Proteolytic processing of thylakoid proteins

Simon J. Thompson

A thesis submitted for the degree of Doctor of Philosophy

December 1998

Department of Biological Sciences,
University of Warwick,
Coventry.

Summary

Nuclear-encoded thylakoid proteins are targeted into and across the thylakoid membrane by four distinct mechanisms. Precursors of lumenal proteins are translocated by either a Sec- or ΔpH-dependent mechanism. In both cases, thylakoidal processing peptidase (TPP) removes this signal peptide to release the mature protein. A structurally similar signal peptide is present in the fourth pathway used by a subset of integral membrane proteins including PSII-W, PSII-X and PsbY. In this case the integration process does not involve an identifiable energy source or known protein transport machinery. This Sec-independent direct insertion mechanism is unlike any other known, with the exception of a single protein, M13 procoat, which similarly inserts in the *Escherichia coli* plasma membrane by means of a signal peptide.

This novel insertion mechanism has been probed by mutagenesis of TPP cleavage sites within pre-PSII-W, -PSII-X and -PsbY, in order to generate intermediates on the insertion pathway. TPP cleaves preferentially after Ala-X-Ala, and the terminal alanine was mutated to threonine by site-specific mutagenesis of the cDNA.

Chapter 3 - In the case of pre-PSII-W and pre-PSII-X, TPP is the only proteinaceous component known to be involved at any stage of the insertion process. The PSII-W cleavage site mutant is imported and inserted into the thylakoid membrane, but cleavage by TPP is inhibited. Import into chloroplasts results in the accumulation of a mature size protein and an intermediate form that accumulates when TPP fails to complete the maturation (the precursor protein is processed to an intermediate in the stroma). Importantly, the intermediate is located exclusively in the thylakoid membrane, confirming that the action of TPP is not required for correct localisation. Protease-topology mapping presented, shows that this intermediate is in the form of a loop intermediate in which both the N- and C-termini are exposed on the stromal face of the membrane with the intervening region on the lumenal face. Preliminary data is also presented for a loop intermediate with the insertion of pre-PSII-X.

Chapter 4 - Precursor PsbY has a more complex insertion mechanism and 6 mutants are presented that show various intermediates on the pre-PsbY insertion pathway. All the intermediates are stable in the thylakoid membrane. Furthermore, using this approach, it can be shown that pre-PsbY is in fact a polyprotein containing two similar single-span proteins, both of which are preceded by cleavable signal peptides. Pre-PsbY thus contains a single chloroplast targeting peptide, an initial thylakoid signal peptide and mature protein followed by a second thylakoid signal peptide and mature protein. This is therefore the first nuclear-encoded polyprotein targeted to the thylakoid membrane in higher plants, and the data suggest that this protein inserts as a double-loop form.

Table of contents

Chapter 2 - *Materials and methods*

List of figures and tables

Acknowledgements

Prof. Colin Robinson - for his humour and innumerable help throughout this PhD.

Thanks go to the group, the Robinson clan, who have helped so much.

A big thank you to my friends and colleagues - Alistair, Erik, Soo, Pete, Wayne, Alex, Susanne, Suzanne, Dave and all the others who have helped in their own ways throughout this PhD.

And finally to my parents, whom I owe my abilities and existence - many, many thanks.

Declaration

Work presented in this thesis is original work, conducted by myself under the supervision of Prof. C. Robinson.

The research was funded by a BBSRC studentship.

All sources of information have been acknowledged by means of reference. None of this work has been used in any previous application for a degree.

The following people are gratefully acknowledged for their help:
S. J. Kim (Warwick) for cDNA clones of *A. thaliana* wild-type pre-PSII-W and -X.
Dr. A. Mant (Warwick) for a cDNA clone of *A. thaliana* wild-type pre-PsbY and also for the use of the construct encoding PsbY pre-A2.

Some of the work presented in this thesis has been published in, or submitted to, the following journals:

Thompson, S. J. and Robinson, C. (1997). "The Role of Thylakoid Processing Peptidase in Chloroplast Protein Import". *Proteolysis in Cell Functions*. Edited by Hopsu-Havu, V.K, Jarvinen, M and Kirschle, H. IOS Press

Thompson, S. J., Kim, S. J. and Robinson, C. (1998). Sec-independent insertion of thylakoid membrane proteins: Analysis of insertion forces and identification of a loop intermediate involving the signal peptide. *J. Biol. Chem.* **273** 18979-18983

Thompson, S. J., Mant, A. and Robinson, C. (1999). Dual signal peptides mediate the SRP/Sec-independent insertion of a thylakoid membrane polyprotein, PsbY. *J. Biol. Chem.* **in press**

Abbreviations

ADP	adenosine diphosphate
amp	ampicillin
Arg	arginine
A. thaliana	*Arabidopsis thaliana*
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
BiP	binding protein
bis-acrylamide	N'N'-methylene-bisacrylamide
bp	base pair/s
BSA	bovine serum albumin
CCCP	carbonyl cyanide *m*-chlorophenylhydrazone
cDNA	complementary DNA
CF_0II	chloroplast F_0 ATP synthase subunit II
Ci	Curie/s
cm	centimetre/s
cpDNA	chloroplast DNA
cpm	count/s per minute
cpn60, cpn10	chaperonins 60 and 10
C. reinhardtii	*Chlamydomonas reinhardtii*
C-terminal	carboxyl-terminus
Cys	cysteine
cyt	cytochrome
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
ddATP	2', 3'-dideoxyadenosine 5'-triphosphate
ddCTP	2', 3'-dideoxycytidine 5'-triphosphate
ddGTP	2', 3'-dideoxyguanosine 5'-triphosphate
ddTTP	2', 3'-dideoxythymidine 5'-triphosphate

DEAE-	diethylaminoethyl-
DHFR	dihydrofolate reductase
$\Delta\mu_{H+}$	membrane electrochemical potential
Δp	protonmotive force
ΔpH	proton concentration gradient
$\Delta\Psi$	membrane potential difference
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNAaseI	deoxyribonuclease I
dNTP/s	deoxyribonucleoside triphosphates/s
ds	double stranded
DTT	1,4-dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid (disodium salt)
ER	endoplasmic reticulum
F_1ATPase	mitochondrial F_1ATPase
Fd	ferredoxin
Fe/S	Rieske iron-sulphur protein
FNR	ferredoxin-NADPH reductase
g	gramme/s
gd	gapped duplex (DNA)
GEP	general export pathway
GTP	guanosine triphosphate
h	hour/s
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
His	histidine
HM	Hepes-$MgCl_2$ buffer
HMC	Hepes-$MgCl_2$-$CaCl_2$ buffer
HME	Hepes-$MgCl_2$-EDTA buffer
hsp	heat shock protein
i-	intermediate
IgG	immunoglobulin G

IM	inner membrane
IMP-1	inner membrane protease 1
IMS	intermembrane space
kb	kilobase/s
kDa	kiloDalton/s
l	litre/s
L-agar	L-agar culture medium (Sambrook *et al.*, 1989)
LB	Luria-Bertani medium (Sambrook *et al.*, 1989)
LP	leader peptidase
Leu	leucine
LHCP	light-harvesting chlorophyll *a/b* binding protein
LSU	rubisco large subunit
Lys	lysine
M	molar
µCi	microCurie/s
µg	microgramme/s
µl	microlitre/s
µM	micromolar
µmol	micromole/s
mA	milliAmpère/s
MBP	maltose binding protein
MDa	megaDalton/s
Mes	2-(morpholino)ethanesulphonic acid
Met	methionine
mg	milligramme/s
min	minute/s
ml	millilitre/s
mM	millimolar
mRNA	messenger RNA
MW	molecular weight
NADH	nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	nicotinamide adenine dinucleotide phosphate (oxidised form)

NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NBD	nucleotide binding domain
N. crassa	*Neurospora crassa*
ng	nanogramme/s
nm	nanometre/s
nM	nanomolar
nmol	nanomole/s
N-terminal	amino-terminal
NTP/s	nucleoside triphosphate/s
OD	'optical density'
OEC	oxygen-evolving complex
OM	outer membrane
OmpA	outer membrane protein A
PC	plastocyanin
PCR	polymerase chain reaction
POR	NADPH-protochlorophyllide reductase
PPO	polyphenol oxidase
PQ	plastoquinone
pre-	precursor
PSI	photosystem I
PSII	photosystem II
Psb	photosystem II
PSB	protein sample buffer
PTS	peroxisomal targeting signal
P680	reaction centre of photosystem II
P700	reaction centre of photosystem I
rATP	adenosine ribonucleotide triphosphate
rCTP	cytidine ribonucleotide triphosphate
rGTP	guanosine ribonucleotide triphosphate
rUTP	uridine ribonucleotide triphosphate
RER	rough endoplasmic reticulum

RNA	ribonucleic acid
RNaseA	ribonuclease A
RNasin	ribonuclease inhibitor
rpm	revolution/s per minute
rubisco	ribulose-1,5-bisphosphate carboxylase/oxygenase
s	second/s
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SP	signal peptidase
SPP	stromal processing peptidase
SRP	signal recognition particle
ss	single-stranded
SSU	rubisco small subunit
TBE	Tris-borate-EDTA
TCA	trichloroacetic acid
TE	Tris-HCl-EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
TPP	thylakoidal processing peptidase
Tris	tris(hydroxymethyl)methylamine
Triton X-100	*iso*-octylphenoxpolyethoxyethanol
u.v.	ultra-violet light
v/v	volume/volume
w/v	weight/volume
[^{3}H]leu	tritium-labelled leucine
[α-^{35}S]dATP	^{35}sulphur-labelled deoxyadenosine 5'-α-thiotriphosphate
[^{35}S]met	^{35}sulphur-labelled methionine
16K	16 kDa OEC protein
23K	23 kDa OEC protein
33K	33 kDa OEC protein
54CP	chloroplast signal recognition particle homologue

