

Chloroplast and Mitochondrial Proteases in Arabidopsis. A Proposed Nomenclature¹

Zach Adam*, Iwona Adamska, Kazumi Nakabayashi, Oren Ostersetzer, Kirsten Haussuhl, Andrea Manuell, Bo Zheng, Olivier Vallon, Steven R. Rodermel, Kazuo Shinozaki, and Adrian K. Clarke

Department of Agricultural Botany, The Hebrew University of Jerusalem, Rehovot 76100, Israel (Z.A., O.O.); Department of Biochemistry, Arrhenius Laboratories for Natural Sciences, Stockholm University, S-10691 Stockholm, Sweden (I.A., K.H.); Department of Biological Sciences, Graduate School of Science, University of Tokyo, Hongo, Bunkyo-ku, Tokyo, 113-0033 Japan (K.N.); Department of Botany, Iowa State University, Ames, Iowa 50011 (A.M., S.R.R.); Department of Plant Physiology, University of Umeå, S-901 87 Umeå, Sweden (B.Z.); Institut de Biologie Physico-Chimique, Centre National de la Recherche Scientifique, Unité Propre de Recherche 1261, F-75005 Paris, France (O.V.); Laboratory of Plant Molecular Biology, The Institute of Physical and Chemical Research, Koyadai, Tsukuba, Ibaraki, 305-0074 Japan (K.S.); and Botanical Institute, Göteborg University, S-405 30 Göteborg, Sweden (A.K.C.).

The identity and scope of chloroplast and mitochondrial proteases in higher plants has only started to become apparent in recent years. Biochemical and molecular studies suggested the existence of Clp, FtsH, and DegP proteases in chloroplasts, and a Lon protease in mitochondria, although currently the full extent of their role in organellar biogenesis and function remains poorly understood. Rapidly accumulating DNA sequence data, especially from Arabidopsis, has revealed that these proteolytic enzymes are found in plant cells in multiple isomeric forms. As a consequence, a systematic approach was taken to catalog all these isomers, to predict their intracellular location and putative processing sites, and to propose a standard nomenclature to avoid confusion and facilitate scientific communication. For the Clp protease most of the ClpP isomers are found in chloroplasts, whereas one is mitochondrial. Of the ATPase subunits, the one ClpD and two ClpC isomers are located in chloroplasts, whereas both ClpX isomers are present in mitochondria. Isomers of the Lon protease are predicted in both compartments, as are the different forms of FtsH protease. DegP, the least characterized protease in plant cells, has the most number of isomers and they are predicted to localize in several cell compartments. These predictions, along with the proposed nomenclature, will serve as a framework for future studies of all four families of proteases and their individual isomers.

Proteolytic processes in chloroplasts (C) and mitochondria (M) have been a subject of interest for many years, but until recently the identity of the proteases involved has remained obscure (for review, see Adam, 1996; Herrmann, 1996; Andersson and Aro, 1997). Genetic, biochemical, and molecular approaches taken in recent years has led to the identification and characterization of several organellar proteases, all of which are homologs of bacterial proteases best characterized in *Escherichia coli* (for

review, see Clarke, 1999; Adam 2000a, 2000b). Four major families of proteases have been discovered and characterized in Arabidopsis to varying degrees: Clp protease in Cs (Sokolenko et al., 1998; Nakabayashi et al., 1999 and refs. therein) and M (Halperin et al., 2001b), FtsH (Lindahl et al., 1996; Chen et al., 2000) and DegP proteases (Itzhaki et al., 1998; Haussühl et al., 2001) in Cs, and Lon protease in M (Sarria et al., 1998).

