Presequence Protease (PreP), a novel Peptidasome in Mitochondria and Chloroplasts: Localization, Function, Structure and Mechanism of Proteolysis

Shashi Bhushan

Department of Biochemistry and Biophysics
Arrhenius Laboratories for Natural Sciences
Stockholm University
S-10691 Stockholm
Sweden
Stockholm, 2007
Cover picture: Dual targeting of the PreP1:GFP fusion protein to mitochondria and chloroplasts (Paper I) and a proposed mechanism for the PreP peptidosome substrate binding, proteolysis and release (Paper III).
- To my family
CONTENTS

Abstract 7
List of publications 8
Abbreviations 10
The mitochondria and chloroplasts 11
   The origin and evolution 11
   Structure and function 12
   Genome 13
   Proteome 15
Organellar import machineries 16
   Mitochondrial protein import machinery 16
      Mitochondrial targeting peptides - the presequences 18
      Cytosolic factors 18
      Translocase of the outer membrane (TOM) 19
         Import receptors 19
         The general import pore 20
      Translocase of the inner membrane (TIM) 21
         TIM23 complex 21
         TIM22 complex 23
      Protein Import to the intermembrane space 23
      Processing peptidases 24
   Chloroplastic protein import machinery 26
      Chloroplastic targeting peptides - the transit peptides 26
      Cytosolic factors 27
      Translocase of the outer envelope membrane (TOC) 28
      Translocase of the inner envelope membrane (TIC) 30
      Stromal processing peptidase 31
Dual targeting to mitochondria and chloroplasts 31
   Studying dual targeting 32
   Mechanisms of dual targeting 33
Proteolytic system in mitochondria 34
   ATP-dependent proteases 34
      The FtsH (AAA) protease 34
      Lon-like protease 35
      Clp-like protease 36
   ATP-independent proteases 37
      Oligopeptidases 37
      OMA protease 38
      Rhomboid protease 38
Proteolytic system in chloroplasts 39
   ATP-dependent proteases 39
      The FtsH (AAA) protease 39
      Clp-like protease 40
      Lon-like protease 41
   ATP-independent proteases 41
      DegP-like protease 42
      SppA-like protease 42
<table>
<thead>
<tr>
<th>The PresequenceProtease, PreP</th>
<th>43</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual targeting of PreP in mitochondria and chloroplasts</td>
<td>44</td>
</tr>
<tr>
<td>Function of PreP in mitochondria and chloroplasts</td>
<td>45</td>
</tr>
<tr>
<td>Expression of the AtPreP1 and AtPreP2 in A. thaliana plants</td>
<td>47</td>
</tr>
<tr>
<td>Crystal structure of AtPreP1, a Peptidasome</td>
<td>47</td>
</tr>
<tr>
<td>Mechanism of proteolysis by PreP Peptidasome</td>
<td>48</td>
</tr>
<tr>
<td>The role of the PreP peptidasome in the degradation of the amyloid β-peptide: A possible link to Alzheimer’s disease</td>
<td>50</td>
</tr>
<tr>
<td>Alzheimer’s disease (AD) and amyloid β-peptide (Aβ)</td>
<td>50</td>
</tr>
<tr>
<td>Amyloid β-peptide in mitochondria</td>
<td>51</td>
</tr>
<tr>
<td>PresequenceProtease in human - The hPreP</td>
<td>52</td>
</tr>
<tr>
<td>A novel function of hPreP in mitochondria: Aβ degradation</td>
<td>53</td>
</tr>
</tbody>
</table>

**Future perspectives** | 55 |
**Acknowledgements** | 57 |
**References** | 59 |
Abstract

Mitochondria and chloroplasts contain several thousand different proteins, almost all of which are synthesized in the cytosol as precursor proteins and imported into the correct organelle. The information for the organellar targeting and import generally resides in the N-terminal part of the protein, called a targeting peptide. The targeting peptide is cleaved off by the organellar processing peptidases after import of a precursor protein. Free targeting peptides generated inside the organelle after import are rapidly degraded by proteolysis as their accumulation can have severe effects on the functional and structural integrity of the organelle since they can penetrate membranes, induce channel formation in membranes, dissipate membrane potential and uncouple respiration. The aim of this thesis has been a thorough investigation of the newly identified targeting peptide degrading protease, the PresequenceProtease (PreP).

We have shown that the two isoforms of Arabidopsis thaliana PresequenceProteases (AtPreP1 and AtPreP1) are dually targeted and localized to both mitochondria and chloroplasts. Dual targeting of the AtPreP1 is due to an ambiguous targeting peptide with a domain organization for mitochondrial and chloroplastic targeting. Both the AtPreP1 and AtPreP2 are expressed in A. thaliana plants in an organ specific manner and they have distinct but overlapping substrate specificity for efficient degradation of a wide variety of peptides.

The crystal structure of the recombinant AtPreP1 E80Q was solved at 2.1 Å resolution. The structure represents the first substrate bound, closed conformation of a protease from the pitrilysin family. The PreP polypeptide folds in a unique peptidasome structure, surrounding a huge cavity of more than 10 000 Å³ in which the active site resides. Cysteine mutants of AtPreP1 designed for locking the PreP in a closed conformation showed no proteolytic activity when disulfide bonds were allowed to form, while activity was normal in absence of disulfide bonds. A novel mechanism for proteolysis is proposed involving hinge-bending motions that cause the PreP protease to open and close in response to substrate binding.

PreP is localized to the mitochondrial matrix in human mitochondria where, beside degradation of targeting peptides, it has a novel function: degradation of amyloid β-peptide (Aβ). Immunoinactivation of PreP in human brain mitochondria resulted in complete loss of the proteolytic activity against Aβ-peptide, showing that under circumstances when Aβ is present in mitochondria, human PreP is the protease responsible for degradation of this toxic peptide. These findings contribute to studies of the mitochondrial component in Alzheimer’s disease.
List of publications included in the thesis


* Both authors have contributed equally to this work.


Additional publications


Abbreviations

AAA    ATPases associated with a number of cellular activities
AAC    ADP/ATP translocator
Aβ     Amyloid β-peptide
ABAD   Amyloid β-binding alcohol dehydrogenase
ABC    ATP-binding cassette
AD     Alzheimer’s disease
AIP    Aryl hydrocarbon receptor interacting protein
APP    Amyloid precursor protein
Clp    Caseinolytic protease
GIP    General import pore
GFP    Green fluorescent protein
GST    Glutathione-S-transferase
Hsp70  Heat shock protein 70
IDE    Insulin degrading enzyme
IM     Inner membrane
IMP    Inner membrane peptidase
IMS    Intermembrane space
LC     Liquid chromatography
MIP    Mitochondrial intermediate peptidase
MOP    Mitochondrial oligopeptidase
MPP    Mitochondrial processing peptidase
MS     Mass spectrometry
MSF    Mitochondrial import stimulating factor
NEM    N-ethylmaleimide
OM     Outer membrane
Oma    Overlapping activity with m-AAA proteases
PBF    Presequence binding factor
PiC    Phosphate carrier
Pim1   Protease in mitochondria 1
PreP   PresequenceProtease
RISP   Rieske Fe-S protein of the cytochrome b̂f complex
SPP    Stromal processing peptidase
TIC    Translocase of the inner envelope membrane of chloroplast
TIM    Translocase of the inner membrane of mitochondria
TM     Transmembrane
TOC    Translocase of the outer envelope membrane of chloroplast
TOM    Translocase of the outer membrane of mitochondria
Yme1   Yeast mitochondrial escape 1
Yta    Yeast tat-binding-like proteins
Zn-MP  Zinc metalloprotease
Δψ     Membrane potential
ΔpH    Proton gradient
The mitochondria and chloroplast

The name mitochondrion arises from Greek words mitos = thread and chondrion = granule. Mitochondria were originally identified as the site of oxidative energy metabolism (Kennedy and Lehninger, 1950). Mitochondria are also the host for enzymes of the Krebs cycle and β-oxidation of fatty acids. In today’s world mitochondria are known not only as the “power station” of the cell, but also for playing a vital role in the transmission of extra- and intracellular signals that activate reaction cascades leading to cellular senescence and programmed cell death (PCD) (Wang, 2001). The discovery of a number of human diseases associated with mitochondrial dysfunctions once again brought mitochondria into the spotlight of biological research.

The name chloroplast is derived from the Greek words chloros = green and plast = form or entity. Chloroplasts are members of a class of plant cell organelles known as plastids that all originate from proplastids. During plant development the proplastids differentiate to form three major groups of plastids, the green chloroplasts, the colored chromoplasts and the colorless leucoplasts. The most abundant and important plastids are the chloroplasts. Chloroplasts harvest energy from sunlight to split water and fix carbon dioxide to produce sugars. This process called photosynthesis also converts harvested solar energy into a conserved form of energy: ATP and NADPH through a complex set of processes.

The origin and evolution

Mitochondria and chloroplasts are not synthesized de novo, but originate from pre-existing organelles by partition in a fission process. The most accepted theory about the origin of these two organelles is the endosymbiotic theory proposed by Lynn Margulis in 1970 (Margulis, 1970). According to this theory mitochondria arose from aerobic prokaryotes that began to live in symbiosis with a primitive, anaerobic eukaryotic cell. The endosymbiotic theory is supported by many facts such as e.g. 1) mitochondria and chloroplasts contain their own genome and divides independently from the cell where they reside, 2) in most organisms mitochondrial and chloroplastic DNA is circular, just like in bacteria, 3) transcriptional and translational machineries of these organelles are very similar to those found in bacteria and 4) the mitochondrial inner membrane (IM) contains certain lipids, such as cardiolipins, only found in the mitochondrial IM and in the plasma membrane of bacteria. Based on the genome sequences from a number of different organisms, phylogenetic reconstitutions of mitochondria have been
made. These analyses show that mitochondrial genome sequences are the descendents of α-proteobacterial homologues (Lang et al., 1999). The closest known ancestor to mitochondria is Rickettsia prowazekii, an obligate intracellular parasite (Andersson et al., 1998, 2003).

Phylogenetic data suggests that the chloroplasts were engulfed after mitochondria in an endosymbiotic event. In a first primary endosymbiosis event about 1.5 billion years ago an ancient cyanobacterium was engulfed by a mitochondria containing eukaryote. The plastids formed in this way are surrounded by two membranes and are found in land plants and in red and green algae (Gray, 1999). However, a growing body of evidence indicates that the chloroplasts of some algae have not been derived by engulfing cyanobacteria in a primary endosymbiosis like those discussed above, but by engulfing photosynthetic eukaryotes. This is called secondary endosymbiosis and these plastids are called secondary plastids characterized by the presence of three or four surrounding membranes. Secondary plastids are found in lineages such as apicomplexa, dinoflagellates and ciliates (as reviewed by Stoebe and Maier, 2002).

**Structure and function**

Mitochondria are present in virtually all eukaryotic cells. They are typically 0.5-1.8 µm wide and 1-2 µm long in size. The number and distribution of mitochondria are dependent on the metabolic activity of the cell. Mitochondria are distinct organelles surrounded by two membranes. The two membranes divide the mitochondrion into two distinct compartments, the intermembrane space (IMS) and the mitochondrial matrix. Electron tomography has provided the three-dimensional (3D) structure of mitochondria and gives new insight to the internal organization (Frey and Mannella, 2000). The outer membrane (OM) is smooth and permeable to ions and molecules smaller than 10 kDa. The inner membrane (IM) is impermeable and highly convoluted, forming folds called cristae. Cristae are connected to the IM by narrow tubular segments, called cristae junctions. The narrowness of the cristae junctions has led to the hypothesis that the IM is further divided into distinct sub compartments (Frey and Mannella, 2000).

The mitochondrial matrix contains the enzymes of the tricarboxylic acid cycle (TCA) and fatty acid oxidation. The important function of mitochondria is to produce the ATP required by the cell. ATP production in mitochondria starts with the oxidative decarboxylation of pyruvate in the matrix to form NADH and acetyl-CoA. Acetyl-CoA enters the TCA cycle to
generate NADH and FADH$_2$ from NAD$^+$ and FAD. The electrons from NADH and FADH$_2$
are funnelled to oxygen through a chain of electron carriers in the respiratory chain. The
transfer of electrons between these carriers releases energy that is used to transfer protons
across the IM from matrix to the IMS side. The proton translocation generates an
electrochemical proton motive force consisting of a membrane potential ($\Delta\psi$) and a trans-
membrane proton gradient ($\Delta p\text{H}$) (Mitchell, 1974). The energy from the proton gradient is
utilized to synthesize ATP from ADP by the ATP synthase in mitochondria. ATP formed in
this way is exported from the mitochondria and provides the energy required for different
purposes in a cell.

Chloroplasts look like as flat discs usually 2 to 10 µm in diameter and 1 µm thick. Chloroplasts are surrounded by two membranes, the outer envelope (OE) membrane and the
inner envelope (IE) membrane. The OE membrane is permeable to small molecules up to 10
kDa, but impermeable to bigger molecules such as proteins and nucleic acids. The IE
membrane is impermeable to most molecules and those that are permeable can only cross this
membrane through specific translocases. The compartment between these two envelopes is
called the interenvelope space (IES). The soluble material within the chloroplast is called
stoma, corresponding evolutionary to the cytosol of a bacterium, and contains one or more
molecules of small circular DNA. Chloroplasts perform the very important function of
photosynthesis within plant cells and contain the chlorophyll molecules that are essential for
this process. All reactions of photosynthesis occur in this organelle including CO$_2$ fixation.
The chloroplasts use photosynthetic chlorophyll pigment and take in sunlight, water and
carbon dioxide to produce glucose and oxygen. An important structure in the chloroplasts is
the inter-connected, flattened, membranous sacs called thylakoids. There are many thylakoid
stacks in a chloroplast, providing a vast surface area within a compact volume for harvesting
light energy to drive photosynthesis. These structures are the site of the photosynthetic light
reactions.

