The Roles of HSP100/Clp Proteins in Living Organisms

Mitsuru AKITA

Faculty of Agriculture, Ehime University, Matsuyama 790-8566, Japan

(Received June 15, 2001)

HSP100/Clp family of proteins, unnoticed until quite recently, involves in various cellular activities in the various cellular locations in the living organisms. Their family proteins have been classified into subfamilies based upon their sequence similarities. Their functions are varied, e.g. to obtain the tolerance to the higher temperature, to enhance the proteolytic activity by the protease, to promote transport, and so on. It seems that even the proteins in the same subfamily in the different organisms have the different features. However, HSP100/Clp proteins have one feature in common, playing a role as a molecular chaperone utilizing ATP to accomplish these activities.

Keywords: heat shock protein, HSP100/Clp protein, molecular chaperone

INTRODUCTION

To maintain their lives, “quality control” has to be managed in the living organisms. For instance, inside the cells, the unnecessary materials must be turned over. Under various stress conditions, notably under the heat stress, a set of proteins is induced. These proteins are usually called “heat shock proteins” (hsp proteins). These proteins are playing important roles in preventing other proteins to fold or helping to degrade misfolded proteins resulted under the stress by interacting with these proteins. The proteins possessing these functions are also called “molecular chaperones.” Since the original meaning of “chaperone” is the lady who accompanies unmarried young lady, the protein accompanying other proteins is regarded as a “chaperone” for them. Even though not all the molecular chaperones are heat inducible, some of these heat non-inducible molecular chaperones are also given the term “hsp,” by custom. HSP100/Clp proteins are among these proteins. HSP100/Clp proteins have molecular mass of 80-100 k and some of them, such as ClpB and Hsp104, are heat inducible (Gottesman and Maurizi, 1992). Therefore, they are called HSP100 proteins. On the other hand, the term “Clp” is derived from caseinolytic protease, since casein was used as a substrate to measure the proteolytic activity of Clp protease. This terminology leads to confusion, since HSP100/Clp proteins by themselves do not have proteolytic activities. The first ever found Clp protein is ClpA/P protease, as known as protease Ti, in Escherichia coli and this Clp protease is composed with two subunits ; one is the proteolytic subunit (ClpP) and the other is the regulatory subunit (ClpA) (Hwang et al., 1988; Katayama et al., 1988). ClpA is the member of HSP100/Clp proteins.

Since ClpA had been reported for the first time, many ClpA homologues have been discovered in various living organisms. Currently, these proteins are classified into nine subfamilies, based upon their sequence similarities (Schirmer et al., 1996; Derré et al., 1999).
The number of subfamily will likely increase and the classification of HSP100/Clp proteins will be slightly modified as the various genome projects progress. The progresses will also lead to the discoveries of the new function of these proteins.

This review article summarizes the diversity of HSP100/Clp proteins and discusses the current understanding of how HSP100/Clp proteins function in the living organisms.

**CLASSIFICATION OF HSP100/Clp PROTEINS**

HSP100/Clp proteins are classified into two subfamilies, class 1 and class 2 (Schirmer et al., 1996). As shown in Fig. 1, class 1 HSP100/Clp proteins have two ATP binding domains, while class 2 HSP100/Clp proteins are amino-terminal (N-terminal) truncated and have single ATP binding domain. The molecular masses of class 1 HSP100/Clp proteins are 80–100 k and further classified into A, B, C, D and E, based upon their sequence similarities. Not only the similarities, these subfamilies are relatively easy to be distinguished by the length of the “spacer” region between two ATP binding domains (indicated as ATP-1 and ATP-2 in Fig. 1). Two ATP binding domains are well conserved among all the subfamilies, since the site for the ATP binding is strictly determined based upon the amino acid sequences, such as Walker A and B motifs (Table 1) (Walker et al., 1982). In the first ATP binding domain, one A motif and two B motifs (ABB in the order) are present. Whereas, the second ATP binding domain contains one A motif and one B motif (AB in the order). Since the clpB gene in *Escherichia coli* and cyanobacteria has the internal initiation site and the shorter B-type HSP100, Hsp78, is found in the mitochondrial matrix, two forms of the B-type HSP100 are shown in Fig. 1 (Squires et al., 1991; Leonhardt et al., 1993; Eriksson and Clarke, 1996).