Clp proteases comprise a large family of ATP-dependent, Ser-type proteases characterized by separation of the two essential functions to two different polypeptides: a small subunit containing the proteolytic active site (ClpP), and a larger regulatory ATPase subunit (for review, see Gottesman, 1996). Lon protease is the first ATP-dependent protease to be described in *E. coli*. It is also a Ser-type protease, but unlike Clp proteases, its catalytic and ATPase domains reside within a single polypeptide (for review, see Gottesman, 1996). In contrast to Clp and Lon, FtsH is the only essential ATP-dependent protease in *E. coli*. It is a membrane-bound metalloprotease that is characterized by an approximately 200-amino acid ATPase domain (AAA motif) and a zinc-binding domain, His-Asp-X-X-His, that serves

¹ This work was supported in part by the Israel Science Foundation, by the U.S.-Israel Binational Agricultural Research and Development Fund, and by the U.S.-Israel Binational Science Foundation (grants to Z.A.); by the Carl Tryggers Foundation for Scientific Research, by the Swedish Natural Science Research Council, and by the Swedish Strategic Foundation (grants to I.A.); by the Japan Society for the Promotion of Science (grant to K.N.); by the Ekströms Fellowship of the Stockholm University (grant to K.H.); and by the Swedish Agricultural and Forestry Resource Council, by the Carl Tryggers Foundation for Scientific Research, and by the Swedish Foundation for International Cooperation in Research and Higher Education (grants to A.K.C.)

* Corresponding author; e-mail zach@agri.huji.ac.il; fax 972-8-946-7763.

the catalytic function (for review, see Gottesman, 1996). Like for the Lon protease, the catalytic and ATPase functions of FtsH reside on the same polypeptide. *E. coli* contains another family of Ser-type proteases known as the Deg proteases of which DegP (or HtrA) was the first described and is currently the best characterized member (Pallen and Wren, 1997). It is a peripheral membrane protein localized to the periplasmic side of the inner membrane (Skorko-Glonek et al., 1997). The other members of the family are DegQ and DegS, also known as HhoA and HhoB, respectively. All three family members share the putative catalytic triad of Ser-type proteases (His-Asp-Ser), and one or two PDZ-like domains at their carboxy-termini (Pallen and Wren, 1997; Ponting, 1997) implicated in protein-protein interactions.

Rapidly accumulating DNA sequences, primarily from Arabidopsis, have revealed that unlike in *E. coli*, plant genomes code for multiple isomers from each of the major protease families. As more member proteins are discovered and studied, the risk increases of conflicting and misleading nomenclature appearing in the literature—a problem that has already materialized for Clp proteases with similar names being given to different gene products (Clarke, 1999; Nakabayashi et al., 1999). To avoid such confusion in the future and facilitate scientific communication we

have searched the Arabidopsis sequence databases and compiled all available data on the different plant organellar proteases and their isomers. The intracellular location and putative processing sites for each isomeric form was predicted using dedicated software, and we propose a standardized nomenclature for future reference to these proteases.

RESULTS AND DISCUSSION

Clp Proteases

Six different ClpP genes encoding the proteolytic subunit of Clp protease are found in Arabidopsis (Table I). All ClpP proteins share a common motif, the highly conserved catalytic triad of Ser-type proteases comprising of Ser-His-Asp residues. One of them, ClpP1, is found in the C genome of all higher plants and most eukaryotic algae, and is the only protease encoded within the plastid genome. The remaining five ClpP isomers are nuclear-encoded, on chromosomes I or V. ClpP4 through 6 are targeted to the C stroma (Sokolenko et al., 1998; K. Nakabayashi, unpublished data), and ClpP2 is targeted to M (Halperin et al., 2001b). It should be noted that the current designation of ClpP isomers is somewhat different from previous reports, where different names were given to the same protein. Thus, ClpP1 was previ-

Table I. Proposed nomenclature for Clp proteolytic subunits from Arabidopsis

Predicted locations are in italics.