**Genome**

Mitochondria and chloroplasts harbor their own genome in form of the circular DNA. The
mitochondrial and chloroplastic DNA exhibit a remarkable variation in terms of structure and
size as well as gene content and expression. Currently known mitochondrial genome sizes
range from 5 966 base pairs in *Plasmodium reichenowi* (malaria parasite) to 569 630 base
pairs in *Zea mays* (maize) (Conway et al., 2000; Clifton et al., 2004). Chloroplasts genome
sizes range from 110,000 base pairs to 160,000 base pairs, depending on the species (Cui et al., 2006). Mitochondrial DNA encodes a limited number of proteins and RNAs that are essential for the formation of functional mitochondria. Generally, mitochondrial DNA encoded proteins are core components of the respiratory chain complexes and the ATP synthase (Boore, 1999). Land plant chloroplast genomes typically contain around 110-120 unique genes. Some algae have retained a large chloroplast genome with more than 200 genes, while the plastid genomes from non-photosynthetic organisms have retained only a few dozen genes (Cui et al., 2006).

Over evolutionary time both mitochondria and chloroplasts have lost most of their genes and transitioned from being free living prokaryotes to organelles with key roles in eukaryotic cellular function. Two distinct modes of genetic loss are responsible for the reduced genome we now see in these two modern organelles (Berg and Kurland, 2000; Kurland and Andersson, 2000). First these organelles reside in a cell so they can import rather than synthesize a number of biomolecules from the host. Second, many genes were transferred to, and are now expressed, in the nuclear genome. The low gene content of mitochondrial DNA implies a rapid and extensive loss or transfer of genetic material during early stages of mitochondrial evolution (Grey et al., 1999). In plants, mitochondrial gene transfer can still be traced (Adams et al., 2000; Daley et al., 2002). The subunit 2 of cytochrome oxidase (CoxII) gene from legumes (Nugent and Palmer, 1991; Adams et al., 1999), the Rps12 gene from Oenothera (Grohmann et al., 1992), Rps10 and Rps19 genes from A. thaliana (Wischmann and Schuster, 1995; Sanchez et al., 1996) and the Rps11 gene from Oryza sativa (Kadowaki et al., 1996) have all been identified as recent gene transfers from the mitochondrial genome to the nuclear genome by comparing gene content in nuclear and mitochondrial DNA.

What are the evolutionary advantages driving gene transfer from mitochondria and chloroplasts to nucleus? One popular hypothesis is Muller’s rachet which suggests that the asexual reproduction of mitochondria and chloroplasts can lead to a faster accumulation of deleterious mutations (Kurland, 1992; Berg and Kurland, 2000). High concentrations of free radicals can be produced in mitochondria and chloroplasts because of the high redox activity of these organelles (Martin and Palumbi, 1992; Allen and Raven, 1996). Increased levels of oxygen free radicals have been shown to cause an increase in DNA mutation rates and therefore it would be advantageous to relocate genetic material from organelles to the nucleus. However, the genome of plant mitochondria tends to be less mutation prone than the nuclear
genome (Martin and Herrmann, 1998). Therefore, free radicals that lead to mutations in the organellar genome might not be the sole driving factor for the organellar gene transfer in plants.

Why are some genes still maintained in these organelles when most of them have been transferred to the nucleus? There is no straight answer to this question. A hydrophobicity hypothesis was proposed by von Heijne, who suggested that genes remaining in these organelles encode proteins that are too hydrophobic to be imported across membranes back into the organelles (von Heijne, 1986a). However, there are some very hydrophobic proteins that have been moved from these organellar genomes to the nucleus and are able to be imported back to these organelles (Gray et al., 1999). A second theory, CORR (Co-location for Redox Regulation) was proposed by Allen in 1992. According to this theory co-location of chloroplast and mitochondrial genes with their gene products is required for rapid and direct regulatory coupling. Redox control of gene expression is suggested as the common feature of those mitochondrial and chloroplastic proteins that are encoded in their own genome (Allen, 1992).

**Proteome**

Mitochondria contain about 800 to 2500 different proteins (Emanuelsson et al., 2000; Zhang et al., 2001). The availability of genome databases and recent advances in proteomics have enabled us to gain a better insight of the mitochondrial proteome. In recent years several proteomic studies of isolated mitochondria have been reported. The aim of these studies is to gain a better understanding of the role of mitochondria and its function. Taylor et al. (2003) have been able to identify 615 proteins of human heart mitochondria using proteomic approaches and identification by mass spectrometric (MS) analysis. 81% of the identified proteins were classified amongst protein families with identified functions, while the function of the remaining 19% proteins remains to be explored. In a separate study Da Cruz et al. (2003) have been able to identify 183 mitochondrial IM proteins from rat mitochondria using liquid chromatography (LC) directly coupled to MS analysis. In another study on isolated *Saccharomyces cerevisiae* mitochondria, Sickmann et al. (2003) were able to identify 750 proteins which they suggested comprise about 90% of the total mitochondrial proteome.

Plant mitochondria resemble mammalian and yeast mitochondria in many ways but still have some additional functions, such as an uncoupled bypass of the electron transport chain.
by the alternative oxidase (AOX) and the synthesis of lipids and vitamins (Rebeille et al., 1997; Bartoli et al., 2000; Gueguen et al., 2000). Heazlewood et al. (2004) have been able to identify 416 mitochondrial proteins from *A. thaliana* using a systematic LC-MS/MS approach. 407 proteins out of those 416 are nuclear encoded and the remaining 9 are encoded in the mitochondrial genome.

Chloroplasts contain about 3500 different proteins as estimated *in silico* using different prediction programs such as TargetP and ChloroP (Emanuelsson et al., 1999, 2000). Kleffmann et al. (2004) have used different fractionation techniques, followed by LC-ESI-MS/MS to identify 687 proteins in isolated *A. thaliana* chloroplasts. 70% of these identified proteins could be assigned to one or more known metabolic pathways, whereas the remaining 30% of the proteins were of unknown function. Surprisingly, 48% of the identified nuclear encoded proteins did not have a predicted targeting peptide when analysed using TargetP. In a separate study Friso et al. (2004) have been able to identify 198 proteins associated with thylakoid membranes of *A. thaliana*.

**Organellar import machineries**

**Mitochondrial protein import machinery**

The mitochondrial protein import machinery has been extensively studied in *S. cerevisiae* and *Neurospora crassa* using both biochemical and genetic approaches (as reviewed by Neupert and Herrmann, 2007). The plant mitochondrial import process has been studied in *Solanum tuberosum*, *Spinacia oleracea*, *Pisum sativum* and *A. thaliana* (as reviewed by Glaser and Whelan, 2007). The plant and yeast mitochondrial import machinery share several aspects, but differs in some respects.

Most of the mitochondrial proteins are nuclear encoded and synthesized in the cytosol as precursor proteins. Mitochondrial protein import generally requires a signal for import, cytosolic factors/chaperones, import receptors on the mitochondrial surface, an import pore, chaperones inside the mitochondria and processing peptidases for maturation (Figure 1). The functioning of the general mitochondrial import pathway and machinery can be briefly described as follows:

- Synthesis of the nuclear encoded precursor proteins in the cytosol.
-Interaction of the newly synthesized precursor proteins with cytosolic factors/chaperones that keep them in an unfolded and import competent conformation.

-Recognition of the precursor proteins by import receptors on the mitochondrial surface.

-Translocation of the precursor proteins across the mitochondrial OM via the general import pore.

Figure 1. Mitochondrial protein import machinery. General overview of the protein import pathway into the matrix. TOM and TIM refer to Translocases of the outer and inner membrane protein complexes in mitochondria. The numbers represent the molecular masses of the components of the TOM and TIM complexes. MPP, Mitochondrial processing peptidase; Hsp70, Heat shock protein 70; MGE and MDJ, mitochondrial co-chaperones, homologs of bacterial DnaJ and GrpE; α- and β-, subunits of MPP; PreP PresequenceProtease.

-Chaperone assisted passage through mitochondrial IMS.

-Interaction of the precursor proteins with import receptors on the mitochondrial IM.
- ATP and Δψ-dependent translocation of the precursor proteins across the mitochondrial IM.
- Proteolytic maturation of the precursor proteins by removal of the presequence by processing peptidases.
- Chaperone assisted folding and assembly of the mature proteins.
- Degradation/removal of the free, cleaved targeting peptides/presequences.

**Mitochondrial targeting peptides - the presequences**

Precursor proteins can be divided into two classes on the basis of their targeting mechanisms used for import. More than half of the precursor proteins carry a cleavable N-terminal extension known as the presequence or targeting peptide. Some integral membrane proteins, such as the metabolic carriers are synthesized without cleavable extensions. These precursor proteins contain internal targeting signals that are distributed throughout the entire length of the proteins. The presequence or targeting peptide carries all of the information required for protein targeting and import into the mitochondria. Most presequences are 20-50 amino acid residues in length. However, they can vary substantially, from 13 residues to 136 residues (Zhang et al., 2001). There is no consensus at the primary structure level among presequences, only a very loosely conserved motif has been found around the processing site (von Heijne et al., 1989; Zhang et al., 2001). Presequences are enriched in basic and hydrophobic residues and are generally deficient in acidic residues (von Heijne et al., 1989). Plant presequences are about 7-9 amino acid residues longer and contain about 2-5 times more serine residues than non-plant presequences (Glaser, et al., 1998). Presequences are known to adopt a positively charged amphiphilic α-helical conformation in membrane mimicking environments, while they are largely unstructured in aqueous solutions (von Heijne, 1986b; Moberg et al., 2004).

**Cytosolic factors**

Mitochondrial protein import is believed to occur in a post-translational manner. This assumption is based on the fact that in vitro synthesized precursor proteins can be imported into isolated mitochondria (Hallermayer et al., 1977; Wienhues et al., 1991). However, cotranslational protein import into mitochondria can not be ruled out in vivo (Suissa and Schatz, 1982; Furuya et al., 1991; Verner, 1993). A genome-wide analysis of mRNA encoding mitochondrial proteins showed that some of the mRNA was closely associated with mitochondrially bound polyribosomes (Marc et al., 2002). Interestingly, genes producing
mRNA that are attached to mitochondria were mainly of ancient bacterial origin, while those producing mRNA that is translated in the cytoplasm were mainly of eukaryotic origin. The 3’UTR of mRNA that were attached to mitochondria carries information required for their targeting and attachments to mitochondria (Corral-Debrinski et al., 2000; Marc et al., 2002).

A majority of precursor proteins are synthesized in the cytosol at some distance from mitochondria and must be targeted through the cytosol to mitochondria. Precursor proteins are also imported into mitochondria in an extended or import competent form, different than their final native conformation. Because of these requirements precursor proteins in the cytosol are kept as complexes with chaperones or other factors that are believed to stabilize them, as they are not in their right conformation and therefore are prone to aggregation and degradation. Many of these factors have been described such as e.g. the aryl hydrocarbon receptor interacting protein (AIP), presequence binding factor (PBF), mitochondrial import stimulating factor (MSF) and cytosolic chaperones Hsp70 and Hsp90 (Mihara and Omura, 1996; Yano et al., 2003; Young et al., 2003).

**Translocase of the outer membrane (TOM)**
The first step of recognition and subsequent translocation of precursor proteins across the OM is accomplished by the multi subunit TOM complex. Virtually all mitochondrial proteins are translocated across the OM by the TOM complex. The TOM complex was purified from yeast as a ~ 490 kDa complex and contained TOM70, TOM40, TOM22, TOM20, TOM7, TOM6 and TOM5 (Ahting et al., 1999). The TOM complex acts as a receptor for the recognition of mitochondrial precursor proteins synthesized in the cytosol and subsequent transfer through import pores across the OM.

**Import Receptors**
There are three main receptors in *S. cerevisiae* mitochondria for precursor protein recognition on the mitochondrial surface. These are TOM70, TOM20 and TOM22, named according to their apparent molecular mass (Hines et al., 1990; Kunkele et al., 1998; Brix et al., 1999; Abe et al., 2000). Each of these receptors is anchored in the OM by a single transmembrane (TM) segment. TOM20 and TOM70 are anchored to the membrane via their N-terminus, while the C-terminal domain is exposed to the cytosol. TOM22 has an inverted membrane topology than TOM20 and TOM70 and also contains an additional small C-terminal domain protruding into the IMS (Neupert, 1997). Precursor proteins with a presequence are first recognized by
TOM20 and subsequently transferred to TOM22. TOM22 also plays a critical role for the general integrity of the TOM complex (van Wilpe et al., 1999). The NMR structure of the cytosolic domain of rat TOM20 complexed with a peptide derived from the aldehyde dehydrogenase (ALDH) presequence revealed that the presequence forms an amphiphilic α-helix when bound to TOM20 (Abe et al., 2000) and their interaction is hydrophobic in nature. Plant TOM20 differs from S. cerevisiae and animal TOM20 in its topology as it is anchored to the membrane via the C-terminal domain (Macasev et al., 2000). TOM22 in plants is also different to animal and S. cerevisiae TOM22. The plant TOM22 homologue lacks the cytosolic acidic domain and is about 9 kDa in size. It has been suggested that the difference is due to the unique environment in plants where both mitochondria and chloroplasts are present and it is possible that TOM22 in plants provides targeting specificity (Macasev et al., 2004).

TOM70 is the main receptor for precursor proteins that contain internal targeting signals, such as the ADP/ATP translocator (AAC) and Phosphate carrier (Pic) of the carrier family. TOM70 is not only an import receptor but can also act as docking platform for cytosolic chaperones (Young et al., 2003). Isolated TOM complexes from plants lack TOM70 and no homologue has been found in the A. thaliana genome. The absence of TOM70 in plants is puzzling since the carrier import pathway has been demonstrated (Lister et al., 2002).

**The general import pore**

After recognition and binding to the receptors, the precursor proteins are inserted into the general import pore. The general import pore of the TOM complex is composed of the TOM40 and three small TOM proteins TOM5, TOM6 and TOM7, through which all precursor proteins cross the OM (Ahting et al., 2001). TOM40 is an integral membrane protein that is believed to form a β-barrel structure. Purified TOM40 forms a cation-selective channel of about 22 Å when inserted in artificial membranes (Hill, et al., 1998; Kunkele et al., 1998; Becker et al., 2005). The pore diameter is large enough to accommodate an α-helical peptide or even a protein loop. TOM40 is the only TOM protein that is essential for cell viability in yeast under all growth conditions (Baker et al., 1990). The loss of individual small TOM proteins does not lead to any major effects, but the simultaneous deletion of all three small TOM is lethal (Sherman et al., 2005). The small TOM6 and TOM7 subunits do not interact with precursor proteins during protein import, rather they modulate the stability of the TOM complex (Alconada et al., 1995; Honlinger et al., 1996). TOM6 has been proposed to support the cooperation between the TOM22 receptor and the general import pore (Alconada et al., 1995; Dekker et al., 1998; van Wilpe et al., 1999).
According to the “acid chain hypothesis” electrostatic interactions are the main driving forces behind the unidirectional protein translocation across the OM (Komiya et al., 1998). Precursors with a presequence successively interact with at least five different TOM subunits (TOM20, TOM22, TOM5, TOM40, and the IMS domain of TOM22) during translocation across the OM. Most of these TOM subunits contain negatively charged patches and it has been proposed that the positively charged presequence is recognized by increasing affinity along the import pathway. However, a report by Muto et al. (2001) showed that the presequence interacts with TOM20 via hydrophobic patches. This suggests that forces other than ionic interaction are important for the interaction of the presequence with TOM subunits and has led to a revised theory named the “binding chain hypothesis”.