The class 2 HSP100/Clp proteins are N-terminal truncated and have only second ATP binding domain (Fig. 1 and Table 1). This ATP binding domain is well conserved among not only the class 2 proteins, but also the class 1 proteins. The class 2 HSP100/Clp proteins are further classified into M, N, X and Y subfamilies. The Y-type HSP100 has an insertion of about 140 amino acids in the middle of the sole ATP binding domain (Missiakas et al., 1996).

---

*Fig. 1 Structure of the HSP100/Clp family of proteins.*

ATP-1 and ATP-2 indicate the two ATP-binding domains with Walker A (A) and B (B) motifs. The thicker boxes with the Roman numerals (I–V) are the greater conserved regions. Other thicker boxes at the N-terminal of the C- and the D-type HSP100 proteins indicate the targeting sequences. The folded lines found in the A-, the C-, the D-, and the E-type HSP100 indicate the gap in the “spacer” region (modified from Schirmer et al. (1996)).
Table 1 The consensus sequences of Walker motifs and those sequences found in HSP100/Clp proteins.

<table>
<thead>
<tr>
<th>Class</th>
<th>A motif</th>
<th>B motif (1)</th>
<th>B motif (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, B</td>
<td>GxGKT</td>
<td>RxGhD</td>
<td>RxGIDhD</td>
</tr>
<tr>
<td>C, D</td>
<td>GxGKT</td>
<td>RxGhD</td>
<td>RxGIDhD</td>
</tr>
<tr>
<td>E</td>
<td>GxGKT</td>
<td>KxGhD</td>
<td>RxGIDhD</td>
</tr>
</tbody>
</table>

A-TYPE HSP100

The A-type HSP100 is found only in Gram-negative bacteria until now. As described above, ClpA protein in E. coli was first found as a regulatory subunit of Clp protease, i.e. to unfold the substrate in an ATP-dependent manner to help ClpP, a proteolytic subunit, to digest the substrate (Katayama et al., 1988). ClpA interacts with ClpP (Katayama et al., 1988). ClpA also interacts with RepA protein of P1 plasmid working as a molecular switch. ClpA converts inactive RepA dimer into active monomer in vitro (Wickner et al., 1994). ClpA exhibits ATPase activity and forms a hexamer in an ATP-dependent manner (Maurizi, 1992). The hexamer formation, as well as the substrate-binding, requires the binding of ATP, but not the hydrolysis of ATP, thus the oligomerization can be observed in the presence of slowly hydrolyzable ATP analogue of ATP-γS (adenosine-5’-O-(3-thiotriphosphate)) (Kessel et al., 1995). One of the commonly used strategies to determine the function of ClpA or the B-type HSP100 is to analyze the mutant forms of HSP100 proteins, which have mutations in the ATP binding domains introduced by site-directed mutagenesis. For instance, lysine residues at the position of 220 in the Walker A motif of the first ATP binding domain and/or 501 in the Walker A motif of the second one were altered to threonine residues (Seol et al., 1995). The first ATP binding domain was important for oligomerization, while the second one was important for ATPase activity (Seol et al., 1995). When the first ATP binding domain was mutated, RepA activation was completely abolished. But the same mutation did not affect the proteolytic activity of ClpA/P complex (Pak et al., 1999). By the electron-microscopic analysis, the hexameric ClpA was observed as a ring-like structure (Kessel et al., 1995). The ring-like structure is also formed in the other HSP100/Clp proteins and molecular chaperones, such as tetradecameric Cpn60 (Chaperonin60, as known as GroEL) (Kim et al., 1994; Parsell et al., 1994; Kim et al., 2000). Combining all of these results, the working model of ClpA is proposed as shown in Fig. 2.