Protein Name	cDNA Accession No.	Precursor	Mature Protein	Transit Peptide	Chromosomal Location and Accession No.	Cellular Location	Reference
		<i>amino acids</i>					
ClpP1 ^a	–	196	–	–	C (AP000423)	C (stroma)	Sato et al. (1999)
ClpP2 ^b	AI996025 ^c	240	211	–	V (AB006708; protein id: BAB09831)	M (<i>matrix</i>)	Halperin et al. (2001b)
ClpP3 ^d	AB022328	309	237	72	I (AC013288; protein id: BAA82067)	C (stroma)	
ClpP4 ^e	AB022329	292	232	60	V (AB018113; protein id: BAA82068)	C (stroma)	K. Nakabayashi (unpublished data)
ClpP5 ^f	AJ012278	298	235	63	I (AC022521; protein id: CAB43488)	C (<i>stroma</i>)	K. Nakabayashi (unpublished data)
ClpP6 ^g	AF016621	271	219	52	I (AC007296; protein id: AAD30248)	C (stroma)	Sokolenko et al. (1998)
ClpR1 ^h	AB022330	387	346	41	I (AC015445; protein id: BAA82069)	C (stroma)	Nakabayashi et al. (1999)
ClpP2 ⁱ	AB022327	279	225	54	I (AC025416; protein id: BAA82066)	C (<i>stroma</i>)	K. Nakabayashi (unpublished data)
ClpR3	AV535047 ^c	292	240	52	I (AC000106; protein id: AAB70396)	C (<i>stroma</i>)	
ClpR4	AV440574 ^c	284	216	68	IV (Z97342; protein id: CAB10484) ^j	C (<i>stroma</i>)	

^a Previously known as pClpP (Clarke, 1999; Nakabayashi et al., 1999). ^b Previously known as nClpP2 (Clarke, 1999). ^c Expressed sequence tag (EST). ^d Previously known as nClpP3 (Nakabayashi et al., 1999) or nClpP4 (Clarke, 1999). ^e Previously known as nClpP4 (Nakabayashi et al., 1999) or nClpP3 (Clarke, 1999). ^f Previously known as nClpP5 (Clarke, 1999) or nClpP1 (Nakabayashi et al., 1999). ^g Previously known as nClpP6 (Nakabayashi et al., 1999) or nClpP1 (Clarke, 1999); sequence conflict in positions 179 through 180. LL in genomic sequence, FF in cDNA. ^h Previously known as nClpP5 (Nakabayashi et al., 1999). ⁱ Previously known as nClpP2 (Nakabayashi et al., 1999). ^j The prediction of splicing in GenBank is incorrect.

ously known as pClpP (Clarke, 1999; Nakabayashi et al., 1999); ClpP3 was previously designated nClpP4 (Clarke, 1999), and vice versa; ClpP5 was known as nClpP1 (Nakabayashi et al., 1999); and ClpP6 was known as nClpP1 (Clarke, 1999).

The Arabidopsis genome contains four other sequences that show homology to ClpP proteins. However, these proteins do not contain the aforementioned catalytic triad, and therefore cannot be considered as bona fide ClpP isomers. Because some of them were previously designated as ClpP, we chose to present them here, although their function and relevance to Clp proteases remains unclear. For their classification we have used the earlier nomenclature of ClpR (Porankiewicz et al., 1999). The previously designated nClpP5 and nClpP2 (Nakabayashi et al., 1999) are now named ClpR1 and ClpR2, respectively. In vitro translation import studies have shown ClpR1 to reside in the C stroma (Nakabayashi et al., 1999), as is also predicted for the other ClpR isomers.

The proteolytic subunit of the Clp protease is dependent on a cognate ATPase for its proteolytic activity. ClpC, a homolog of the *E. coli* ClpA containing two ATP-binding sites, is found in the C stroma of many different plant species (Moore and Keegstra, 1993; Shanklin et al., 1995; Halperin and Adam, 1996; Nakabayashi et al., 1999). Two near-identical ClpC isomers exist in Arabidopsis (Table II). ClpC is found associated with ClpP in the stroma (Desimone et al., 1997; Sokolenko et al., 1998) in an ATP-dependent manner (Halperin et al., 2001a). Another ATPase containing two ATP-binding domains, ClpD, is also found in the stroma (Kiyosue et al., 1993; Nakashima et al., 1997). Although closely related, ClpD differs from ClpC1 and -2 by specific signature sequences (Schirmer et al., 1996) and by its differential expression characteristics (Nakabayashi et al., 1999). ClpD has previously been referred to as ERD1 (Kiyosue et al., 1993; Nakashima et al., 1997; Nakabayashi et al., 1999) and SAG15 (Lohman et al., 1994).

