorded on Kodak Ektachrome P1600 film using FITC and Rhodamine filter sets, respectively.

Immunoprecipitation and western blotting

To detect binding of the tobacco GrpE homologue to mitochondrial Hsp70, Flag peptide epitope (DYKD-DDDK; [16]) tagged GrpE was expressed in BY2 protoplasts using a tomato leaf curl virus (ToLCV) expression vector [31]. Expression of β -glucuronidase (GUS) under the coat protein (CP) promoter in the ToLCV vector results in at least ten times more GUS enzyme activity compared to GUS enzyme activity produced from a plant expression vector with duplicated 35S promoter (unpublished observations). Sequences coding for the Flag peptide epitope were added at the 3' end of the Type 1 tobacco

GrpE homologue sequence (following the codon for amino acid 299) by PCR amplification using forward primer MP256 and reverse primer MP283 (5'-TAGAATTCTTACTTGTGCATCGTCATCTTTGTAGTCAGTGTTTGTACTTTGATC-3'; *EcoRI* site is underlined). The amplified fragment was digested with *AflIII* and end filled at the *EcoRI*, and subsequently ligated into *AflIII* and end-filled *HindIII* sites of ToLCV DNA T251te. The cloning resulted in replacement of ToLCV AV2 and CP genes with tobacco GrpE homologue sequence.

The ToLCV expression vector was introduced into BY2 protoplasts as described above. Protoplasts were collected 48 h after transfection and lysed with a hand-held Polytron in TBSN buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP40) containing a cocktail of protease inhibitors (Boehringer Mannheim, catalog number 1697498) and with or without 5 mM ATP. Lysates (15 000 \times g supernatants) were immunoprecipitated with anti-Flag monoclonal M2 antibody

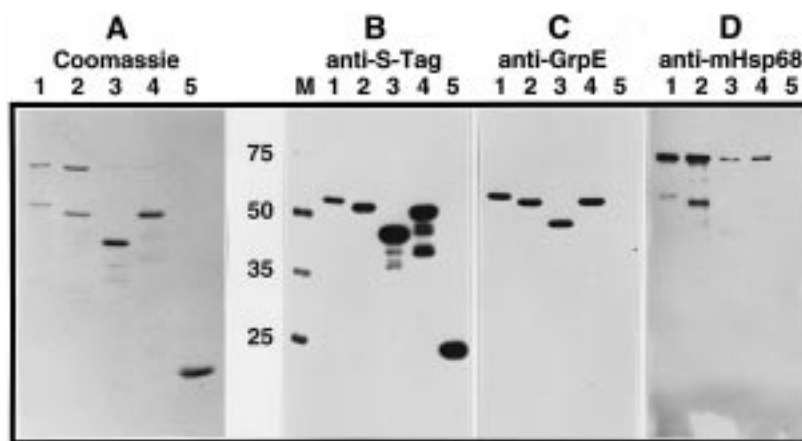


Figure 2. Identification of GrpE homologues from tobacco and *Arabidopsis*, and DnaK from *E. coli*. A. Coomassie blue-stained SDS-PAGE gel. B. Western blot probed with S-protein. C. Western blot probed with anti-*E. coli* GrpE antisera. D. Western blot probed with anti-tomato mHsp68 antisera. Lane 1, tobacco Type 1; lane 2, tobacco Type 2; lane 3, tobacco Type 2 mature protein; lane 4, *Arabidopsis* AR192; lane 5, phage M13 gene 5 protein, and lane M, protein size marker (molecular masses in kDa are indicated at the left). Each lane contained 1 mg of purified fusion protein from *E. coli*. Protein purifications and western blot analyses were carried out as described in Materials and methods.

covalently linked to agarose (Sigma). Immune complexes were washed four times with TBSN and once with TBS. The samples were then heated in Leammli sample buffer (90 °C for 5 min), fractionated by SDS-PAGE (13%), and transferred onto PVDF membrane (Schleicher & Schuell). Immunoprecipitated proteins were visualized with anti-Flag antibody (Sigma) or anti-tomato mHsp68 antisera and ECL-western blot reagents (Pierce).

Modeling of the tobacco GrpE homologue three-dimensional structure

The three dimensional model for tobacco GrpE was generated based on homology modeling [19, 36], using the *E. coli* GrpE-DnaK complex structure as a template [14]. Initially the complete sequence of the Type 1 protein was aligned with the GrpE sequences from various sources using the alignment program (based on MaxHom method [34, 39]) available at Predict Protein server on the world wide web (www.embl-heidelberg.de/predictprotein). The tobacco GrpE homologue sequence was mapped onto the template *E. coli* GrpE structure, by mutating the residues as suggested by the multiple sequence alignment using the crystallographic graphics program 'O' [17]. Subsequently, the model of the tobacco GrpE homologue was subjected to energy minimization using the program CHARMM [8]. Initially, harmonic restraints (force constant 25 kcal/mol Å) were imposed on all the non-hydrogen atoms, and the force constant was

reduced by 5 kcal/mol Å at the end of 1000 cycles of minimization. All hydrogen atom parameters and topologies from version 22 of CHARMM were used in all energy calculations. A total of six rounds of minimization were performed by imposing decreasing amounts of harmonic constraints at the end of each round. At the end of the minimization, the root mean squared deviation for all the non-hydrogen atoms (1483) between the starting and final models is 1.4 Å, and ca. 1.0 Å for all the C α atoms (187).