Chapter 1

Literature review

Chapter 1

Literature Review

1.1 *Introduction*

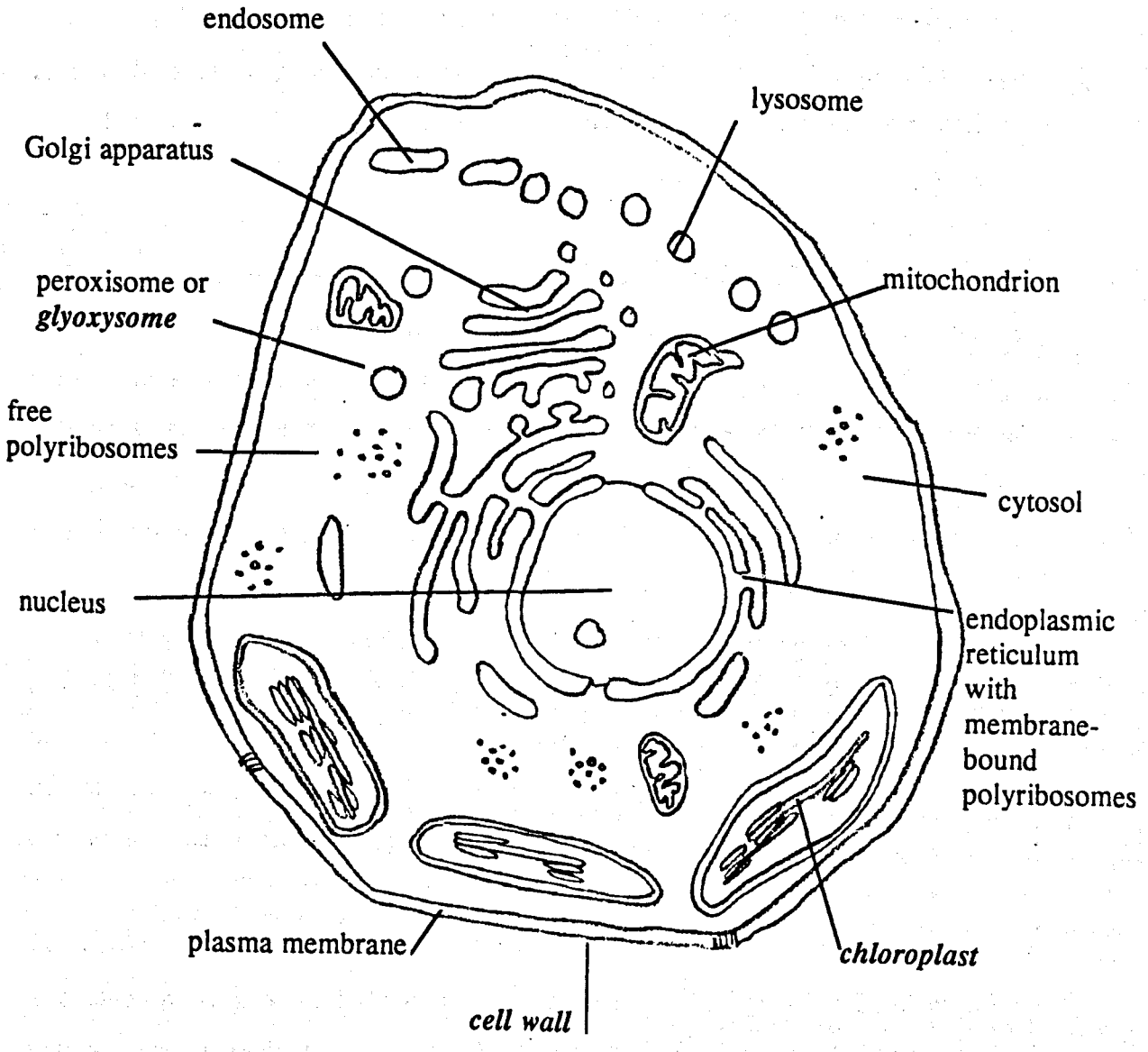
1.1.1 *Major compartments in a eukaryotic cell*

Figure 1 illustrates the major compartments found in the eukaryotic cell. The nucleus is bounded by a double membrane envelope, and contains the major genome of the cell. Surrounding the nucleus is the cytoplasm, which is made up of the cytosol and the organelles which are suspended in it. The cytosol is where most proteins are synthesised, and also where most metabolic interconversions involving small molecules take place. Continuous with the nuclear envelope is the rough endoplasmic reticulum (RER), which then gives way to smooth endoplasmic reticulum (ER). Many ribosomes are bound to the cytosolic surface of the RER, and these are responsible for synthesising integral membrane proteins or proteins which will be modified in the ER and/or secreted from the cell. ER also produces lipid for the cell. Lipids and proteins from the ER may continue their journey *via* the Golgi apparatus cisternae, which make covalent modifications to the molecules before sending them to other intracellular destinations. Endosomes are vesicles which enclose endocytosed molecules and transport them to lysosomes, which in turn degrade the particles, and also digest (recycle) expired organellar components. Peroxisomes are vesicles bounded by a single membrane, and contain enzymes which perform oxidative reactions. In plant cells, another type of peroxisome exists, termed the glyoxysome, responsible for converting fatty acids to carbohydrates in germinating seeds. Mitochondria and chloroplasts are the generators of nearly all energy currency (i.e. ATP) required by the cell to drive biosyntheses. These organelles are surrounded by a double membrane and each contain minor genomes, which encode some of the proteins required for mitochondrial and chloroplast function.

Often the most prominent compartment of the plant cell is the vacuole, which is surrounded by the tonoplast and is structurally and functionally related to lysosomes. It also acts as a metabolite store, pigment depository and support for the plant.

Figure 1 *Compartments of the eukaryotic cell.*

Legends in ***bold italics*** denote those compartments not found in animal cells



Despite its relationship with the organelles of the secretory pathway, the vacuole is not usually considered part of the cytoplasm.

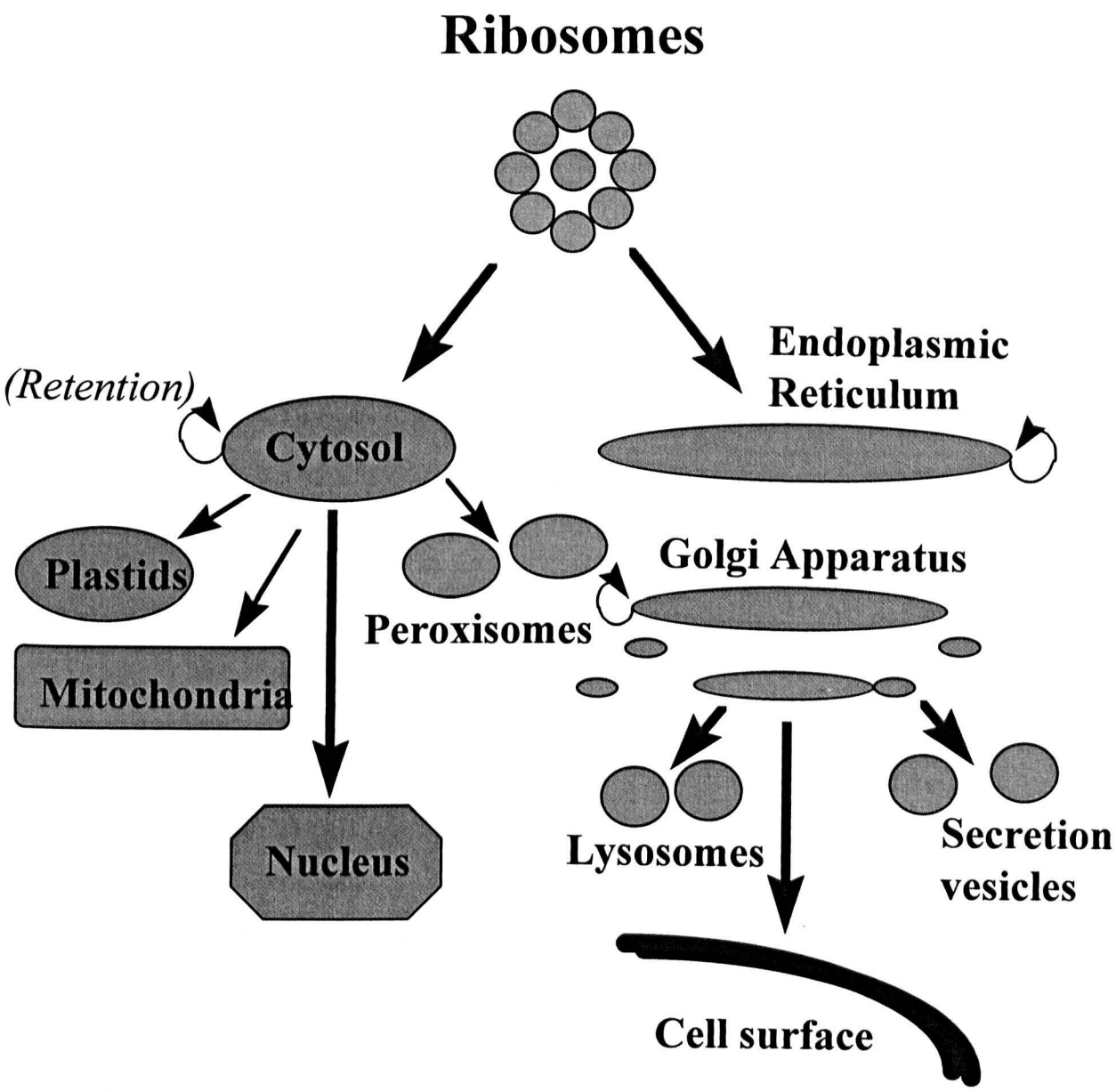
1.1.2 *Intracellular protein sorting*

Since every cellular compartment has a certain specialised function, each must contain a suite of enzymes and other proteins which are specific for that function. Nearly all proteins are synthesised on cytosolic ribosomes, however, and so the cell must direct proteins from their site of synthesis to their final destination. A simplified scheme of the protein traffic flow is shown in **figure 2**. Basically, if a protein is destined for secretion (or any part of that pathway), it is synthesised upon ribosomes which are bound to the ER (Blobel and Dobberstein, 1975), and is then transported co-translationally into the ER lumen. The subsequent journey through the secretory membrane system is carried out *via* transport vesicles between the ER, Golgi apparatus and plasma membrane. Transmembrane proteins destined for the ER, or other membranes remain anchored in the ER membrane by virtue of hydrophobic, membrane-spanning regions. Retention in the ER, Golgi apparatus or sorting to lysosomes and secretory vesicles (for regulated, as opposed to constitutive secretion) are all decisions likely to require specific signals - constitutive secretion is a default process. For proteins translated on cytosolic ribosomes, retention in the cytosol is the default process. Targeting to the nucleus, peroxisomes, chloroplasts and mitochondria all require signals contained within the protein, which are either cleaved by specific processing peptidases, or retained.