In *E. coli*, ClpP can associate with another member of the Clp-ATPase family, ClpX, which contains only a single ATP-binding site as opposed to the two in ClpA, ClpC, and ClpD. As for ClpC, Arabidopsis has two distinct forms of ClpX encoded in chromosome V (Table II). ClpX1 is absent from Cs and is instead localized in M (Halperin et al., 2001b). Similar targeting to M is also predicted for the second ClpX isomer. The localization of ClpX and ClpP2 to M suggests that an active Clp protease is found not only in Cs, but also in M.

Lon Protease

Three genes encoding Lon proteases are found in Arabidopsis (Table III). Lon1 was localized to M (Sarria et al., 1998), matching the mitochondrial localization of a Lon homolog in yeast (Suzuki et al., 1994; van Dyck et al., 1994). A second isomer, Lon2, deduced from a genomic sequence, is predicted by the Predotar program to be chloroplastic, although the ChloroP program fails to detect a corresponding transit peptide. This situation is reversed for a third isomer, Lon3. ChloroP suggests Lon3 is chloroplastic, whereas Predotar predicts it is not chloroplastic or mitochondrial. Immunoblot analysis with an antibody specific to Lon supports the presence of at least one Lon protease in both plant organelles (O. Osterseker and Z. Adam, unpublished data), although this observation needs to be further substantiated.

FtsH Protease

Arabidopsis contains nine isomers of the FtsH protease (Table IV). Two of these isomers, FtsH1 and -2 were found in Cs as integral proteins within the thylakoid membrane, with their ATP-binding domain and catalytic zinc-binding site facing the stroma (Lindahl et al., 1996; Chen et al., 2000). Mutations in FtsH2 were linked with a variegated phenotype, having green- and white-sectored leaves (Chen et al.,

Table II. Proposed nomenclature for Clp ATPase subunits from Arabidopsis

Predicted locations are in italics.							
Protein Name	cDNA Accession No.	Precursor	Mature Protein	Transit Peptide	Chromosomal Location and Accession No.	Cellular Location	Reference
			<i>amino acids</i>				
ClpC1	AF022909	928	836	92	V (AB017063; protein id: AAC04687)	C (stroma)	
ClpC2	AB022324	952	839	113	III (AL132963; protein id: AAC04687)	C (stroma)	Nakabayashi et al. (1999)
ClpD ^a	D17582	945	867	78	V (AB023044; protein id: P42762)	C (stroma)	Kiyosue et al. (1993)
ClpX1	AF036328	579	533	46	V (AB013388; protein id: AAB88706)	M (<i>matrix</i>)	Halperin et al. (2001b)
ClpX2	T14143/ AV440312 ^b	608	545	63	V (AB025612)	M (<i>matrix</i>)	

^a Previously known as Erd1 (Kiyosue et al., 1993).

^b Partial cDNAs (ESTs).

Table III. Proposed nomenclature for Lon proteases from Arabidopsis

Predicted locations are in italics.							
Protein Name	cDNA Accession No.	Precursor	Mature Protein	Transit Peptide	Chromosomal Location and Accession No.	Cellular Location	Reference
<i>amino acids</i>							
Lon1	U88087	941	875	66	V (AF007270; protein id: AAB61060)	M	Sarria et al. (1998)
Lon2	AV524160 ^a	888	–	–	V (AF033862 ^b ; protein id: AAC05085)	C	
Lon3	–	924	841	83	III (AC012393; protein id: AAF26080)	C	

^a EST. ^b The coding sequence spans nucleotides 67,179 through 61,860.

