Characterisation of the Clp proteins in
Arabidopsis thaliana

BO ZHENG

Department of Plant Physiology
Umeå Plant Science Center
Umeå University
Sweden 2003
Characterisation of the Clp proteins in

Arabidopsis thaliana

Bo Zheng (郑波)
Dissertation Abstract

Unlike in the greenhouse, plants need to cope with many environmental stresses under natural conditions. Among these conditions are drought, waterlogging, excessive or too little light, high or low temperatures, UV irradiation, high soil salinity, and nutrient deficiency. These stress factors can affect many biological processes, and severely retard the growth and development of higher plants, resulting in massive losses of crop yield and wood production. Plants have developed many protective mechanisms to survive and acclimate to stresses, such as the rapid induction of specific molecular chaperones and proteases at the molecular level. Molecular chaperones mediate the correct folding and assembly of polypeptides, as well as repair damaged protein structures caused by stress, while proteases remove otherwise non-functional and potentially cytotoxic proteins.

The Clp/Hsp100 family is a new group of chaperones that consists of both constitutive and stress-inducible members. Besides being important chaperones, many Clp/Hsp100 also participate in protein degradation by associating with the proteolytic subunit ClpP to form the Clp protease complex. Higher plants have the greatest number and complexity of Clp proteins than any other group of organisms, and more than 20 different Clp isomers in plants have been identified (Paper I). Because of this diversity, we have adopted a functional genomics approach to characterise all Clp proteins in the model plant Arabidopsis thaliana. Our ongoing research strategy combines genetic, biochemical and molecular approaches. Central to these has been the preparation of transgenic lines for each of the chloroplast Clp isomers. These transgenic lines will be analysed to understand the function and regulation of each chloroplast Clp protein for plant growth and development.

In Paper II, an Arabidopsis thaliana cDNA was isolated that encodes a homologue of bacterial ClpX. Specific polyclonal antibodies were made and used to localise the ClpX homologue to plant mitochondria, consistent with that predicted by computer analysis of the putative transit peptide. In addition to ClpX, a nuclear-encoded ClpP protein, termed ClpP2, was identified from the numerous ClpP isomers in Arabidopsis and was also located in mitochondria. Relatively unchanged levels of transcripts for both clpX and clpP2 genes were detected in various tissues and under different growth conditions. Using β-casein as a substrate, plant mitochondria possessed an ATP-stimulated, serine-type proteolytic activity that could be strongly inhibited by antibodies specific for ClpX or ClpP2, suggesting an active ClpXP protease.

In Paper III, four nuclear-encoded Clp isomers were identified in Arabidopsis thaliana: ClpC1 and ClpP3-5. All four proteins are localized within the stroma of chloroplasts, along with the previously identified ClpD, ClpP1 and ClpP6 proteins. Potential differential
regulation among these Clp proteins was analysed at both the mRNA and protein level. A comparison between different tissues showed increasing amounts of all plastid Clp proteins from roots to stems to leaves. The increases in protein were mirrored at the mRNA level for most ClpP isomers but not for ClpC1, ClpC2 and ClpD and ClpP5, which exhibited little change in transcript levels. Potential stress induction was also tested for all chloroplast Clp proteins by a series of brief and prolonged stress conditions. The results reveal that these proteins, rather than being rapidly induced stress proteins, are primarily constitutive proteins that may also be involved in plant acclimation to different physiological conditions.

In Paper IV, antisense repression transgenic lines of clpP4 were prepared and then later characterised. Within the various lines screened, up to 90% of ClpP4 protein content was specifically repressed, which also led to the down-regulation of ClpP3 and ClpP5 protein contents. The repression of clpP4 mRNA retarded the development of chloroplasts and the differentiation of leaf mesophyll cells, resulting in chlorotic phenotypes. The chlorosis was more severe in young than in mature leaves due likely to the developmental expression pattern of the ClpP4 protein. Chlorotic plants eventually turned green upon aging, accompanied by a recovery in the amount of the ClpP4 protein. The greening process could be affected by the light quantity, either by altering the photoperiod or light intensity.
List of Papers


Paper IV  Zheng B, MacDonald TM, Sutinen S, Hurry V, Clarke AK (2003) Antisense repression of clpP4 gene retards development of chloroplasts in Arabidopsis thaliana and can be relieved by leaf maturation or more light irradiation (Manuscript)
Acknowledgements

I wish to acknowledge the guidance and assistance given throughout my five-year PhD education by all colleagues from Department of Plant Physiology, Umeå University. I am very grateful especially to the individuals listed below.

Adrian K. Clarke, Supervisor
Vaughan Hurry, Co-supervisor
Göran Samuelsson, Head of Department
Per Gardeström, Reference Group
Stefan Jansson, Reference Group

The study was supported in part by grants from the following organizations and foundations.

Akademiens samordnade fonder, Kungl. Skogs- och Lantbruksakademiers (KSLA)
Carl Tryggers Foundation for Science Research
Knut och Alice Wallenbergs Foundation
Stiftelsen J C Kempes Minnes Stipendiefond (SJCKMS)
Swedish Agricultural and Forestry Resource Council
Swedish Foundation for International Cooperation in Research and Higher Education (STINT)
Table of Contents

Abstract 3
List of papers 5
Acknowledgements 6
Part I. Protein Quality Control by Molecular Chaperones and Proteases 8
1-1. Protein triage model for quality control
1-2. Molecular chaperones in protein folding
   - Folding newly synthesised polypeptides by Hsp70s and chaperonins
   - Chloroplast import complexes, a pathway of molecular chaperone
1-3. Protein degradation
   - Ubiquitin/26S proteasome pathway
   - Proteases in mitochondria
   - Proteases in chloroplasts
Part II. Clp/Hsp100 Chaperone and Clp Proteases 22
2-1. Classification of the Clp family
   - Clp/Hsp100, the regulatory subunits
   - ClpP, the proteolytic subunits
2-2. Structure and mechanism of Clp proteasome
Part III. Characterisation of Clp Proteins in Arabidopsis thaliana 29
3-1. Preparation of molecular tools
3-2. Localisation of newly found Clp proteins
3-3. Tissue-specific expression of Clp proteins
3-4. Expression of Clp proteins under different conditions
   - Brief severe stresses
   - Prolonged moderate stress: 1. high light stress response
   - Prolonged moderate stress: 2. cold acclimation
3-5. Expression of Clp proteins during leaf senescence
3-6. Screening for T-DNA insertion mutants and antisense transgenic lines
   - Characterisation of clpD antisense lines
   - Characterisation of clpP4 antisense lines
3-7. Perspectives and future work
References 43
Part I. Protein Quality Control by Molecular Chaperones and Proteases

1-1. Protein triage model for quality control

Proteins constitute a large percentage of living cells and play vital roles in cell growth and development. To carry out their biological functions, the polypeptide chain needs to be accurately synthesized and folded to form a precise three-dimensional structure. Living cells store genetic information for protein biosynthesis in the form of double-stranded DNA. Single stranded RNA is transcribed from the DNA template. Some RNA molecules undergo specific processing reactions to produce mature mRNAs for translation, to yield proteins with specific amino acid sequences.

Besides the fidelity of transcription and translation, functional proteins within cells require the participation of molecular chaperones – proteins that catalyse the folding of other proteins, to fulfil the initial folding and assembly of multi-protein complexes (reviewed in Wickner et al., 1999). Polypeptides emerging from the ribosome must fold into stable native three-dimensional structures to obtain biological functions. However, hydrophobic patches of amino acid aggregate during translation, trapping the protein in an incorrect conformation or allowing it to aggregate with other proteins. Aggregation can be prevented by temporarily binding stretches of amino acid to a peptide-binding groove on Hsp70 (Fig. 1-1), a molecular chaperone family that maintains polypeptides in an unfolded state (reviewed in Spremulli, 2000). Misfolded proteins caused by mutation in

![Figure 1-1. Molecular chaperone Hsp70 prevents nascent and newly synthesized polypeptide chains from improper folding or aggregation (Spremulli, 2000).](image-url)
the encoding gene or lack of fidelity in transcription or translation face a number of possible fates: rescue by molecular chaperone, destruction by energy-dependent cytoplasmic proteases, or aggregation (Wickner et al., 1999).

A protein triage model was proposed for partitioning of non-native proteins between chaperones for remodelling and proteases for degradation (Gottesman et al., 1997). Triage is defined as “sorting and allocation of treatment to patients”. The patients here are cellular proteins. The first level of triage is to identify whether the proteins are damaged and require treatment. The quality control system must be able to distinguish native proteins from everything else that might be considered as nonnative, due to errors in protein folding, modification, assembly or sorting. Once damaged proteins have been identified, a second level of triage follows: Can these proteins be saved? Chaperones or chaperone components of proteases attempt first to correct misfolded proteins to restore their activity. Hopeless cases in which structural damage cannot be repaired however need to be degraded by the cytoplasmic ATP-dependent proteases. Aggregation is the likely default outcome when both the chaperone and protease systems fail, possibly leading to cytotoxic effects.

The protein triage model is further illustrated in Figure 1-2. As hydrophobic regions of polypeptides are exposed, either as the newly made proteins emerge from the ribosome or because of subsequent misfolding or failure to assemble properly, they are subject to binding by chaperones or the ATP-dependent proteases. Chaperone binding and release of folding intermediates may allow proteins to reach their native conformation or may return them to the pool of nonnative proteins that can rebind chaperones or proteases. Protease binding followed by ATP-dependent unfolding and subsequent degradation removes the protein

**Figure 1-2.** Protein triage model for quality control (Wickner et al., 1999)
from the pool of nonnative proteins. Some misfolded or partially folded proteins will eventually aggregate. Although the chaperones act most generally to prevent aggregation, they are also able to dissolve aggregates. Because proteins in aggregates are relatively resistant to proteolysis, chaperones promote proteolysis indirectly by maintaining misfolded proteins in an unaggregated state. Steps in boxes represent major processes at which competition takes place for protein triage.

1-2. Molecular chaperones in protein folding

The term “molecular chaperone” was originally introduced to describe the function of nucleoplasmin, a nuclear protein that facilitates chromatin assembly by preventing improper interactions between histones and DNA (Laskey et al., 1978). The term was later generalized to include a range of functionally related, but diverse proteins that assist the folding and assembly of other proteins (Ellis, 1987). A current definition describes a molecular chaperone as “a protein which transiently binds to and stabilizes an unstable

![Diagram](https://via.placeholder.com/150)

**Figure 1-3.** Model for molecular chaperone Hsp70 and co-chaperone Hsp40 in folding newly synthesized polypeptides (Spremulli, 2000)
conformer of another protein, and through regulated binding and release, facilitates its correct fate in vivo: be it folding (following de novo synthesis, transit across a membrane, or stress-induced denaturation), oligomeric assembly, interaction with other cellular components, switching between active and inactive conformations, intracellular transport, or proteolytic degradation, either singly or with the help of co-factors” (Hendrick and Hartl, 1993). Chaperones are distributed ubiquitously in both prokaryotes and eukaryotes. More than 20 different families of chaperone appear to be conserved in evolution. Many of them are constitutively expressed and essential under normal growth conditions, while approximately a quarter of them are stress proteins (Hartl, 1996).