B-TYPE HSP100

The B-type HSP100 proteins are most widely found, both in prokaryote and in eukaryote. Most of B-type HSP100 proteins are found in the cytoplasm, while Hsp78 in Saccharomyces
Hexameric ClpA and tetradecameric ClpP are assembled to form ClpA/P complex (1). Substrates interact with ClpA alone or ClpA/P complex (2). The first and the second steps require the binding of ATP to ClpA, not the hydrolysis. Therefore, the slowly hydrolyzable ATP homologue of ATP-γS has the same effect as ATP at these stages. ATP-γS, instead of ATP, is shown in this figure to explain the further steps in more detail. ATP is substituted for the ClpA-bound ATP-γS, which promotes substrate unfolding (3). By the hydrolysis of ATP, unfolded substrates are released for remodeling (4a) or translocating through ClpA pore to ClpP part where substrates are degraded (4b) (modified from Wickner and Maurizi (1999) based on the results present in Pak et al. (1999)).

cerevisiae is found in the mitochondrial matrix (Leonhardt et al., 1993). As described above, in the case of E. coli and cyanobacteria, the B-type HSP100 gene possesses the internal initiation site, thus the N-terminal truncated ClpB coexists with the full length ClpB in the cytoplasm (Squires et al., 1991; Eriksson and Clarke, 1996). The significance of the coexistence of both ClpB is not clear. Hsp78 is regarded as the homologue of the shorter form of ClpB (Leonhardt et al., 1993). The cytoplasmic B-type HSP100 proteins are heat inducible and essential for cells to survive under higher temperature in E. coli (Squires et al., 1991), yeast (Sanchez and Lindquist, 1990), cyanobacteria (Eriksson and Clarke, 1996), and the higher plant (Hong and Vierling, 2000; Queitsch et al., 2000). Unlike ClpA, the B-type HSP100 proteins do not interact with ClpP (Parsell and Lindquist, 1993). Instead, the B-type HSP100/Clp proteins interact with other molecular chaperones to disaggregate other proteins. For instance, HSP104 interacts with Hsp70 and Hsp40 (Glover and Lindquist, 1998), whereas, ClpB in bacteria interacts with DnaK (Hsp70 homologue), DnaJ (Hsp40 homologue), and GrpE (Motohashi et al., 1999).

Though the oligomerization of ClpB in E. coli occurs in the presence of ATP at low protein concentration as ClpA, it can be occurred at high protein concentration even in the absence of ATP (Zolkiewski et al., 1999). It has been thought that all the B-type HSP100 proteins form a hexamer as a result of oligomerization, as ClpA does. However, the number of subunits in ClpB in E. coli has become in question. Kim et al. (2000) proposed that ClpB existed as a heptamer. Regardless of the number of subunits, the B-type HSP100 proteins form a ring-like structure (Parsell et al., 1994; Kim et al., 2000).

The role of two ATP-binding domains of B-type HSP100 are also examined with ClpB in E. coli and Hsp104 by analyzing the mutant forms of these proteins, introducing mutations in their two ATP binding domains. Interestingly, the results are not the same as the results of ClpA. Furthermore, even the same B-type HSP100 from the different origins does not share the same results. In case of Hsp104, the first ATP-binding domain is essential for...
ATPase activity, while the second ATP-binding domain is required for oligomerization (Schirmer et al., 1998). But in ClpB in *E. coli*, ATPase activity was retained when one of the two ATP-binding domains was mutated and lost when both of them were mutated (Kim et al., 1998).