2000). Reported biochemical activities of FtsH protease include degradation of unassembled or oxidatively damaged membrane proteins (Ostersetzer and Adam, 1997; Lindahl et al., 2000). Four additional homologs, FtsH5 through 8, are predicted as chloroplastic, whereas FtsH3 is predicted to localize in M. The location of FtsH4 cannot be predicted with available programs, although a phylogenetic analysis (O. Ostersetzer and Z. Adam, unpublished data) shows that its closest homolog is one of the three mitochondrial homologs found in yeast (Leonhard et al., 1996). FtsH9 is a partial cDNA whose localization cannot be predicted from the available sequence, although this may be resolved by the eventual identification of its genomic sequence. It should be noted that additional FtsH-like sequences can be found among Arabidopsis genomic sequences, but the lack of a catalytic zinc-binding domain currently excludes them from our nomenclature.

DegP Protease

Thirteen different genes encoding proteins related to DegP, all containing the catalytic triad His-Asp-Ser, are present in the Arabidopsis genome (Table V). Although some of them are more similar to one *E. coli* protein than to the other (i.e. DegP, DegQ, or DegS), they differ in the number of PDZ domains. Because of this discrepancy we chose to designate all plant members of this protease family as DegP.

DegP1 and -2 are located in C thylakoid membranes (Itzhaki et al., 1998; Peltier et al., 2000; Haus-sühl et al., 2001). Other members of the family are likely to be expressed because they are represented among Arabidopsis ESTs, but their predicted locations vary from Cs (DegP5 and -8) to M (DegP3, -4, and -10⁻¹²), nuclei (DegP9), cytoplasm (DegP7), and the endoplasmic reticulum or plasma membrane (DegP13).

Table IV. Proposed nomenclature for FtsH proteases from Arabidopsis

Predicted locations are in italics.							
Protein Name	cDNA Accession No.	Precursor	Mature Protein	Transit Peptide	Chromosomal Location and Accession No.	Cellular Location	Reference
<i>amino acids</i>							
FtsH1	X99808	709	661	48	I (AC007980; protein id: AAD50055)	C	Lindahl et al. (1996)
FtsH2	AF135189 ^a	695	648	47	II (AC004669; protein id: AAC20729)	C	Chen et al. (2000)
FtsH3	AV550192 ^b	807	753	54	II (AC005315; protein id: AAC33234)	M	
FtsH4	AI994723 ^b	627	–	–	II (AC004747; protein id: AAC31223)	–	
FtsH5	AV521115 ^b	761	703	58	V (AB023032) ^c	C	
FtsH6		687	612	75	V (AL353993; protein id: CAB89335)	C	
FtsH7	BE523403 ^b	802	747	55	III (AL133292; protein id: CAB61952)	C	
FtsH8	AI995330 ^b	662	625	37	I (AC007592; protein id: AAF24819)	C	
FtsH9	Y12780 ^d	–	–	–	–	–	

^a Designated also VAR2. ^b EST. ^c The coding sequence spans nucleotides 41,203 through 43,350. ^d Partial cDNA.