Outer mitochondrial membrane proteins are first imported through TOM complex and later inserted into the outer membrane using SAM (Sorting and Assembly Machinery) complex (as revived by van der Lann, et al., 2005).

**Translocase of the inner membrane (TIM)**

After crossing the OM with the TOM complex, precursor proteins interact with either of two TIM complexes in the IM, the TIM23 complex or the TIM22 complex. Matrix targeted or presequence carrying precursor proteins are recognized and translocated by TIM23, while polytopic IM proteins are inserted into the IM by the TIM22 complex. IM also contains the OXA1 complex that mediates insertion of precursor proteins from the matrix side into the IM (Hell et al., 1998; Jensen and Dunn, 2002).

**TIM23 complex**

The TIM23 complex is the main precursor protein translocase in the IM of mitochondria. The TIM23 complex is responsible for translocating all precursors of matrix proteins, most inner membrane proteins and many of the IMS proteins. Translocation by TIM23 requires an electrical membrane potential ($\Delta\psi$) across the IM and the hydrolysis of ATP. The TIM23 complex is composed of two parts, the protein conducting channel and the protein import motor or presequence translocase-associated motor (PAM). The protein conducting channel consists of TIM50, TIM23, TIM17 and TIM21 (Bauer at al., 1996; Dekker et al., 1997; Chacinska et al., 2005). The core of the translocase is formed by TIM17 and TIM23 with a molecular mass of 90 kDa (Dekker et al., 1997). TIM23 and TIM17 are integral membrane
proteins with four TM segments. TIM23 and TIM17 are believed to form a pore through which proteins are translocated into the matrix. TIM23 also exposes an N-terminal hydrophilic region to the IMS. This part has been proposed to place the TIM23 complex in the proximity to the OM (Donzeau et al., 2000). TIM50 is anchored into the IM by an N-terminal TM segment. TIM50 interacts with incoming proteins as they come out from the TOM complex and can pass them to other subunits of the TIM23 complex. In this way TIM50 acts as a receptor for the TIM23 complex. TIM50 has also been proposed to have a role in regulation of import channel’s permeability (Meinecke et al., 2006). TIM21 was found to be directly interacting with the IMS domain of TOM22, suggesting a direct interaction between TIM23 and the TOM complex (Chacinska et al., 2005; Mokranjac et al., 2005).

Most of the presequence carrying precursor proteins are imported into the matrix by the combined action of the TIM23 protein conducting channel and the protein import motor. The protein conducting channel can only transfer the presequence part of the precursors, which requires $\Delta \psi$. After the presequence emerges from the TIM23 pore, PAM has to take over. For a long time it was thought that the ATP-dependent import motor consists of three proteins, the peripheral inner membrane protein TIM44, the mitochondrial chaperone Hsp70 and the nucleotide exchange factor Mge1. Recently two new essential co-chaperones have been identified, Pam18 and Pam16 (Li et al., 2004; van der Laan et al., 2005). TIM44 recruits Hsp70 in its ATP bound form, which then immediately can grasp the incoming unfolded polypeptide as its substrate binding site is open (Schneider et al., 1994). After binding to the emerging precursor, ATP is hydrolyzed, the substrate binding site closes and Hsp70 is released. This release of Hsp70 requires the nucleotide exchange factor Mge1. Mge1 removes the bound nucleotide and allows cycling of the ATP bound Hsp70 to the PAM (Schneider et al., 1994).

Two mechanisms for the PAM have been proposed and evidence suggested that the mechanisms co-operate in translocating the precursor proteins across the IM. The Brownian ratchet mechanism (Neupert and Brunner, 2002) suggests that the precursors are prevented from sliding back upon binding to the Hsp70 in the matrix. The matrix-bound Hsp70 biases spontaneous oscillations of the incoming polypeptide chain toward the matrix, and makes new Hsp70 binding sites accessible. Thus, by successive binding of Hsp70, the precursor protein is trapped into the matrix. In the pulling mechanism (Matouschek et al., 1997) Hsp70 plays an active role in the translocation. Upon ATP hydrolysis a conformational change of Hsp70 pulls
the polypeptide into the matrix. By using a model protein, Huang et al. (2002) showed that simple trapping of precursor protein segments by Hsp70 was enough to import loosely folded precursor proteins, while partially folded precursors also required more efficient Hsp70-TIM44 cycling, suggesting that pulling is needed.

**TIM22 complex**

The TIM22 complex is required for the import and insertion of the carrier proteins and of the hydrophobic TIM proteins (Kerscher et al., 1997; Kurz et al., 1999). Import via the TIM22 complex only requires membrane potential (\(\Delta\psi\)) and not ATP (Jensen and Dun, 2002). The TIM22 complex is about 300 kDa in size and consists of the three integral membrane proteins TIM22, TIM54 and TIM18 (Sirrenberg et al., 1996). TIM22 is a homologue of the TIM23 and TIM17 and is the only essential protein of the TIM22 complex (Kovermann et al., 2002). TIM22 forms the essential core of the TIM22 complex that can mediate the insertion of carrier proteins without TIM54 and TIM18 (Kovermann et al., 2002). The precise role of TIM54 and TIM18 is not known, although both of these TIM proteins are required for the formation and stability of the TIM22 complex.

A complex of the small TIM subunits comprising of the TIM8, TIM9, TIM10, TIM12 and TIM13 in the IMS also interacts with the TIM22 complex. These complexes are proposed to act as chaperones by transporting the hydrophobic IM proteins from the TOM complex to the TIM22 complex through the aqueous IMS and preventing their aggregation (Koehler et al., 1998a; Koehler et al., 1998b). TIM54 is believed to act as a binding site for the small TIM complex because its interaction with TIM22 was destabilized in a TIM54 deletion mutant (Kovermann et al., 2002). The small TIM complex can also play an important role in substrate recognition by the TIM22 complex.

**Protein Import to the intermembrane space**

All of the proteins residing in the intermembrane space (IMS) are nuclear encoded and imported from the cytosol. Bigger IMS proteins contain bipartite signal sequences consisting of a matrix-targeting presequences followed by a hydrophobic sorting signal (Hartl et al., 1987). Bipartite signal sequences direct the proteins to the IM before they are proteolytically cleaved, thereby releasing the mature part of the protein into the IMS (Glick et al., 1992). Small IMS proteins with sizes 7-16 kDa (e.g. small TIM) carry a characteristic “twin CX\(_9\)C motif”. Import of the Cox17 protein into the IMS was abolished upon mutating one of the
cysteine residues present in the CX$_2$C motif (Heaton et al., 2000), indicating critical importance of cysteine residues in import. Cysteine residues of the twin CX$_2$C motif have been also shown to be required for stability and folding of the proteins in the IMS (Lu et al., 2004). Mesecke et al. (2005) have identified a “disulfide relay system” for protein import into the IMS. It was shown that the newly arrived TIM13 and Cox17 proteins are entrapped in the IMS by forming disulfide bonds with Mia40, a component of disulfide relay system. It was also shown that the sulphhydryl oxidase Erv1 directly interacts and is required for maintaining Mia40 in an oxidized state. Depletion of either Erv1 or Mia40 in _S. cerevisiae_ resulted in no import of Cox17 and TIM13 into the IMS (Mesecke et al., 2005).

**Processing peptidases**

Once a precursor has been imported into the mitochondrial matrix, the presequence has fulfilled its function and is no longer needed. The presequence may actually interfere with further sorting and protein folding or assembly. There are three types of processing peptidases in mitochondria that mediate removal of the presequence from the precursors. These are Mitochondrial Processing Peptidase (MPP), Mitochondrial Intermediate Peptidase (MIP) and Inner Membrane Peptidase (IMP) (as reviewed by Gakh et al., 2002).

MPP is an essential protein in _S. cerevisiae_ and processes precursors that are fully translocated to the matrix as well as precursors in transit to the IM or the IMS. MPP has been purified and characterized from different sources including fungi, mammals and plants (Glaser and Dessi, 1999; Gakh et al., 2002). MPP is an integral part of the cytochrome bc$_1$ complex in plant mitochondria, while it is a soluble protein in fungal and mammalian mitochondria (Braun et al., 1993; Eriksson et al., 1994). MPP is a heterodimeric protein composed of α- and β-subunits of about 50 kDa each. The catalytic site is present in the β subunit with a characteristic inverted zinc binding motif (HXXEH), while substrate recognition and binding is mediated by the α-subunit (Luciano et al., 1997). MPP is classified as a member of the pitrilysin family of proteases on the basis of the zinc binding motif (Kitada et al., 1995). The crystal structure of the recombinant _S. cerevisiae_ MPP in complex with a synthetic presequence peptide has been determined (Taylor et al., 2001). The crystal structure showed that the presequence peptide was bound in an extended conformation at the active site present in a large polar cavity. It was suggested that the presequences adopt context-dependent conformations through mitochondrial import and processing, helical for
A number of mitochondrial precursors destined to the mitochondrial matrix or the IM are processed in two sequential steps by MPP and MIP. These precursors carry a characteristic R-XΨ(F/L/I)-X-X(T/S/G)-X-X-X-XΨ (first arrow indicates cleavage by MPP and the second by MIP) motif at the C-terminus of the presequence. The first cleavage is made by MPP one residue downstream from the arginine that yields a processing intermediate with a typical N-terminal octapeptide that is sequentially cleaved by MIP producing a mature size protein (Gakh et al., 2002). MIPs from different species are soluble monomers of about 70-75 kDa. MIP is a thiol-dependent metalloprotease and belongs to the thimet (thiol and metal dependent) oligopeptidase family (Barret et al., 1995). Deletion of MIP in \textit{S. cerevisiae} causes loss of respiratory competence, suggesting that MIP is involved in the biogenesis of some of mitochondrial proteins. A number of substrates for MIP have been identified in \textit{S. cerevisiae} including CoxIV, ubiquinol-cytochrome c reductase iron sulphur protein (Fe/S) and malate dehydrogenase (MDH) (Branda and Isaya, 1995). The biological significance of the processing by MIP is not certain. It is known that the N-terminal region of MIP processed precursor proteins is incompatible with cleavage by MPP and octapeptides may have evolved to overcome this problem (Isaya et al., 1991).

Some of the proteins imported into the IMS carry a bipartite N-terminal targeting signal consisting of a matrix-targeting signal, (Hartl et al., 1987), that is cleaved by MPP, followed by a hydrophobic signal that is cleaved by IMP. In \textit{S. cerevisiae}, IMP exists as a heterodimeric protein composed of two different subunits: Imp1 and Imp2, both of these subunits possess catalytic activity (Schneider et al., 1994). Each subunit is bound to the outer face of the IM through an N-terminal membrane spanning domain and exposes the C-terminus with the catalytic site into the IMS (Daum et al., 1982). The catalytic sites of both subunits are characterized by a conserved serine/lysine dyad (Chen et al., 1999). Interestingly, each subunit recognizes different substrates. Imp1 is involved in the maturation of at least three proteins; CoxII, Cyt b$_2$ and NADH-cytochrome b$_5$ reductase (Mcr1), whereas Imp2 cleaves the targeting signal of cyt c$_1$ (Nunnari et al., 1993). Deletion of the \textit{Imp1} or \textit{Imp2} gene in \textit{S. cerevisiae} leads to no growth on-non fermentable carbon sources, indicating a role of IMP in mitochondrial biogenesis. Both the Imp1 and Imp2 are homologous to the signal peptidases of the bacterial and ER membrane (Dalbey et al., 1991). A potential third subunit
of the IMP complex, Som1, has been identified using co-immunoprecipitation and cross linking experiments (Jan et al., 2000). Som1 is required for the processing of two of the three known substrates of Imp1 (CoxII and Mcr1), but not for the processing of the Cyt b2 (Esser et al., 1996).

**Chloroplastic protein import machinery**

The majority of the chloroplastic proteins are nuclear encoded and post translationally imported into the chloroplasts in a similar way as for mitochondria (Figure 2). Import into the chloroplast involves transit peptides, cytosolic factors and two translocases present at the outer and inner envelope of the chloroplasts. Being a recent organelle in the modern eukaryotic plant cell, the chloroplastic import machinery possesses unique features to ensure the targeting specificity. Isolated *P. sativum* chloroplasts have been used as a model system to identify and characterize the components of the chloroplastic protein import machinery using a variety of biochemical techniques (Perry and Keegstra, 1994; Schnell et al., 1994). Sequencing of the *A. thaliana* and *O. sativa* genomes has enabled the use of more advanced genetic techniques in search of the import machinery components.

**Chloroplastic targeting peptides – the transit peptides**

Chloroplastic targeting peptides called transit peptides do not show any sequence consensus at the primary structure level and vary greatly in length from 13 to 146 amino acid residues with an average length of about 60 residues. Generally, transit peptides are longer than mitochondrial presequences (Zhang and Glaser, 2002; Bhushan et al., 2006). Interestingly, transit peptides are very similar in overall amino acid composition to the presequences. They are enriched in hydroxylated and hydrophobic amino acids, have some positively charged residues and lack negatively charged amino acids. In comparison to presequences, positively charged amino acids are usually lacking in the very N-terminal part of transit peptides (Peeters and Small, 2001; Zhang and Glaser, 2002; Bhushan et al., 2006). Transit peptides are mainly unstructured in an aqueous environment and it has been proposed that transit peptides have evolved to maximize the potential to form a random coil (Bruce, 2000). Both ferrodoxin (Fd) and Rubisco activase transit peptides (Lancelin et al., 1994; Krimm et al., 1999) from *Chlamydomonas reinhardtii* were shown to contain a helix and a random coil structure as determined by NMR. NMR structural data available for the higher plant transit peptide from the *Silene* Fd shows that addition of micelles to Fd transit peptide induced N- and C-terminal helical formation in the transit peptide. However, induced helices were short and contained
only 3-4 amino acid residues indicating that the major part of the transit peptide remained unstructured even in the presence of membrane mimicking environment (Wienk et al., 2000).