1.2 *Chloroplast structure and function*

The chloroplast harvests light energy and is the site of light-stimulated electron transport leading to the synthesis of NADPH and ATP. By capturing light energy, the chloroplast is capable of increasing the total free energy content of living material. It also assimilates carbon from carbon dioxide, transforming the carbon to carbohydrate. This may be stored in the organelle as starch, or exported as a three-carbon compound used by the cell as an energy source, or a starting point for a wide range of biosyntheses. In addition to the unique reactions of light-harvesting and carbon fixation, the chloroplast is able to reduce nitrite to ammonia (for the synthesis

Figure 2 *Movement of proteins between cellular compartments*



of amino acids and nucleotides). All cellular fatty acids, and some amino acids are produced using enzymes in the chloroplast stroma.

The metaphyte chloroplast is a structurally complex organelle, shown by electron microscopy to consist of six different compartments: the outer- and inner- envelope membranes which bound the intermembrane space, the stromal compartment, the thylakoid membrane and the thylakoid lumen. **Figure 3** shows a very generalised view of the various compartments.

1.2.1 *The envelope*

The two envelope membranes contain a mixture of phosphatidylcholine and glycolipids, and are pigmented with carotenoids (carotene and xanthophylls), but not chlorophylls. There are differences, both in composition and in function between the two membranes - the inner membrane is the more dense, due to a greater number of protein particles. The latter membrane is also enriched in monogalactosyldiacylglycerol and devoid of phosphatidylcholine, compared to the outer membrane in which the lipids digalactosyldiacylglycerol and phosphatidylcholine predominate (Douce and Joyard, 1990). The outer membrane is freely permeable to molecules of 10 kDa or less, but the inner membrane acts as a permeability barrier, and as such, contains specific translocators, probably accounting for much of the difference in particle number between the outer and inner membranes. One important example is the phosphate translocator (the gene for which was cloned by Flügge *et al.*, 1989), which catalyses the counter exchange of cytosolic P_i for dihydroxyacetone phosphate, glyceraldehyde-3-phosphate and 3-phosphoglycerate.

1.2.2 *The stroma*

In electron micrographs, the stroma appears granular, and many contain inclusions such as starch granules (up to 2 μm long), droplets of lipid or crystals of rubisco (the major stromal enzyme), which can form under conditions of environmental stress. All the dark reaction enzymes, for fixing carbon dioxide and converting it to carbohydrate, are found in the stroma, as well as some enzymes for the synthesis of

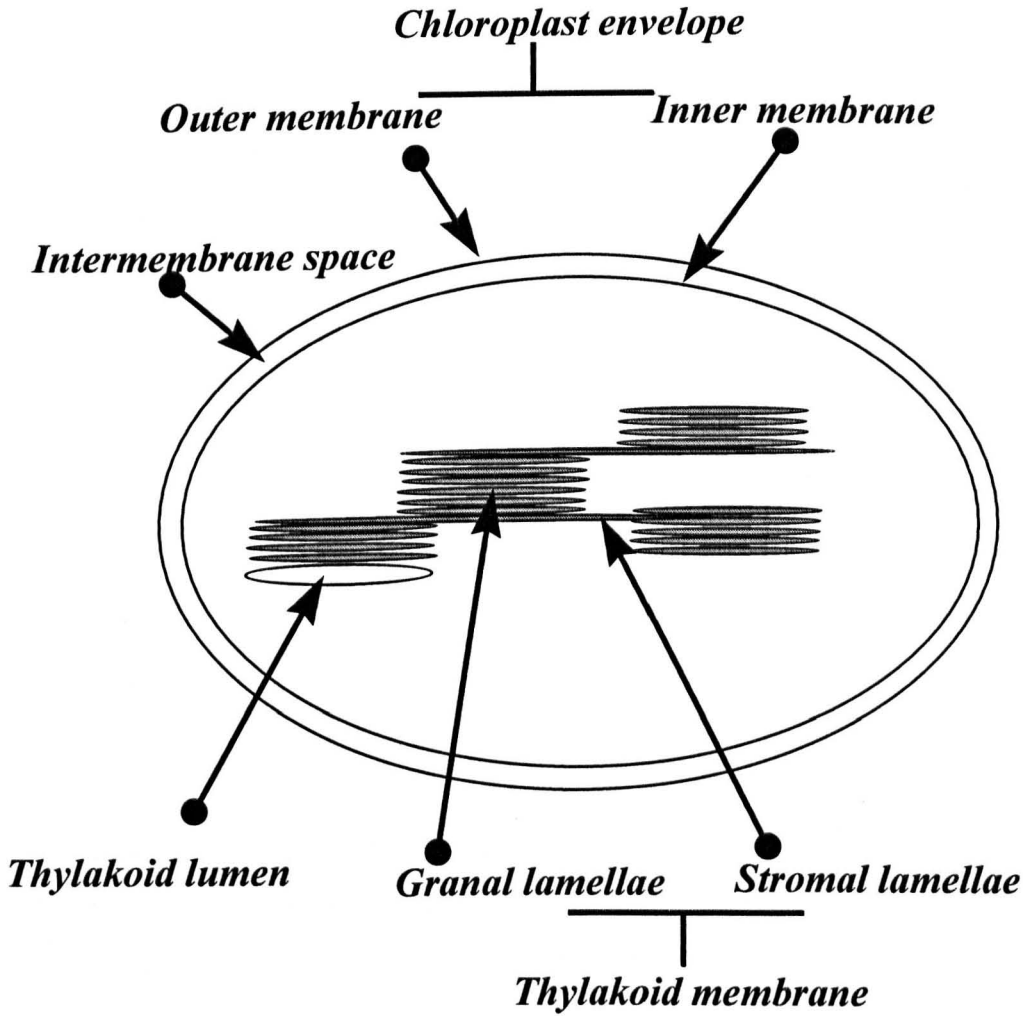


Figure 3 *Generalised structure of the metaphyte chloroplast*

In a 'typical' C3 leaf (such as pea or barley), chloroplasts are 3-10 μm in length, and occupy approximately 40% of the cytosolic volume. The total thylakoid membrane area per m^2 leaf is estimated at between 300 and 835 m^2. The total thylakoidal volume represents around 20% of the chloroplast volume.

The envelope membranes are about 5 - 6nm thick, seperated by an intermembrane space of about 10 nm width.

The thylakoidal membranes are 5-7 nm thick, and the distance between two membranes enclosing a lumenal space is about 8 nm.

Data obtained from Lawlor (1993)

fatty acids and amino acids. The stroma is also the site of chloroplast DNA replication, transcription and protein synthesis. Molecules of chloroplast DNA may often be linked to the thylakoid membranes, forming nucleoid structures, and ribosomes may either function in the stroma, or when attached to the thylakoid membranes.

1.2.3 *The thylakoid network*

The thylakoid network is organised as a series of stacked, flattened disks (granal lamellae), interconnected by pairs of membranes (stromal lamellae). The appressed membranes within the granal stacks are not in contact with the stroma. The membranes enclose the thylakoid lumen, which is thought to be one, single continuous phase (hence thylakoid 'network'). The entire thylakoid system appears to develop by the folding and joining of sheets of membrane which derive from a single origin, the prolamellar body, in proplastids. The continuous nature of the thylakoid lumen is significant for electron transport and generation of the proton gradient required for photosynthesis.

The thylakoid membranes are composed of 50% lipid by mass, of which the majority (80%) is monogalactosyldiacylglycerol and digalactosyldiacylglycerol (Lawlor, 1993). These two abundant lipids have highly unsaturated fatty acid tails, resulting in an extremely fluid membrane at physiological temperatures. This fluidity allows lateral diffusion of pigment-protein complexes within the membrane, essential for optimising the rates of photosynthetic processes.