**Folding newly synthesised polypeptides by Hsp70s and chaperonins**

Newly synthesized polypeptides must fold and assemble into stable native three-dimensional structures to obtain biological functions. Incompletely folded polypeptides tend to aggregate during synthesis on polyribosomes. The tendency is strongly enhanced in vivo by the high concentration of macromolecules in the cellular solution and by the close proximity of nascent emerging polypeptides. In many cases, molecular chaperones are able to assist newly synthesized polypeptides to complete correct folding, preventing off-pathway reactions that lead to premature misfolding or aggregation (reviewed by Agashe and Hartl, 2000).

The major classes of chaperones acting in cytoplasmic protein folding are the Hsp70s and the chaperonins (Hsp60s). The Hsp70s are found in bacteria and in most compartments of eukaryotic cells. All Hsp70s consist of the same working parts: a highly conserved NH2-terminal ATPase domain of 44 kDa and a COOH-terminal region of 25 kDa, divided into a conserved substrate-binding domain of 15 kDa and a less-conserved immediate COOH-terminal domain of 10 kDa (reviewed by Bukau and Horwich, 1998). The Hsp70s bind to polypeptide chains emerging from the ribosome during protein synthesis (Fig. 1-1), shielding the hydrophobic regions with the assistance from co-chaperone Hsp40. Substrates undergo repeated cycles of binding to/release from Hsp70/Hsp40 (Szabo et al., 1994; Buchberger et al., 1996) and kinetic partitioning among folding to native state, aggregation, rebinding to Hsp70, and binding to other chaperones or proteases as part of a multidirectional folding network (Fig. 1-3).

Unlike Hsp70s, the chaperonins provide a sequestered environment, a cylinder structure, in which folding can proceed unimpaired by intermolecular interactions
between non-native polypeptides. Chaperonins are the most varied and structurally complex group of the molecular chaperone. They are divided into two families: the GroEL group (i.e., the Hsp60, chaperonin60, or cpn60 group), and the TRIC (TCP-1 ring) family. The GroEL protein of *Escherichia coli* was the first chaperone to be studied on a molecular level and is the most thoroughly investigated chaperone system, both functionally and structurally. Two host proteins, GroEL (57 kDa) and GroES (10 kDa), are required for the correct assembly. The GroEL molecule comprises 14 identical subunits forming two heptameric rings. A model of GroE-assisted protein folding in a three-step process (Fig. 1-4) is widely accepted: an aggregation-prone folding intermediate is first captured by GroEL and thereby becomes protected from aggregation.

**Figure 1-4.** Model for the roles of GroEL and GroES in protein folding (Spremulli, 2000)
Upon binding of ATP and GroES to the GroEL/polypeptide complex, the polypeptide is ejected into a closed compartment formed by the GroE chaperone, where folding is initiated. After hydrolysis of ATP, both GroES and the polypeptide are released. The transitions between the different functional states of the chaperone are triggered by a set of domain movements, which in turn are controlled by the binding of ATP and the cochaperone GroES (currently reviewed by Bukau et al., 2000; Feldman and Frydman, 2000; Hartl and Hayer-Hartl, 2002; Walter, 2002).

**Chloroplast import complexes, a pathway of molecular chaperone**

Chloroplasts are believed to originate from cyanobacteria as a result of endosymbiosis (Mcfadden, 1999). In *Arabidopsis*, there are 87 protein-coding genes in the 154-Kb plastome (Sato et al., 1999; Abdallah et al., 2000), compared to more than 25K protein-coding genes in the 125-Mb genome (The Arabidopsis Genome Initiative, 2000). Nuclear-encoded chloroplast proteins need to be synthesized in the cytosol as precursor proteins, with an N-terminal signal peptide for targeting into different compartments of chloroplasts. Chloroplast precursors are imported into the organelle through the TOC and TIC complexes (Schleiff and Soll, 2000; Keegstra and Cline, 1999). Precursors are phosphorylated in the cytosol and form a guidance complex with the Hsc70 and 14-3-3 proteins, both molecular chaperones, and then are transported to the outer envelope (OE) (May and Soll, 2000). There they are recognized by Toc64, transferred

---

**Figure 1-5.** Model for chloroplast import complexes (Zhang and Glaser, 2002. Reprinted with permission)
to Toc34 or Toc60, then dephosphorylated and fed into the Toc75 channel (Waegemann and Soll, 1996). The TIC complex that forms a channel through the inner envelope (IE) comprises Tic110, Tic 55, Tic 40, Tic 22 and Tic20 (Fig. 1-5).

Transport of cytoplasmically synthesized precursor proteins into chloroplasts appears to require the action of molecular chaperones. These molecules are likely to be the sites of the ATP hydrolysis required for precursor proteins to bind to and be translocated across the two membranes of the chloroplast envelope. Three Hsp70 homologues are involved in the import process from the cytosol to the stroma: the outer envelope Com70, located on the cytosolic side of the membrane; the inter-space IAP70 chaperone, located at the inner leaflet of the outer envelope membrane; and the stromal Hsp70 (sHsp70). Com70 might be involved in unfolding of precursors, whereas IAP70 might provide the driving force for translocation across the outer and inner envelopes (reviewed by Jackson-Constan and Keegstra, 2001; Zhang and Glaser, 2002). The function of sHsp70 might be similar to those Hsp70 homologues helping the folding of newly synthesized polypeptides emerging from ribosomes. ClpC, a molecular chaperone that belongs to the Clp/Hsp100 family, is associated with the chloroplast import machinery with unknown function (Keegstra and Cline, 1999; Nielsen et al., 1997).

1-3. Protein degradation

Protein degradation involves a protease (or a group of proteases) digesting proteins that are no longer needed or must be removed to start new biological processes. Protein degradation is required in many respects of a plant’s life cycle: growth, development, differentiation, reproduction, stress response, and disease elimination. Protein degradation removes proteins that are abnormal or damaged arising from a variety of mechanisms including mutations, biosynthetic errors, spontaneous denaturation and free radical-induced damage. Such damage can be accelerated by environmental stresses such as heat shock, desiccation, high-fluency light, disease, nutrient deprivation, and exposure to heavy metals. Protein degradation also removes proteins that are improperly processed or mistargeted, as well as proteins controlling metabolic pathways and various regulatory processes, including signal reception and transduction, homeosis, transcription, and cell growth and division (reviewed by Gatenby and Viitanen, 1994; Hochstrasser, 1995; Varshavsky, 1992, Vierstra, 1993).
General features of protein degradation in plants have been previously reviewed by Vierstra (1996). First, most *in vivo* proteolysis requires energy for peptide bond hydrolysis. Second, protein degradation is a rapid process, making the detection of partial breakdown products very difficult. The third feature is that most proteases are specific to certain amino acids sequences and/or cleavage sites within the protein. And fourthly, substrates degraded by certain proteases are highly selective. Moreover, the turnover of individual proteins can vary dramatically depending on the conformational state and location of the protein, or on the developmental and physiological state of the cell. Evidence has emerged that several distinct pathways exist in plants, with one or more in each cellular compartment.

**Ubiquitin/26S proteasome pathway**

Recent analyses of the near-complete *Arabidopsis thaliana* genome identified more than 1300 genes, about 5% of the proteome, involved in the ubiquitin (a 76 amino acid peptide) dependent 26S proteasome pathway, making it one of the most elaborate regulatory mechanisms in plants. The Ubiquitin/26S proteasome pathway rivals transcription complexes and protein kinase cascades as the main player in cell regulation of almost all aspects of plant biology, including the cell cycle, embryogenesis, circadian rhythms, hormone signalling, homeosis, disease resistance and senescence (reviewed by Vierstra, 2003). The 26S proteasome predominates protein turnover in the nuclear and cytoplasmic compartments. In the ER, misfolded proteins are retro-translocated to the cytosol and degraded by the proteasome. The 26S proteasome is also involved in degradation of peptides leaking from gerontoplast during chloroplast senescence (reviewed by Arnim, 2001; Ellgaard et al., 1999; Hörtensteiner and Feller, 2002).

The 26S proteasome is a 2-megadalton protein complex consisting of a proteolytically active cylindrical 20S core particle (CP) and up to two 19S regulatory particles (RP) (Fig. 1-6). The CP is an ATP- and Ub-independent peptidase with an assembly of four, stacked heptameric rings of related α and β subunits in a α₁β₁β₁α₁β₁β₁α₁β₁ configuration. The RP presumably helps identify appropriate substrates for breakdown, releases the attached ubiquitins, opens the subunit ring gate, and directs entry of unfolded protein into the CP lumen for degradation. Each 19S RP consists of a cylindrical “base” with ATPase activity, and a “lid” complex. Substrates to be degraded are recognized by the lid, unfolded and fed the 20S cylinder by the base (reviewed by Arnim, 2001; Vierstra, 2003).
Figure 1-6. Ubiquitin-dependent proteolytic pathway by 26S proteasome (Callis and Vierstra, 2000; Hershko and Ciechanover, 1998)


To prevent accidental capture of non-substrate proteins by the proteasome, target proteins are flagged by covalent attachment of poly-ubiquitin chains. The ubiquitin (Ub) conjugation cascade requires the sequential action of three enzyme families: E1, E2 and E3. The biochemical steps in the Ub pathway have been reviewed previously and are
illustrated in Figure 1-6 (Hershko and Ciechanover, 1998; Callis and Vierstra, 2000). E1 (or Ub-activating enzyme) activates the Ub by coupling ATP hydrolysis to the formation of an E1-Ub intermediate in which the C-terminal Gly of Ub is linked via a thiolester bond to a Cys residue of E1. Activated Ub is then transferred to an E2 (or Ub-conjugating enzyme) by transesterification. This intermediate delivers the Ub moiety to the substrate, usually using an E3 (or Ub-protein ligase) as the catalyst. The E3 recruits the substrate, positions it for optimal transfer of the Ub moiety, and then initiates conjugation. In the final product, the C-terminal Gly carboxyl group of Ub is attached via an iso-peptide bond to a free lysl-ε-amino group in the target. Through rounds of conjugation, a multiple Ub chain is assembled to the target substrate.