Results

Cloning of GrpE homologues from tobacco

During our efforts to purify plant cell cycle-related proteins from tobacco BY2 cells we identified a protein that cross-reacted with an antibody raised against mammalian cell cycle inhibitor protein p21 [13]. The protein was resistant to heat denaturation and did not precipitate even after boiling for 10 min. It had an apparent molecular mass of 27 kDa on SDS-PAGE gel but eluted along with ca. 160 kDa molecular mass marker on Sephacryl S-200 size exclusion chromatography, both before or after boiling (data not shown). The latter result suggested that the protein cross-reacting with the anti-p21 antibody is a homomultimeric or heteromeric complex. The protein was purified as described in Materials and methods. Limited N-terminal sequence analysis of purified protein yielded the amino acid sequence SSAAP/GQ/PP/NN/PE and

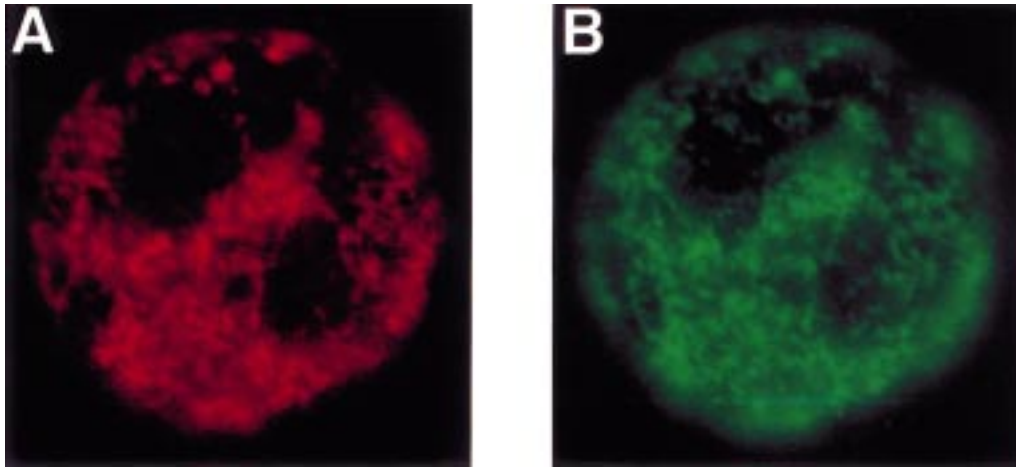


Figure 3. Mitochondrial localization of tobacco GrpE homologue and green fluorescent protein (GFP) fusion in BY2 protoplasts. The protoplast in A shows fluorescence due to Rodamine 6G chloride which stains mitochondria and the same protoplast in B shows fluorescence due to GFP. Protoplasts were transfected with a vector designed to express the putative mitochondrial transit peptide of the Type 1 homologue (64 residues) fused to GFP. Living cells were observed 24 h after transfection.

was not suitable for designing an oligonucleotide primer. Three internal peptides were then sequenced and the resulting peptide sequences are shown in Figure 1. Full-length cDNA clones were obtained by a combination of 5' and 3' RACE PCR procedures using the oligonucleotide primers based on the peptides. Sequencing of the 8 randomly selected clones revealed that 6 clones encoded protein that corresponded to the original peptide sequence and 2 clones encoded a protein that has 80% amino acid identity to the sequenced protein. However, they had 91% sequence similarity when conserved substitutions are considered (Figure 1). The two types of cDNA clones had no nucleotide sequence similarity in the 3'-untranslated region (data not shown). The clones that encoded a protein that is identical to the sequenced protein were designated Type 1 and the others were designated Type 2. The deduced Type 1 and Type 2 proteins have molecular masses of 33.64 and 34.05 kDa, respectively. Analysis of 35 additional clones by PCR using oligonucleotide primers specific to Type 1 or Type 2 clones showed that 29 clones were of Type 1 and 6 clones were of Type 2.