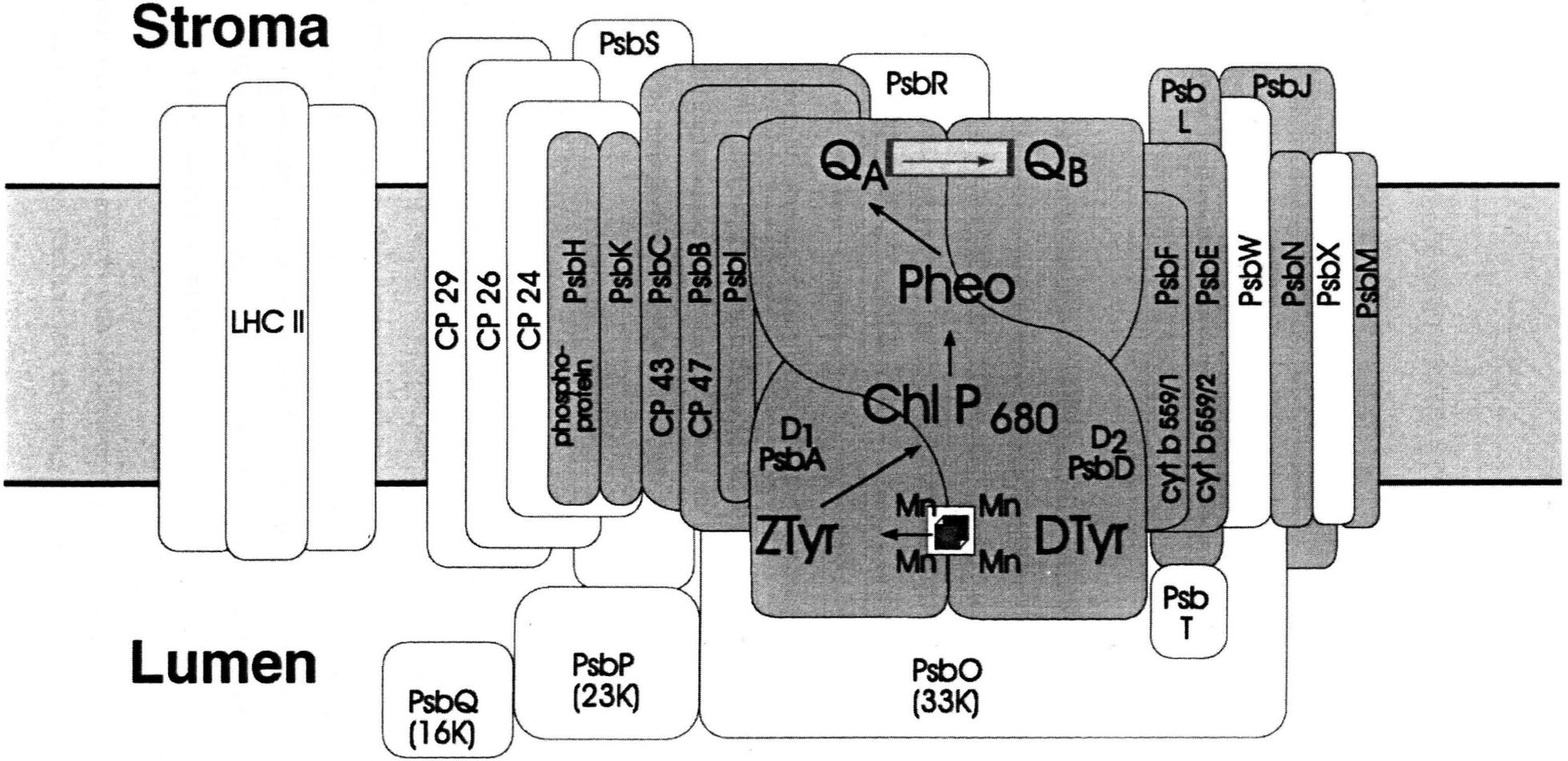
The four major protein complexes involved in photosynthesis are photosystem II (PSII; **figure 4**) with associated antenna complexes and the oxygen-evolving complex, cytochrome b_6f complex, photosystem I (PSI) with its associated antennae and finally, the coupling factor responsible for ATP synthesis.

Figure 4 *The subunits of photosystem II*

In this diagram, shaded subunits are those synthesised within the chloroplast; the remainder are those which have been imported post-translationally from the cytosol. "Psb" is the nomenclature for the subunits of PSII in this particular diagram; "*psb*" is also the name of the relevant gene.

LHC II	light-harvesting complex II
psbO	33 kDa protein of the oxygen-evolving complex
psbP	23 kDa protein of the oxygen-evolving complex
psbQ	16 kDa protein of the oxygen-evolving complex
ZTyr	tyrosine residue 160 on the D1 protein; electron transport donor to P680*
P680	reaction centre, consisting of a pair of closely-interacting chlorophyll *a* molecules
Pheo	Phaeophytin; primary electron acceptor
QA	quinone bound to D2; secondary electron acceptor
QB	quinone; electron acceptor which exchanges with free plastoquinone when reduced

This diagram was kindly provided courtesy of Prof. C. Robinson.


Stroma
Lumen
LHC II
CP 29
CP 26
CP 24
PsbS
PsbR
PsbH
PsbK
PsbC
PsbB
PsbI
phospho-protein
CP 43
CP 47
Q_A
Q_B
Pheo
Chl P_{680}
D1
PsbA
D2
PsbD
ZTyr
DTyr
Mn
Mn
Mn
Mn
PsbL
PsbJ
PsbF
PsbE
cyt b 559/1
cyt b 559/2
PsbW
PsbN
PsbX
PsbM
PsbT
PsbQ
(16K)
PsbP
(23K)
PsbO
(33K)

1.3 *Photosynthesis*

Photosynthesis is the conversion of solar energy into the chemical energy of organic molecules. For metaphytes, the process may be summarised by the following equation:-

$$CO_2 + H_2O + \text{light energy} \xrightarrow[\text{Plants}]{\text{Chlorophyll-containing}} (CH_2O) + O_2 + \text{chemical energy}$$

The chemical energy in question is contained within the intramolecular bonds of ATP and NADPH. Photosynthesis, however, is not only the assimilation of carbon dioxide with concomitant oxygen evolution - other organisms may use different sources of reducing electrons. For example, green and purple sulphur bacteria do not evolve oxygen, and utilise hydrogen sulphide, sulphur, hydrogen and organic compounds as their source of reductant. Cyanobacteria are capable of assimilating atmospheric nitrogen to form ammonia, a process which cannot be carried out by eukaryotic cells. As the protein substrates used during this project originate from metaphyte chloroplasts, the remainder of the discussion will concentrate upon the structure and function of the metaphyte photosynthetic apparatus.

1.3.1 *The light reaction*

The light reaction of photosynthesis is a series of electron transfers (stimulated by light energy), from water to $NADP^+$, with concomitant evolution of oxygen, and the transfer of protons across the thylakoid membrane into the lumen to form a protonmotive force (Δp). This Δp drives protons back across the thylakoid membrane through the chloroplast ATP synthase complex, to produce ATP. The electron transfers are summarised, in a simplified form, in **figure 5**. An alternative, cyclic form of electron transport may occur without splitting water and releasing oxygen - this is discussed later in this section.

Electron transport commences with the capture of photons of light by chlorophylls and accessory pigments (carotenoids, chlorophylls *a* and *b*). Excitation energy is

Figure 5 *The Z-scheme of photosynthetic electron transport*

A simplified diagram of electron transport in relation to the redox potentials of the electron carriers, during the process of non-cyclic photophosphorylation. ATP synthesis is shown coupled to the stages where protons are either generated in the thylakoid lumen, or pumped into that compartment.

P680 represents the reaction centre of photosystem II

Q represents quinone acceptors QA and QB

PQ is plastoquinone

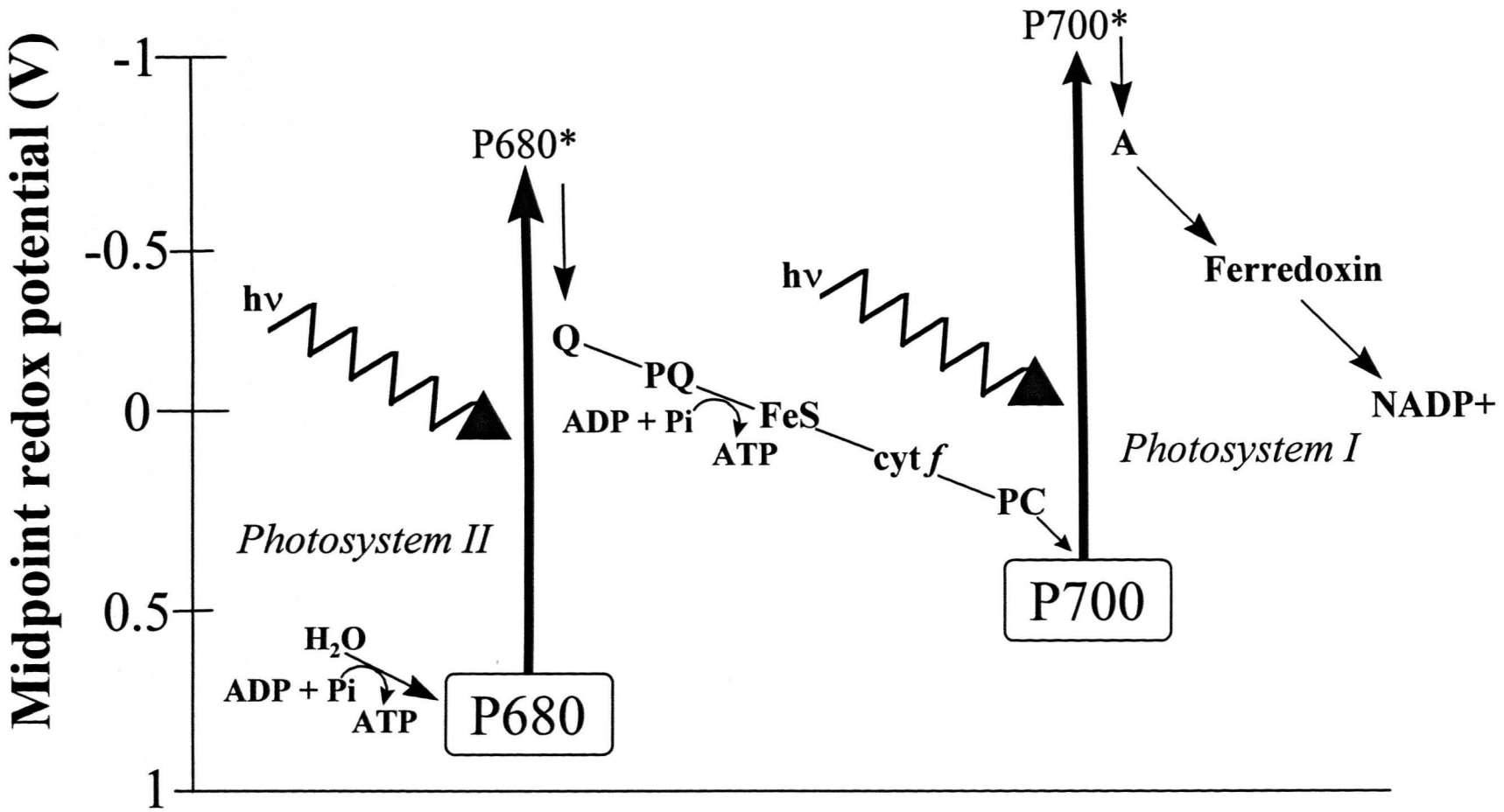
Fe-S is the Rieske iron-sulphur centre from the cytochrome b_6f complex

cyt*f* represents cytochrome f

PC represents plastocyanin

P700 represents the reaction centre of photosystem I

A denotes the combination of electron-carrying intermediates A_0, A_1 and F_x, which are a chlorophyll a monomer, a phylloquinone and a 4Fe-4S cluster, respectively.