The importance of each of the three enzymatic steps in target selection is evident by the discovery of a variety of their corresponding genes: the *Arabidopsis* genome encodes two E1s, more than 45 E2 or E2-like proteins and around 1200 E3 components (Table 1-1). Five E3 types have been described so far based on subunit composition and mechanism of action: HECT, SCF, VBC-Cul2, Ring/U-box and APC.

### Table 1-1. Genomic organization of the *Arabidopsis* Ub/26S proteasome pathway

<table>
<thead>
<tr>
<th>Protein</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ub</td>
<td>16</td>
</tr>
<tr>
<td>E1</td>
<td>2</td>
</tr>
<tr>
<td>E2 and E2-like</td>
<td>~45</td>
</tr>
<tr>
<td>E3</td>
<td></td>
</tr>
<tr>
<td>HECT</td>
<td>7</td>
</tr>
<tr>
<td>SCF F-box</td>
<td>694</td>
</tr>
<tr>
<td>RBX-Cullin-ASK</td>
<td>33</td>
</tr>
<tr>
<td>Ring finger</td>
<td>~387</td>
</tr>
<tr>
<td>U-box</td>
<td>37</td>
</tr>
<tr>
<td>APC</td>
<td>&gt;20</td>
</tr>
<tr>
<td>DUBs</td>
<td>32</td>
</tr>
<tr>
<td>26S proteasome</td>
<td></td>
</tr>
<tr>
<td>20S CP</td>
<td>23</td>
</tr>
<tr>
<td>19S RP</td>
<td>31</td>
</tr>
<tr>
<td>Total &gt; 1327</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: APC, anaphase-promoting complex; CP, core protease; DUBs, deubiquitination enzymes; E1, Ub-activating enzyme; E2, Ub-conjugating enzyme; E3, Ub-protein ligase; HECT, homology to E6-AP C-terminus; RP, regulatory particle; SCF, SKP1, CDC53 and F-box protein complex; Ub, Ubiquitin (Vierstra, 2003).
Protein degradation in mitochondria

Mitochondria peptidases can be categorized into three groups: processing peptidases, oligopeptidases and ATP-dependent proteases (reviewed by Käser and Langer, 2000). Like plastidic proteins, most mitochondrial proteins are nuclear encoded and need proteolytic processing of N-terminal presequences to reach their final destination in mitochondria. Processing enzymes have been identified in various mitochondrial sub-compartments. The mitochondria processing peptidase (MPP) removes N-terminal mitochondrial targeting sequences of nuclear encoded precursor proteins in the mitochondrial matrix space. The mitochondria intermediate peptidase (MIP) cleaves off N-terminal octa-peptides from some matrix-localized proteins. Maturation of intermembrane space proteins with a bipartite presequence requires cleavage by both MPP in the matrix and the inner membrane protease (IMP) in the intermembrane space. The mitochondrial oligopeptidase MOP in the intermembrane space represents the only identified oligopeptidase in mitochondria so far.

Figure 1-7. The proteolytic system of mitochondria (Käser and Langer, 2000)
Unlike the limited proteolytic activity and substrate specificity of processing peptidases, ATP-dependent proteases mediate the complete degradation of dispensable mitochondrial proteins. Several ATP-dependent proteases have been identified in different sub-compartments of mitochondria, including Lon-like and Clp-like proteases in the matrix, and AAA proteases in the mitochondrial inner membrane. Lon-like proteases comprise a conserved protein family that is present in archaebacteria, eubacteria, and eukaryotic cells. Lon-like proteases have several distinguishable domains, including an ATPase domain indispensable for proteolysis, and a proteolytic domain containing the catalytically active serine residue in the C-terminus. The ATPase domain of Lon proteases exhibits chaperone-like activity, promoting substrate unfolding and ensuring the specificity of proteolysis. Clp-like proteases will be described in detail in Part II of the thesis. AAA proteases form a conserved class of ATP-dependent proteases that mediate the degradation of membrane proteins in eubacteria, mitochondria and chloroplasts (Langer 2000). They are named after their ATPase domain, which is characteristic of a large superfamily of Walker-type P-loop ATPase, the AAA family (for ATPases Associated with a number of cellular Activities). Two proteolytic complexes with a native molecular mass of more than one mega-Dalton are present in the inner membrane and expose their catalytic sites to opposite membrane surfaces: the m-AAA protease, active at the matrix side of the inner membrane; and the i-AAA protease, which is presumably a homooligomeric complex in the inner membrane whose catalytic sites are exposed to the intermembrane space. AAA proteases are thought to be involved in mitochondria biogenesis and turnover of membrane proteins (reviewed by Käsger and Langer, 2000; Langer et al., 2001).

Proteases in chloroplasts

Proteolysis is involved in a wide range of processes during the biogenesis and maintenance of chloroplasts. There are defined proteases within each of the major chloroplast compartments: the ATP-dependent Clp and FtsH proteases in the stroma and stroma-exposed thylakoid membranes, respectively, the ATP-independent DegP proteases within the thylakoid lumen and on both sides of thylakoid membranes, and the SppA protease on the stromal side of the thylakoid. All of these proteases have homologues in bacteria, but most have many paralogues in higher plants (reviewed Adam and Clarke, 2002).
The bacterial FtsH protease, one of the AAA proteins, is a membrane-bound ATP-dependent metalloprotease (Tomoyasu et al., 1995) involved in the degradation of unassembled proteins and different regulators of gene expression (Gottesman, 1996). FtsH proteases are anchored to their respective membranes by one or two trans-membrane hydrophobic domains located in their amino-terminus, followed by a conserved ATP-binding domain containing the Walker A and B motifs (Patel and Latterich, 1998). The proteolytic site, the zinc-binding domain HExxH (x represents any amino acid), is located near the carboxy-terminus of the protein (Adam, 2000). The chloroplast FtsH homologues (Lindahl et al., 1996; Adam et al., 2001) are localized to the stromal-exposed regions of the thylakoid membrane, with the ATP- and zinc-binding domains exposed to the stroma.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1-8}
\caption{A model for the degradation of the D1 protein of photosystem-II (PSII) reaction center (Adam and Clarke, 2002; Estelle, 2001)}
\end{figure}

\textit{E. coli} has a family of serine-type ATP-independent proteases known as Deg, the members of which are DegP (HtrA), DegQ (HhoA) and DegS (HhoB). DegP is the first and the best characterized to date, with PDZ-like domains at the carboxy-terminus (Pallen and Wren, 1997). PDZ domains participate in protein-protein interaction in various biological processes and apparently mediate substrate recognition and/or binding in the context of proteases (Levchenko et al., 1997). \textit{E. coli} DegP is a hexamer composed of two
staggered trimeric rings. The proteolytic sites are located in a central cavity and the two PDZ domains of each monomer form the sidewalls (Krojer et al., 2002). Like FtsH and Clp proteases, DegP incorporates both chaperone and proteolytic activities; the chaperone activity dominates at low temperatures and the proteolytic activity is present at elevated temperatures (Spiess et al., 1999). In Arabidopsis, DegP is encoded by multiple genes, of which at least four are putatively targeted to chloroplasts (Adam et al., 2001). The DegP1, DegP5 and DegP8 isomers were found in the lumen and a fourth Deg isomer, DegP2, was identified as a peripheral protein attached to the stromal side of the thylakoid membrane (Haussuhl et al., 2001; Schubert et al., 2002).

Other proteases have also been found in chloroplasts, including SppA and Lon-like proteases. SppA is a serine-type, ATP-independent protease tightly bound to the stromal side of the thylakoid. Plant SppA is a constitutively expressed protease, but can be induced by increasing light intensity (Lensch et al., 2001). Four Lon-coding genes have been found in Arabidopsis genome, of which products of two are located in mitochondria, whereas the other two are predicted to be chloroplast proteins (Adam et al., 2001). Unfortunately, the structure and function of these proteases remain unknown.

Despite of the variety of chloroplast proteases, our knowledge about their function and regulation is still very limited. However, evidence does show their involvement in the degradation of several photosystem-II (PSII), such as the reaction-centre D1 protein, chlorophyll-a/b-binding proteins and the early light-inducible protein (ELIP) (Adam, 1996; Andersson and Aro, 1997). The model for degradation of D1 protein by cooperation of FtsH and DegP proteases is the most established (Fig.1-8). DegP2 cleaves light-damaged D1 protein into 23-kDa and 10-kDa fragments; and the 23-kDa fragments are further digested by FtsH in the stroma and DegP1, DegP5 and DegP8 in the lumen. The mechanism involved in the degradation of the 10 K-Dalton fragments has not yet been solved (Haussuhl et al., 2001; Lindahl et al., 2000; Spetea et al., 1999).
Part II. Clp/Hsp100, a Novel Family of Molecular Chaperones

2-1. Classification of the Clp family

The Clp/Hsp100 protease system was first identified in *E. coli*, and consists two non-homologous components: a proteolytic subunit known as ClpP and an ATPase subunit (ClpA). The ATPase subunit is required for binding and unfolding of the protein substrates and then transferring it to the ClpP protease for degradation (Squires and Squires, 1992). Two distinct ATPase subunits in *E. coli*, ClpA (83 kDa) and ClpX (46 kDa), are known to associate transiently with ClpP (21.5 kDa). Due to their difference in substrate specificity, the subsequent proteases degrade different types of proteins (Hwang et al., 1987; Katayama et al., 1987). To date, discovery of ClpA has led to many new additions to this family.

Clp/Hsp100, the regulatory subunits

The Clp family of ATPases consists of at least nine different subfamilies (ClpA, ClpB, ClpC, ClpD, ClpE, ClpM, ClpN, ClpX and ClpY) that can be classified into two general
categories: the first includes ClpA-E and consists of relatively large proteins ranging from 68 to 110 kDa. Each of them contains two characteristic ATP-binding domains (ATP-1 and ATP-2) which are highly conserved, whereas the flanking and spacer regions differ in size and amino acid composition. The second group includes relatively new members: ClpM, ClpN, ClpX and ClpY. These Clp proteins are considerably smaller, and notably only have one ATP-binding domain, which is more similar to the ATP-2 domain of ClpA-E proteins (Fig. 2-1) (Schirmer et al., 1996).

ClpA belongs to the AAA superfamily, a ubiquitous family of ATP-dependent molecular machines with one or two 230-residue ATP-binding domains containing well conserved sequence and structural motifs (Neuwald et al., 1999). ClpA alone has no proteolytic activity, but instead it has ATPase activity (Hwang, et al., 1987; Hwang, et al., 1988; Katayama et al., 1988). It may associate with ClpP, which functions as a peptidase on its own, but can cleave denatured proteins in concert with ClpA (Woo et al., 1989). It has been discovered that ClpA can also act as a molecular chaperone, in that it can reactivate partially denatured, misfolded proteins by disaggregating them into monomers (the active form) (Pak and Wickner, 1997).