Comparison of the putative Type 1 and Type 2 protein products with entries in the GenBank and SwissProt databases revealed homology with a cDNA clone from *Arabidopsis* (AR192) and *E. coli* GrpE homologues from prokaryotes and eukaryotes (Figure 1). *Arabidopsis* cDNA clone AR192 was isolated by its ability to complement a yeast pheromone receptor-deficient mutant (GeneBank accession D88745). The

deduced Type 1 and Type 2 proteins displayed 45% identity to AR192 and 31–38% to GrpE homologues. However, both tobacco and *Arabidopsis* proteins contain 5 invariable residues present in 20 GrpE sequences (Figure 1). Members of GrpE family exhibit a low degree of sequence similarity (20–40%) even though they have the same function [28].

The sequence comparisons with entries in the GenBank and SwissProt databases using the BLAST search protocol [1] did not identify any sequence homology between the tobacco Type 1 protein and p21 even though anti-p21 antibody cross-reacted with the tobacco protein. Careful examination identified amino acids 107IDAEDLSRD115 of the Type 1 protein and amino acids 25VDSEQLSRD33 of p21 to be homologous. It is possible that the anti-p21 antibody cross-reacted with amino acids 107–115 of the Type 1 protein.

The predicted molecular mass of 34 kDa for the tobacco proteins is much higher than the apparent size of the purified protein (27 kDa). Comparison of the deduced amino acid sequence with the N-terminal amino acid sequence obtained from purified protein revealed a putative cleavage site after amino acid 64 for the Type 1 protein and after amino acid 66 for the Type 2 protein (Figure 1), suggesting a posttranslational cleavage to generate the mature protein. The putative leader peptide is rich in Ser residues and contains many basic amino acids while the predicted mature protein contains many acidic residues (Figure 1). The predicted isoelectric points (pI) for the putative

leader peptides are 10.56 (Type 1) and 11.58 (Type 2) while for the mature proteins the predicted pIs are 4.64 (Type 1) and 4.74 (Type 2). Many mitochondrial and chloroplast targeting peptides are rich in basic, hydroxylated, and hydrophobic residues [45]. Yeast, *Drosophila*, mice, rat, and human GrpE homologues are nuclear-encoded mitochondrial proteins and the rat mitochondrial GrpE was shown to contain a typical 27 residue N-terminal mitochondrial targeting sequence [28]. The predicted proteolytic removal of the leader peptides generated mature proteins with a predicted molecular mass of 26 kDa, which is similar to the observed molecular mass of 27 kDa. Taken together, the above results suggest that the tobacco and *Arabidopsis* genes encode proteins that are homologous to GrpE.

Tobacco and Arabidopsis GrpE homologues cross-react with anti-E. coli GrpE antisera and bind to DnaK

To confirm that the tobacco proteins are homologues of *E. coli* GrpE we expressed and purified Type 1, Type 2, and Type 1 mature (amino acids 65–299) proteins from *E. coli* and assayed their cross-reactivity to anti-*E. coli* GrpE antisera and binding to *E. coli* DnaK protein. It was shown previously that the yeast and mammalian GrpE homologues bind with high affinity to *E. coli* DnaK and this binding is inhibited by ATP [50]. We also cloned and expressed the *Arabidopsis* homologue (clone AR192) and gene 5 from *E. coli* phage M13. Gene 5 from M13 encodes a single-stranded DNA binding protein [37] and is used here as a negative control. The proteins were purified on a S-Tag affinity column; proteins recovered from the column were subjected to SDS-PAGE, and the gel was stained with Coomassie blue (Figure 2A). Each lane contained a protein of about the expected molecular mass for the fusion protein, i.e., ca. 53 kDa for Type 1, ca. 52 kDa for Type 2, ca. 45 kDa for Type 1 amino acids 65–299, ca. 50 kDa for *Arabidopsis* AR192, and ca. 25 kDa for gene 5. Western blotting of the gel shown in Figure 2A with S-protein (which binds to S-Tag) identified the fusion proteins (Figure 2B). The protein bands that migrated more rapidly than the expected molecular weights of fusion proteins are presumed to be degraded products. The tobacco and *Arabidopsis* proteins, but not the gene 5 protein, cross-reacted with antisera raised against *E. coli* GrpE (Figure 2C).

While purifying each fusion protein several other proteins also co-purified. Significantly, a protein of

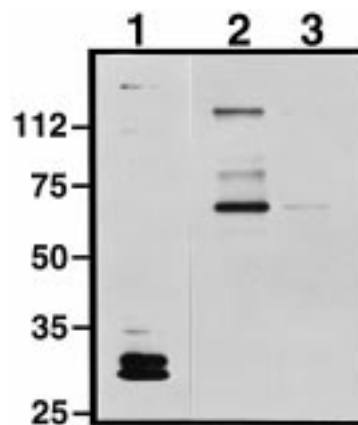


Figure 4. Interaction of the tobacco GrpE homologue with the mHsp70. BY2 protoplasts were transfected with a vector designed to express the Type 1 tobacco GrpE homologue tagged with a Flag epitope peptide. Protoplasts were lysed 48 h after transfection and used in immunoprecipitations with anti-Flag antibody linked to agarose. The immunoprecipitated proteins were detected on a western blot with either anti-Flag antibody (lane 1) or anti-tomato mHsp68 antisera (lanes 2 and 3). Extracts prepared for immunoprecipitation shown in lane 3 contained 5 mM ATP in the buffer while those in lanes 1 and 2 did not contain ATP. Molecular mass standards (kDa) are shown at the left.