**Figure 2. The chloroplastic protein import machinery.** General overview of the protein import into stroma. TOC and TIC refer to Translocases of the outer and inner envelope membrane of chloroplasts. The numbers represent the molecular masses of the components of the TOC and TIC complexes. SPP, Stromal processing peptidase; Hsp, Heat shock protein; Cpn60, Chaperonin 60; PreP, PresequenceProtease (modified from Bedard and Jarvis, 2005).

**Cytosolic factors**

Chloroplastic precursor proteins are synthesized in the cytosol and have to be imported into the organelle post-translationally and therefore they need to be protected from aggregation and degradation. During or after translation in the cytosol most of these precursor proteins associate with cytosolic factors or chaperones. This interaction is believed to be non selective and due to the unfolded protein exposing hydrophobic amino acids. *In vitro* import of
bacterial overexpressed and urea denatured light-harvesting chlorophyll-binding protein (LHCP) precursor protein into chloroplasts was greatly stimulated by cytosolic factors (Waegemann et al., 1990). One of these factors could be replaced by purified Hsp70. Some of the chloroplastic transit peptides contain a motif that can be phosphorylated on a serine or threonine residue by a protein kinase. Some of the abundant chloroplastic precursor proteins such as the precursors of small subunit of ribulose bisphosphate carboxylase/oxygenase (SSU), LHCP and outer envelop 23 (OE23) have been shown experimentally to become phosphorylated (Waegemann and Soll, 1996), while many more are predicted to contain the potential phosphorylation motif. The phosphorylated precursor protein interacts with the 14-3-3 protein and Hsp70 to form a guidance complex (May and Soll, 2000). 14-3-3 proteins belong to a ubiquitous protein family of regulatory proteins with their main function being molecular chaperones mediating protein-protein interaction (Aitken et al., 1992). Binding of precursor proteins to the guidance complex stimulated the import rate about 4-5 fold into chloroplasts when compared to the free precursor (May and Soll, 2000). However, removal of the phosphorylation site does not result in loss of the targeting specificity (Waegemann and Soll, 1996). Martin et al. (2006) have recently isolated a serine/threonine protein kinase from Arabidopsis that is able to phosphorylate chloroplast targeted precursor proteins.

**Translocase of the outer envelope membrane (TOC)**

Like the TOM complex of mitochondria the TOC complex is involved in both recognition and translocation of chloroplastic precursor proteins across the outer envelope membrane of chloroplasts. Unlike the mitochondrial TOM complex, translocation through the TOC complex is an energy dependent process (Jarvis and Soll, 2001). The core of the TOC complex isolated from Petunia chloroplasts consists of three proteins, TOC34, TOC75 and TOC159, named according to their molecular masses (Perry and Keegstra, 1994; Waegemann and Soll, 1995). The molecular stoichiometry of TOC75, TOC34 and TOC159 in the complex was determined to be 4:4:1 (Schleiff et al., 2003).

TOC34 is anchored to the outer envelope membrane by a C-terminal tail, while a large N-terminal domain is exposed into the cytosol (Seedorf et al., 1995). The N-terminal domain possesses GTP binding and GTPase activity of TOC34 (Kessler et al., 1994). Becker et al. (2004) have suggested that TOC34 acts as an initial receptor on the chloroplastic surface. TOC34 binds precursor proteins with high affinity in its GTP bound form. The precursor functions as a GTPase activating factor and stimulates the GTP hydrolysis of TOC34 by about
40-50 fold (Jelic et al., 2002). TOC34-GDP has a much lower affinity for the precursor, which continues its path to the next translocon subunit, most likely TOC159. There are two TOC34 homologues present in *A. thaliana*, named *At*TOC34 and *At*TOC33. *At*TOC33 is expressed predominantly in photosynthetic and meristematic tissue, while *At*TOC34 is expressed in all tissues, but at a relatively lower level (Jarvis et al., 1998; Gutensohn et al., 2000).

TOC159 was the first TOC component to be identified (Waegemann and Soll, 1991; Perry and Keegstra, 1994; Ma et al., 1996). TOC159 is proposed to be the main receptor for the chloroplastic import machinery (Kessler et al., 1994; Perry and Keegstra, 1994). TOC159 is composed of the three domains: an N-terminal A-domain that contains many acidic amino acid residues, a central G-domain containing a GTP binding domain with sequence homology to TOC34 (Hirsch et al., 1994, Kessler at al., 1994) and a C-terminal M domain that is essential for targeting and anchoring to the membrane (Lee at al., 2003). TOC159 is essential as *A. thaliana* seedlings with TOC159 knockout die early during development (Bauer et al., 2000).

TOC75 is the most abundant outer envelope protein. TOC75 forms the pore through which precursor proteins cross the outer envelope membrane. Overexpressed and purified TOC75 forms a cation-selective channel when inserted into the lipid bilayer (Hinnah et al., 2002). TOC75 is predicted to be a β-barrel protein with 16 TM β-sheets (Sveshnikova et al., 2000). Calculation of the pore diameter indicates that the channel is approximately 15-25 Å wide (Hinnah et al., 2002). This is wide enough to accommodate a polypeptide chain with some secondary structure (Hinnah et al., 2002). TOC75 has a protein binding site at the cytosolic face of the channel that can discriminate between the precursor and mature form of the protein (Ma et al., 1996; Hinnah et al., 1997). There are four TOC75 homologues present in *A. thaliana*, however only one isoform is dominantly expressed.

The role of the fourth TOC component, TOC64, is not well defined. TOC64 exposes tetratricopeptide repeats in the cytosol, like the peroxisomal receptor Pex5 or the mitochondrial receptor TOM70 (Sohrt and Soll, 2000). On the basis of similarity to TOM70 it has been proposed that TOC64 has a similar role to that of TOM70 in recognition of polytopic membrane proteins (Soll and Schleiff, 2004). Becker et al. (2004) have identified a new TOC component, TOC12. TOC12 carries a J-domain and stimulates the ATPase activity of Hsp70.
Translocase of the inner envelope membrane (TIC)

After their translocation through the TOC complex into the IES, the precursor proteins are transferred to the TIC complex for translocation across the inner envelope membrane. ATP is required for translocation across the inner envelope membrane (Flugge and Hinz, 1986). Several TIC components have been identified, however their role in import is less well defined. The TIC translocase is a multi subunit complex consisting of TIC110, TIC62, TIC55, TIC40, TIC32, TIC22 and TIC20 (as reviewed by Gutensohn et al., 2006).

TIC110 is an abundant protein in the inner envelope and has one or two TM segments in its N-terminal region (Kessler and Blobel 1996; Lubeck et al., 1996). TIC110 is believed to form a pore in the inner envelope and can form a cation-selective channel when inserted in the lipid bilayer (Heins et al., 2002). The pore diameter was estimated to be between 15-20 Å, which is the same as for TOC75. TIC40 is an integral membrane protein tightly associated with TIC110 (Stahl et al., 1999). The exact role of TIC40 is not known, but it shares some sequence similarity with Hsp70-interacting protein (Hip) in its C-terminal domain (Chou et al., 2003). Hip is a mammalian co-chaperone that regulates nucleotide exchange by Hsp70 (Hohfeld et al., 1995; Frydman and Hohfeld, 1997) and it may be possible that TIC40 has a role in chaperone recruitment at the TIC complex during protein import into chloroplasts. A role for TIC40 as a chaperone recruitment factor is further supported by the demonstration that Hsp93 and TIC40 can be immunoprecipitated together (Chou et al., 2003). Three of the subunits of the TIC translocon, TIC62, TIC55 and TIC32 are redox components of the TIC complex. TIC55 contains a Rieske iron sulphur center and a mononuclear iron binding site, which indicates the potential for electron transfer (Caliebe et al., 1997). TIC62 contains a conserved NAD/NADP binding site and a C-terminal motif, which interacts with stromal ferredoxin-NAD/NADP reductase (Kuchler et al., 2002). Ferredoxin-NADP-reductase connects photosynthetic electron transfer with metabolically required reducing power. TIC62 might therefore represent a link between the metabolic redox status of the chloroplasts and TIC translocon (Hirohasi et al., 2001). TIC32 belongs to the family of short chain dehydrogenases, which also use NAD/NADP as a cofactor. TIC22 is localized to the IES and has been proposed to be a link between TOC and TIC complexes or in the transfer of proteins across the IES (Ma et al., 1996; Kouranov and Schnell, 1997). TIC20 is another integral subunit of the TIC complex with homology to bacterial amino acid transporters and TIM17 of mitochondria (Kouranov et al., 1998; Rassow et al., 1999), and has been suggested to take
part in the channel formation. There is need of more biochemical work on these individual TIC subunits to define their exact role in chloroplastic protein import.

**Stromal processing peptidase**

Stromal Processing Peptidase (SPP) is the protease that cleaves off the transit peptides from precursor proteins after their import into the stroma. SPP is responsible for cleaving off the transit peptide from a number of different precursor proteins involved in different biosynthetic pathways and destined for different locations in the chloroplasts (Richter and Lamppa, 1999). SPP was initially purified from *P. sativum* chloroplasts as a soluble metalloprotease of about 100 kDa (Oblong and Lamppa 1992). SPP contains an inverted zinc binding motif (HXXEH) characteristic of members of the metallopeptidase family of pitrilysin proteases such as pitrilysin, insulin degrading enzyme (IDE) and the catalytic β-subunit of mitochondrial MPP (Rawlings et al., 2006). Down regulation of SPP in *A. thaliana* yielded many lines that were seedling lethal. Import of a model precursor protein was defective in surviving plants, indicating a critical function for SPP in the chloroplast protein import pathway (Zhong et al., 2003). SPP initially recognizes a precursor by binding to the transit peptide and then cuts it off in a single proteolytic event. The mature form of the protein is then released, while SPP remains bound to the transit peptide. Before the release from SPP, transit peptides are further cleaved into sub fragments by a second proteolytic event (Richter and Lampaa, 2000, 2003).

**Dual targeting to mitochondria and chloroplasts**

Plant cells contain both mitochondria and chloroplasts and therefore require more efficient sorting mechanism than non plant cells. The existence of a higher order protein sorting is evident from in vivo studies where protein import into these two organelles was shown to be highly specific (Boutry et al., 1987; Schmitz and Lonsdale 1989; Silva-Filho et al., 1997). There are a number of proteins present in these two organelles with similar functions that are encoded by a distinct gene for each organelle. However, there are some proteins encoded by a single gene but targeted to both mitochondria and chloroplasts, referred to as dual targeted proteins (Peeters and Small, 2001). Since the first report of dual targeting of *P. sativum* glutathione reductase (GR) by Creissen et al. (1995), 33 dually targeted proteins have been identified and it is expected that there will be many more (Silva-Filho, 2003). In silico analysis of the *A. thaliana* genome predicted that as many as 160 proteins may be dually targeted to both mitochondria and chloroplasts (Small I, personal communication).
The mystery of dual targeting lies in the targeting peptide by which the precursors are targeted and imported into both mitochondria and chloroplasts. Analysis of the dual targeting peptides has revealed that they are intermediate in length and have an overall amino acids composition similar to that of mitochondrial and chloroplastic targeting peptides. However, they contain fewer alanines and a greater abundance of phenylalanine and leucine, suggesting that dual targeting peptides are more hydrophobic (Peeters and Small, 2001). This implies that they have potential to be targeted and imported simultaneously to both of these organelles.

**Studying dual targeting**

Targeting of proteins to mitochondria and chloroplasts has been studied using a number of different experimental approaches including both **in vivo** and **in vitro** methods. When it comes to studying the targeting of a dual targeted protein, none of these methods alone is ideal. The most commonly used method to study subcellular localization of dually targeted proteins is an **in vivo** method expressing a chimeric construct consisting of a reporter protein such as green fluorescence protein (GFP) fused to the full length precursor or targeting peptide (Peeters and Small, 2001). **In vivo** methods use an intact cellular system and are the best system to study the **in vivo** targeting capacity of a targeting peptide. However, there are some limitations of this system: 1) fusion construct often use a small part of the protein coupled to a reporter protein and therefore the role of the mature protein is ignored, 2) fusion proteins are usually under a strong promoter and overexpressed at very high level which can affect targeting and 3) it is not possible to study the kinetics and efficiency of protein recognition and import.

Another method is to import **in vitro** synthesized radiolabelled precursor proteins into isolated organelles. This **in vitro** method can be useful sometimes but has other disadvantages: 1) isolated organelles lack an intact cellular system and other factors required for protein sorting, 2) protein can be miss-targeted to an incorrect organelle and 3) there is no competition between organelles. Rudhe et al. (2002a) established an **in vitro** dual import system enabling the simultaneous import of radiolabelled precursor proteins into both mitochondria and chloroplasts minimizing the miss-targeting associated with the classical single organellar **in vitro** import system. A combination of different complementary methods should be applied in order to study the targeting of a dual targeted protein.
Mechanisms of dual targeting

The majority of reports on single gene products that are targeted to more than one subcellular location in a cell are related to mitochondria and chloroplasts. Dual targeting to these organelles can be achieved in two ways whereby a single gene product can be targeted to both mitochondria and chloroplasts: either through an ambiguous targeting signal or via a twin targeting signal (Peeters and Small, 2001). The precursor proteins with an ambiguous targeting signal are synthesized as a single polypeptide, but can be recognized and transported by the import machinery of both mitochondria and chloroplasts (Small et al., 1998). The majority of the known dually targeted proteins carry an ambiguous targeting signal. Most of these proteins are involved in gene expression e.g. most of the aminoacyl-tRNA synthetases and RNA polymerases. Other dually targeted proteins are involved in various processes such as: biosynthetic pathways, phosphoribosyl aminoimidazole synthase (Smith et al., 1998) and phosphatidylglycerophosphate synthase I (Babiychuk et al., 2003); protein modification function such as methionine amino-peptidase (Giglione et al., 2000) and; anti-oxidant activities such as GR. It has been suggested by Chew et al. (2003) that the enzymes involved in entire enzymatic cycles may be dually targeted. They have shown that main components of the ascorbate glutathione cycle in A. thaliana, ascorbate peroxidase, monodehydroascorbate reductase (MDAR) and GR were dually targeted to mitochondria and chloroplasts both in vitro and in vivo (Chew et al., 2003). See also page 46 in this thesis and papers I and II.