Unlike ClpA, ClpB is found in both prokaryotes and eukaryotes, and in all cases except one is a stress-inducible protein. ClpB is distinguishable from other group I Clp proteins by its relatively long intervening region between the two conserved ATP-binding domains (Squires and Squires, 1992). ClpB functions as a molecular chaperone, although the existing form and function of ClpB in bacteria, yeast and higher plants differ somewhat (Parsell et al., 1991; Leonhardt et al., 1993; Porankiewicz and Clarke, 1997). In almost all organisms, there are two forms of ClpB. In bacteria, these different forms (78 and 94 kDa) arise from two translational initiation sites within the same *clpB* transcript (Park et al., 1993). In contrast, eukaryotes possess separate nuclear *clpB* genes, with the larger protein (100-110 kDa) found in the nucleus and cytoplasm (Sanchez and Lindquist, 1990) and the smaller form in the mitochondria (Leonhardt et al., 1993). In *E. coli*, loss of ClpB synthesis will significantly reduce cell viability after a sudden and extreme heat shock but does not affect the cell’s ability to develop thermotolerance (Squires et al., 1991). Contrary to this, induction of cytoplasmic ClpB is essential for *Saccharomyces cerevisiae* to acquire thermotolerance, but has little influence on cell survival after heat shock (Sanchez and Lindquist, 1990). Both the cytosolic and mitochondrial ClpB proteins in yeast function as molecular chaperones during heat shock. The cytosolic ClpB (Hsp104)
promotes the resolubilization of unfolded proteins that have aggregated during heat stress (Parsell et al., 1994), and the smaller mitochondrial ClpB (Hsp78) prevents protein denaturation at high temperatures (Schmitt et al., 1995). The ClpB homologue in the cyanobacterium *Synechococcus* sp. strain PCC 7942 is also essential for sustained thermotolerance, similar to the yeast Hsp104 protein (Eriksson and Clarke, 1996).

ClpC is a relatively new member of the Clp/Hsp100 family. It has a spacer domain of an intermediate length about 60 amino acid residues (Gottesman et al., 1990). ClpC is present in Gram-positive bacteria, cyanobacteria, algae and higher plants, but its function is so far unclear. Unlike the heat-inducible ClpB, ClpC is constitutively expressed with little induction during heat stress in plants and cyanobacteria. In *Bacillus subtilis*, however, the basal level of *clpC* expression is relatively low, and increases dramatically under certain types of stresses, particularly high temperatures (Kruger et al., 1994; Schirmer et al., 1994). Genetic evidence from both plants and cyanobacteria suggests that the ClpC protein has evolved a constitutive role essential for normal growth (Shanklin et al., 1995; Clarke and Eriksson, 1996). ClpC in higher plants is nuclear encoded, translated in the cytosol, and subsequently imported into the chloroplast where it localises preferentially in the stroma (Moore and Keegstra, 1993; Shanklin et al., 1995). The additional role of ClpA as a molecular chaperone also raises the possibility that ClpC might also participate in both chaperone and proteolytic process. For example, ClpC might participate in the folding or degradation of metabolically important polypeptides such as those related to oxygenic photosynthesis (Halperin and Adam, 1996). It is established that ClpC (ATPase) and a chloroplast ClpP homologue can be co-immuno-precipitated from the stromal protein extract of a higher plant (Desimone et al., 1997). ClpC is proposed to function as a molecular chaperone, co-operating with other components to accomplish the transport of precursor proteins into chloroplasts, by means of forming protein transport complexes in the chloroplast envelope membranes (Akita et al., 1997).

During the investigation of molecular responses to dehydration in *Arabidopsis*, a novel protein, termed Erd1 (early response to dehydration), was discovered (Kiyosue et al., 1993), which was later identified as a new form of Clp protein (ClpD). ClpD has the two distinct ATP-binding domains characteristic of ClpA-like proteins, as well as a putative chloroplast transit peptide at the amino-terminus. The transcription of *clpD* is not only induced by dehydration, but also developmentally upregulated during senescence (Nakashima et al., 1997; Nakashima et al., 1999). It has been shown that the expression of
the *erd1* gene is not affected at the transcription level by heat, cold or heavy metal stress (Kiyosue et al., 1993). To date, the specific function of ClpD during dehydration stress is unknown, or whether it possesses the dual chaperone/proteolytic regulatory activities like ClpA.

ClpE is a new member of Clp family initially described in *B. subtilis* during the genome-sequencing project (Kunst et al., 1997), and has since then been identified in many Gram-positive bacteria, including several pathogens (Derré et al., 1999; Nair et al., 1999). Besides the two nucleotide-binding domains of Class I Clp/Hsp100 proteins, ClpE has a unique feature at the short N-terminal domain, a putative zinc-binding domain (-CX₂₅CX₂₀CX₂₃C-). It is suggested ClpE, along with ClpP, which was shown to participate in the degradation of randomly folded proteins in *Lactococcus lactis*, could be necessary for degrading proteins generated by certain types of stress, such as heat shock (Ingmer et al., 1999). It is also shown that ClpE together with ClpC plays an important role in cell division (Nair et al., 1999; Nair et al., 2000).

Of the group II members, the ClpM and ClpN subfamily is distinguished from that of ClpX and ClpY by having greater homology with group I Clp proteins (Schirmer et al., 1996). ClpX and ClpY were both originally found in bacteria, with ClpM and ClpN, respectively, their eukaryotic counterparts. ClpM differs from ClpN by the presence of an additional amino-terminal extension (Schirmer et al., 1996). ClpX was first found in *E. coli*, where it was located downstream of the *clpP* gene as part of a *clpP-clpX* operon (Gottesman et al., 1993). ClpX enables ClpP to efficiently degrade specific protein substrates that are not recognised by ClpA (Mhammedi, 1994). Like ClpA, ClpX has dual functions - alone as a molecular chaperone, and as part of a proteolytic complex when associated with ClpP (Wawrzynow, 1995). With ClpP, ClpX degrades some proteins involved in replication transcription, and translational control, such as the λ-0 protein (Wojtkowiak et al., 1993), the phage P1 protein, Phd (Lehnherr and Yarmolinsky, 1995), and the σ factor RpoS (Schweder et al., 1996). ClpY is the most divergent of the group in amino acid sequence, with a spacer region dividing the ATP-2 domain. ClpY in *E. coli* is present downstream of the *clpQ* gene as part of a *clpQ-clpY* operon, and is thought to have a similar function to that of ClpX (Kessel et al., 1996; Bochtler et al., 1997).
**ClpP, the proteolytic subunit**

Although there is no homology between ClpP and the Clp/Hsp100 family, it is important to understand how the two proteins complex together and function in protein degradation. In *E. coli*, the Clp protease consists of ClpA (regulatory ATPase subunit) and ClpP (proteolytic subunit) (Maurizi, et al., 1990). Later, it was found that in addition to ClpA, other large Clp proteins such as ClpX (Wojkowiak et al., 1993) and probably ClpC (Shanklin et al., 1995) could associate with ClpP for protein degradation. In contrast, no evidence exists so far for ClpB being able to complex or to support the proteolytic activity of ClpP (Woo et al., 1992; Parsell et al., 1994).

ClpP is a unique serine protease (Hwang et al., 1987), which alone has only limited peptidase activity (Woo et al., 1989). This suggests ClpP possesses within its own primary structure the essential elements for peptide bond cleavage, a particularly interesting finding given that it requires intramolecular, autocatalytic cleavage of its precursor form to produce the mature ClpP protein (Maurizi et al., 1990). With the help of ClpA, it can degrade specific protein substrates by a pathway based on N-end rule in prokaryotes (Tobias et al., 1991), similar to those proteins degraded by the ubiquitin system in eukaryotes. In *E. coli*, ClpX/ClpP degrades a different set of protein substrates that are involved in DNA replication, transcription, and translational control (Wojkowiak et al., 1993).

In higher plants, the ClpP homologue was originally found as an open reading frame in the chloroplast genomes of several species such as tobacco, liverwort, maize, rice, wheat and pine (Gray et al., 1990; Maurizi et al., 1990; Clarke et al., 1994). The amino acid sequences of these chloroplast ClpP proteins are well conserved, especially the Ser, His and Asp residues that comprise the catalytic site previously identified for the *E. coli* ClpP. For most plant species, chloroplast-encoded ClpP is expressed as a polycistronic mRNA of variable length, spliced post-transcriptionally to give the mature monocistronic message (Gray et al., 1990; Clarke et al., 1994). It is unclear at this stage whether this chloroplast ClpP also requires post-translational processing of the nascent polypeptide as does the *E. coli* homologue to produce a proteolytically active mature protein.

Most chloroplast *clpP* genes are part of an operon that also includes the gene for the ribosomal proteins Rps12 and Rpl20, with the position of the first exon of *rps12* in relation to *clpP* being conserved (Clarke et al., 1994). This arrangement suggests both a
common evolutionary origin for these chloroplast clpP genes, and a possible functional relationship between ClpP proteolysis and chloroplast protein translation. Despite this, not all chloroplast clpP genes are organised in such operons, with the Arabidopsis homologue being monocistronic and divided into three separate exons. Furthermore, since the ClpA homologue, ClpC, is also present in chloroplasts of higher plants, it is likely a similar proteolytic complex to that in E. coli is functional in chloroplasts (Shanklin et al., 1995; Halperin and Adam, 1996).

2-2. Structure and mechanism of Clp proteasome

The Clp protease has a structural arrangement similar to the Ubiquitin/26S proteasome in archaebacteria and eukaryotes, and it is now considered the equivalent protease in eubacteria (Vierstra, 1996). Both are composed of a proteolytic core particle, in which the active sites are compartmentalized, and an ATP-dependent regulatory particle, with ATPase and chaperone activity (Kessel et al., 1995). ClpAP and ClpXP from E. coli are the best-characterized Clp protease systems. Structural studies indicate that the Clp protease consists of two heptameric rings of ClpP and up to two flanking hexameric rings of ClpA, forming a complex with a molecular mass of around 750-kDa (Fig. 2-2) (Flanagan et al. 1995; Kessel et al. 1995).

![Figure 2-2. Model for structure of the Clp proteasome (Flanagan et al. 1995; Kessel et al. 1995).](image)

In the absence of nucleotides, regulatory subunits form monomers, dimers and trimers, according to a variety of analytical techniques. However, in the presence of ATP or ATP analogues, they assemble into hexamers (Schirmer et al., 1996). Electron microscopy studies show that ClpA has a two-tiered structure with topologically separate rings: One interacts with ClpP, whereas the other binds protein substrates (Ishikawa, 2001). The six
subunits of the ClpA hexamer enclose a vase-shaped central cavity with two spacious compartments connected by a narrow passage and an opening on each end (Guo et al., 2002). The regulatory subunits are responsible for substrate selection, unfolding and translocation to the proteolytic core, and allosteric modulation of the ClpP activity. Alone they also have chaperone activity, such as catalyzing limited structural remodeling and disassembly of stable protein complexes in the absence of the proteolytic component (Wickner et al., 1994).