molecular mass of ca. 72 kDa was co-purified along with the tobacco and *Arabidopsis* proteins but not with the gene 5 protein (Figure 2A). The molecular mass of this co-purified protein is similar to the molecular mass (72 kDa) of *E. coli* DnaK protein. The 72 kDa protein was not co-purified when 5 mM ATP was included in the purification buffer (data not shown). To determine if the 72 kDa protein is indeed DnaK, the gel shown in Figure 2A was blotted onto PVDF membrane and bound proteins were reacted with antisera raised against tomato mHsp68. Anti-mHsp68 antiserum was shown to recognize DnaK previously [29]. As expected, anti-mHsp68 antisera detected the co-purified 72 kDa protein (Figure 2D). Collectively, these results indicated that the tobacco Type 1 and Type 2 clones and *Arabidopsis* clone AR192 encode plant homologues of *E. coli* GrpE and, like GrpE, bind DnaK.

Type 1 tobacco GrpE homologue is localized to plant mitochondria and interacts with mHsp70

The difference in molecular mass between the purified and the deduced proteins, putative cleavage sites in the N-terminal region of the deduced proteins, and the similarity of the putative leader sequences to plant mitochondrial and chloroplast transit peptides sug-

gested that the tobacco GrpE homologues are targeted to organelles. To establish the cellular localization of the tobacco GrpE homologue, the putative leader sequence (encoding amino acids 1–64) of the Type 1 clone was fused to green fluorescent protein (GFP) at the N-terminus and cloned behind a duplicated 35S promoter. The fusion constructs were introduced into BY2 protoplasts and fluorescence due to GFP was analyzed 24 h after transfection.

Figure 3 shows mitochondrial localization of the GFP fusions. Mitochondria in transfected cells were identified by staining with the dye Rhodamine 6G chloride (Figure 3A). GFP was localized to mitochondria when fused to the putative leader sequence (Figure 3B) demonstrating that this sequence is sufficient to transport a protein into mitochondria.

Binding of the tobacco GrpE homologue to mHsp70 was demonstrated by transfecting BY2 protoplasts with the Type 1 clone tagged with a Flag peptide epitope (at the C-terminus) and determining if mHsp70 co-immunoprecipitates along with the epitope-tagged GrpE homologue. Tobacco mHsp70 was detected with antisera raised against tomato mHsp68 [29]. Anti-tomato mHsp68 antisera did not cross-react with Hsp70 homologues in the tomato cytosol and chloroplasts [29] and in tobacco BY2 cytosol (data not shown).

Protoplasts were transfected with tomato leaf curl virus (ToLCV) vector designed to express the Type 1 tobacco GrpE homologue with Flag peptide tag. As expected, the GrpE homologue (double bands at ca. 29 kDa position) was immunoprecipitated with anti-Flag agarose from protoplasts transfected with the ToLCV vector expressing the Flag-epitope-tagged GrpE homologue (Figure 4, lane 1). Similarly, mHsp70 co-immunoprecipitated along with Flag-epitope-tagged GrpE homologue (Figure 4, lane 2). Very little of mHsp70 co-immunoprecipitated when ATP was included in the immunoprecipitation buffer (Figure 4, lane 3) showing the specificity of binding. Thus, the tobacco GrpE homologue which is imported into mitochondria also interacts with mHsp70.

Modeling of the three-dimensional structure of the tobacco GrpE homologue

The three-dimensional model for the tobacco GrpE homologue (Type 1, amino acids 65–299) was constructed based on the X-ray structure of *E. coli* GrpE complexed with the ATPase domain of DnaK [14]. The extent of identical residues between the tobacco