The shorter form of the B-type HSP100, Hsp78 could interact with protein translocation intermediate in the mitochondrial matrix under the condition that matrix Hsp70 (mt-Hsp70) is limited (Schmitt et al., 1995). It is also shown that Hsp78 could increase the thermotolerance of the mutant *S. cerevisiae* that lacked the *hsp104* gene and expressed the *hsp78* gene in the cytosol. N-terminal truncated ClpB in cyanobacteria also showed the thermotolerance (Clarke and Eriksson, 2000). Even though the basal ATPase activity was not affected by the truncation, the ATPase activity of the truncated form of ClpB in *E. coli* was not stimulated by other proteins (Park et al., 1993). On the other hand, there is a study of the carboxy-terminal (C-terminal) end deleted mutant form of ClpB in *E. coli* (Barnett et al., 2000). For the C-terminal deleted ClpB, self-association was inhibited. From these studies, N-terminal portion of ClpB is important for substrate binding, while C-terminal portion is important for oligomerization.

### C-TYPE HSP100

The C-type HSP100 is found in Gram-positive bacteria, in cyanobacteria, and in chloroplasts of algae and higher plants (Schirmer et al., 1996). Even though they are included in the same protein family, ClpC from the various organisms have important differences. ClpC in *Synechococcus* sp. PCC 7942 was shown to be essential for survival (Clarke and Eriksson, 1996). ClpC in Gram-positive bacteria is a heat shock protein (Krüger et al., 2001). On the other hand, chloroplastic ClpC proteins are constitutively expressed and may even be slightly decreased during heat stress (Clarke and Eriksson, 1996). The chloroplastic ClpC protein is encoded in the nuclear genome synthesized in the cytosol as a precursor form with the chloroplastic targeting sequence at the N-terminal end and is imported into chloroplasts (Moore and Keegstra, 1993). As shown in Fig. 3 and reviewed by Keegstra and Froehlich (1999), chloroplast targeting proteins, including ClpC, are imported into chloroplasts through the proteinaceous machinery embedded in the double membranes of the outer and the inner envelope in an ATP-dependent manner. Once precursors reach to the stromal space, the targeting sequences are cleaved by the specific enzyme. Based on the sequence comparison, ClpC is more closely related to ClpA than to ClpB. Because of the fact that ClpC in Gram-positive bacteria interacts with ClpP (Turgay et al., 1998) and ClpA is not found in chloroplasts, it is postulated that ClpC substitutes for ClpA in chloroplasts. Immunological studies indicated the possible interaction of chloroplastic ClpC and ClpP in barley (Desimore et al., 1997). ClpC in pea (Hsp93) stimulated the proteolytic activity of *E. coli* ClpP in vitro, but did not stimulate the recombinant pea ClpP overexpressed in *E. coli* (Shanklin et al., 1995).

Except for the ClpC in Gram-positive bacteria, the exact role of ClpC is still unclear. In cyanobacteria, by deleting *clpB* gene, the expression of ClpC increased twice when the mutant was exposed to heat (Eriksson and Clarke, 1996). It is not known yet whether ClpA can substitute for ClpB in certain circumstances in *E. coli*. From this result, ClpC was to compensate for the loss of ClpB, therefore, the role of ClpC in cyanobacteria maybe similar to ClpB which does not interact with ClpP. Akita et al. (1997) and Nielsen et al. (1997) have demonstrated that Hsp93 was present in the protein translocation machinery in chloroplastic envelope membranes and postulated that it has a role in protein translocation (Fig. 3). However, it is still unknown how Hsp93 functions during protein translocation.
Fig. 3 Schematic model of protein import into chloroplasts.

After chloroplastic protein is synthesized in the cytosol as a precursor form with targeting sequence at the N-terminal end, precursor protein interacts with the lipid of the outer membrane (a), followed by the receptor (b). Then precursors proteins form a translocation intermediate (c). During the processes of “b” and “c,” GTP and low concentration of ATP (about 100 μM) are required. In the presence of high concentration of ATP (greater than 1 mM), precursor protein starts to translocation through the protein translocation machinery embedded in both the outer and inner envelope membranes (d). During translocation, ATP is postulated to be utilized by Hsp93. In the stromal space, targeting sequence is cleaved and the newly imported protein is headed to the further destination. The numbers indicate the approximate molecular weight (kDa) of the postulated components of the protein translocation machinery (modified from Keegstra and Froehler (1999)).

mitochondrial protein import, mt-Hsp70 is thought to interact with the protein being imported to provide the driving force for protein translocation (Glick, 1995). Since Hsp70 in the stromal space was not found in translocation complexes (Akita et al., 1997; Nielsen et al., 1997), it is possible that chloroplastic Hsp93 plays a similar role as mt-Hsp70. It is also possible that Hsp93 also forms a ring-shaped oligomer in the inner membrane and to promote newly translocating proteins into the stromal space, as is suggested in the case of ClpA (Fig. 2) and Hsp104 (Parsell et al., 1994; Kessel et al., 1995).