The proteolytic core consists of 14 identical subunits of ClpP in two back-to-back stacked heptameric rings (Wang et al., 1997). The proteolytic active sites are located within a central, roughly spherical chamber ~51Å in diameter. Two axial pores control access to the proteolytic chamber, each having a minimum diameter of ~10 Å. The full size of tetradecameric ClpP in both height and diameter is about 90Å (Flanagan et al., 1995; Kessel et al., 1995; Wang et al., 1997). ClpP displays limited serine-peptidase activity against peptides of less than six amino acid residues (Woo et al., 1989; Maurizi et al., 1994). Degradation of peptides longer than six residues requires ClpA to form an active ClpAP protease and hydrolyse ATP. Protein substrates are degraded in a highly processive manner, producing peptides of seven to ten residues. The cleavage does not show any sequence preference (Maurizi et al., 1994; Thompson and Maurizi, 1994; Thompson et al., 1994).
Part III. Characterisation of Clp Proteins in Arabidopsis thaliana

Arabidopsis is now one of the most commonly studied higher plants for several reasons. It is easy to culture and has a short life cycle. It also has the advantage of being readily transformable which enables the construction of specific mutants. Another advantage is the large number of available EST clones and the recently completed DNA sequencing of the chromosomal genome (Arabidopsis Genome Initiative, 2000). When we started to characterise Arabidopsis Clp family in 1997, there were only seven Clp proteins found in this organism: two ClpB, one ClpC, ClpD, ClpX, chloroplast-encoded ClpP (ClpP1, previously named as pClpP) and a single nuclear-encoded ClpP (ClpP6, previously named as nClpP). We first obtained cDNA or genomic clones coding for four distinct chloroplast Clp proteins - ClpC, ClpD, ClpP1 and ClpP6 from Arabidopsis. The aim was to characterise the complement of Clp proteins in the chloroplast of Arabidopsis, and to examine changes in the level of these proteins in leaves exposed to different stresses, based on previous observation of apparent stress responses of Clp proteins in cyanobacteria and other organisms (Porankiewicz et al., 1999). Once the Arabidopsis genome sequencing was completed, however, a variety of new Clp proteins were identified (see Table 3-1, Paper I, and Adam and Clarke, 2002). Together with previously known Clp proteins, more than 20 Clp proteins have now been found in Arabidopsis, including four ClpB, two ClpC, ClpD, two ClpX, six ClpP, four ClpR, two ClpS and several as yet undefined Clp homologues. Alignments of the six ClpP proteins, and ClpD with the two ClpC proteins from Arabidopsis are shown in Figures 3-1 and 3-2, respectively. Asterisks (*) above the sequence indicate the Ser, His and Asp residues of the proteolytic active site of ClpP. All the four ClpR proteins have sequence similarity to ClpP, but they don’t have the serine-type proteolytic site. A phylogenetic tree illustrating the evolutional distance of all ClpP and ClpR proteins from Arabidopsis can be found on the website http://www.genebee.msu.su (Fig. 3-3).
Table 3-1. Changes in nomenclature and number of Clp proteins in *Arabidopsis thaliana*

<table>
<thead>
<tr>
<th>Year 1997</th>
<th>Year 2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClpB1-2</td>
<td>ClpB1-4</td>
</tr>
<tr>
<td>ClpC</td>
<td>ClpC1-2</td>
</tr>
<tr>
<td>ClpD</td>
<td>ClpD</td>
</tr>
<tr>
<td>ClpX</td>
<td>ClpX1-2</td>
</tr>
<tr>
<td>PClpP and nClpP</td>
<td>ClpP1-6</td>
</tr>
<tr>
<td></td>
<td>ClpR1-4</td>
</tr>
<tr>
<td></td>
<td>ClpS1-2</td>
</tr>
<tr>
<td>Other Clp homologues with unknown function</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-1. Alignment of the ClpP proteins from *Arabidopsis thaliana*

Asterisks (*) above the sequence indicate the Ser, His and Asp residues of the proteolytic active site of ClpP.
Figure 3-2. Alignment of the ClpD and the two ClpC proteins from *Arabidopsis thaliana*

Dark shading denotes residues conserved among all three proteins, and gray shading is applied for conserved residues in two proteins.
3-1. Preparation of molecular tools

To understand the function and regulation of Clp proteins in *Arabidopsis*, studies on both translational and transcriptional levels were carried out. Specific polyclonal antibodies were prepared (Table 3-2) to detect each Clp protein for various analyses such as determination of relative protein contents, subcellular localisation, and the composition of native Clp complexes. At the same time, specific primers were also prepared to quantify the *clp* mRNAs under optimised conditions by RT-PCR (Table 3-3).

**Table 3-2.** Specific polyclonal antibodies against Clp proteins

<table>
<thead>
<tr>
<th>Target Proteins</th>
<th>Antigens</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClpP1</td>
<td>C-terminal fusion proteins with the maltose-binding protein (MBP)</td>
</tr>
<tr>
<td>ClpP2</td>
<td>KLH conjugated synthetic peptide MVIEHSSRGERAYDIC</td>
</tr>
<tr>
<td>ClpP3</td>
<td>KLH conjugated synthetic peptide KVEGKKDNTNLPSERSMTQ</td>
</tr>
<tr>
<td>ClpP4</td>
<td>KLH conjugated synthetic peptide FEELDTTNMLRQRI</td>
</tr>
<tr>
<td>ClpP5</td>
<td>KLH conjugated synthetic peptide DIDIQANMLHELHKKANLYNGYL</td>
</tr>
<tr>
<td>ClpP6</td>
<td>Entire protein fused to MBP</td>
</tr>
<tr>
<td>ClpC</td>
<td>N-terminal domain of the <em>Synechococcus</em> ClpC fused to MBP</td>
</tr>
<tr>
<td>ClpD</td>
<td>C-terminal domain fused to MBP</td>
</tr>
<tr>
<td>ClpX</td>
<td>BSA and KLH Conjugated synthetic peptide LTEPKNALGQYKKM</td>
</tr>
</tbody>
</table>
### Table 3.3. Primers designed for RT-PCR

<table>
<thead>
<tr>
<th>Target Genes</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>clpP1</strong></td>
<td>Antisense: 5'-CAT ATA GGT TTG CCC GTT CTT TGT AC-3'&lt;br&gt; Sense: 5'-CCT GGA GAA GGA GAT ACA TCT TGG G-3’</td>
</tr>
<tr>
<td><strong>clpP2</strong></td>
<td>Antisense: 5'-AAG CAT CGC CAG TGG TAG AAG AAG C-3'&lt;br&gt; Sense: 5'-ACA TAC AAT TCA TTT AAC GCA TCC C-3’</td>
</tr>
<tr>
<td><strong>clpP3</strong></td>
<td>Antisense: 5'-TAC GGC TTC CCA GTG ATT CTA GAG AAG A-3’&lt;br&gt; Sense: 5'-AAA CCA CCC AGA CAA ACC TTA TCT AGT A-3’</td>
</tr>
<tr>
<td><strong>clpP4</strong></td>
<td>Antisense: 5'-TGC TGG TGA CAT TGT TCT TGT TAT GC-3'&lt;br&gt; Sense: 5'-TCA ATC GAG ATG TCG CAG ACA CAG GAA-3’</td>
</tr>
<tr>
<td><strong>clpP5</strong></td>
<td>Antisense: 5'-CAT TTC ATT TGC TCT AAT GTC AAT GTC-3'&lt;br&gt; Sense: 5'-AGA TTC CTC TCT CTC AAG GAG TTT GGT C-3’</td>
</tr>
<tr>
<td><strong>clpP6</strong></td>
<td>Antisense: 5'-GGC TTC ATT GAC TCT CCT CAC GTC-3'&lt;br&gt; Sense: 5'-GAT GAC CCC TGG AGG ACC TTT AGA CC-3’</td>
</tr>
<tr>
<td><strong>clpC1</strong></td>
<td>Antisense: 5'-CTG TGG TAA CTG TCT TGG TCC TCC TC-3'&lt;br&gt; Sense: 5'-GTA TCT GCA ATC CAA GCG AAA-3’</td>
</tr>
<tr>
<td><strong>clpC2</strong></td>
<td>Antisense: 5'-CTG TGG TAA CTG TCT TGG TCC TCC TC-3'&lt;br&gt; Sense: 5'-GAT GAC TAA TGT TTT ATC TCG-3’</td>
</tr>
<tr>
<td><strong>clpD</strong></td>
<td>Antisense: 5'-GCT CAT CGT CGC TCT TGG CAG GTG GC-3'&lt;br&gt; Sense: 5'-CGA GCC AAG AGA AGA AGA AGA TGC GGT G-3’</td>
</tr>
<tr>
<td><strong>clpX</strong></td>
<td>Antisense: 5'-CTG CCT GTA CCC TCC TC-3'&lt;br&gt; Sense: 5'-GAA TCA TCA CAG AAA CGG TC-3’</td>
</tr>
</tbody>
</table>

### 3.2. Localisation of newly found Clp proteins

ClpC and ClpP1 were both detected in the stroma of chloroplasts (Shanklin et al., 1995), but the localisation of ClpP6, ClpD and ClpX was still unknown at the time of this study. With more and more EST sequences emerging and the near-complete genome sequencing, members of the Clp family increased rapidly, making their localisation an important characteristic towards understanding the biological functions of all plant Clp proteins. Predictions of intracellular locations and transit peptide processing sites were done using TargetP version 1.01 and ChloroP version 1.01 (Paper I). All five nuclear encoded ClpP proteins had very high scores as being chloroplast-localised. Leaves of six-week-old *Arabidopsis* plants were harvested for chloroplast isolation as previously described (Kunst, 1998), followed by sub-fractionation to separate stromal and thylakoid membrane proteins. Whole leaf protein extracts, whole chloroplast protein extracts, stromal and thylakoid fractions were separated by denaturing-PAGE and analysed immunologically. ClpC and ClpP1, as previously observed, were both localised in the stroma of...
chloroplasts, which served as a positive control for the stroma fraction; LHCb2, a known thylakoid membrane protein, was used as the control for the thylakoid fraction. ClpD and ClpP3-6 were all detected in the stroma, with no signal in the thylakoid fraction (Paper III). Levels for each Clp protein in the leaf extract, whole chloroplast extract and the stroma fraction were almost identical, suggesting these Clp proteins are present mainly in the stroma of chloroplasts. Despite these results, some Clp components have also been detected on the stromal face of thylakoid membrane (Peltier et al., 2001). As acknowledged by the authors, however, the amount of Clp protein attached to the thylakoid surface is likely to be minor compared to that in the stroma. One possible explanation for the differences in Clp protein localisation observed the two studies could be the severity of the rupture conditions used to lyse the isolated chloroplasts.