protein and the GrpE homologues was 31–38% and the identity with *E. coli* GrpE is 34%. The conserved residues in the alignment of the various sequences of GrpE are found throughout the protein (Figure 1), and is a prerequisite for successful homology-based structural modeling. There are only two insertions of amino acids in the tobacco protein compared to *E. coli*, which also contributes to the success of the modeling. The structure of the *E. coli* GrpE monomer consists of primarily two domains: a long N-terminal α -helix domain and a domain comprising two short helices involved in the GrpE dimer formation followed by a C-terminal globular domain made up of β -structure (Figure 5A). The first insertion of 10 residues (Leu127-Arg136) is located in the middle of the long α -helix and was modeled as a continuing helix based on the secondary structure predictions. The second insertion of 13 amino acids (Ser192-Gly204) occurs on the loop connecting the two short helices of the second domain. The latter was modeled as a random coil structure that forms part of the loop connecting the two helices. The side-chain conformations of the mutated/replaced residues were chosen from one of the preferred rotamer configurations such that the side chain conformation of the replaced residue of the tobacco GrpE homologue fills the space occupied by the residue of *E. coli* GrpE. The side chain conformations of the conserved residues were kept unaltered. The mutation of residues and the selection of different side chain rotomers were carried out using the 'lego' option in the program 'O' [17]. The ribbon diagrams of the *E. coli* GrpE-DnaK complex and the modeled tobacco GrpE homologue with DnaK are shown in the Figure 5. The apparent insertions in the tobacco protein are shown in magenta. The location of the insertions suggests that they neither interfere with the dimer formation nor disrupt the interaction with the ATPase domain of DnaK.

There are 11 residues of the *E. coli* GrpE that are involved in establishing the GrpE-DnaK complex. Five of the eleven residues are conserved in the tobacco GrpE homologue while the other residues are replaced by homologous residues, which are predicted to provide the equivalent interactions with the DnaK. It is interesting to note that the only salt link found in the *E. coli* GrpE-DnaK complex formed between residues Lys-82 and Arg-183 of GrpE and Glu-28 of DnaK, is apparently maintained in the tobacco GrpE-DnaK complex. The salt bridge in the tobacco complex is formed between residues Arg-183 of tobacco GrpE and Glu-28 of DnaK, as Lys-82 of tobacco GrpE is

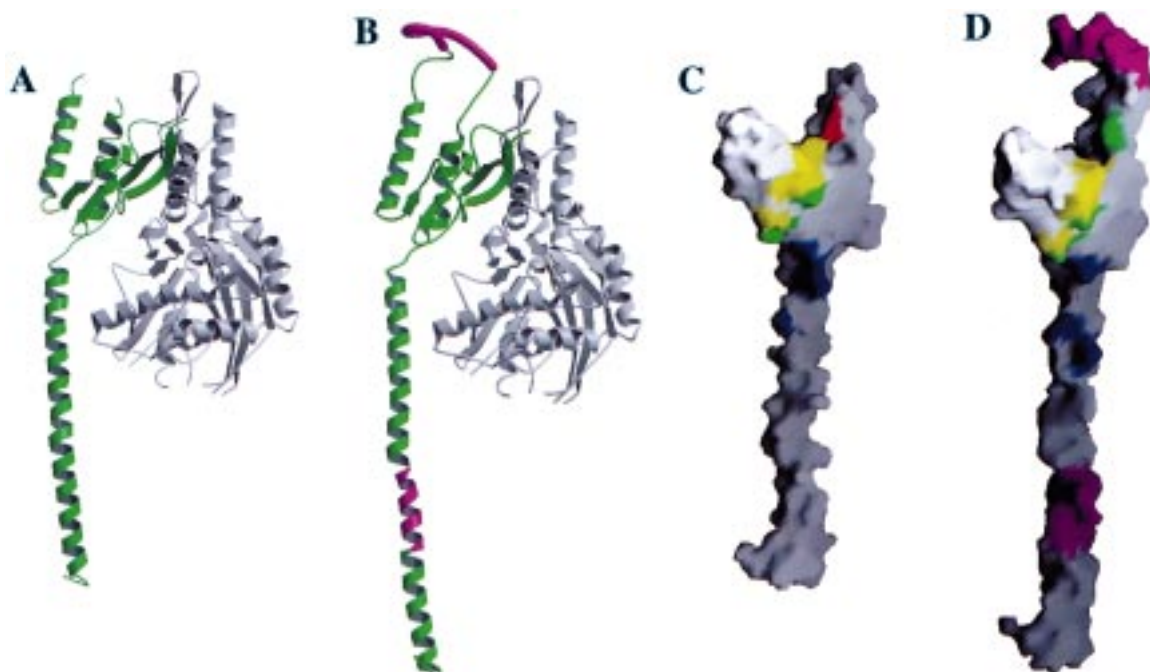


Figure 5. Schematic illustrations of the *E. coli* GrpE and the tobacco GrpE homologue structures. **A.** Ribbon drawing of the complex of *E. coli* GrpE and the ATPase domain of DnaK. GrpE is shown in green and the ATPase domain is shown in gray. **B.** Ribbon representation of the three-dimensional model of the Type 1 tobacco GrpE homologue generated using homology modeling. The ATPase domain is shown in gray, positioned in the same relative orientation as in *E. coli* GrpE-DnaK complex. The two insertions that occur in the tobacco GrpE homologue relative to the *E. coli* protein are shown in magenta. **C.** The molecular surface of the *E. coli* GrpE with the position of the residues that interact with the ATPase domain are highlighted in different colors. The color coding refers to the amino acid residue type: blue, basic residues; red, acidic residues; green, polar residues; and yellow, non-polar residues. **D.** The molecular surface for the tobacco GrpE homologue and the corresponding residues that interact with the *E. coli* complex are highlighted. The same color coding is maintained as in **C** and the magenta color indicates the regions where the insertions in the tobacco protein occur relative to the *E. coli* protein.