Twin targeting signals have two separate targeting signals for mitochondria and chloroplasts in tandem and at a given time only one targeting signal is present in the precursor protein. Twin targeting signals may arise by either alternative transcription or translational initiation, alternative splicing or via post translational modifications resulting in the formation of two different precursors with distinct targeting specificity. Twin targeting signals seem not to be common among dual targeting proteins. Protox, a protein involved in the biosynthesis of chlorophyll and heme, and THI1 involved in thiamine biosynthesis are dually targeted using alternative translational initiation, with the longer form of protein targeted to chloroplasts and the shorter to the mitochondria (Chabregas et al., 2003; Watanabe et al., 2001). A. thaliana MDAR is dually targeted using alternative transcription start sites, producing two forms of mRNAs, the longer form of mRNA is translated with a mitochondrial targeting signal, while the shorter one is translated with a chloroplastic targeting signal (Obara et al., 2002). A domain structure was proposed for the dually targeted P. sativum GR targeting peptide,
indicating that targeting information for mitochondria and chloroplasts is located in different domains (Rudhe et al., 2002b).

**Proteolytic system in mitochondria**

Proteolysis is important for the biogenesis, morphology and homeostasis of mitochondria. There are about 40 proteases predicted to be present in mitochondria, but only a very few of them have been characterized so far (Esser et al., 2002). Non-selective degradation of mitochondrial proteins occurs in the lysosome after autophagy of the whole organelle, whereas selective degradation is mediated by proteases within the mitochondrion (as reviewed by Kaser and Langer 2000). Mitochondrial proteases can be classified into two classes based on their requirements for ATP: ATP-dependent and ATP-independent.

**ATP-dependent proteases**

ATP-dependent proteases are involved in the assembly of mature proteins by regulation of the stoichiometric amount of polypeptides in protein complexes and are also required for the removal of miss-folded and damaged proteins. These proteases catalyze the first step of degradation by cleaving the substrate polypeptide into peptides that are later cleaved to free amino acids by ATP-independent proteases. ATP-dependent proteases do not require ATP for hydrolysis, but rather for unfolding of target polypeptides and to regulate their proteolytic activity. Mitochondria contain a few ATP-dependent proteases including the membrane bound FtsH protease and the soluble Lon and ClpP proteases (Kaser and Langer, 2000; Adam and Clarke, 2002; Urantowka et al., 2005).

**The FtsH (AAA) protease**

The FtsH proteases, also called AAA-proteases, are membrane bound, ATP-dependent metalloproteases. FtsH proteases are required for the assembly of the newly imported proteins into their native protein complexes by degrading superfluous subunits (Langer, 2000). These proteases are present in eubacteria and in mitochondria and chloroplasts. *S. cerevisiae* mitochondria contain two classes of AAA-proteases with different topologies, named the m-AAA and i-AAA proteases (Leonhard et al., 1996, 2000; Klanner et al., 2001). The m-AAA protease catalytic site faces the matrix side, forms a mega complex of 1 Mega Dalton (MDa), and is composed of Yta10 and Yta12 subunits (Yeast Tat binding like proteins). The i-AAA protease catalytic site is exposed to the IMS and also forms a 1 MDa complex (Langer, 2000). Whereas there is a single gene of *FtsH* protease in bacteria, there are three homologues
present in *S. cerevisiae* and humans (Arnold and Langer, 2002). It has been suggested that AAA-proteases function in the degradation of the non-assembled membrane proteins. A non-assembled CoxII protein and protein inhibitors of the ATP synthase have been identified as native substrates of the i-AAA type of FtsH protease (Pearce and Sherman, 1995; Kominsky, et al., 2002). The cellular function of mitochondrial FtsH protease has been studied in detail in *S. cerevisiae*, where severe phenotypes were shown to be associated with mutations in these proteases (Leonhard et al., 1996; Weber et al., 1996). The deletion of either Yta10 or Yta12 subunits resulted in inhibition of respiration and impaired the degradation of none assembled IM proteins and assembly of the respiratory chain complexes and ATP synthase (Arnold and Langer, 2002). It has been also shown that the mutations in m-AAA proteases affected splicing of *CoxI* and *Cob*, both of which are encoded by genes in the mitochondrial genome (Arlt et al., 1998). Mitochondrial dysfunction such as impaired respiration and changes in morphology were shown in *S. cerevisiae* lacking the i-AAA protease subunit Yme1 (Yeast mitochondrial escape) (Thorsness et al., 1993; Campbell et al., 1994). Mutations in the human homologue of the AAA-protease, paraplegin, cause Hereditary Spastic Paraplegia (HSP), a neurodegenerative disorder (Casari et al., 1998).

In plants, the *A. thaliana* genome harbors a total of 12 *FtsH* like genes (Sokolenko et al., 2002) with four suggested to be localized to mitochondria and rest eight to chloroplasts (Adam and Clarke, 2002). Our knowledge of FtsH proteases in plants is limited. A plant m-AAA protease (*PsFtsH*) has been identified and studied from *P. sativum* (Kolodziejczak et al., 2002). *PsFtsH* was shown to complement respiratory defects of *S. cerevisiae* lacking the m-AAA protease, indicating a conserved function of AAA-protease from fungi to plants. On the basis of Blue-native gel electrophoresis, it has been suggested that plants, in contrast to yeast, have more than one i-AAA protease complex (Urantowka et al., 2005).

**Lon-like Protease**

Lon Like proteases belong to a conserved protein family with members present in eubacteria, archaeabacteria and eukaryotes (Van Dyck and Langer, 1999). These are ATP-dependent, serine proteases present in both the mitochondrial matrix and the chloroplastic stroma (Suzuki et al., 1994). Functional conservation between various members of the Lon protease family has been demonstrated by complementation studies in *S. cerevisiae* (Barakat et al., 1998; Teichmann et al., 1996). Mitochondria in *S. cerevisiae* contain the Lon homologue, Pim1 (Protease in mitochondria), which is a homo-oligomer composed of 7 subunits and is about
800 kDa in molecular mass (Stahlberg et al., 1999). Pim1 consists of two catalytic domains, an ATPase domain and a protease domain. Pim1 in *S. cerevisiae* mitochondria cleaves several non assembled proteins such as the β-subunit of MPP, and the α-, β- and γ-subunits of ATP synthase (Kaser and Langer, 2000). Inactivation of Pim1 in *S. cerevisiae* causes severe phenotypes in mitochondria indicating specific regulatory functions of Pim1. Cells lacking Pim1 are unable to maintain mitochondrial DNA and are respiratory deficient. Pim1 mutants are defective in splicing of CoxI and Cob mRNAs (Van Dyck et al., 1994). This role of Pim1 in splicing is similar to that of m-AAA proteases and indicates that Pim1 can also be required for the processing of enzymes required in mRNA maturation. Down-regulation of human Lon protease resulted in disruption of mitochondrial structure and function and eventually cell death within four days. Cell death in the majority of these cells was a result of caspase 3 activated apoptosis (Bota et al., 2005). It has been suggested that the Lon proteolytic system plays an important role in the degradation of oxidized proteins in the mitochondrial matrix and in the maintenance of mitochondrial structure and functional integrity (Bulteau et al., 2006).

There are four Lon homologues present in the *A. thaliana* genome. Lon1 is localized to mitochondria, Lon2 to peroxisomes, the remaining two are predicted to be localized to the chloroplasts (Adam et al., 2001). A gene encoding Lon protease has been reported from *Z. mays* (Barakt et al., 1998). *Z. mays* Lon1 can complement mitochondrial DNA maintenance in Pim1 deficient *S. cerevisiae*. In legume mitochondria, Lon1 has been identified as at least one of the proteases involved in the degradation of orf239, a cytoplasmic male sterility (CMS) associated protein in mitochondria (Sarria et al., 1998).

**ClpP-like protease**

Lon and Clp (Caseinolytic protease) are the two major proteases in *E. coli* accounting for about 80% of protein degradation. Homologues of Clp proteases are present in mammalian and plant mitochondria, but are absent in *S. cerevisiae* (Corydon et al., 1998; Santagata et al., 1999; Halperin et al., 2001a). Most of our knowledge about Clp proteases is based on *E. coli* Clp protease. Clp-like proteases form hetero-oligomeric complexes with an interior chamber for proteolysis and consist of proteolytic (ClpP) and regulatory (ClpA or ClpX) subunits. The catalytic ClpP subunit has an active site similar to serine proteases, a catalytic triad of Ser-His-Asp, whereas regulatory subunits have a chaperone like function. The crystal structure of *E. coli* Clp protease has been solved at 2.3 Å resolution (Wang et al., 1997). The proteolytic
chamber is composed of two central heptameric rings of ClpP that form a hollow chamber of 50 Å, flanked by one or two hexameric rings of ClpA or ClpX. The catalytic chamber has two narrow openings of about 10 Å at either end, indicating that the substrate needs to be unfolded before entering into the proteolytic chamber. ClpP shares similar structural features to the *S. cerevisiae* 20S proteasome and the *E. coli* Hs1V (Lowe et al., 1995; Bochtler et al., 2000). Bacterial Clp protease is involved in degradation of specific regulatory proteins, aggregated and mis-folded proteins, nascent peptide chains that are stalled on the ribosomes, and proteins involved in the stress and starvation responses (Gottesman et al., 1998). A cellular quality control role has been suggested for the Clp proteases (Adam et al., 2006).

Plants have a battery of Clp like proteases. There are at least 30 Clp-related genes in the *A. thaliana* genome, 15 of these genes encoding plastid localized proteins, 5 serine type ClpP proteases containing catalytic triad motifs, 4 ClpP related ClpR proteins lacking catalytic triad motifs, 3 ClpA homologues (ClpC1, ClpC2 and ClpD), and 3 Clp proteins with unknown function (ClpS1, ClpS2 and ClpT) (Halperin et al., 2001b; Peltier et al., 2004; Adam et al., 2006). ClpP2, ClpX1 and ClpX2 are predicted to be located in mitochondria. Plant Clp proteases are believed to play a housekeeping role since they are constitutively expressed under different growth and environmental conditions (Halperin et al., 2001b; Adam et al., 2006).

**ATP-Independent proteases**

Beside the processing peptidases MPP, MIP and IMP, mitochondria also contain a few ATP-independent proteases, which are oligopeptidases. These include: mitochondrial oligopeptidases, MOP and MOP112, in the IMS of yeast (Buchler et al., 1994; Kambacheld et al., 2005); Oma, a membrane bound metalloprotease (Kaser et al., 2003) with overlapping functions with the m-AAA protease and; a serine rhomboid membrane protease responsible for the degradation of intermembrane space proteins (Van der Bliek and Koehler, 2003).

**Oligopeptidases**

Oligopeptidases are endopeptidases that act on shorter peptides of about 6-18 amino acid residues (Barrett et al., 1995). In *S. cerevisiae*, MOP is mainly located in the cytosol, but some of the activity has also been reported in mitochondria. MOP contains an N-terminal presequence and is targeted to different subcellular locations by using alternative promoters (Buchler et al., 1994; Serizawa et al., 1995; Kato et al., 1997). Deletion of MOP in *S.
cerevisiae does not lead to any major defects, but cells exhibit a decrease in the intracellular degradation of a collagen-like substrate, indicating that MOP plays a role in the late stages of protein degradation (Serizawa et al., 1995).

**OMA protease**

OMA (overlapping function with the m-AAA proteases) was identified as a conserved metallopeptidase, a novel component of the quality control system in the inner membrane of mitochondria from *S. cerevisiae* (Kaser et al., 2003). OMA1 was shown to have functions overlapping with the m-AAA protease and cleaves a mis-folded polytopic membrane protein (OXA1) in an ATP-independent manner at multiple sites. Proteins homologous to OMA1 comprise a large protein family with members present in higher eukaryotes, including plants, as well as in eubacteria and archaeabacteria. Although different in their domain structure, all of them are predicted to be integral membrane proteins and contain a metallopeptidase domain characteristic of the M48 family of proteases, suggesting that OMA1 represents a novel enzyme class capable of degrading membrane proteins (Kaser et al., 2003). It was proposed that OMA1 functions under conditions of limited m-AAA proteolytic activity (Kaser et al., 2003).

**Rhomboid protease**

Rhomboid protease is an interesting integral membrane protease that cleaves the substrate within the membrane. Rhomboid protease was initially identified in *Drosophila* Golgi bodies. Here it regulates epidermal growth factor receptor signaling by cleaving the transmembrane domain of Spitz, the principal ligand for the receptor in flies, and promoting its release from signal-sending cells (Urban et al., 2001). Rhomboid protease is a serine protease with a characteristic Ser-His-Asn catalytic triad. Rhomboid proteases are present in all three domains of life, archea, bacteria and eukaryotes (Koonin et al., 2003). Wang et al. (2006) have recently solved a 2.1 Å resolution crystal structure of the rhomboid core domain. The structure contained six transmembrane segments. Catalytic site residues, the Ser-His dyad, and several water molecules were found at the protein interior at a depth below the membrane surface. These observations indicated that, in intramembrane proteolysis, the scission of peptide bonds takes place within the hydrophobic environment of the membrane bilayer.

Whereas most prokaryotes have a single gene encoding the rhomboid protease, *Drosophila* has seven and *A. thaliana* has eight. One of the *S. cerevisiae* homologues of rhomboid, Rbd1
has been shown to be located in mitochondria. Rbd1 has been shown to be involved in processing of the bipartite signal peptides of cytochrome c peroxidase, Ccp1 and a Dynamin-like GTPase, Mgm1 (Esser et al., 2002; Herlan et al., 2003, 2004; Mc Quibban et al., 2003; van der Blick and Koehler, 2003). It has been suggested that cleavage of Mgm1 by the rhomboid protease regulates mitochondrial membrane remodeling indicating a role of proteases in regulating membrane biogenesis (McQuibban et al., 2003). Cipolat et al. (2006) have also shown that rhomboid protease directly regulates apoptosis by controlling cytochrome c release from mitochondria. There is currently no information on the role of the rhomboid protease in plants, but two of them are predicted to be localized to mitochondria in A. thaliana and possibly have a role in processing the N-terminal extension present on few carrier proteins (Murcha et al., 2004).