Although most ClpC protein is soluble in chloroplasts, Nielsen et al. (1997) showed that less than 10% of ClpC protein co-precipitated with IEP110, a component of chloroplast import complexes (see Figure 1-5). Using our specific antibodies, we examined if a proportion of ClpD protein could also be associated to the import complex along with ClpC. Intact chloroplasts were isolated from Arabidopsis leaves, treated with thermolysin to remove cellular proteins attaching to the outer envelop of chloroplasts, and then fractionated. Without thermolysin treatment, ClpD was detected as the expected 95-kDa protein in the fraction of whole chloroplast, whereas after treatment only a 33-kDa band was detected (unpublished data). However, the isolation of an envelope fraction from Arabidopsis was extremely difficult, and the existence of ClpD in other organisms more amenable for such fractionations, such as pea, has not been well studied as yet. Whether ClpD binds to the import complexes or not remains a challenge for the future.

Surprisingly, ClpP2 and ClpX were not found in the stroma of chloroplasts despite the predictions based on early transit peptide sequence analysis. Evidence from studies of the only ClpP protein in humans showed it localised in mitochondria (Corydon et al., 1998), and interestingly Arabidopsis ClpP2 closely aligned with the human ClpP in phylogenetic studies (Porankiewicz et al., 1999). Using more refined programs (PSORT and ChloroP), ClpX was later predicted to be located in mitochondria and not in chloroplasts. Using specific antibodies, we eventually localised both ClpX and ClpP2 to mitochondria and not to chloroplasts (Paper II), with known chloroplastic (OE33, 33 kDa) and mitochondrial (SHMT, 53 kDa) proteins used as markers. The mitochondrial location of ClpX and
ClpP2 was later confirmed in *Arabidopsis* by a proteomic study of total mitochondrial proteins (Kruft et al., 2001).

**3-3. Tissue-specific expression of Clp proteins**

Tissue from roots, stems and leaves were collected to test for potential differential gene expression. Transcripts for all *clp* genes could be detected in these tissues. The relative level of the transcripts of *clpP2, clpX, clpC1, clpC2* and *clpD* was similar in all tissues, whereas the corresponding protein contents were relatively low in roots but steadily increased from stems to leaves (Paper II and III). Unfortunately, maybe due to relatively low cellular contents, ClpX and ClpP2 proteins failed to be detected in whole leaf extracts using the specific polyclonal antibodies. In contrast, transcripts and protein levels for all plastid ClpP were both low in roots and increased steadily from stems to leaves (Paper III). The relative abundance of plastid Clp proteins in leaves may be due to their functions related to photosynthesis and chloroplast development. However, their presence in roots, although in lower concentration, suggests they may also perform housekeeping duties within the plastids and be involved in turnover of proteins unrelated to photosynthesis.

**3-4. Expression of Clp proteins under different conditions**

Clp proteins in bacteria are involved in many stress responses and the induction can often be very strong after brief exposures. Heat shock induction of the ClpP protein was first shown in *E. coli* (Kroh and Simon, 1990), and was later also observed in *B. subtilis* (Völker et al., 1994). The *B. subtilis* ClpP can also be induced during salt and oxidative stress, or glucose and oxygen deprivation. The ClpP1 protein from the cyanobacterium *Synechococcus* is also induced strongly by high light, low temperature and UV-B exposure (Porankiewicz et al., 1998). To examine if any Clp protein could be involved in stress responses, wild type *Arabidopsis* plants were exposed to different stress treatments.

**Brief severe stresses**

Stresses including dehydration, high salt concentration, heat and cold shock, oxidation, wounding and high light intensity were briefly applied to detached *Arabidopsis* leaves (Paper III). Plastid Clp proteins were examined by immunoblot analysis on the basis of equal chlorophyll content. There were no significant, reproducible induction of these proteins under any of the stresses tested, whereas a decreased level was observed for some
Clp proteins under certain stress conditions, especially for ClpP3 after wounding and high light treatments.

Expression of mitochondrial Clp proteins was studied by RT-PCR Whole plants were transferred to either 40 or 15°C for 4 h to test the effect of temperature, or shifted to either 40 or 700 µmol photons m⁻² s⁻¹ of irradiation for 3 d to assess the effect of light intensity. Neither clpX nor clpP2 exhibited any significant changes in transcript levels under all treatments.

**Prolonged moderate stress: 1. High light stress response**

Unlike in cyanobacteria, the level of Clp proteins in chloroplasts seemed relatively unaffected by brief and severe stresses. As a consequence, prolonged stresses were introduced to investigate potential long-term responses of Clp proteins. High light intensity, an easy-to-set-up and relevant stress condition, was considered as the first model stress (*Paper III*). Wild type plants were shifted from a day irradiation of 150 to 850 µmol photons m⁻² s⁻¹ for 7 d. Samples were harvested prior to, and 3 and 7 d after the shift, and analysed immunologically. The amount of ClpC and ClpP6 did not change significantly, but that of ClpP3 and ClpP4 slightly increased. ClpP1 and ClpP5 protein levels rose dramatically and steadily during the 7 d high light regime, while ClpD protein content returned to control levels by 7 d after 3-4 d of strong induction.

**Prolonged moderate stress: 2. Cold acclimation**

The lack of induction of Clp proteins during brief stresses and the tendency to be induced instead under prolonged high light stress suggests that Clp proteins function more during acclimation than during rapid stress responses. To test this proposal further, expression of Clp proteins was studied under another prolonged moderate stress - subfreezing temperatures (*Paper III*). Samples were collected just prior to the shift and then after 3, 7, 15, 21 and 42 d at 5°C. Another set of plants was kept at warm growth condition as an extra control. The amounts of all plastid Clp proteins, except for ClpP4, significantly increased after 7 to 15 d at 5°C, but not in the control leaves. Instead, the control leaves had a developmental expression pattern of Clp proteins, which will be mentioned later.
3-5. Expression of Clp proteins during leaf senescence

Expression of Clp proteins during leaf senescence was previously studied at both transcriptional and protein levels (Weaver et al., 1999; Nakabayashi et al., 1999). The expression of clpD increased during both dark-induced and natural senescence, while that of clpC decreased (Nakashima et al., 1997). In contrast to the clpD mRNA, the ClpD protein first increased before leaves fully expanded, and then strongly declined in abundance with natural leaf yellowing (unpublished data). The amount of mRNA for all plastid-localised ClpP proteins decreased during dark-induced senescence but showed varied expression patterns during natural senescence – higher levels of clpP3-5 transcripts but unchanged levels for clpP1 and clpP6 (Nakabayashi et al., 1999). All plastid ClpP protein levels decreased during both dark-induced and natural senescence, with the exception of ClpP3, ClpP5 and ClpP6 that remained unchanged during dark-induced senescence (unpublished data).

3-6. Screening for clp mutants and other transgenic lines

To further understand the function and regulation of Clp proteins during plant growth and development, mutant lines for each of the chloroplast Clp isomers were screened from many of the existing T-DNA insertion libraries. Unfortunately, no viable homozygous clpP mutants were later found, despite obtaining seeds for several potential mutant lines from the following libraries: Feldmann and Jack lines, Syngenta SAIL lines/GARLIC lines, and Salk Institute lines. Screening for antisense repression transgenic lines was hence started and has been proved very successful (Table 3-4). The plant expression vector pSJ10 was used for all antisense constructs (Ganeteg et al., 2001). Full length or partial clp genes were amplified by PCR with two specific primers incorporating EcoRI sites at the 5’ ends. Fragments were digested with EcoRI, and then ligated into the EcoRI-digested pSJ10 vector. Correct orientation of the insert was confirmed by sequencing. Arabidopsis transformation was performed according to Bechtold and Pelletier (1998). Transformed seeds were selected on MS (Murashige and Skoog) plates containing 50 mg/mL of Kanamycin and 200 mg/mL Cefotaxime (KM50 + CT200). Kanamycin-resistant plants were collected, screened by immunoblotting to verify the efficiency of the antisense repression (for the example, see Figure 3-4), and allowed to self-pollinate for further study.
Table 3-4. Screening for clp antisense repression transgenic lines

<table>
<thead>
<tr>
<th>Target Genes</th>
<th>Antisense construct</th>
<th>Kanamycin positives</th>
<th>Lines with ≥ 60% of repression</th>
<th>Morphological changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>clpP3</td>
<td>whole protein</td>
<td>&gt; 100</td>
<td>10 – 20%</td>
<td>chlorosis</td>
</tr>
<tr>
<td>clpP4</td>
<td>whole protein</td>
<td>&gt; 100</td>
<td>10 – 20%</td>
<td>chlorosis</td>
</tr>
<tr>
<td>clpP5</td>
<td>whole protein</td>
<td>&gt; 100</td>
<td>10 – 20%</td>
<td>chlorosis</td>
</tr>
<tr>
<td>clpP6</td>
<td>whole protein</td>
<td>&gt; 100</td>
<td>10 – 20%</td>
<td>chlorosis</td>
</tr>
<tr>
<td>clpD</td>
<td>the spacer</td>
<td>&gt; 100</td>
<td>~90%</td>
<td>ND</td>
</tr>
</tbody>
</table>

Characterisation of clpD antisense lines

As we observed from the study of Clp proteins during cold stress in Arabidopsis, ClpD is strongly inducible during the acclimation process to non-freezing temperatures. This observation attracted much interest in examining the possible function of the ClpD protein during the cold shift. Does it act as a molecular chaperone to facilitate correct folding and assembly of other proteins at low temperatures? Does it in association with ClpP help to remove protein aggregates caused by the cold? Does it ultimately increase the freezing tolerance of plants subject to the subfreezing temperatures? To address these questions, antisense repression lines were prepared. Although about 90% of the Kanamycin-resistant

Figure 3-4. Screening for clpP4 antisense repression transgenic lines by immunoblotting.

Lines with more than 60% of repression were marked with a solid triangle. M: protein standards.
lines had up to 90% repression of ClpD under standard growth conditions, when shifted to the 5/5°C temperature regime many induced ClpD protein to levels similar to that in wild type plants. Therefore, a second round screening by immuno-blotting was performed to identify lines that maintained strong repression of ClpD protein even in the cold acclimated leaves (Figure 3-5). However, when the ability of these strong antisense lines to acclimate to low temperatures was tested, no significant difference in freezing tolerance was observed relative to wild type *Arabidopsis*. Other physiological parameters were also measured without any obvious discrepancies, including the photosynthesis efficiency (Fv/Fm), growth rate, flowering time, shoot biomass (fresh weight) and leaf area.