D- AND E-TYPE HSP100

The D-type HSP100, ClpD, is another chloroplastic HSP100 protein found in Arabidopsis thaliana and is also known as ERD1 (Nakashima et al., 1997). ClpD protein is encoded in the nuclear genome, like Hsp93, ClpD is synthesized in the cytoplasm with the targeting sequence at the N-terminal end and is imported into the stromal space of chloroplasts (Nakashima et al., 1997). The clpD gene is under the regulation of heat inducible element and chloroplastic localized ClpD is playing a major role during senescence(Nakashima et al., 1997).

The E-type of HSP100, ClpE, is the newest discovered subfamily. ClpE is found in Gram-positive bacteria and its expression is heat regulated, though the clpE gene does not appear to be required for stress tolerance (Derré et al., 1999). ClpE has a zinc-finger motif at its N-terminal end, which suggest that ClpE is playing an important role in DNA binding or, more likely, protein-protein interaction. The overall amino acid sequence of ClpE is the most homologous to the C-type of HSP100 analyzed by standard protein-protein BLAST search (NCBI; http://www.ncbi.nlm.nih.gov/).
CLASS 2 HSP100

As shown in Fig. 1, the class 2 HSP100 includes M-, N-, X-, and Y-type HSP100. The X-type HSP100, ClpX, in E. coli is heat inducible and interacts with ClpP (Wawrzynow et al., 1996). ClpX/P protease has different substrate specificity from the ClpA/P protease and affects different cellular activities (Wojtkowiak et al., 1993). ClpX involves in the ATP-dependent degradation of a bacteriophage λ replication protein. ClpX also has a zinc-finger motif at its N-terminal end as ClpE does (Derré et al., 1999).

The Y-type HSP100, ClpY, interacts with another proteolytic subunit, ClpQ (Kessel et al., 1995). The clpQ gene and the clpY gene in E. coli formed an operon (Missiakas et al., 1996). Mutations in the ATP-binding domain of ClpY impaired the ability of the protein to interact with ClpQ (Missiakas et al., 1996). ClpA/P, ClpX/P, and ClpY/Q proteases have substrate specificities (Porankiewicz et al., 1999).

As for other class 2 HSP100 proteins, the M-type and the N-type HSP100 proteins, not many observations have become public yet, except that the genetical information of both proteins can be found somewhere (M-type HSP100: M76451, X77160 in Genbank; N-type HSP100: THC88419, THC79885 in TIGR database).

CONCLUSION

HSP100/Clp family of proteins is playing important roles in the various cellular processes beyond the difference of species, not only under the stress conditions, but also under normal condition. Every HSP100/Clp proteins work as molecular chaperones in different ways. One of the important role is their involvement in “quality control,” that is essential to eliminate nonfunctional proteins or to prevent proteins to become nonfunctional. In prokaryote, Clp proteases, such as ClpA/P, ClpX/P, ClpY/Q, are a complex of the regulatory subunit (ClpA, ClpX, ClpY) and the proteolytic subunit (ClpP, ClpQ) and degrade the specific substrates (Schirmer et al., 1996). The topologies of these complexes are similar to those of 26 S proteasomes in archaebacteria and eukaryote (Kessel et al., 1995; Grimaud et al., 1998). This indicate that both prokaryote and eukaryote share the common mechanism of proteolysis. The fact that each protease has a substrate specificity seems to be important to control the quality of proteins. The recent study by Hoskins et al. (2000) showed that the N-terminal 15 amino acids of RepA contained the ClpA recognition signal. It is expected that more signals will be found in other substrates in the future.