Midpoint redox potential (V)
-1
-0.5
0
0.5
1
P680*
hν
Q
PQ
ADP + Pi
ATP
FeS
cyt f
PC
P700*
A
Ferredoxin
NADP+
hν
Photosystem I
Photosystem II
H2O
ADP + Pi
ATP
P680
P700

transferred to the reaction centres of PSII and PSI, which contain special dimers of chlorophyll molecules. From these are ejected electrons which are rapidly removed from the donor by a series of acceptors, which become alternately reduced and oxidised as the electrons are passed along. Excitation of P680 (the reaction centre of PSII) yields an activated reaction centre, P680*, and a high-energy electron, which is passed to the primary acceptor, a phaeophytin molecule, which is bound to the core proteins D1 and D2. The rapid transfer of the electron to a secondary acceptor, Q_A, contributes to preventing the electron recombining with 680^+. Q_A is a quinone molecule bound to a polypeptide in PSII, and it passes the electron to another quinone, Q_B, which is protein-bound when oxidised, but exchanges freely with the pool of plastoquinone when reduced. Q_B is an important carrier for two reasons: it can transfer two electrons at a time, and is the means by which the high-energy electrons may be spatially removed from the reaction centre, thus completing the light energy trap.

The $P680^+$ reaction centre is a powerful oxidising agent, capable of removing electrons from water. It does this by first oxidising a tyrosine residue (termed Z) on the D1 protein, which in turn is reduced by electrons from the water-splitting/oxygen evolving complex (OEC). This complex consists of the transition metal manganese (four atoms per PSII complex), stabilised on the photosystem core by an extrinsic lumenal 33kDa protein (33K); and two other extrinsic proteins of 23kDa and 16kDa (23K and 16K respectively). The OEC functions by feeding electrons to Z and cycling through four intermediate states (S_1 to S_4), which are likely to represent oxidation states of the manganese atoms (Lawlor, 1993). Four photons must be captured by the light-harvesting apparatus (and therefore, four electrons ejected from P680) before the S_4 state is reached. At this point, S_4 reacts with two molecules of water to yield S_0, four protons and one molecule of gaseous oxygen. Therefore, S acts as if it were a charge accumulator, although this is not literally the case, since protons are released during the cycle, having the effect of neutralising accumulated positive charges on S (Lawlor, 1993).

Plastoquinone is a mobile electron carrier, capable of collecting protons on the stromal side of the thylakoid membrane when it is reduced, and releasing them into

the thylakoid lumen when it is oxidised. Thus, it contributes to the formation of the proton gradient across the thylakoid membrane. Plastoquinone and the cytochrome b_6f complex may transport electrons by what is known as the "Q-cycle". A pair of electrons is thought to be donated from plastoquinone to the cytochrome b_6f complex; one being transferred to plastocyanin (another mobile electron carrier) *via* a Rieske iron-sulphur protein and cytochrome *f*, and the other one transferred to cytochrome b_{563} (b_6). The latter electron may then be transported to the stromal side of the thylakoid membrane, where two electrons and two protons from the stroma reduce a bound plastoquinone molecule, which dissociates and joins the pool of reduced plastoquinone. In this way, the cycling of electrons should double the number of protons transferred across the membrane per electron transported to PSI, although so far, both one and two protons per electron have been measured (Lawlor, 1993).

Plastocyanin, a copper-containing protein, forms the connection between cytochrome b_6f and PSI. The protein accepts a single electron from cytochrome *f* and can donate it to PSI. This photosystem has a P700 reaction centre, which is likely to be a dimer of chlorophyll *a* epimers (*a*'), as judged by difference spectrometry and reconstitution experiments (Maeda *et al.*, 1992). Excitation energy from antennae leads to oxidation of P700; the primary electron acceptor A_0 is a chlorophyll *a* monomer. The secondary electron acceptor is phylloquinone (or vitamin K_1), and subsequent carriers in the complex are iron-sulphur clusters. The high-energy electron from PSI is transferred to ferredoxin, which is a type of iron-sulphur protein, catalysing the transfer of single electrons. Ferredoxin may either reduce plastoquinone in the electron transport chain (cyclic electron transport), or it may reduce $NADP^+$, to give NADPH *via* the enzyme ferredoxin-$NADP^+$ oxidoreductase (non-cyclic electron transport).

As long as there is a sink for NADPH, such that the pool of $NADP^+$ does not become depleted, then non-cyclic electron transport is coupled to ATP synthesis, requiring PSII and PSI. However, there is an alternative flow of electrons, cyclic in nature, which does not result in the photolysis of water, because there is no net transport of electrons through the system. Cyclic electron transport requires PSI only; the electron ejected from P700 is probably passed to ferredoxin, and then to cytochrome b_{563} *via*

plastoquinone. A proton gradient is generated by the movement of protons by reduction-oxidation of plastoquinone, therefore ATP synthesis can take place.

1.3.2 *The dark reaction*

Of the energy yielded by the light reactions, the vast majority is utilised in the carbon fixation process, which is summarised in **figure 6**. The fundamental chemical reaction, which is the carboxylation of ribulose-1,5-biphosphate to yield two molecules of 3-phosphoglycerate, is catalysed by ribulose-1,5-biphosphate carboxylase/oxygenase (rubisco). This enzyme is often referred to as the most abundant protein on the planet, for it comprises half of all the soluble leaf protein. In metaphytes, rubisco consists of eight large subunits (LSU) and eight small subunits (SSU), where the former is synthesised in the chloroplast. As in all cases of protein complexes which consist of nuclear- and chloroplast-encoded subunits, the co-ordination of gene expression and enzyme assembly must be strictly controlled. This is demonstrated by the observation that during meristem development, there is a parallel increase (six-fold) in the levels of both LSU and SSU (Lawlor, 1993). Much of the initial work on protein targeting to chloroplasts was carried out using pre-SSU as the subject for investigation (for example, Highfield and Ellis, 1978; Chua and Schmidt, 1978).

1.4 *The chloroplast as a member of the plastid family*

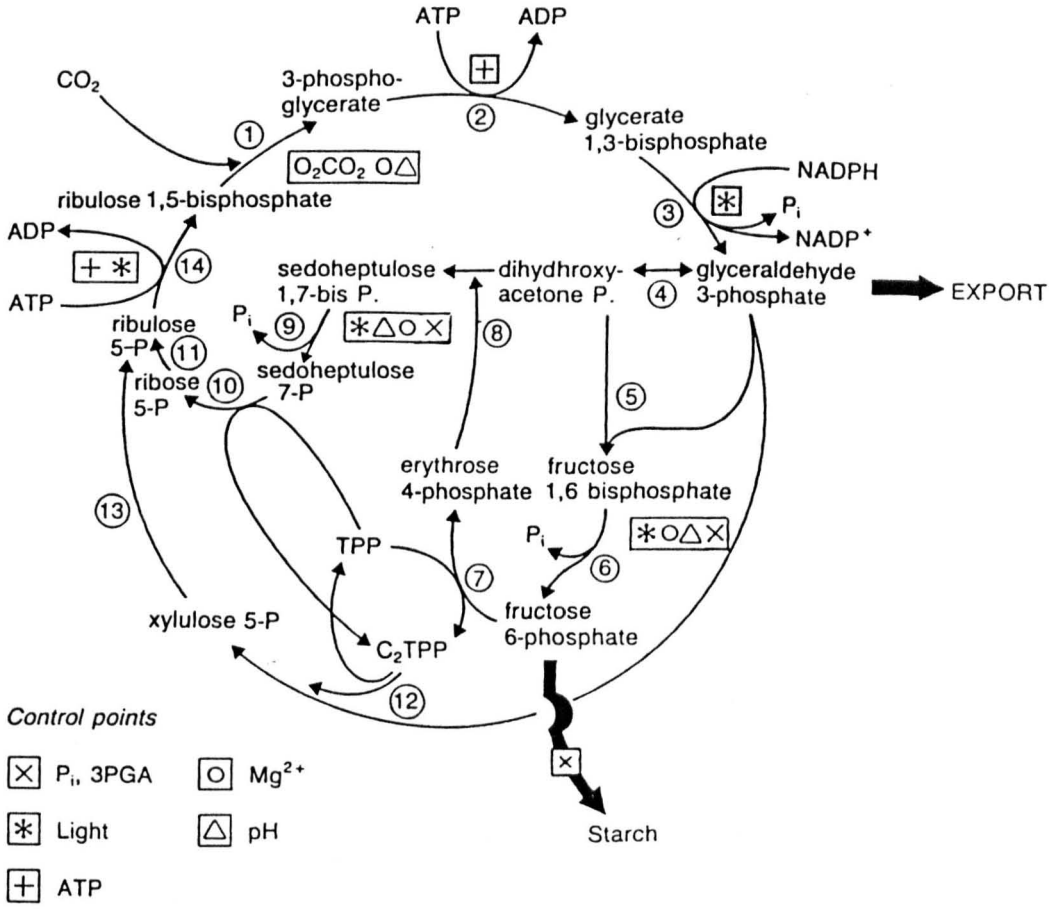
The chloroplast belongs to a group of organelles, called the plastids, which are to some extent, capable of interconversion between the various forms, depending on developmental and environmental cues (Raven *et al.*, 1986). Proplastids are small, unpigmented (or very pale green), undifferentiated plastids which are the precursors of more highly differentiated plastids, such as chloroplasts, amyloplasts or chromoplasts. They occur in the meristems of roots and shoots. Should the development of these proplastids into chloroplasts be blocked by the absence of light, they develop into etioplasts, which are characterised by the possession of the semi-crystalline prolamellar bodies. Despite darkness, fully-developed etioplasts contain proteins from the ATP synthase, rubisco, other stromal enzymes, and the 33kDa oxygen-evolving complex protein (the latter, from Sutton *et al.*, 1987). Upon

Figure 6 *The reactions of the Calvin Cycle for carbon fixation*

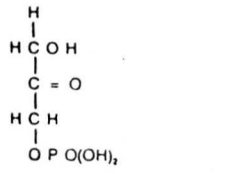
The numbers between intermediates correspond to the enzymes which catalyse the reactions

1.	Rubisco
2.	Phosphoglycerate kinase
3.	Triosephosphate dehydrogenase
4.	Triosephosphate isomerase
5., 8.	Aldolase
6.	Fructose-1, 6-bisphosphatase
7., 10., 12.	Transketolase
9.	Sedoheptulose bisphosphatase
11.	Ribose-5-phosphate isomerase
13.	Ribulose-5-phosphate 3 epimerase
14.	Ribulose-5-phosphate kinase

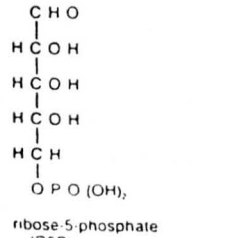
The diagram is taken from Lawlor (chapter 7), 1993.