replaced by an Asn residue. In addition, the strong polar and nonpolar interactions involving a short stretch of residues, namely Pro-151, Asn-152, Gln-155, Ala-156, Ile-157, and Ala-158 of *E. coli* GrpE, and the residues Leu-49, Arg-6, Leu-257, and Glu-264 of DnaK are very well maintained in the tobacco GrpE-DnaK complex (Figure 5). The estimated buried accessible surface area, calculated using probe radius of 1.4 Å, between the modeled tobacco GrpE and the ATPase domain of DnaK is 2344 Å². This is slightly larger than the value 1970 Å² calculated for the *E. coli* complex. Relatively more surface area buried in the tobacco GrpE-DnaK complex could be due to changes in the primary sequence as well as the relaxed nature of the energy minimized modeled structure versus compact nature of the crystal structure of the *E. coli* complex. The conserved set of interactions between the tobacco GrpE and DnaK support the experimental results (Figure 2) demonstrating binding of the tobacco GrpE to DnaK.

Discussion

Our results provide the molecular, biochemical, and structural information of a mitochondrial GrpE homologue from plants. We isolated two types of cDNA clones from tobacco based on peptide sequences of a protein purified from tobacco BY2 cells. Both cDNAs encoded proteins of 34 kDa and shared 80% amino acid sequence identity. We conclude that these proteins are mitochondrial homologues of *E. coli* GrpE based on several criteria. First, sequence alignment of deduced protein sequences of the two cloned cDNAs with *E. coli* GrpE showed conservation of amino acid residues throughout most of the protein. Second, the tobacco proteins and a similar protein from *Arabidopsis* were shown, by expressing in *E. coli*, to associate with DnaK and the association was inhibited by ATP. Third, the tobacco proteins contained typical mitochondrial targeting sequence, and the targeting sequence of Type 1 protein localized GFP-fusion protein to mitochondria. Fourth, the tobacco protein interacted

with mHsp70 and ATP prevented this interaction. Fifth, it was possible to model a three-dimensional structure of the tobacco protein based on the X-ray structure of *E. coli* GrpE complexed with DnaK. We designate the characterized proteins as tobacco mitochondrial GrpE (NtmGrpE) with Type 1 as NtmGrpE1 and Type 2 as NtmGrpE2.

Type 1 tobacco GrpE appears to be the most predominant form: purification from BY2 cells yielded only Type 1 protein and a majority of the cDNA clones obtained were of Type 1. We cloned Type 2 because Type 1 and Type 2 have an identical nucleotide sequence at the sites at which the forward primer for PCR amplification annealed. Note that the Type 1 and Type 2 differed in the deduced amino acid sequence. The clones also are different within the 3'-untranslated sequences (data not shown). The 80% amino acid sequence identity between Type 1 and Type 2, and the 45% identity between tobacco and *Arabidopsis* proteins suggest that plant GrpE homologues have diverged in sequence while maintaining function.

The interactions that were demonstrated to occur between NtmGrpE and DnaK and mHsp70 suggest that the essential features of the bacterial DnaK/DnaJ/GrpE chaperone system are maintained in plant mitochondria. We also suggest that structural features of the bacterial chaperone are conserved in plant mitochondria. Although the sequence identity between NtmGrpE and *E. coli* GrpE is low (34%), the predicted structure of NtmGrpE1 is similar to the structure of *E. coli* GrpE. Amino acid residues of *E. coli* GrpE that contact the DnaK are conserved in the NtmGrpE. Furthermore, the extent of the buried surface areas between the modeled NtmGrpE-DnaK complex is similar to that calculated for the *E. coli* GrpE-DnaK complex. Plant mHsp70 proteins share 60% sequence identity with DnaK and are more similar to DnaK than to nuclear/cytoplasmic Hsp70 from eukaryotes [10, 20, 22, 29, 41, 48]. Unlike nuclear/cytoplasmic Hsp70, mHsp70 from yeast and mammals requires GrpE and DnaJ for its function [6, 15, 43, 49]. Mitochondrial Hsp70 is a key component in import and folding of mitochondrial proteins in yeast and mammals [3, 23, 35, 43], and it is likely that tobacco mHsp70 and mGrpE have similar functions in import and folding of mitochondrial proteins.