The B-type HSP100 proteins, such as ClpB in prokaryote (Motohashi et al., 1999), Hsp104 in yeast (Glover and Lindquist, 1998), and Hsp101 in A. thaliana (Hong and Vierling, 2000; Queitsch et al., 2000) contribute to the quality control in the different manner as ClpA, ClpX, and ClpY do. The B-type HSP100 proteins are heat inducible and increase the thermotolerance in these organisms, but they do not work with proteases (Woo et al., 1992). The B-type HSP100 proteins prevent other proteins to aggregate by cooperating with other molecular chaperones, such as Hsp70 and Hsp 40 (Glover and Lindquist, 1998; Motohashi et al., 1999).

Some types of HSP100/Clp proteins involve in regulating gene expression in the indirect manner. ClpA and ClpC in Bacillus subtilis works as a molecular switch to control the quality. For instance, ClpC even regulate its own expression by degrading the transcription element, CtsR, in cooperation with ClpP under stress condition (Krüger et al., 2001).

The other important role of HSP100/Clp proteins is their involvement in protein translocation. It was shown that the shorter form of B-type HSP100, Hsp78, could supplement the role of mitochondrial matrix Hsp70 (Schmitt et al., 1995). In chloroplasts of higher plants,
the C-type HSP100 protein seems to involve in protein translocation into chloroplasts (Fig. 3). Hsp93, a C-type HSP100 in pea chloroplasts, was found associating with the protein import machinery (Akita et al., 1997; Nielsen et al., 1997). The exact function of Hsp93 during protein translocation has not yet determined. By comparison with other HSP100/Clp proteins, it is likely that importing proteins translocate through Hsp93 pore in an ATP-dependent manner like substrates translocate through ClpA pore as shown in Fig. 1.

Regarding the C-type HSP100 proteins, they are found only in the Gram-positive bacteria, in cyanobacteria, in chloroplasts of algae and higher plants (Schirmer et al., 1996). Even though the function of C-type HSP100 seems different between species, from the evoluntional point of view, the fact that C-type HSP100 proteins can be found in the limited species supports the symbiotic theory about the origin of chloroplasts.

Lastly, the other interesting feature of HSP100/Clp proteins that should be noted is about the interaction between Hsp104 and yeast prion-like protein, Sup35 (Schirmer and Lindquist, 1997). Hsp104 was also shown to interact with mammalian prion protein, PrP and β-amyloid peptide (Schirmer and Lindquist, 1997). It was not yet reported whether mammalian B-type HSP100 protein interacts with PrP or β-amyloid as Hsp104 does, the observation that Hsp104 interacts with PrP and β-amyloid peptide will lead to the better understanding of the diseases caused by them.

Most of the results described in this review have been obtained from prokaryote, from yeast and, to lesser extent, from higher plants. With the completion of the genome projects of Homo sapiens and A. thaliana, the investigations on HSP100/Clp proteins is expected to emerge in mammalian and higher plants.

REFERENCES


(和文抄録)
生体内における分子シャペロン HSP100/Clp 蛋白質の働き
秋 田 充
愛媛大学農学部

最近になってその存在が広く知られるようになった HSP100/Clp に属する蛋白質は、生体内において様々な場で様々な細胞内活動に関わっている。HSP100/Clp に属する蛋白質は、アミノ酸配列の相違性によってさらに細かく分類されるに至っている。それぞれの蛋白質の機能は複数にわたっており、例えば、高温耐性の獲得、蛋白質分解酵素の酵素活性の促進、物質輸送等に関与している。分類上同じ蛋白質同士であっても、その蛋白質の由来する生物種が異なれば、その特徴にも差異があたり、しかし、どの HSP100/Clp 蛋白質にも共通の特徴がある。それは、分子シャペロンとして機能していることである。分子シャペロンとして機能しながためく細胞内活動を遂行するためには、ATP を利用する必要がある。