Table V. Proposed nomenclature for DegP proteases from *Arabidopsis*

Protein Name	cDNA Accession No.	Precursor	Mature Protein	Transit Peptide	Chromosomal Location and Accession No.	Cellular Location ^a	Reference
			<i>amino acids</i>				
DegP1	AF028842	437	334	103	III (AP001302/AP000371 ^b ; protein id: BAB00408)	C (thylakoids)	Itzhaki et al. (1998)
DegP2	AF245171 ^c	607	538	69	II (AC005309; protein id: AAC63648)	C (thylakoids)	Haussühl et al. (2001)
DegP3	–	559	511	48	I (AC007234; protein id: AAF23847)	<i>M (matrix)</i>	
DegP4	–	518	492	26	I (AC007234; protein id: AAF23846)	<i>M (membranes)</i>	
DegP5	AF114386 ^c	321	309	12	IV (AL021710; protein id: CAA16717)	<i>C (stroma)</i>	
DegP6	–	252 ^d	207 ^d	45	I (AC006085; protein id: AAD30628)	<i>M/C</i>	
DegP7	AV530177 ^e	627	–	–	III (AC009895; protein id: AAF01594)	<i>Cy</i>	
DegP8	AI992881 ^e	448	422	26	V (AB024023; protein id: BAA98101)	<i>C (stroma)</i>	
DegP9	AV560841 ^e	592	–	–	V (AB010699; protein id: BAB10901)	<i>N</i>	
DegP10	AA713240 ^e	586	564	22	V (AB016877; protein id: BAB11638)	<i>M (matrix)</i>	
DegP11	–	565	500	65	III (AP000373; protein id: BAB01153)	<i>M (membranes)</i>	
DegP12	–	499	475	24	III (AP000373; protein id: BAB01154)	<i>M (matix)</i>	
DegP13	–	486	–	–	V (AB009052; protein id: BAB08526)	<i>ER/PM</i>	

^a Cy, Cytoplasm; ER, endoplasmic reticulum; N, nucleus; PM, plasma membrane. Predicted locations are in italics. ^b The genomic sequence is divided between two bacterial artificial chromosomes. ^c Genomic sequence does not match cDNA sequence due to mistakes in exon prediction. ^d Partial N-terminal sequence; does not contain the region where the catalytic Ser should be. ^e EST.

Another predicted difference within the plant DegP family is the apparent lack of PDZ domains in some of the isomers (i.e. DegP5, -6, -8, -9, -11, and -13). As PDZ domains mediate not only protein-protein interaction, but also influence the catalytic properties of the enzyme, the presence or absence of these domains in the different isomers is an important feature. DegP1–4, -7, -10, and -12 are all predicted to contain a single PDZ domain at their carboxy termini. Thus, the absence of the PDZ domain might have implications on the function of these proteins. It is also interesting to note that *E. coli* DegP contains two such domains, whereas none of the plant isomers have more than one. However, the functional significance of this difference is not yet clear.

CONCLUDING REMARKS

Up until a few years ago the identity of chloroplastic and mitochondrial proteases was a mystery that defied resolution by existing experimental approaches. In recent years, however, the advent of broad-scale EST/genomic sequencing and functional genomics have revealed the identity and structural characteristics of several distinct families of proteases in higher plants. Also revealed have been the many different isomeric forms that exist in plants for each of these protease families. To date, our understanding of the different protease families remains in its infancy, especially in terms of their molecular structure, substrate specificity, expression patterns, regulation of activity, and physiological roles. Such characterizations are further complicated by the numerous isomers that exist for each of the protease families. For example, it remains unclear whether isomers of the same protease have overlapping activities, and thus a degree of redundancy, or instead

have distinct properties such as different substrate specificity. As a consequence, the future challenge to elucidate the roles of proteases in the biogenesis and functions of Cs and M will require taking into consideration the multiple isomers of each.

To address the complexity of C and M proteases and facilitate scientific exchange, a standard nomenclature is urgently required such as the one proposed here derived from the analysis of *Arabidopsis* sequence data. Our nomenclature has been based on that used for the well-characterized *E. coli* proteases and has been adapted for the numerous predicted homologs in the nearly completed *Arabidopsis* genome. We propose that this nomenclature be used as a foundation for defining homologous proteases in other plant species, with extra numbers used if additional homologs than those in *Arabidopsis* are later discovered.