**Proteolytic system in chloroplasts**

Proteolysis is required for a number of processes during the biogenesis and maintenance of chloroplasts. Proteolysis should therefore be considered a vital homeostatic factor that influences metabolic functions such as photosynthesis under both normal and adverse growth conditions (Adam, 2000, 1996). Chloroplasts contain defined proteases within each compartment; the ATP-dependent Lon, Clp in the stroma and FtsH in stroma exposed thylakoid membranes; the ATP-independent DegP protease within the thylakoid lumen and on both sides of thylakoids membrane and the SppA protease on the stromal side of the thylakoid. All five of these chloroplastic proteases are homologous to bacterial proteases and are present in multiple copies in higher plants.

**ATP-dependent proteases**

**The FtsH protease**

FtsH proteases in chloroplasts were first identified by immunoblot analysis from S. oleracea leaves. They were characterized as integral thylakoid membrane proteins expressed in a light dependent manner (Lindahl et al., 1996). The ~70 kDa protein is bound to the stroma exposed lamellae, with its metal binding and ATP binding sites facing into the soluble stroma. There are a total of sixteen FtsH genes present in the A. thaliana genome with four of them having incomplete metal binding site (Sokolenko et al., 2002). These four FtsH with incomplete metal binding sites might be involved in chaperone like functions because they retain the AAA-domain. In chloroplasts FtsH1 is involved in the degradation of unassembled subunits...
of membrane complexes, such as the Rieske Fe-S protein of the cytochrome $b/f$ complex and the degradation of oxidatively damaged proteins such as the D1 protein of the photosystem II (PSII) reaction centre (Lindahl et al., 2000; Ostersetzer and Adam, 1997).

Mutations in the FtsH2 homologue in A. thaliana leads to a variegated phenotype (leaves with green and yellow sections) (Chen et al., 2000). Ultrastructural analysis of the variegated leaves shows underdeveloped plastids in the yellow sections and normal chloroplasts in the green sections (Chen et al., 2000; Takechi et al., 2000). This suggests a role of FtsH protease in chloroplasts development. However, the patchy phenotype implies that the loss of FtsH2 can be compensated for, at least partially within the green leaf sections. Because up to eight different FtsH proteases are believed to be localized to chloroplasts, some of them can probably compensate for each other. The role of FtsH proteases in biogenesis is supported by a report from cyanobacteria where the loss of four FtsH genes resulted in reduced level of functional PSI by 60%, whereas PSII and phycobilisome levels remained unchanged (Mann, et al., 2000). Recently, the FtsH11 homologue in A. thaliana has been implicated as playing a critical role in A. thaliana thermo-tolerance (Chen et al., 2006).

**Clp-like protease**

Clp is a multi-subunit enzyme complex, in which the catalytic and ATPase domain are present on different subunits. Clp proteases are believed to be the main proteases responsible for most of the protein degradation in stroma. There are about two dozen genes in A. thaliana encoding four Clp subunits (Zheng et al., 2006) with ClpP1 encoded in the chloroplast genome. Chloroplasts do not contain ClpX homologue; instead ClpC and ClpD are suggested to function as the regulatory subunits of the chloroplastic Clp protease. The identified core complex of the Clp protease is about 325-350 kDa in size and consists of five ClpP isomers (ClpP1, ClpP3, ClpP4, ClpP5 and ClpP6), four proteolytically inactive or regulatory subunits (ClpR1 to ClpR4) and two plant specific subunits ClpS1 and ClpS2 (Peltier et al., 2001). Although the function of ClpR remains unknown, it might control access of the substrate to the catalytic site in a chaperone like manner. In A. thaliana, most of the Clp genes showed an increase in transcript level under high light conditions, but were less responsive to temperature shifts (Zheng et al., 2002; Sinvany-Villalobo et al., 2004). These results together with the proteome analysis imply that all isomers constituting the protease core complex
substantially accumulate in all plastid types and that regulation of the proteolytic activity of
the Clp complex may depend on regulatory subunits.

The exact role of Clp proteases is unknown, however they are essential for chloroplast
function. Inactivation of the ClpP1 (encoded in the chloroplasts genome) in *Nicotiana* and
*Chlamydomonas* shows that ClpP1 is essential for cell viability (Huang et al., 1994; Kuroda
and Maliga, 2003). Knockdown of ClpP4 in *Nicotiana* by antisense methods caused severely
reduced growth with chlorotic leaf tissues (Shen et al., 2007). This may mean that each
isomer is required for the formation of the functional complex, supporting the hetero-
oligomeric composition as proposed by Peltier et al. (2004). As for other subunits, visible
phenotypes such as yellow and pale green leaves have been reported in the knockout or
knockdown mutants of ClpR1, ClpR2, ClpR4, ClpC1 and ClpB3, whereas no clear
phenotypes were observed for ClpC2, ClpD and ClpT (Sjögren et al., 2006). Interestingly,
mutations in ClpC2 have been shown to suppress a variegated phenotype caused by loss of an
FtsH (Park and Rodermel, 2004). Although the mechanism of this genetic suppression
remains unclear, it demonstrates the interaction of the Clp regulatory components with other
protease family members.

**Lon-like protease**

Lon was the first ATP-dependent protease found in *E. coli*. *E. coli* cells lacking in Lon shows
accumulation of abnormal proteins and increased sensitivity to DNA damage (Gottesman and
Zipser, 1978). A characteristic property of Lon is its affinity to DNA, although the domain
required for DNA binding is not known (Rotanova et al., 2004). There are four homologues of
the Lon present in the *A. thaliana* genome and two of them are predicted to be chloroplastic.
A gene encoding Lon is not present in *Synechocystis*. Thus, whether Lon is present in
chloroplasts remains controversial. Based on transient assay and immunoblot analysis Lon4
has been shown to be present in chloroplasts (Sakamoto, 2006).

**ATP-independent proteases**

Besides Clp, FtsH and Lon, there are few ATP-independent proteases present in the
chloroplasts. These are SPP, DegP and SppA like protease.
**DegP-like protease**

DegP homologues are found in most organisms including bacteria, human and plants (Spies et al., 1999). DegP is a serine protease that forms a homotrimeric oligomer in *E. coli* and humans (Clausen et al., 2002). Two trimers further dimerize to form a hexamer (Krojer et al., 2002). There are three Deg proteases in *E. coli*, which are named DegP, Q and S. Each of these three protease has one or two characteristic PDZ domains in the C-terminus that are necessary for protease-protease interaction and that possibly regulate recognition of substrates. DegP also has a chaperone activity at higher temperature (Spiess et al., 1999; Ehrmann and Clausen, 2004).

There are 16 DegP homologues present in the *A. thaliana* genome. Four of the *A. thaliana* DegP homologs (Deg 1, 2, 5 and 8) are located in chloroplasts (Haussuhl et al., 2001; Chassin et al., 2002), where they are peripherally attached to the thylakoid membrane. DegP1, 5 and 8 are located on the luminal side, whereas DegP2 is on the stromal side (Haussuhl et al., 2001; Itzhaki et al., 1998). There is no report on the effects of knockout mutations in any DegP genes in *A. thaliana*. Expression of DegP proteins is increased by abiotic stresses such as salt, light and temperature. DegP2 has been proposed to perform the primary cleavage of photo-damaged D1 protein from the PSII complex, prior to its complete degradation by FtsH2 (Haussuhl et al., 2001). Degradation and removal of damaged D1 protein is crucial for the integration of a new D1 copy and for restoring a functional PSII.

**SppA-like protease**

SppA, a homologue of *E. coli* protease IV has been identified in *A. thaliana* chloroplasts. There is only one SppA homologue present in the *A. thaliana* genome (Lensch et al., 2001). SppA is an ATP-independent serine protease. SppA is tightly bound to the stromal side of the thylakoid membrane. SppA is present as a 270 kDa complex suggesting a homotetrameric structure of the SppA complex. SppA is up-regulated by light and might be involved in the light dependent degradation of antenna and PSII complexes (Lensch et al., 2001).
The PresequenceProtease, PreP

After or during import into mitochondria and chloroplasts, most precursor proteins are proteolytically processed by MPP in mitochondria and SPP in chloroplasts, resulting in production of the mature protein and free targeting peptides. In general, targeting peptides do not seem to be required inside organelles. Subunit 9 of the mammalian cytochrome $bc_1$ complex is the only known example of a targeting peptide integrated as a functional subunit of an oligomeric protein complex. Subunit 9 corresponds to the targeting peptide of the Rieske iron sulphur protein (RISP) (Brandt et al., 1993; Iwata, et al., 1998). The RISP precursor is targeted and assembled into the $bc_1$ complex before its targeting peptide is cleaved off (Deng et al., 1998).

Targeting peptides are potentially harmful for the integrity of the structure and function of mitochondria and chloroplasts. They can perturb natural and artificial lipid bilayers. Addition of presequences to mitochondria results in membrane lysis, uncoupling of respiration and dissipation of the membrane potential (Roise et al. 1986; Glaser and Cumsky, 1990a; Glaser and Cumsky, 1990b; Hugosson et al., 1994; Nicolay et al., 1994; van't Hof et al., 1991, 1995). The mechanism of action of presequences on the mitochondrial membrane is not clear, but it has been proposed that the presequence peptides induce channel opening (Lu et al., 1997), or that the peptides themselves form a pore (Matsuzaki et al., 1996). Furthermore, mitochondrial presequences have been shown to possess antimicrobial activity (Hugosson et al., 1994). Therefore, free targeting peptides generated inside the mitochondria and chloroplasts have to be rapidly removed, either by proteolysis or export. Mitochondrial export of peptides has been reported in S. cerevisiae mitochondria. It was shown that the ABC (ATP-binding cassette) Mdh1 protein of S. cerevisiae transports small peptide fragments, generated by AAA-proteases, across the mitochondrial IM (Young et al., 2001), although the efficiency of the export was very low. The low export efficiency of targeting peptides to the outside of mitochondria and the energy requirement for the export of positively charged targeting peptides against the membrane potential, suggests that the degradation of targeting peptides inside mitochondria is the most likely system for their disposal.

In agreement with these observations, it was shown by Ståhl et al. (2000) in our laboratory, that mitochondrial targeting peptides (presequences) are rapidly degraded by a matrix
localized ATP-independent protease after import into mitochondria, while the mature part of the protein remained stable. PresequenceProtease (PreP), the protease responsible for this degradation was isolated from potato tuber mitochondrial matrix and identified by mass spectrometric analysis, ESI MS/MS. (Ståhl et al., 2002). There are two homologues of PreP present in the A. thaliana genome, an A. thaliana zinc metalloprotease (renamed, AtPreP1) (AAL90904, on chromosome 3: Zn-MP) and an A. thaliana putative zinc metalloprotease (renamed, AtPreP2) (AAG13049, on chromosome 1: putZn-MP). Both proteases display high amino acid sequence similarities, with most differences occurring in their predicted organelar targeting peptides. Both AtPreP1 and AtPreP2 harbor a characteristic inverted zinc-binding motif, HILEHX96E, and are classified to the pitrilysin protease family (Ståhl et al., 2002).

**Dual Targeting of PreP in Mitochondria and Chloroplasts (Paper I and II)**

Subcellular prediction programs did not yield conclusive results for the intracellular localization of the AtPreP1 (AtZn-MP) and the AtPreP2 (AtputZn-MP) (Predotar: http://www.inra.fr/predotar/ and TargetP: http://www.cbs.dtu.dk/services/TargetP/). However, the targeting peptide of AtPreP1 and AtPreP2 was predicted to be 85 amino acid residues, by both MitoProt (http://ihg.gsf.de/ihg/mitoprot.html) and ChloroP (http://www.cbs.dtu.dk/services/ChloroP/). In order to investigate the subcellular localization of AtPreP1 and AtPreP2, a number of different, but complementary techniques were employed. In the in vivo import assay, the predicted targeting peptide plus 40 amino acid residues from the mature part of the protein for AtPreP1, and 70 residues for AtPreP2, were fused to the GFP reporter protein. Transient expression of the GFP fusion constructs was carried out in N. tabacum protoplasts and leaves. Protoplasts transformed with the AtPreP1-GFP (Zn-MP-GFP) or AtPreP2-GFP (putZn-MP-GFP) construct showed GFP fluorescence in two distinct locations, with both punctuated and large round shaped structures. Fluorescence in the punctuated shape structures co-localized with Mitotracker, whereas in the large round shaped structures, GFP fluorescence co-localized with chlorophyll autofluorescence. This indicated that the targeting peptide of both PrePs is capable of dually targeting the GFP to both mitochondria and chloroplasts (Paper I and II). In another complementary assay, where Agrobacterium tumefaciens mediated transient expression was carried out with intact N. tabacum leaves, the targeting peptide of AtPreP1 and AtPreP2 again dually targeted GFP to both organelles (Paper I and II). These results indicated that both the AtPreP1 and the AtPreP2 could be members of a dually targeted protein family. The presence of PreP in both organelles was also verified by Western blot analysis.
Interestingly, there is a second methionine at position 29 in the predicted targeting peptide of *AtPreP1* that can be used as a second translational initiation site in *AtPreP1* mRNA, producing two forms of the protein differing in targeting specificity. We therefore used a deletion construct starting from the second methionine (Δ1-28PreP1-GFP), and a mutant with the second methionine changed to leucine ([M29L]PreP1-GFP). The truncated targeting peptide in the Δ1-28PreP1-GFP construct targeted the GFP to chloroplasts only, while the mutant construct, [M29L]PreP1-GFP targeted GFP to both mitochondria and chloroplasts. Furthermore, no fluorescence was detected when a mutated version of *AtPreP1*-GFP construct that contained a frame shift mutation in between the first and the second ATG of the *AtPreP1*-GFP construct, was introduced into the *N. tabacum* leaves (Paper I). These results excluded the dual targeting of the *AtPreP1* as a result of alternative translational initiation *in vivo* and also suggested the presence of a domain organization in *AtPreP1* targeting peptide. In conclusion, the *AtPreP1* harbors an ambiguous targeting signal with a distinct domain present for mitochondrial and chloroplastic targeting that is recognized and transported by both mitochondrial and chloroplastic import machineries.