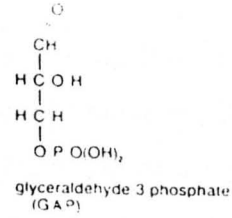
Some important chemical formulae of cycle intermediates



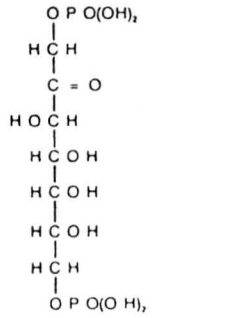
dihydroxyacetone phosphate (DHAP)



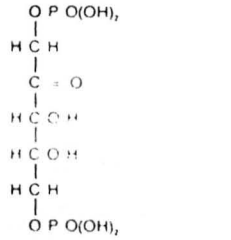
ribose-5-phosphate (R5P)



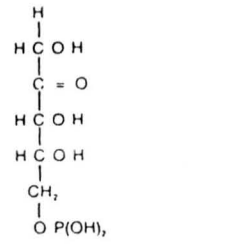
glyceraldehyde 3 phosphate (GAP)



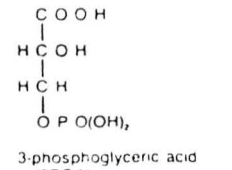
sedoheptulose-1, 7-bisphosphate (SBP)



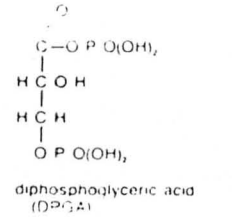
ribulose-1 5 bisphosphate (RuBP)



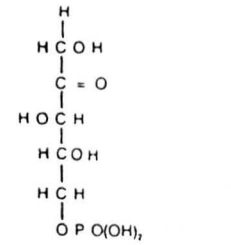
ribulose-5-phosphate (Ru5P)



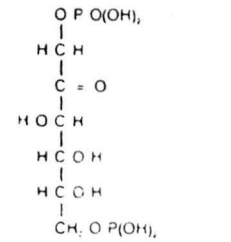
3-phosphoglyceric acid (3PGA)



diphosphoglyceric acid (DPGA)



xylulose-5-phosphate (X5P)



fructose-1,6-bisphosphate (FBP)

illumination, the membranes of the prolamellar body develop into thylakoids, chlorophyll synthesis progresses and the rest of the photosynthetic apparatus is assembled.

Chromoplasts are the plastids usually responsible for the yellow, orange or red colouration of ripening fruits, flower petals, senescing leaves and some roots. They lack chlorophyll, but synthesise and store carotenoid pigments. Leucoplasts do not contain pigments, but may synthesise a variety of other compounds. Starch-synthesising leucoplasts are termed amyloplasts - others may produce oils and protein.

1.5 *The chloroplast genome and protein synthesis*

Chloroplasts are genetically semi-autonomous, containing multiple, identical copies of their genome, which are replicated independently. Chloroplast DNA (cpDNA) is double-stranded, covalently closed and circular. In the plastid it forms a tight coil, which may aggregate with other cpDNA molecules to form nucleoid structures linked to the thylakoid membrane. In this, and many other respects, cpDNA resembles that of prokaryotes, including the presence of the Pribnow promoter box, polycistronic transcription and homologous RNA polymerases. Similarly, the protein-synthesising machinery is for the most part prokaryotic (for example, 70S ribosomes, instead of the eukaryotic 80S ones), although there are non-prokaryotic ribosomal proteins, indicating the evolving relationship between the endosymbiotic partners (Subramanian, 1993). The vast topic of chloroplast protein synthesis is reviewed in detail by Harris *et al.* (1994).

To date, the cpDNA of the three land plants has been completely sequenced: tobacco (*Nicotiana tabacum*, Shinozaki *et al.*, 1986), rice (*Oryza sativa*, Hiratsuka *et al.*, 1989) and that of a liverwort (*Marchantia polymorpha*, Ohyama *et al.*, 1986). The sequences show a high degree of conservation in the basic structure and gene order of the chloroplast genome in land plants, the most outstanding feature of which is the separation of two unique sequence regions (one larger than the other) by a pair of inverted repeats which are identical, due to the action of copy correction. The three genomes sequenced contain 110-120 genes; isolated chloroplasts will synthesise

around 100 polypeptides, which constitute around 20-30% of the structural and functional proteins needed by the chloroplast (Lawlor, 1993). Proteins encoded by the chloroplast genome include subunits of CF_0CF_1 ATPase, cytochrome *f*, cytochrome b_6, subunits of photosystems I and II, the large subunit of rubisco, NADH dehydrogenase, 30/31 different tRNAs, 20 different ribosomal proteins, 4 rRNAs and 3RNA polymerases. During the evolution of the symbiotic association between the plant cell and its plastids, there is evidence for extensive transfer of genes from the chloroplast to the nucleus (discussed by Jacobs and Lonsdale, 1987; Subramanian, 1993), partly accounting for the number of nuclear-encoded proteins which must be imported to the organelle post-translationally.

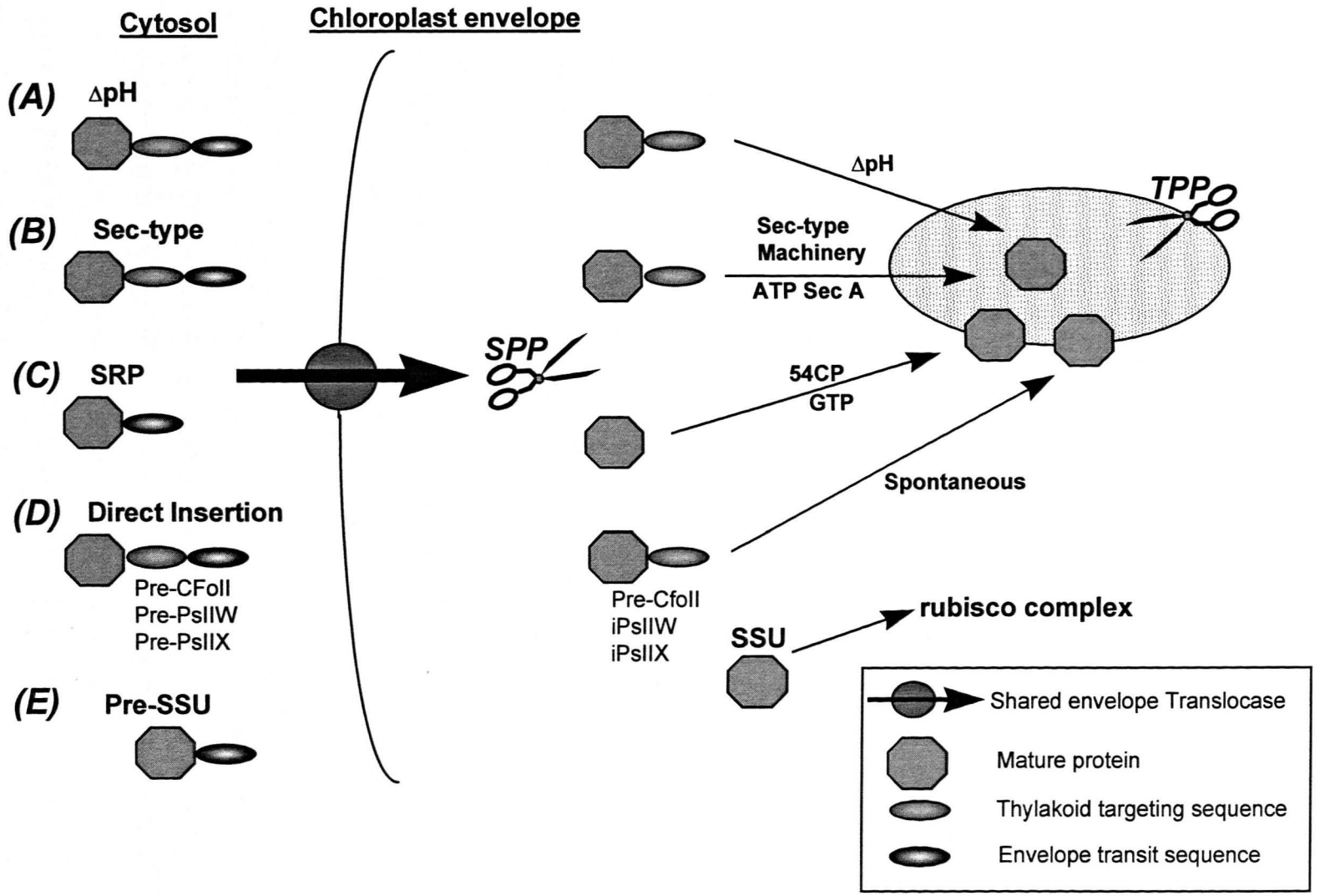
1.6 *Targeting of cytosolically synthesised proteins into and within chloroplasts*

The majority of chloroplast proteins are synthesised in the cytosol from mRNAs which have been transcribed from nuclear genes (Chua and Gilham, 1977). These proteins are then imported post-translationally to the organelle (**figure 7**), where they may be found in any of the chloroplast compartments. Transport of proteins into chloroplasts may be assayed by incubating isolated organelles with *in vitro*-translated, radiolabelled precursor proteins, which was first demonstrated by Highfield and Ellis (1978) and Chua and Schmidt (1978). Successful transport of added proteins can be monitored by protease-treatment of the chloroplasts post-import (Highfield and Ellis, 1978) or by incorporation into active complexes (Chua and Schmidt, 1979; Smith and Ellis, 1979).