MATERIALS AND METHODS

DNA Sequence Analysis

Four families of proteases were analyzed, three of which are ATP-dependent (Clp, Lon, and FtsH) and one is ATP-independent (DegP-like). Peptidases were excluded from this analysis due to their large number and the relative lack of functional information currently available. Identification of family members was done using the BLAST program (Altschul et al., 1997). Protein sequences of members of each family were compared with six-frame translations of DNA sequences available in GenBank (as of October 15, 2000). Although in some cases there were more *Arabidopsis* homologous sequences than we present here, specific selection criteria were applied for inclusion in each of the families, as detailed in the "Results and Discussion." Protein sequences were further analyzed. Prediction of intracellular locations were performed using the programs Tar-

getP version 1.01 (Emanuelsson et al., 2000) and Predotar version 0.5 (<http://www.inra.fr/Internet/Produits/Predotar/>). In cases where contradictory predictions arose from TargetP and Predotar, we used the prediction with the highest probability value. Transit peptide-processing sites for C proteins were predicted by ChloroP version 1.01 (Emanuelsson et al., 1999). Each protein entry includes accession number to the corresponding full-length cDNA clone (if available) or EST clone (if available), predicted sizes of the precursor, mature and transit proteins, chromosomal location with its accession number and protein id number, proven or predicted cellular location, and a reference to the relevant publications. For proteins deduced from genomic sequences that are not fully annotated, and hence do not have protein id number, the nucleotides that span the coding sequence are indicated.

Received August 24, 2000; returned for revision October 23, 2000; accepted November 21, 2000.

LITERATURE CITED

- Adam Z** (1996) Protein stability and degradation in chloroplasts. *Plant Mol Biol* **32**: 773–783
- Adam Z** (2000a) Chloroplast proteases and their role in photosynthesis regulation. In E-M Aro, B Andersson, eds, *Regulatory Aspects of Photosynthesis*. Kluwer Academic Publishers, Dordrecht, The Netherlands
- Adam Z** (2000b) Chloroplast proteases: possible regulators of gene expression? *Biochimie* **82**: 647–654
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zheng Z, Miller W, Lipman DJ** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402
- Andersson B, Aro E-M** (1997) Proteolytic activities and proteases of plant chloroplasts. *Physiol Plant* **100**: 780–793
- Chen M, Choi Y, Voytas DF, Rodermel S** (2000) Mutations in the *Arabidopsis* VAR2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. *Plant J* **22**: 303–313
- Clarke AK** (1999) ATP-dependent Clp proteases in photosynthetic organisms: a cut above the rest! *Ann Bot* **83**: 593–599
- Desimone M, Weiss-Wichert W, Wagner E, Altenfeld U, Johanningmeier U** (1997) Immunochemical studies on the Clp-protease in chloroplasts: evidence for the formation of a ClpC/P complex. *Bot Acta* **110**: 234–239
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G** (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* **300**: 1005–1016
- Emanuelsson O, Nielsen H, von Heijne G** (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* **8**: 978–84
- Gottesman S** (1996) Proteases and their targets in *Escherichia coli*. *Annu Rev Genet* **30**: 465–506
- Halperin T, Adam Z** (1996) Degradation of mistargeted OEE33 in the chloroplast stroma. *Plant Mol Biol* **30**: 925–933
- Halperin T, Ostersetzer O, Adam Z** (2001a) ATP-dependent association between subunits of Clp protease in pea chloroplasts. *Planta* (in press)
- Halperin T, Zheng B, Itzhaki H, Clarke AK, Adam Z** (2000b) Plant mitochondria contain proteolytic and regulatory subunits of the ATP-dependent Clp protease. *Plant Mol Biol* **45**: 461–468
- Hausühl K, Andersson B, Adamska I** (2001) A chloroplast DegP2 protease performs the primary cleavage of the photodamaged D1 protein in plant photosystem II. *EMBO J* **20**: 713–722
- Herrmann RG** (1996) Photosynthetic research: aspects and perspectives. In B Andersson, A Salter, J Barner, eds, *Frontiers in Molecular Biology: Molecular Genetics in Photosynthesis*. Oxford University Press, Oxford
- Itzhaki H, Naveh L, Lindahl M, Cook M, Adam Z** (1998) Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane. *J Biol Chem* **273**: 7094–7098
- Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K** (1993) Characterization of cDNA for a dehydration-inducible gene that encodes a CLP A, B-like protein in *Arabidopsis thaliana* L. *Biochem Biophys Res Commun* **196**: 1214–1220
- Leonhard K, Herrmann JM, Stuart RA, Mannhaupt G, Neupert W, Langer T** (1996) AAA proteases with catalytic sites on opposite membrane surfaces comprise a proteolytic system for the ATP-dependent degradation of inner membrane proteins in mitochondria. *EMBO J* **15**: 4218–4229
- Lindahl M, Spetea C, Hundal T, Oppenheim AB, Adam Z, Andersson B** (2000) The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein. *Plant Cell* **12**: 419–431
- Lindahl M, Tabak S, Cseke L, Pichersky E, Andersson B, Adam Z** (1996) Identification, characterization, and molecular cloning of a homologue of the bacterial FtsH protease in chloroplasts of higher plants. *J Biol Chem* **271**: 29329–29334
- Lohman K, Gan S, John M, Amasino RM** (1994) Molecular analysis of natural leaf senescence in *Arabidopsis thaliana*. *Physiol Plant* **92**: 322–328
- Moore T, Keegstra K** (1993) Characterization of a cDNA clone encoding a chloroplast-targeted Clp homologue. *Plant Mol Biol* **21**: 525–537
- Nakabayashi K, Ito M, Kiosue T, Shinozaki K, Watanabe A** (1999) Identification of clp genes expressed in senescing *Arabidopsis* leaves. *Plant Cell Physiol* **40**: 504–514
- Nakashima K, Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K** (1997) A nuclear gene, *erd1*, encoding a chloroplast-targeted Clp protease regulatory subunit homologue is not only induced by water stress but also developmentally up-regulated during senescence in *Arabidopsis thaliana*. *Plant J* **12**: 851–861
- Ostersetzer O, Adam Z** (1997) Light-stimulated degradation of an unassembled Rieske FeS protein by a