Dual targeting of the *AtPreP1* and *AtPreP2* was also supported by results from *in vitro* import assays using full length radiolabelled precursor proteins of both PreP isoforms. Both full length PreP precursors were imported and processed into mitochondria and chloroplasts, while the Δ1-28PreP1 precursor was only imported to isolated chloroplasts, with no import detected into mitochondria. [M29L]PreP1 precursor was again imported and processed in both mitochondria and chloroplasts. Incubation of the *AtPreP1* or *AtPreP2* precursors simultaneously with isolated mitochondria and chloroplasts in a dual import system, followed by reisolation of the organelles also resulted in import and processing of the precursor inside both organelles (Paper I and II).

**Function of PreP in Mitochondria and Chloroplasts (Paper I and II)**

The predicted mature part of *AtPreP1* and *AtPreP2* was cloned as a fusion protein with glutathione-S-transferase (GST), and the GST-PreP1 and GST-PreP2 fusion proteins were overexpressed in *E. coli*. Recombinant *AtPreP1* and *AtPreP2* completely degraded the mitochondrial and chloroplastic targeting peptides as well as a fluorescent P1 peptide. The proteolytic activity was not dependent on ATP and was specifically inhibited by ortho-
phenanthroline (o-ph). (Paper I, II and VI). Furthermore, immunoinactivation studies on isolated mitochondrial matrix and chloroplastic stroma resulted in complete inhibition of the proteolytic activity against a mitochondrial and chloroplastic targeting peptide indicating the function of PreP in both organelles.

The substrate specificity of the recombinant *At*PreP1 was investigated using degradation of a mitochondrial targeting peptide, a chloroplastic targeting peptide, unstructured peptide (insulin B chain and galanin (Duckworth et al., 1998) and a folded *de novo* peptide ala-α3w (Dai et al., 2002). *At*PreP1 completely degraded all peptides except for the tightly folded peptide ala-α3w. Based on these results, it can be concluded that PreP is not specific for mitochondrial and chloroplastic targeting peptides and does not recognize amino acid residues *per se*, but degrades unstructured peptides and is not active against folded substrates (Paper VI). What are the advantages of having two proteases with the same function? Is it to enhance the proteolysis? To answer these questions, cleavage specificity of *At*PreP1 and *At*PreP2 was investigated by cleavage of a specific fluorescent peptide P1, the mitochondrial presequence peptide and the chloroplastic transit peptide. After incubation of the peptides for 30 min at 30°C, an intermediate product was produced by *At*PreP2 with both the synthetic fluorescent peptide and mitochondrial targeting peptide, while no such intermediate product was seen when these two peptides were incubated with *At*PreP1 (Paper II and VII). The intermediate could not be detected upon degradation with *At*PreP1 even when much shorter incubation times and different concentrations of *At*PreP1 were used. Accumulation of intermediate products of the peptides after incubation with *At*PreP2 shows that the *At*PreP1 and *At*PreP2 have different cleavage specificity. Degradation by *At*PreP1 and *At*PreP2 was also investigated with the chloroplastic transit peptide and its mutants (Paper II). Mutants were designed to study the effect of changing the polypeptide chain flexibility of the transit peptide on import and processing. *At*PreP2 had the capacity to degrade both the transit peptide and all the mutants, whereas *At*PreP1 could not degrade the chloroplastic targeting peptide mutant (P36A), which has decreased polypeptide chain flexibility as proline has been changed to alanine. Thorough substrate specificity studies of the *At*PreP1 and *At*PreP2 proteases using mass spectrometric analysis of degradation products of mutants of mitochondrial targeting peptides, as well as a number of other synthetic peptides show differences in amino acid recognition and the cleavage efficiency (Paper VII). In conclusion, *At*PreP1 and *At*PreP2 may have overlapping, but complementary proteolytic specificity, allowing a wide variety of substrate peptides to be efficiently degraded.
Expression of the AtPreP1 and AtPreP2 in A. thaliana plants

Expression levels of AtPreP1 and AtPreP2 transcripts in A. thaliana were studied using semi quantitative RT-PCR, under carefully optimized conditions for the quantitative measurements of the transcripts. Both the AtPreP1 and the AtPreP2 transcript were detected in young seedlings, however in varying amounts. The AtPreP1 transcript was detected in silique and flower tissues, although the transcript level was much higher in flowers. In contrast to the AtPreP1 transcript, the AtPreP2 transcript was found to be present in leaf, flower and root tissues with no transcript detected in shoot and silique tissues (Paper II). These results showed that both AtPreP1 and AtPreP2 are expressed in an organ-specific manner in A. thaliana plants (Paper II). It will be interesting to investigate the functional importance of the higher transcript level of AtPreP1 present during flower development.

Crystal structure of AtPreP1, a Peptidasome (Paper III)

The crystal structure of the inactive mutant, AtPreP1 E80Q, was solved at 2.1Å resolution (Paper III). The asymmetric unit is composed of two protein molecules, each with a zinc atom and a six-residue peptide substrate entrapped in the active site. Magnesium ions facilitated crystallization of AtPreP1 and two hydrated ions were found in each protein molecule coordinated by acidic residues. The PreP polypeptide folded into four topologically similar domains that formed two bowl-shaped halves. Comparison of the four domains showed that in spite of only 6-11% sequence identity, the root-mean-square deviation (rmsd) after superimposing the main chains is 2 Å or higher, despite a similar topology. A unique hinge region of 82 residues joins the two enzyme halves. The arrangement of the active site is that of a typical metalloprotease, but the zinc-binding motif is inverted. The first domain forms the major part of the active site containing the inverted zinc-binding motif (HXXEH) where His77 and His81 coordinate the zinc and Glu80 (substituted for Gln in the crystallized mutant) acts as a base catalyst. The inactive mutants H77L, E80Q and H81L generated in a separate study (Paper VI) confirm the importance of these residues in proteolysis. However, the third zinc ligand, distal Glu177 was previously unknown. It is also essential for catalysis as demonstrated by the inactive E177Q mutant. Despite the fact that the protease was crystallized in the absence of a substrate, the electron density revealed a peptide of six residues bound in the active site. This substrate peptide was helpful in order to identify amino acid residues participating in the substrate binding and catalysis (Paper III).
An interesting finding was the presence of C-terminal residues, Arg848 and Tyr854 at the active site, residues separated by almost 800 residues in sequence from the inverted zinc binding residues. Arg848 forms a hydrogen bond to the main chain oxygen of P2’, while Tyr854 binds to the main chain oxygen of the scissile bond. Substitution of Arg848 for Ala or Lys and Tyr854 to Phe resulted in the abolition of the AtPreP1 activity. The catalytic site is present inside a large proteolytic chamber surrounded by the two enzyme halves. The chamber has a volume of more than 10 000 Å³. The chamber appears spacious enough to hold peptide substrates such as the targeting peptide, but is sufficiently small to exclude larger, folded proteins. Since the active site includes residues from both the N- and C-terminal part of the protein, proteolysis can occur only when the chamber is closed. Thus the proteolytic chamber protects against unwarranted degradation by preventing folded proteins from entering the active protease and limiting the size of the substrate. This is also supported by another study showing that AtPreP1 only recognized and cleaved peptides of approximately 10 and 65 amino acids residues in length (Paper VII). The way the substrates are cleaved inside a chamber is reminiscent of the proteasome and therefore PreP was named a Peptidasome.

**Mechanism of Proteolysis by PreP Peptidasome (Paper III and IV)**

The PreP structure presented the first, substrate-bound conformation of an M16 protease, where proteolysis of the substrate takes place inside a chamber formed by two halves of the protease connected by a hinge region. There is no hole or cavity present in the structure. How do substrate peptides as long as 65 amino acid residues access the active site that is situated inside the proteolytic chamber? An open conformation of the AtPreP1 was modeled using the known crystal structure of *S. cerevisiae* MPP (Taylor et al., 2001). The modeled conformation suggested that the closed and open states may differ by as much as a 38° rotation of the two halves around the hinge. A hinge bending mechanism is suggested where the unbound state is open and substrate binding triggers a movement that brings the two halves of the enzyme together so that Arg848 and Tyr854 can complete the active site and stabilize the transition state. While the opening and closing of the peptidasome revolves around the hinge, it seems that electrostatic forces could be driving these movements. The electrostatic surface potential for AtPreP1 E80Q showed that the protease is very acidic, especially around the active site. As the two negatively charged halves repel each other, the peptidasome stays open. However, when a substrate containing basic residues is bound, some of the negative charge is neutralized allowing the chamber to close. It is possible that cations such as magnesium
reduce the negative charge and further favour the closed conformation. After proteolysis, the cleaved products are released from the active site and the interactions with Arg848 and Tyr854 are broken, signaling adoption of the unbound state to the hinge region. The two negatively charged inner surfaces then contribute to push the two halves open by repulsion.

The mechanism involving the opening and closing of the enzyme was investigated by introducing disulfide bonds between the two halves and locking the enzyme in a closed conformation. Disulfide bonds are expected to form under oxidizing conditions, thus locking the enzyme in a closed conformation, whereas it would be able to open and close normally under reducing conditions due to the absence of these disulfide bonds. Consequently, mutants with appropriately positioned cysteine pairs would be inactive under oxidizing conditions and fully active under reducing conditions. Four double mutants (C1-C4) of AtPreP1 with a pair of cysteine residues, predicted to form the disulfide bridges, were created by site directed mutagenesis. Catalytic activity was tested under both conditions using the mitochondrial targeting peptide and the P1 peptide as substrates. The proteolytic activity of the wild type was normal under all of the given conditions. The activity of all four cysteine double mutants was normal under reducing conditions. Interestingly, under oxidizing conditions three of the four mutants (C1-C3) were catalytically inactive, whereas the C4 mutant had almost normal activity. The disulfide bond in the C4 mutant is far away from the active site, in a position where no significant conformational changes are expected during opening and closing. This may explain why this mutant form showed no effect on catalytic activity when locked by a disulfide bond. The remaining mutations are located in regions where larger conformational changes are expected and the suppression of proteolytic activity by the introduced disulfide bonds supported this. Under oxidizing conditions, the disulfide bonds restricted the opening of the protease so that the substrate is unable to reach the active site. Modification of the cysteine residues by N-ethylmaleimide (NEM) showed that the change in activity is directly linked to the formation of disulfide bonds. Taken together, these results verify that neither the cysteine substitutions per se, nor the addition of oxidizing or reducing agents, have a negative impact on proteolysis. In conclusion, these results supported the proposed mechanism involving opening and closing of the enzyme in response to substrate binding. The crystal structure of human IDE has also been recently solved at 2.25 Å resolution (Shen et al., 2006). The active site of IDE is also located inside a cavity, as in the PreP Peptidasome. The structure and mechanism of proteolysis of IDE closely resembles that of PreP and a similar opening and closing mechanism is suggested (Shen et al., 2006).
The effect of magnesium or calcium ions was investigated on the proteolytic activity of \textit{AtPreP1}, since it was not possible to crystallize \textit{AtPreP1} without these ions. In the crystal structure, two hydrated Mg\textsuperscript{2+} ions were found in each protein molecule, nicely coordinated by acidic residues. The position of the Mg\textsuperscript{2+}-binding sites suggested that binding of magnesium ions might have an effect on the conformation of the fourth domain. The position of this domain is crucial for proteolytic activity. In order to determine the impact of cations for the proper function of \textit{AtPreP1}, the proteolytic activity was studied using various concentrations of MgCl\textsubscript{2} and CaCl\textsubscript{2}. \textit{AtPreP1} was inactive in the absence of MgCl\textsubscript{2} or CaCl\textsubscript{2}. The addition of MgCl\textsubscript{2} or CaCl\textsubscript{2} resulted in restoration of the proteolytic activity with full activation at about 10 mM concentrations of MgCl\textsubscript{2} or CaCl\textsubscript{2} (Paper III). These results suggested the requirement of additional metal ions beside Zn\textsuperscript{2+} for the activity of \textit{AtPreP1}. Substituting residues involved in Mg\textsuperscript{2+}-binding sites in the \textit{AtPreP1} shows severe effects on the proteolytic activity of the \textit{AtPreP1} (Bäckman and Bhushan, unpublished results).

**The role of the PreP peptidasome in the degradation of the amyloid β-peptide: A possible link to the Alzheimer’s disease**

**Alzheimer’s disease (AD) and amyloid β-peptide**

Alzheimer’s disease (AD), also known simply as Alzheimer’s is a neurodegenerative disease characterized by progressive cognitive deterioration together with declining activities of daily living and neuropsychiatric symptoms or behavioral changes. It is the most common type of dementia. The most striking early symptom is the loss of short term memory (amnesia), which usually manifests as minor forgetfulness that becomes steadily more pronounced with illness progression, with relative preservation of older memories. As the disorder progresses, cognitive (intellectual) impairment extends to the domains of language (aphasia), skilled movements (apraxia), recognition (agnosia), and those functions (such as decision-making and planning) closely related to the frontal and temporal lobes of the brain as they become disconnected from the limbic system, reflecting extension of the underlying pathological process. These changes affect essential human qualities, and thus AD is sometimes described as a disease where the victims suffer the loss of the very qualities that define human existence (as reviewed by Thomas and Fenech, 2006).

Alzheimer’s disease has been identified as a protein misfolding disease due to the accumulation of abnormally folded amyloid beta (Aβ) peptide in the brain of AD patients (Hashimoto, et al., 2003). Amyloid beta is a short peptide that is an abnormal proteolytic
byproduct of the transmembrane protein amyloid precursor protein (APP) (Figure 3), whose function is unclear, but thought to be involved in neuronal development (Kerr and Small, 2005). The $\gamma$-secretase complex is involved in APP processing and production of A$\beta$-peptides (Cai et al., 2003). Although A$\beta$ monomers are soluble and harmless, they undergo a dramatic conformational change at sufficiently high concentration to form a beta sheet-rich tertiary structure that aggregates to form amyloid fibrils (Ohnishi and Takano, 2004) that deposit outside neurons in dense formations known as senile plaques or neuritic plaques, in less dense aggregates as diffuse plaques, and sometimes in the walls of small blood vessels in the brain in a process called amyloid angiopathy or congophilic angiopathy. AD is also considered a tauopathy due to abnormal aggregation of the tau protein, a microtubule-associated protein expressed in neurons that normally acts to stabilize microtubules in the cell cytoskeleton. Like most microtubule-associated proteins, tau is normally regulated by phosphorylation however, in AD patients, hyperphosphorylated tau accumulates as paired helical filaments (Goedert et al., 2006) that in turn aggregate into masses inside nerve cell bodies known as neurofibrillary tangles and as dystrophic neurites associated with amyloid plaques.