Chaperones in plant mitochondria are not well characterized compared to those in other plant cell compartments [7]. An understanding of how plant mitochondrial Hsp70/DnaJ/GrpE chaperones recognize their substrates and the nature of substrate binding

and import into mitochondria is needed. Elucidating the mechanism of the mitochondrial chaperone system will require biochemical assays with purified Hsp70, DnaJ, and GrpE proteins. Availability of well characterized mitochondrial GrpE from tobacco and DnaJ from *Arabidopsis* [21] will facilitate increased understanding of how mHsp70 functions during transport and folding of proteins in plant mitochondria.

Acknowledgements

We thank Drs. B. Bukau and D. Neumann for kindly providing antisera and Drs. L. Brill and C. Reichel for critical reading of the manuscript. This work was supported by financial assistance from United States Agency for International Development (DAN-4197-A-00-1126-00) and ORSTOM, Paris.

References

1. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* 25: 3389–3402 (1997).
2. Azem A, Oppliger W, Lustig A: The mitochondrial Hsp70 chaperone system. Effect of adenine nucleotides, peptide substrate, and mGrpE on the oligomeric state of mHsp70. *J Biol Chem* 272: 20901–20906 (1997).
3. Beissinger M, Buchner J: How chaperones fold proteins. *Biol Chem* 379: 245–259 (1998).
4. Bessoule J: Occurrence and sequence of a DnaJ protein in plant (*Allium porrum*) epidermal cells. *FEBS Lett* 323: 51–54 (1993).
5. Bessoule JJ, Testet E, Cassagne C: Cloning of a new isoform of a DnaJ protein from *Allium porrum* epidermal cells. *Plant Physiol Biochem* 32: 723–727 (1994).
6. Bolliger L, Deloche O, Glick BS: A mitochondrial homolog of bacterial GrpE interacts with mitochondrial Hsp70 and is essential for viability. *EMBO J* 13: 1998–2006 (1994).
7. Boston RS, Viitanen PV, Vierling E: Molecular chaperones and protein folding in plants. *Plant Mol Biol* 32: 191–222 (1996).
8. Brooks B, Broccoleri B, Olafson D, States D, Swaminathan S, Karplus M: CHARMM: a program for macromolecular energy minimization and dynamics calculation. *J Comp Chem* 4: 187–217 (1983).
9. Bunting JR, Phan TV, Kamali E, Dowben RM: Fluorescent cationic probes of mitochondria; metrics and mechanism of interaction. *Biophys J* 56: 979–993 (1989).
10. Craig EA, Gambill BD, Nelson RJ: Heat shock proteins: molecular chaperones of protein biogenesis. *Microbiol Rev* 57: 402–414 (1993).
11. Deloche O, Kelley WL, Georgopoulos C: Structure-function analyses of the Ssc1p, Mdj1p, and Mge1p *Saccharomyces cerevisiae* mitochondrial proteins in *Escherichia coli*. *J Bacteriol* 179: 6066–6075 (1997).