- thylakoid-bound protease: the possible role of the FtsH protease. *Plant Cell* **9**: 957–965
- Pallen MJ, Wren BW** (1997) The HtrA family of serine proteases. *Mol Microbiol* **26**: 209–221
- Peltier JB, Friso G, Kalume DE, Roepstorff P, Nilsson F, Adamska I, van Wijk KJ** (2000) Proteomics of the chloroplast: systematic identification and targeting analysis of luminal and peripheral thylakoid proteins. *Plant Cell* **12**: 319–342
- Ponting CP** (1997) Evidence for PDZ domains in bacteria, yeast, and plants. *Protein Sci* **6**: 464–468
- Porankiewicz J, Wang J, Clarke AK** (1999) New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond. *Mol Microbiol* **32**: 449–458
- Sarria R, Lyznik A, Vallejos CE, Mackenzie SA** (1998) A cytoplasmic male sterility-associated mitochondrial peptide in common bean is post-translationally regulated. *Plant Cell* **10**: 1217–1228
- Schrimer EC, Glover JR, Singer MA, Lindquist S** (1996) HSP100/Clp proteins: a common mechanism explains diverse functions. *TIBS* **21**: 289–295
- Shanklin J, Dewitt ND, Flanagan JM** (1995) The stroma of higher plant plastids contain ClpP and ClpC, functional homologs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease. *Plant Cell* **7**: 1713–1722
- Skorko-Glonek J, Lipinska B, Krzewski K, Zolese G, Bertoli E, Tanfani F** (1997) HtrA heat shock protease interacts with phospholipid membranes and undergoes conformational changes. *J Biol Chem* **272**(14): 8974–8982
- Sokolenko A, Lerbs-Mache S, Altschmied L, Herrmann RG** (1998) Clp protease complexes and their diversity in chloroplasts. *Planta* **207**: 286–295
- Suzuki CK, Suda K, Wang N, Schatz G** (1994) Requirement for the yeast gene *lon* in intramitochondrial proteolysis and maintenance of respiration. *Science* **264**: 273–276
- van Dyck L, Pearce DA, Sherman F** (1994) *PIM1* encodes a mitochondrial ATP-dependent protease that is required for mitochondrial function in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* **269**: 238–242