1.7 *Chloroplast protein presequences and targeting signals*

A typical imported chloroplast protein is synthesised as a larger precursor with an N-terminal presequence, termed transit peptide. This contains the information necessary for the post-translational targeting of the precursor protein to the chloroplast and into the stroma (*via* a translocase in the chloroplast envelope membranes). Once the precursor protein has been transported into the stroma, its transit peptide is removed by a stromal processing peptidase (SPP). If the protein is destined for the thylakoid membrane, then it either integrates into this membrane by means of information

Figure 7 *A summary of protein targeting to and within chloroplasts.* The diagram illustrates the basic features of cytosolically synthesised stromal and thylakoid proteins into and within the chloroplast. In the case of thylakoid protein pathways, two are utilised by lumenal proteins (A and B) and two by integral membrane proteins (C and D). All nuclear-encoded proteins for the stroma and thylakoid are synthesised with N-terminal presequences containing envelope transit signals; proteins on pathways A, B and D contain in addition a cleavable thylakoid-targeting signal behind the envelope transit signal. Entry into the chloroplast is by means of a common mechanism involving translocation machinery in the envelope membranes, after which the envelope transit signal is usually removed by stromal processing peptidase (SPP) to form either a mature protein (in the case of stromal proteins or thylakoid proteins on the SRP-dependent pathway, C) (SRP, 'stromal signal recognition particle'). Lumenal proteins are then targeted across the thylakoid membrane *via* either the Sec-dependent translocase in the thylakoid membrane, in a process requiring stromal SecA and ATP, or by a ΔpH-driven translocase. The latter pathway exhibits no requirement for stromal factors. Integration of the multi-spanning membrane protein, LHCP, requires SRP and GTP; note that this process (pathway C) also relies on unidentified membrane-bound components, possibly the Sec machinery in the thylakoid membrane. A group of single-span proteins integrate by a direct mechanism (pathway D) that does not rely on known protein transport apparatus or any identifiable energy source. Proteins that have thylakoid-targeting signals have them removed either during or after the translocation process across the thylakoid membrane by thylakoidal processing peptidase (TPP).


Cytosol
Chloroplast envelope
(A) ΔpH
(B) Sec-type
(C) SRP
(D) Direct Insertion
Pre-CFoII
Pre-PsIIW
Pre-PsIIX
(E) Pre-SSU
SPP
ΔpH
TPP
Sec-type
Machinery
ATP Sec A
54CP
GTP
Spontaneous
Pre-CfoII
iPsIIW
iPsIIX
SSU
rubisco complex
Shared envelope Translocase
Mature protein
Thylakoid targeting sequence
Envelope transit sequence

contained within the mature protein domain or by means of a thylakoid transfer signal (**section 1.13**). Only a very few inner envelope membrane proteins have been identified, but these appear to be targeted to the stroma by an envelope transit peptide, and then re-directed to the inner membrane by information contained within the mature protein domain (**section 1.15**). Most, but not all outer envelope membrane proteins characterised to date are not synthesised as larger precursors: the targeting information is contained within the mature protein (**section 1.14**). For proteins which must traverse all six chloroplast compartments - those destined for the thylakoid lumen - the presequence is bipartite in nature, consisting of an envelope transit domain in tandem with a thylakoid transfer domain (**section 1.12**). The former targets the precursor to the stroma, where it is processed by SPP to form an intermediate-sized protein. The thylakoid transfer domain then directs the stromal intermediate across the thylakoid membrane and into the lumen, where thylakoidal processing peptidase (TPP) acts to remove the remaining portion of the presequence. **Figure 8** summarises the organisation of targeting information for proteins from the various chloroplast compartments.

1.7.1 *Envelope transit peptides*

The envelope transit peptide is both necessary and sufficient for the import of precursor proteins by chloroplasts. The first study to address this point was carried out by van den Broeck *et al.* (1985), who created a chimeric gene consisting of the envelope transit peptide from SSU fused to the neomycin phosphotransferase II (*npt*II) gene. This gene construct was introduced into, and expressed in tobacco plants, whereupon the *npt*II gene product was found, processed, in the chloroplast stroma. Similar results were obtained when the chimeric protein was synthesised in *E. coli* and incubated *in vitro* with isolated, intact pea chloroplasts. Since then, there have been several demonstrations of successful targeting of foreign proteins to chloroplasts merely by attachment of a chloroplast transit peptide to the protein in question. Conversely, mature proteins lacking transit peptides cannot be imported into chloroplasts (Mishkind *et al.*, 1985; Anderson and Smith, 1986). Examples of targeting to both mitochondria and chloroplasts by the same presequence are few and far between (for example, Hurt *et al.*, 1986; Huang *et al.*, 1990), and may instead reflect mis-sorting by heterologous systems (for example, a yeast mitochondrial

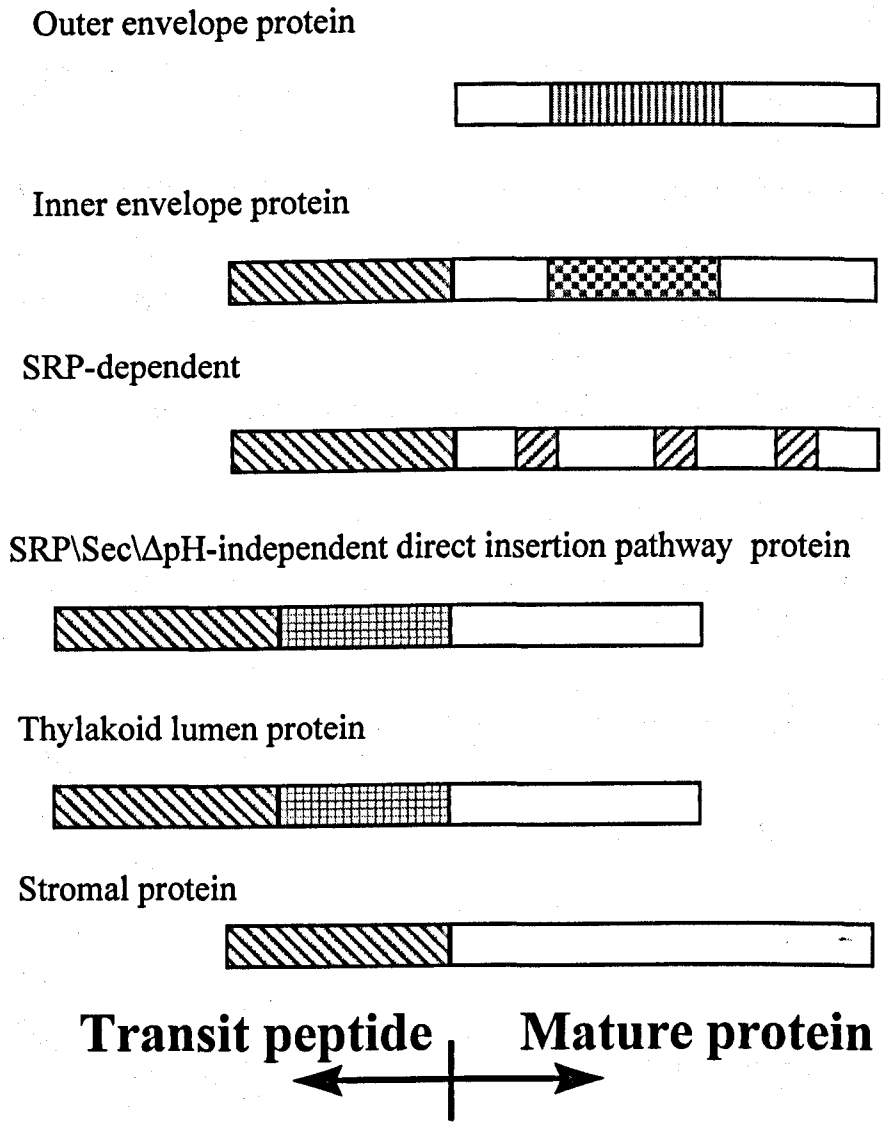
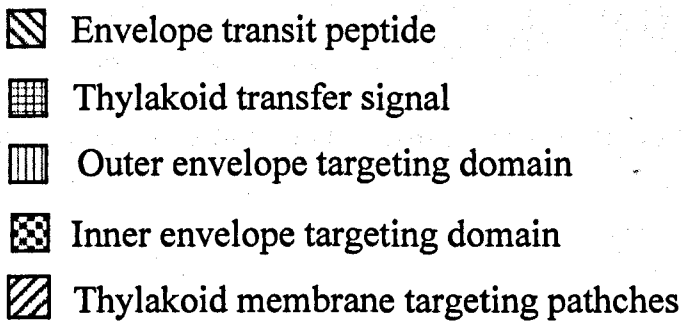


Figure 8. *Targeting information in chloroplast precursor proteins*
The diagram shows the relative positions of targeting signals within the primary sequence of precursors. Note that the Rieske iron-sulphur protein, which may reside in the thylakoid lumen, contains its thylakoid targeting information in the N-terminal region of the mature protein (Madueno et al., 1994)

Adapted from Theg and Scott, 1993



protein presequence directs a chloramphenicol acetyltransferase reporter protein to tobacco mitochondria and chloroplasts *in vivo*).

The demonstration that envelope transit peptides contain all the information necessary for the uptake of precursor proteins by chloroplasts, prompted a search for consensus sequences which would specify a chloroplast destination and therefore prevent mis-targeting to mitochondria, or other compartments. Initially, blocks of homology were identified within the presequences of SSU and light-harvesting chlorophyll *a/b* binding protein (LHCP) from various species (Schmidt and Mishkind, 1986; Karlin-Neumann and Tobin, 1986), but as the range of precursor sequences available became wider, it became evident that there was actually very little conservation in primary structure.