12. Gething MJ, Sambrook J: Protein folding in the cell. *Nature* 355: 33–45 (1992).
13. Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ: The p21 Cdk-interacting protein is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75: 805–816 (1993).
14. Harrison CJ, Hayer-Hartl M, Liberto MD: Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* 276: 431–435 (1997).
15. Hohfeld J, Minami Y, Hartl FU: Hip, a novel cochaperone involved in the eukaryotic Hsc70/Hsp40 reaction cycle. *Cell* 83: 589–598 (1995).
16. Hopp TP: Protein surface analysis: methods for identifying antigenic determinants and other interaction sites. *J Immunol Meth* 88: 1–18 (1986).
17. Jones TA, Cowan S, Zou JY, Kjeldgaard M: Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 57: 110–119 (1991).
18. Judd RC: Peptide mapping. *Meth Enzymol* 182: 613–626 (1990).
19. Khan AR, Johnson KA, Braam J, James MNG: Comparative modeling of the three dimensional structure of the calmodulin related TCH2 protein from *Arabidopsis*. *Proteins* 27: 144–153 (1997).
20. Ko K, Bornemisza O, Kourtz L, Ko ZW, Plaxton WC, Cashmore AR: Isolation and characterization of a cDNA clone encoding a cognate 70-kDa heat shock protein of the chloroplast envelope. *J Biol Chem* 267: 2986–2993 (1992).
21. Kroczyńska B, Zhou R, Wood C, Miernyk JA: AtJ1, a mitochondrial homologue of the *Escherichia coli* DnaJ protein. *Plant Mol Biol* 31: 619–629 (1996).
22. Marshall JS, Keegstra K: Isolation and characterization of a cDNA clone encoding the major Hsp70 of the pea chloroplastic stroma. *Plant Physiol* 100: 1048–1054 (1992).
23. Martinus RD, Ryan RT, Naylor DJ, Herd SM, Hoogenraad NJ, Hoj PB: Role of chaperones in the biogenesis and maintenance of the mitochondrion. *FASEB J* 9: 371–378 (1995).
24. Matsudaira P: Limited N-terminal sequence analysis. *Meth Enzymol* 190: 602–613 (1990).
25. Mayer MP, Bukau B: Hsp70 chaperone systems: diversity of cellular functions and mechanism of action. *Biol Chem* 379: 261–268 (1998).
26. Morimoto R, Tissieres A, Georgopolous C: *The Biology of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor Press, Cold Spring Harbor, NY (1994).
27. Nagata T, Nemoto Y, Hasezawa S: Tobacco BY-2 cell line as the 'HeLa' cell in the cell biology of higher plants. *Int Rev Cytol* 132: 1–31 (1992).
28. Naylor DJ, Hoogenraad NJ, Hoj PB: Isolation and characterisation of a cDNA encoding rat mitochondrial GrpE, a stress-inducible nucleotide-exchange factor of ubiquitous appearance in mammalian organs. *FEBS Lett* 396: 181–188 (1996).
29. Neumann D, Emmermann M, Thierfelder J-M, zur Nieden U, Clericus M, Braun HP, Nover L, Schmitz UK: HSP68 – a DnaK-like heat-stress protein of plant mitochondria. *Planta* 190: 32–43 (1993).
30. Nover L: *Heat Shock Proteins*. CRC Press, Boca Raton, FL (1991).
31. Padidam M, Beach RN, Fauquet CM: The role of AV2 ('pre-coat') and coat protein in viral replication and movement in tomato leaf curl geminivirus. *Virology* 224: 390–404 (1996).
32. Palleros DR, Reid KL, Shi L, Welch WJ, Fink AL: ATP-induced protein-Hsp70 complex dissociation requires K⁺ but not ATP hydrolysis. *Nature* 365: 664–666 (1993).
33. Reichel C, Mathur J, Eckes P, Langenkemper K, Koncz C, Schell J, Reiss B, Maas C: Enhanced green fluorescence by the expression of an *Aequorea victoria* green fluorescent protein mutant in mono- and dicotyledonous plant cells. *Proc Natl Acad Sci USA* 93: 5888–5893 (1996).
34. Rost B, Sander C: Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 232: 584–599 (1993).
35. Rudiger S, Buchberger A, Bukau B: Interaction of Hsp70 chaperones with substrates. *Nature Struct Biol* 4: 342–349 (1997).
36. Sali A: Modeling mutations and homologous proteins. *Curr Opin Biotechnol* 5: 437–451 (1995).
37. Salstrom J, Pratt D: Role of coliphage M13 gene 5 in single-stranded DNA production. *J Mol Biol* 61: 489–501 (1971).
38. Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).
39. Sander C, Schneider R: Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* 9: 56–58 (1991).
40. Schlicher T, Soll J: Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. *FEBS Lett* 379: 302–304 (1996).
41. Schlicher T, Soll J: Chloroplastic isoforms of DnaJ and GrpE in pea. *Plant Mol Biol* 33: 181–185 (1997).
42. Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA: An improved method for walking in uncloned genomic DNA. *Nucl Acids Res* 23: 1087–1088 (1995).
43. Stuart RA, Cyr DM, Craig EA, Neupert W: Mitochondrial molecular chaperones: their role in protein translocation. *Trends Biochem Sci* 19: 87–92 (1994).
44. Szabo A, Langer T, Schroder H, Flanagan J, Bukau B, Hartl FU: The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70-DnaK, DnaJ, and GrpE. *Proc Natl Acad Sci USA* 91: 10345–10349 (1994).
45. von Heijne G, Steppuhn J, Herrmann RG: Domain structure of mitochondrial and chloroplast peptides. *Eur J Biochem* 180: 535–545 (1989).
46. Vidal V, Ranty B, Dillenschneider M, Charpentau M, Ranjeva R: Molecular characterization of a 70 kDa heat-shock protein of bean mitochondria. *Plant J* 3: 143–150 (1993).
47. Vierling E: The roles of heat shock proteins in plants. *Annu Rev Plant Physiol Plant Mol Biol* 42: 579–620 (1991).
48. Watts FZ, Walters AJ, Moore AL: Characterization of PHSP1, a cDNA encoding mitochondrial HSP70 from *Pisum sativum*. *Plant Mol Biol* 18: 23–32 (1992).
49. Ziegelhoffer T, Lopez-Bueasa P, Craig EA: The dissociation of ATP from Hsp70 of *Saccharomyces cerevisiae* is stimulated by both Ydj1p and peptide substrates. *J Biol Chem* 270: 10412–10419 (1995).
50. Zylic M, Ang D, Georgopoulos C: The GrpE protein of *Escherichia coli*. *J Biol Chem* 262: 17437–17442 (1987).