**Amyloid $\beta$-peptide in mitochondria**

Extracellular plaque formation of A$\beta$ has been the main focus of molecular studies associated with AD (Selkoe, 1999). However, there are reports indicating intracellular events including the mitochondrial role in AD (Glabe, 2001). There are many links between mitochondrial dysfunctions and AD (Hashimoto et al., 2003; Yan and Stern, 2005). Impairment of mitochondrial energy metabolism and altered cytochrome c oxidase activity are among the earliest detectable defects in AD (Anandatheerthavarada et al., 2003; Cardoso et al., 2004). It has been shown that Alzheimer’s amyloid precursor protein 695 (APP) is not only targeted to the plasma membrane, but also to mitochondria (Anandatheerthavarada et al., 2003). Accumulation of APP in the outer mitochondrial membrane caused dysfunctions and impaired energy metabolism. The active $\gamma$-secretase, which cleaves APP to generate A$\beta$, has been shown to be present in the mitochondrial membrane (Hansson et al., 2004). Furthermore, the occurrence of A$\beta$ in mitochondria of AD patients and its direct binding to amyloid $\beta$-binding alcohol dehydrogenase (ABAD) induces apoptosis and free radical generation in neurons (Lustbader et al., 2004). A recent study demonstrated that A$\beta$ is present in the mitochondrial matrix in AD brains and in brains from transgenic mice overexpressing mutant human APP that impairs neuronal function and contributes to cellular dysfunction in AD (Caspersen et al., 2005).
Figure 3. Amyloid β-peptide production. APP is a membrane protein producing a number of isoforms which range in size from 695–770 amino acids. Proteolysis of the APP protein involves α-, β- and γ- secretases. APP cleavage by α-secretase releases sAPPα from the membrane leaving an 83 amino acid APP fragment. Cleavage of the APP protein by β-secretase releases sAPPβ from the membrane and leaves behind a 99 amino acid fragment which can be further cleaved by γ-secretase to produce Aβ40/42 fragments extracellularly (Thomas and Fenech, 2006).

PresequenceProtease in human - The hPreP

Interestingly, PreP is an organellar functional analogue of the human insulin degrading enzyme (IDE) that also belongs to the pitrilysin family. IDE has been implicated in AD as it cleaves Aβ-peptide before insoluble amyloid fibers are formed (Tanzi et al., 2004). These findings led us to investigate the degradation of Aβ by human PreP (hPreP). hPreP comprises 1037 amino acid residues with a predicted 29 amino acid residue N-terminal mitochondrial targeting signal. It has been previously identified and referred to as a metalloprotease, hMP1 (Mzhavia et al., 1999).
A novel function of hPreP in mitochondria: Aβ degradation (Paper V)

The predicted mature part of the human PreP (hPreP) was cloned as a fusion protein with GST, and the GST-hPreP fusion protein was overexpressed in E. coli. Recombinant hPreP was purified to homogeneity on GSTrap FF column after cleavage with PreScission protease (Paper V). The recombinant hPreP completely degraded both Aβ-(1-40) and Aβ-(1-42) as well as Aβ Arctic peptide (1-42 E22G). We compared the proteolytic activity of the recombinant hPreP to IDE and found that, unlike IDE, hPreP was not able to degrade insulin. This makes hPreP very interesting and a better candidate over IDE to use for Aβ-peptide degradation in situ. We show that neither PMSF nor bestatin (i.e. serine or aminopeptidase type protease inhibitors) affected proteolytic activity of the recombinant hPreP, whereas NEM, a cysteine-type protease inhibitor, showed some inhibitory effect (about 10%) and the metalloprotease inhibitor ortho-phenanthroline (o-Ph) completely inhibited degradation of Aβ, demonstrating that hPreP is a thiol-sensitive metalloprotease. Apyrase had no effect on the proteolytic activity showing that Aβ degradation is independent of ATP. hPreP contains an inverted zinc-binding motif, H75ILE78H79. The importance of the metal-binding motif for proteolytic activity was investigated by studying the recombinant hPreP mutant, hPreP(E78Q), in which the catalytic base Glu78 was changed to Gln. Overexpressed and purified hPreP(E78Q) did not degrade the Aβ-peptide, confirming the importance of the inverted zinc-binding site for the proteolytic activity.

PreP was shown to be localized to the mitochondrial matrix in plants and mammals (Paper I, II and V). Kambacheld et al. (2005) have shown that PreP homologue in yeast, MOP112 is localized to the IMS where it functions in degradation of the shorter peptide fragments generated after ATP-dependent proteolysis. This indicates that PreP has different subcellular localizations in different organisms; it is present in the mitochondrial matrix in mammals and plants, while it is present in the IMS in S. cerevisiae. In situ, immunoinactivation of PreP in human brain mitochondria, using anti-hPreP antibodies, revealed complete inhibition of the proteolytic activity against Aβ. These results show that under circumstances when Aβ is present in the mitochondria, hPreP is the protease responsible for degradation of this toxic protein. In conclusion, PreP is localized to the mitochondrial matrix in mammalian mitochondria where, beside presequence and other unstructured peptide degradation, it has a novel function: the degradation of Aβ. These findings contribute to studies of the mitochondrial component in AD.
The structure of \textit{AtPreP1} (Paper III) allowed molecular modeling of hPreP and the identification of important amino acid residues involved in substrate binding and proteolysis (Paper V). Unexpectedly, hPreP contains two native cysteine residues in close proximity to each other at position Cys90 in the first domain and position Cys527 at the hinge region. These cysteine residues are conserved in all known mammalian PreP sequences as well as in the \textit{S. cerevisiae} PreP homologue. Interestingly, proteolytic activity of hPreP was almost abolished against the Aβ-peptide in the presence of oxidizing agents, indicating a disulfide bridge formation between the Cys90 and Cys527 that locks the enzyme in a closed conformation and limits substrate access to the active site. This finding was further supported by the full proteolytic activity of hPreP(C90S) mutant under oxidizing conditions. Furthermore, the importance of the status of cysteine residues for the proteolytic activity of hPreP may explain the partial inhibitory effect of NEM on hPreP activity. As NEM is a cysteine-modifying agent, its substitution may cause steric hindrance during proteolysis. The importance of the redox status of cysteine residues for proteolytic activity implies a possible inhibition of the enzyme in mitochondria under conditions of elevated reactive oxygen species (ROS) production that would additionally increase mitochondrial dysfunction. The physiological consequences of this finding require further study.
Future perspectives

Targeting peptides are essential for directing a protein to its final destination in mitochondria and chloroplasts. However, once a protein has reached its final destination the targeting peptide is no longer needed and may also be harmful to the organelle because of its membrane disrupting properties. PresequenceProtease (PreP) was initially identified as the protease responsible for clearing the mitochondria from free presequence peptides generated after the import. PreP was also found to be present in the chloroplasts to take care of free transit peptides. (Paper I and II). We show that PreP is not specific for targeting peptides, but is a general protease that degrades shorter and unstructured toxic peptides in these organelles (Paper VI). The crystal structure of \textit{At}PreP1 presented the first substrate-bound, closed conformation of a protease from the pitrilysin family (Paper III). Based on the structure and proteolytic activity of cysteine mutants designed to lock the PreP in a closed conformation, a novel hinge bending opening/closing mechanism of proteolysis is proposed (Paper III and IV). PreP is localized to the mitochondrial matrix in mammals where, beside degradation of presequences, it has a novel function i.e. degradation of toxic Aβ-peptides present in Alzheimer’s patients (Paper V).

We have been able to carry out a thorough investigation of the PreP by showing its dual localization and function in mitochondria and chloroplasts, the structure and mechanism of proteolysis and a possible link to Alzheimer’s disease. However, there are still as many or even more interesting questions that remain to be answered as there were prior to beginning this work.

How is a dual targeting peptide sorted and imported in both mitochondria and chloroplasts? It seems that dual targeting peptides have some common features of both the mitochondrial and chloroplastic targeting peptides, but there is no structural information available as yet for these targeting peptides. It will be interesting to solve the 3D structure for a dual targeting peptide alone, and in complex with the mitochondrial and chloroplastic import receptors, in order to understand the molecular mechanism of dual targeting. A short dual targeting peptide has been identified with promising results to be used for 3D structural studies by NMR (Bhushan and Berglund, unpublished results).

What is the significance of higher transcript level of \textit{At}PreP1 during flowering? Expression of the many proteases can be induced during stress conditions. During the late stage of flowering, the state resembles desiccation, which is also a type of stress and the higher level of \textit{At}PreP1 transcripts may have an important role under these conditions. It will be
interesting to study the transcript level of PreP under different stress conditions. Stefan Nilsson and Hans Bäckman in our laboratory have succeeded in generating a double knock out mutant of PreP1 and PreP2 in *A. thaliana*, which will enable to investigate the role of PreP in detail. How distribution of PreP between mitochondria and chloroplasts is regulated, is another important issue which should be investigated.

Another interesting issue is the different sub-mitochondrial localization of PreP in yeast compared to mammals and plants. This can be studied either by shuffling the presequences between plants, mammals and yeast, or by directing yeast PreP to the matrix. It will be also interesting to study whether yeast PreP can be imported into the mammalian and plant mitochondria and vice versa.

One basic question is what are the end products of proteolysis catalysed by PreP. Is it single amino acids or small peptide fragments? From the mass spectrometric analysis it seems that the main products of proteolysis are small fragments. It is also possible that these are the remaining fragments that are resistant to further degradation because of their physical properties. In order to solve this problem a detailed MS detection of degradation products of a number of different substrate peptides should be performed.

Next question is how the PreP structure will look in an open conformation. We have modelled an open conformation of the PreP protease on the basis of the known structure of *S. cerevisiae* MPP. However, it is a model and the native PreP structure in an open conformation is not known. It will be of great interest to crystallize and solve the structure of PreP protease in a substrate-free, open conformation.

Another interesting question will be to study the possible link of human PreP (hPreP) in Alzheimer’s disease. hPreP is capable of degrading toxic Aβ-peptides of various lengths as shown by *in vitro* and *in situ* studies (Paper V). However, these are preliminary results and more *in vivo* studies are needed before coming to a satisfactory conclusion. One possibility is to overexpress Aβ and hPreP together in cell culture and test whether PreP has the capability to degrade the overproduced Aβ *in vivo*. It will be also interesting to knockdown the PreP in cell culture using RNAi, and study the physiological consequences of PreP’s absence on mitochondrial functioning. Since we have been able to solve the structure of *At*PreP1 it should not be difficult to solve the crystal structure of hPreP alone, and in a complex with Aβ-peptides. There are plenty of opportunities when it comes to the AD. One can also try to find the activators of hPreP to be used for stimulating the PreP activity *in vivo.*
Acknowledgements

I would like to express my sincere gratitude to all of you who have directly or indirectly been involved in the making of my thesis. It has been great to work with all of the people in the DBB. It is difficult to remember and thank all the people who have helped me during all these years, but I will try my best.

First of all, I would like to express my gratitude to Prof Elzbieta Glaser, my supervisor, for letting my PhD dream come true. She is the best supervisor one can have. Thanks for being such a generous, inspiring and encouraging guide. I have learned much from you and it will help me a lot in days ahead of me.

Patrick Dessi is one of the nicest persons, you will never forget him. He is always willing to help everyone. Thanks for introducing me to Elzbieta’s group and going through the thesis on such a short notice. You are not only a great person, but a saint.

To Prof Mark Boutry, my thanks for accepting me in your lab and helping during my stay in Belgium. Those experiments turned out to be very important for the thesis. Thanks for your encouragement and ideas. Thank you to Benoit also, for teaching me GFP system in such a short time.

To Dr Therese Eneqvist and Dr Ken Johnson my thanks for such a wonderful collaboration and sharing of ideas. It was not possible to come so far in this work without you guys. Thanks a lot!

To Prof Stefan Nordlund, Dean for the graduate studies, my thanks for making DBB a nice place to work. You are always there to help all the PhD students. Thanks for evaluating me so many times.

My thanks to present and former members of EG group: Stefan for your interesting ideas. It was fun to play Poker with you and Tatiana in Linköping. To Anna-Karin and Nyosha for being such cheerful people even when nothing work in the lab. Thanks for making such a nice environment to work in. To Hans for our discussions of stock markets though we never agreed on any thing. To Charlotta, for helping with dual import system. Former EG group members, thanks to Annelie and Per for teaching me how import works. It was so much fun to collaborate with you and share the ideas. Pavel, though we never collaborate directly, you were always ready to discuss anything in science. It was nice to have an experienced person like you around all of these years.

My thanks to Prof Gunnar von Heijne, Prof Åke Wieslander and Prof Mikael Oliveberg for each spending your time evaluating my work and being my examiners.
To Prof Birgitta Norling my thanks for giving me shelter in your house when I did not have any place to live; to Dr Inger and Dr IngMarie for helping in teaching courses; to Dr Jan-Willem, for being my co-supervisor though I never discussed the project with you.

Thanks Dan, Joy, Tara, Mikaela, Filippa, Carolina, Mirjam, Yoko, Karl and rest of the Gunnar’s group for being such nice neighbours. Thank you for letting me borrow so many reagents.

Bogos, thanks for helping me out in everything including translating some of the Swedish documents. Thank you for being around. And thanks to the Secretaries, Anki, Ann, Lotta, Sofia and Peter and Eddie for making sure that everything works in the DBB.

I am indebted to my parents in India. Thank you for all your help and support. It would have never been possible to come to this point without your support. I would also like to thank all of my relatives in India for their understanding and support.

My wife Anu, for sharing life with me and giving me the most precious jewel of the world, our little Dixie. I can not imagine life without the two of you.
References


Corral-Debrinski, M., Blugeon, C. and Jacq, C. (2000) In yeast, the 3' untranslated region or the presequence of AT1M is required for the exclusive localization of its mRNA to the vicinity of mitochondria. Mol Cell Biol, 20, 7881-7892.


Glick, B.S., Brandt, A., Cunningham, K., Muller, S., Hallberg, R.L. and Schatz, G. (1992) Cytochromes c1 and b2 are sorted to the intermembrane space of yeast mitochondria by a stop-transfer mechanism. Cell, 69, 809-822.