**Characterisation of clpP4 antisense lines**

In Paper IV, antisense repression lines of clpP4 were prepared and then later characterised. Within the various lines screened, up to 90% of ClpP4 protein content was specifically repressed, which also led to the down-regulation of ClpP3 and ClpP5 protein contents. The repression of clpP4 mRNA retarded the development of chloroplasts and the differentiation of leaf mesophyll cells, resulting in chlorosis phenotypes. The chlorosis was more severe in young than in mature leaves due likely to the developmental expression pattern of the ClpP4 protein. Chlorotic plants eventually turned green upon aging, accompanied by a recovery in the amount of the ClpP4 protein. The greening process could be affected by the light quantity, either by altering the photoperiod or light intensity.

**3-7. Perspectives and future work**

The existence of Clp proteins in plants has been recognised now for more than ten years, but yet our knowledge still remains rudimentary, due mainly to their large number and complexity of isomeric forms relative to those in non-photosynthetic organisms. Although a fundamental understanding of their structures, functions, regulation and
mechanism can be provided by studies of the Clp proteins from bacteria like E. coli; this knowledge now appears inadequate for characterising the vast diversity of Clp proteins in higher plants. The best available bacterial model system, however, are cyanobacteria. It is almost certain that a study of Clp proteins in these photosynthetic bacteria will provide many important insights into those in chloroplasts of higher plants. Given that cyanobacteria are believed to be the progenitors of plant chloroplasts via the endosymbiotic theory, it now seems in hindsight not that surprising that both organisms share similar Clp proteins. For example, the unicellular cyanobacterium Synechococcus has several Clp proteins closely related to those in chloroplasts of Arabidopsis: highly conserved and functionally vital ClpC proteins that are ca. 90% similar; multiple ClpP isomers, and; the presence of the ClpP variant of unknown function, ClpR. The case of ClpR in particular highlights the advantage of studying the cyanobacterial homologue, of which only one exists, compared to the four paralogues that occur in Arabidopsis. Solving the structure and function of the single ClpR protein in Synechococcus will vastly improve our understanding of the plant ClpR forms. However, despite the advantages of using cyanobacteria, studies on higher plants are still essential for understanding the importance of plant Clp proteins because of their added complexity, especially for those localised within chloroplasts. Studies of plant Clp proteins may also have potential value for eventually improving the plant’s ability to survive and acclimate to various harsh environmental conditions caused by the seasons, climates, geographic locations or increasing human activities, thus leading possibly to higher crop yields, greater wood production, or better environmental conservation.

As mentioned previously, Arabidopsis has become the most commonly used model plant today, for reasons that include ease of cultivation and genetic transformation, relatively small genome, completely sequenced genome, and the availability of many genetic resources. Up to twenty different Clp proteins have been identified in Arabidopsis, with most of them localised in chloroplasts. Most of our efforts so far have been to generally characterise these plastidic Clp proteins, including their subcellular localisation, expression within different tissues, responses to various stress regimes, and expression during senescence and diurnal cycles. We have also tried to complement these early characterisation studies with the identification/preparation of transgenic lines for further functional studies. After failing to obtain T-DNA insertion knockouts of all four nuclear clpP genes that encode the plastid ClpP isomers, effort was moved to the preparation of
gene-specific antisense repression lines. The clp4 antisense repression lines are so far the best characterised, exhibiting a marked chlorotic phenotype that is strongly affected by the growth light environment and the developmental status of the plant. Putative antisense repression lines for clp3, clp5 and clp6 were also prepared but these need to be further verified to ensure gene-specific repression. Preliminary studies of these other clp antisense lines, however, have revealed chlorotic phenotypes of varying degrees, although such an effect would seem unsurprising given the co-down-regulation of these other ClpP isomers in the clp4 antisense repression lines and that similar phenotypes have been observed by other groups when studying mutants of the plastid-encoded ClpP in tobacco (as ClpP1 in Arabidopsis) (Kuroda and Maliga, 2002; Shikanai et al., 2001). As was done for the clp4 antisense lines, any future characterisation of the other clp antisense lines will first need to confirm the transcriptional levels of the non-targeted clp genes (and even those for clpR) to avoid non-specific cross-repression.

Although the successful preparation of antisense lines for clpD produced plants with high levels of ClpD protein repression (i.e., >90% decrease in wild type levels), no prominent differences occurred relative to wild type Arabidopsis. This lack of phenotypic differences was observed in the photosynthetic efficiency, growth rate, flowering time, shoot fresh weight and leaf morphology of plants grown under short day conditions at both 23/18°C or 5/5°C day/night temperature regimes. Despite this, two T-DNA insertion knockout lines for clpD recently obtained from the SALK library may provide new opportunities in the future to better study the function of ClpD and its interaction with other Clp proteins, especially during cold acclimation when ClpD protein levels dramatically rise in wild type Arabidopsis.

Unlike our success with clpD in Arabidopsis, attempts to antisense repress clpC failed to produce viable lines in tobacco (Shanklin et al., 1995). At the time, only one clpC gene was thought to exist in higher plants, and the apparent lack of viable clpC antisense lines was presumably due to the vital importance of the constitutively expressed ClpC for normal chloroplast function. Since then, a second clpC gene has been identified in several plant species including Arabidopsis, which encodes an almost identical form of ClpC to that derived from clpC1 (Adam and Clarke et al., 2002). Recently, we have identified two viable independent T-DNA insertion mutants of clpC1 in Arabidopsis, which show no change in the amount of clpC2 transcripts but a 65-70% decrease in total ClpC protein. These mutants exhibit a pleiotropic phenotype, including a homogeneous chlorosis on the
leaves distinct from the variegated one in the \textit{clpP4} antisense lines, a significant decrease in growth rate and photosynthetic performance, and marked changes in the levels of other chloroplast Clp proteins (Sjögren et al., in preparation). It is hoped that further characterisation of these mutants will provide valuable insights into the specific importance of ClpC to various chloroplast processes like protein import, as well as its likely interaction with other Clp proteins such as ClpD and various ClpP isomers. We have also recently obtained seeds for a putative \textit{clpC2} knockout in \textit{Arabidopsis} that will enable us to reveal any functional differences between the two ClpC isomers despite their high amino acid identity. Moreover, as there are only three Clp/Hsp100 proteins (\textit{i.e.}, ClpC1, ClpC2 and ClpD) that may act as the regulatory subunits for possible plastid Clp protease complexes, it will be interesting to complement the biochemical elucidation of these protein complexes with crossing the various single-gene knockout lines we already have to prepare double mutant lines.

One big challenge that remains in the field of plant Clp proteins is how the numerous plastid Clp proteins interact with each other and the types of complexes that result. There are three candidates for the regulatory subunits and five for the proteolytic subunits, plus four ClpR that are also likely to be involved. At this stage it is hard to make too many presumptions about possible combinations of Clp proteins within various complexes. Studies so far have identified possible complexes between ClpC and ClpP, and among the ClpP and ClpR isomers (reviewed in Adam and Clarke, 2002), but much work still remains to resolve the precise composition and occurrence of each Clp protein complex. Study on the Uniquitin/26S proteasome in \textit{Arabidopsis} revealed more than one thousand regulatory subunits involved in substrate recognition, therefore it is also possible that the Clp protease recruits several, as yet unknown regulatory components. Candidates for such regulatory components have already been identified within proposed oligomeric ClpP/ClpR complexes attached to the surface of thylakoid membranes (Peltier et al., 2001). Complicating the resolution of different Clp protein complexes is the fact that many are likely involved in a broad range of biological processes in chloroplasts, including some that are indispensable. Any attempts to impair the Clp protease in chloroplasts will therefore directly retard the development of chloroplasts or even lead to lethality.

Although the characterisation of plant Clp proteins has been focused primarily on those located inside chloroplasts, it should not be forgotten that several Clp proteins also exist in
mitochondria. The fact that fewer Clp proteins occur in mitochondria (i.e., two ClpX and one ClpP [ClpP2]) does not necessarily lessen their importance to various mitochondrial functions, especially since homologous Clp proteins also exist in the mitochondria of mammals (Corydon et al., 1998). Of course it will always be a challenge to isolate mitochondria from photosynthetic leaves, which often includes contamination from the thylakoid membranes where other ClpP isomers are apparently extrinsically attached on the stromal face (Peltier et al., 2001). However, with the powerful molecular tools we have prepared, including specific antibodies for most of the plastid-localised Clp isomers from Arabidopsis, study on the mitochondrial Clp proteins now seems a much more feasible prospect. It will be interesting to examine whether the plant mitochondrial Clp proteins share similar functions to their counterparts in mammals, especially humans, or instead are involved in processes or enzymes unique to plant mitochondria, such as the photorespiratory pathway and several enzymes in the respiratory chain. Furthermore, does fewer Clp members mean simpler regulation and structure? Will knockouts of ClpP2 or either one (or both) of the ClpX proteins produce viable mutants, and if so what kinds of phenotypic changes will result? Can any of the plastid ClpP isomers substitute for ClpP2 within a mitochondrial Clp protease associated to ClpX? Many questions remain unresolved for the role of mitochondrial Clp proteins in higher plants, and this will be another demanding challenge for future research.
References

Abdallah F, Salamini F, Leister D.  
A prediction of the size and evolutionary origin of the proteome of chloroplasts of Arabidopsis.  

Adam Z.  
Protein stability and degradation in chloroplasts.  

Adam Z.  
Chloroplast proteases: possible regulators of gene expression?  

Chloroplast and mitochondrial proteases in Arabidopsis. A proposed nomenclature.  

Adam Z, Clarke AK.  
Cutting edge of chloroplast proteolysis.  

Agashe VR, Hartl FU.  
Roles of molecular chaperones in cytoplasmic protein folding.  

Akita M, Nielsen E, Keegstra K.  
Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking.  

Andersson B, Aro EM.  
Proteolytic activities and proteases of plant chloroplasts  

Arabidopsis Genome Initiative.  
Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.  

Bechtold N, Pelletier G.  
In planta Agrobacterium-mediated transformation of adult Arabidopsis thaliana plants by vacuum infiltration.  

Borchert M, Ditzel L, Grob M, Huber R.  
Crystal structure of heat shock locus V (HslV) from Escherichia coli.  

Buchberger A, Schroder H, Hesterkamp T, Schonfeld HJ, Bukau B.  
Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding.  

Bukau B, Deuerling E, Pfund C, Craig EA.  
Getting newly synthesized proteins into shape.  

Bukau B, Horwich AL.  
The Hsp70 and Hsp60 chaperone machines.  

Callis J, Vierstra RD.  
Protein degradation in signalling.  

Clarke AK, Eriksson MJ.  
The cyanobacterium Synechococcus sp. PCC 7942 possesses a close homologue to the chloroplast ClpC protein
of higher plants.

**Clarke AK, Gustafsson P, Lidholm JA.**
Identification and expression of the chloroplast clpP gene in the conifer Pinus contorta.

**Corydon TJ, Bross P, Holst IU, Neve S, Kristiansen K, Gregersen N, Bolund L.**
A human homologue of *Escherichia coli* ClpP casinolytic protease: recombinant expression, intracellular processing and subcellular localization.

**Derre I, Rapoport G, Devine K, Rose M, Msadek T.**
ClpE, a novel type of HSP100 ATPase, is part of the CtsR heat shock regulon of *Bacillus subtilis*.

**Derre I, Rapoport G, Msadek T.**

**Desimone M, Weiss-Wichert C, Wagner E, Altenfeld U, Johanningmeier U.**
Immunochemical studies on the Clp-protease in chloroplasts: Evidence for the formation of a ClpC/P complex.

**Eriksson MJ, Clarke AK.**
The heat shock protein ClpB mediates the development of thermotolerance in the cyanobacterium *Synechococcus* sp. strain PCC 7942.

**Estelle M.**
Proteases and cellular regulation in plants.

**Feldman DE, Frydman J.**
Protein folding *in vivo*: the importance of molecular chaperones.

**Flanagan JM, Wall JS, Capel MS, Schneider DK, Shanklin J.**
Scanning transmission electron microscopy and small-angle scattering provide evidence that native *Escherichia coli* ClpP is a tetradecamer with an axial pore.

**Ganeteg U, Strand A, Gustafsson P, Jansson S.**
The properties of the chlorophyll a/b-binding proteins Lhca2 and Lhca3 studied *in vivo* using antisense inhibition.

**Gatenby AA, Viitanen PV.**
Structural and functional-aspects of chaperonin-mediated protein folding.

**Gottesman S.**
Proteases and their targets in *Escherichia coli*.

**Gottesman S, Clark WP, de Crecy-Lagard V, Maurizi MR.**
ClpX, an alternative subunit for the ATP-dependent Clp protease of *Escherichia coli*. Sequence and *in vivo* activities.

Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes.
Proc Natl Acad Sci U S A. 1990 May; 87(9): 3513-7.
Gottesman S, Wickner S, Maurizi MR.
Protein quality control: triage by chaperones and proteases.

Gray JC, Hird SM, Dyer TA.
Nucleotide sequence of a wheat chloroplast gene encoding the proteolytic subunit of an ATP-dependent protease.

Guo F, Maurizi MR, Esser L, Xia D.
Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease.

Halperin T, Adam Z.
Degradation of mistargeted OEE33 in the chloroplast stroma.

Hartl FU.
Molecular chaperones in cellular protein folding.

Hartl FU, Hayer-Hartl M.
Molecular chaperones in the cytosol: from nascent chain to folded protein.

Haussuhl K, Andersson B, Adamska I.
A chloroplast DegP2 protease performs the primary cleavage of the photodamaged D1 protein in plant photosystem II.

Hendrick JP, Hartl FU.
Molecular chaperone functions of heat-shock proteins.

Hershko A, Ciechanover A.
The ubiquitin system.

Hochstrasser M.
Ubiquitin, proteasomes, and the regulation of intracellular protein degradation.

Hortensteiner S, Feller U.
Nitrogen metabolism and remobilization during senescence.

Hwang BJ, Woo KM, Goldberg AL, Chung CH.

Hwang BJ, Park WJ, Chung CH, Goldberg AL.
Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La.

Ingmer H, Vogensen FK, Hammer K, Kilstrup M.
Disruption and analysis of the clpB, clpC, and clpE genes in Lactococcus lactis: ClpE, a new Clp family in gram-positive bacteria.

Ishikawa T, Beuron F, Kessel M, Wickner S, Maurizi MR, Steven AC.
Translocation pathway of protein substrates in ClpAP protease.

Jackson-Constan D, Keegstra K.
Arabidopsis genes encoding components of the chloroplastic protein import apparatus.
Kaser M, Langer T.
Protein degradation in mitochondria.

Katayama-Fujimura Y, Gottesman S, Maurizi MR.
A multiple-component, ATP-dependent protease from Escherichia coli.

Katayama Y, Gottesman S, Pumphrey J, Rudikoff S, Clark WP, Maurizi MR.
The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component.

Keegstra K, Cline K.
Protein import and routing systems of chloroplasts.

Kessel M, Maurizi MR, Kim B, Kocsis E, Trus BL, Singh SK, Steven AC.
Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome.

Kessel M, Wu W, Gottesman S, Kocsis E, Steven AC, Maurizi MR.
Six-fold rotational symmetry of ClpQ, the E. coli homolog of the 20S proteasome, and its ATP-dependent activator, ClpY.

Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K.
Characterization of cDNA for a dehydration-inducible gene that encodes a CLP A, B-like protein in Arabidopsis thaliana L.

Kruft V, Eubel H, Jansch L, Werhahn W, Braun HP.
Proteomic approach to identify novel mitochondrial proteins in Arabidopsis.

Kroh HE, Simon LD.
The ClpP component of Clp protease is the sigma 32-dependent heat shock protein F21.5.

Krojer T, Garrido-Franco M, Huber R, Ehrmann M, Clausen T.
Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine.

Kroger E, Volker U, Hecker M.
Stress induction of clpC in Bacillus subtilis and its involvement in stress tolerance.

The complete genome sequence of the gram-positive bacterium Bacillus subtilis.

Kunst F.
Preparation of physiologically active chloroplasts from Arabidopsis.

Kuroda H, Maliga P.
Overexpression of the clpP 5'-untranslated region in a chimeric context causes a mutant phenotype, suggesting competition for a clpP-specific RNA maturation factor in tobacco chloroplasts.

Langer T.
AAA proteases: cellular machines for degrading membrane proteins.

Langer T, Kaser M, Klanner C, Leonhardt K.
AAA proteases of mitochondria: quality control of membrane proteins and regulatory functions during
mitochondrial biogenesis.

Laskey RA, Honda BM, Mills AD, Finch JT.
Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA.

Lehnher H, Yarmolinsky MB.
Addiction protein Phd of plasmid prophage P1 is a substrate of the ClpXP serine protease of Escherichia coli.

Lensch M, Herrmann RG, Sokolenko A.
Identification and characterization of SppA, a novel light-inducible chloroplast protease complex associated with thylakoid membranes.

Lehnhardt SA, Farray K, Danese PN, Mason TL.
HSP78 encodes a yeast mitochondrial heat shock protein in the Clp family of ATP-dependent proteases.

Levchenko I, Smith CK, Walsh NP, Sauer RT, Baker TA.
PDZ-like domains mediate binding specificity in the Clp/Hsp100 family of chaperones and protease regulatory subunits.

Lindahl M, Speta C, Hundal T, Oppenheim AB, Adam Z, Andersson B.
The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein.

Lindahl M, Tabak S, Cseke L, Pichersky E, Andersson B, Adam Z.
Identification, characterization, and molecular cloning of a homologue of the bacterial FtsH protease in chloroplasts of higher plants.

Maurizi MR, Clark WP, Katayama Y, Rudikoff S, Pumphrey J, Bowers B, Gottesman S.
Sequence and structure of ClpP, the proteolytic component of the ATP-dependent Clp protease of Escherichia coli.

Maurizi MR, Clark WP, Kim SH, Gottesman S.
ClpP represents a unique family of serine proteases.

Maurizi MR, Thompson MW, Singh SK, Kim SH.
Endopeptidase Clp: ATP-dependent Clp protease from Escherichia coli.

May T, Soll J.
14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants.

McFadden GI.
Endosymbiosis and evolution of the plant cell.

Mhammedi-Alaoui A, Pato M, Gama MJ, Toussaint A.

Moore T, Keegstra K.
Characterization of a cDNA clone encoding a chloroplast-targeted Clp homologue.

Nair S, Dere I, Msadek T, Guillot O, Berche P.
CtsR controls class III heat shock gene expression in the human pathogen Listeria monocytogenes.
Nair S, Frehel C, Nguyen L, Ensuyer V, Berche P.
ClpE, a novel member of the HSP100 family, is involved in cell division and virulence of *Listeria monocytogenes*.

Nakabayashi K, Ito M, Kiyosue T, Shinozaki K, Watanabe A.
Identification of clp genes expressed in senescing *Arabidopsis* leaves.

Nakashima K, Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K.
A nuclear gene, erd1, encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not only induced by water stress but also developmentally up-regulated during senescence in *Arabidopsis thaliana*.

Neuwald AF, Aravind L, Spouge JL, Koonin EV.
AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes.

Nielsen E, Akita M, Davila-Aponte J, Keegstra K.
Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone.

Pak M, Wickner S.
Mechanism of protein remodeling by ClpA chaperone.

Park SK, Kim KI, Woo KM, Seol JH, Tanaka K, Ichihara A, Ha DB, Chung CH.
Site-directed mutagenesis of the dual translational initiation sites of the clpB gene of *Escherichia coli* and characterization of its gene products.

Parsell DA, Kowal AS, Singer MA, Lindquist S.
Protein disaggregation mediated by heat-shock protein Hsp104.

Parsell DA, Sanchez Y, Stitzel JD, Lindquist S.
Hsp104 is a highly conserved protein with two essential nucleotide-binding sites.

Patel S, Latterich M.
The AAA team: related ATPases with diverse functions.

Pelletier JB, Ytterberg J, Liberles DA, Roepstorff P, van Wijk KJ.
Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*.

Porankiewicz J, Clarke AK.
Induction of the heat shock protein ClpB affects cold acclimation in the cyanobacterium *Synechococcus* sp. strain PCC 7942.

Porankiewicz J, Schelin J, Clarke AK.
The ATP-dependent Clp protease is essential for acclimation to UV-B and low temperature in the cyanobacterium *Synechococcus*.

Pallen MJ, Wren BW.
The HtrA family of serine proteases.

Sanchez Y, Lindquist SL.
HSP104 required for induced thermotolerance.


Thompson MW, Singh SK, Maurizi MR. 

Tobias JW, Shrader TE, Rocap G, Varshavsky A. 


Varshavsky A. 

Vierstra RD. 

Vierstra RD. 

Vierstra RD. 


von Arnim AG. 

Waegemann K, Soll J. 

Walter S. 

Wang J, Hartling JA, Flanagan JM. 

The ClpX heat-shock protein of Escherichia coli, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. EMBO J. 1995 May 1; 14(9): 1867-77.

Weaver LM, Froehlich JE, Amasino RM. 

Wickner S, Maurizi MR, Gottesman S. 

Wickner S, Gottesman S, Skowrya D, Hoskins J, McKenney K, Maurizi MR. 