A wide selection of chloroplast presequences available at the time (26 in all) was analysed in detail by von Heijne *et al.* (1989). The first conclusion reached was that envelope transit peptides do not contain any regions of highly conserved amino acids, but may be divided into three structurally distinct domains. The N-terminal region consists of a stretch of around 10 amino acids which are mainly uncharged, and deficient in proline and glycine. The second, central region varies in length, contains very few acidic amino acids and is highly enriched for serine, compared to the amino acid composition of mature-sized chloroplast proteins. Threonine is also relatively abundant in this central domain. Finally, the C-terminal region is enriched for arginine, relative to mature-sized chloroplast proteins, and hydrophobic moment analysis indicated that the domain capable of folding into an amphiphilic β-strand. Interestingly, Theg and Geske (1992) analysed the presequence of *Chlamydomonas reinhardtii* chloroplast ATP synthase subunit by circular dichroism methods, and found it to be devoid of any secondary structure even in the presence of detergent micelles. On the other hand, Horniak *et al.* (1993) reported an induction of α-helical structure (at the expense of random coil and β-strand) when the ferredoxin (Fd) transit peptide was inserted into vesicles containing anionic lipids (which form 15-20% of the total chloroplast outer membrane lipid). Therefore, it is still far from clear as to which structural motifs in the envelope transit peptide, if any, are responsible for correct targeting to the chloroplast stroma.

1.7.2 *Thylakoid signal peptides*

The availability of the first cDNA clone for a thylakoid lumen protein enabled Smeekens *et al.* (1986) to compare targeting properties of the presequence of a stromal protein (Fd) with that of a lumenal one (plastocyanin; PC). Chimaeric precursors, consisting of the Fd transit peptide connected to the PC presequence, were synthesised *in vitro*, and their import into chloroplasts analysed. The results demonstrated that the Fd transit peptide is capable of re-directing a lumenal protein to the stroma, and that the PC transit peptide can target Fd at least some of the way to the thylakoid lumen, hindered probably by the lack of an intact protease processing site in the chimaera. A processed intermediate form of both pre-PC and pre-PC/Fd was observed in the stroma when each of these precursors were imported to intact chloroplasts, suggesting that there is an intermediate event between crossing the envelope and sorting to the thylakoid lumen. Based on all these observations, a two domain hypothesis was proposed for the PC presequence, which stated that the first (N-terminal) domain was functionally equivalent to the envelope transit peptide of imported stromal proteins, and that the second (C-terminal) domain mediated transfer to the thylakoid lumen after removal of the envelope transit peptide. Further evidence for this scheme was supplied by James *et al.* (1989), Ko and Cashmore (1989) and Hageman *et al.* (1990), the latter two of whom showed, by creating deletion, insertion or substitution mutants of pre-33K and pre-PC respectively (both proteins require Sec machinery for translocation), that import to the chloroplast and sorting inside the organelle are independent processes, mediated by the two domains of the protein presequence.

A comparison of a number of thylakoid transfer signals from imported chloroplast protein presequences (von Heijne *et al.*, 1989) revealed a strong similarity to signal peptides from both the eukaryotic and prokaryotic secretory pathways. Such signal peptides are composed of three sub-domains: a short, N-terminal positively-charged region, a central apolar region and a C-terminal region, containing small, uncharged amino acids at the -3 and -1 positions, relative to the start site of the mature protein. A turn-inducing residue (such as proline, glycine, serine, aspartate or asparagine) usually precedes the cleavage site. The similarity of thylakoid transfer signals to

bacterial signal peptides is sufficient that when the chloroplast pre-33K is expressed in *E. coli*, it is exported to the periplasm and processed to the mature size by (presumably) bacterial leader peptidase (Meadows and Robinson, 1991). Despite these similarities, however, the apolar regions of thylakoid transfer signals have lower leucine, and greater alanine contents than either eukaryotic or prokaryotic signal peptides.

The question that still has to be answered is: why are precursor proteins on the ΔpH-dependent pathway (e.g. 16K and 23K) not recognised by the Sec machinery despite the presence of a signal peptide (see **figures 7 and 8**)? Analysis using chimaeric constructs has led to the finding that pathway specificity is contained primarily, if not wholly, by the thylakoid-targeting signal. For example, the signal peptides of 16K and 23K are able to direct the efficient targeting of mature-size PC (wild-type PC is a Sec-dependent protein) by the ΔpH-dependent pathway (Henry *et al.*, 1994; Robinson *et al.*, 1994). Signals for the ΔpH-dependent pathway have been shown to contain a twin-arginine (twin-Arg) motif immediately prior to the hydrophobic region and Chaddock *et al.* (1995) demonstrated that this twin-Arg motif is crucial for targeting by the ΔpH-dependent system. They also showed, substitution of either arginine, even with lysine, resulted in a complete halt of translocation across the thylakoid membrane. In contrast, signal peptides for the Sec system similarly contain one or more basic residues at this position (and bacterial studies have shown this property to be important), but both lysine and arginine are equally permissible.

However, the twin-Arg motif, does not appear to be the only factor required for targeting by the ΔpH-dependent mechanism. Mutation of pre-PC with a twin-Arg motif, did not divert the precursor from the Sec pathway (Chaddock et al., 1995), which indicates that other signals are required to direct lumenal proteins along the correct targeting pathway. Work carried out by Bogsch *et al.* (1997) demonstrated that the 23K signal peptide contains three basic residues around the hydrophobic region, a twin-Arg before and lysine at the C-terminal end, and this charge distribution is sufficient to prevent the peptide from functioning as a Sec-type signal peptide. Removal of the C-domain lysine generates an apparently perfect signal peptide capable of targeting mature-size PC by the Sec pathway with wild-type

efficiency (Bogsch *et al.*, 1997). Consistent with this 'Sec-avoidance' role, it was found that the insertion of either twin-Arg before or lysine after the hydrophobic region of the PC signal peptide had little effect, but the introduction of both features completely blocked translocation across the thylakoid membrane by the Sec route. Bogsch *et al.* (1997) therefore suggested that pathway specificity is dictated to a large extent by factors present in the N- and C-domains, at least in the case of pre-23K.

1.8 *Precursor binding to the chloroplast envelope*

1.8.1 *Possible involvement of membrane lipids*

In contrast to mitochondrial signal sequences, the chloroplast-transit peptides usually show no secondary structure in aqueous solution, but behave as ideal random coils (von Heijne *et al.*, 1991). A specific presequence secondary structure may only form after the interaction of the precursor with the lipid bilayer at the chloroplast surface. Those lipids found in the outer envelope membrane, such as the anionic phospholipid phosphatidylglycerol and the chloroplast-specific glycolipids monogalactosyldiglyceride and sulphoquinovosyldiglyceride, were observed to interact with transit peptides (Van t'Hoff *et al.*, 1993; Pinnaduwage *et al.*, 1996). The interaction resulted in defined folding patterns and a specific topological arrangement of both the transit peptide and the associated lipids (Van t'Hoff *et al.*, 1993). Monolayer experiments with mutant precursor proteins indicated that the N-terminus of the transit peptide predominantly mediates the interaction with monogalactosyldigylceride, whereas the C-terminus has a high affinity for phosphatidylglycerol (Pilon *et al.*, 1995). These observations strongly suggest that specific interactions between the transit peptide and the membrane lipids in the outer envelope membrane of the chloroplast affect the targeting and membrane translocation of precursor proteins.

1.8.2 *Proteinaceous receptors and early translocation events*

In general, proteins destined for different chloroplast compartments are synthesised as precursor proteins with a cleavable, N-terminal extension - termed the presequence or transit peptide. These peptides are both necessary and sufficient to mediate precise chloroplast recognition, envelope transfer, stroma targeting and productive import of

passenger proteins (Cline *et al.*, 1996). Stroma-targeting domains generally lack acidic amino acids and have many basic and hydroxylated residues. A cytosolic protein kinase, which recognises transit sequences of chloroplast precursor proteins, but not transit sequences of mitochondrial preproteins, catalyses the phosphorylation of one specific serine or threonine residue within the stroma-targeting, envelope-transfer domain of the transit sequence (Waegemann and Soll, 1996). Phosphorylated and non-phosphoryated precursor proteins bind with similar efficiency to the organelle surface, but translocation of the former into chloroplasts is inhibited. Dephosphorylation catalysed by a protein phosphatase, which is probably localised in the outer envelope, is required to allow complete import of the precursor into the stroma. This cycle of phosphorylation and dephosphorylation could be a control system between the chloroplasts and the rest of the cell, ensuring the correct uptake of chloroplast precursors (Heins *et al.*, 1998). No further cytosolic proteins have been shown to be involved in the guidance of chloroplast preproteins or in retaining them in an import-competent conformation (e.g. by interaction with molecular chaperones *in vivo*). *In vitro* molecular chaperones of the hsp70 family of heat-shock proteins stimulated the import efficiency of the light-harvesting, chlorophyll-*a/b*-binding protein, but are not required for the import of soluble stromal proteins such as ferredoxin or the small subunit of Rubisco (SSU) (Waegemann *et al.*, 1990; Pilon *et al.*, 1992).

1.9 *Translocation across the envelope membrane*

The transit peptide is recognised by protease-sensitive components at the chloroplast surface in a reversible, energy-independent step (see **figure 9**). The hydrolysis of low concentrations of ATP (10-100μm) results in tight binding of the precursor protein to the import machinery (Waegemann and Soll, 1991) (i.e. docking) and is probably regulated by guanidine nucleotides (Filetici, 1996). Under these conditions, complete translocation into the organelle does not take place, although movement of the precursor protein into the translocation machinery is sustained (Filetici, 1996). The transfer of the precursor proteins from the chloroplast surface to Toc75 requires phospholipids on the cytosolic leaflet of the outer envelope membrane (Kerber and Soll, 1992). Furthermore, when the precursor protein inserts into the outer envelope membrane, components of the import machinery of the inner envelope membrane

have been isolated, supporting the contention that the import apparatus of both the outer- and inner envelope membranes are close together (Lübeck *et al.*, 1996; Nielsen *et al.*, 1997). Recently, components of the outer and inner envelope translocon were shown to co-immunoprecipitate with precursor proteins crosslinked to the protein-import machinery, indicating that joint translocation sites might be formed because of the interaction of a precursor protein with the Toc complex (Lübeck *et al.*, 1996; Nielsen *et al.*, 1997). These results suggest that precursor translocation proceeds simultaneously across the outer and inner envelope membrane (see **figure 9**). Complete translocation into the organelle is obtained only in the presence of higher ATP concentrations (0.1-3mM) in the stroma (Hirel *et al.*, 1997). Subsequently, the transit peptide is removed by a stromal processing peptidase (SPP).

1.10 *Model for the chloroplast envelope import process*

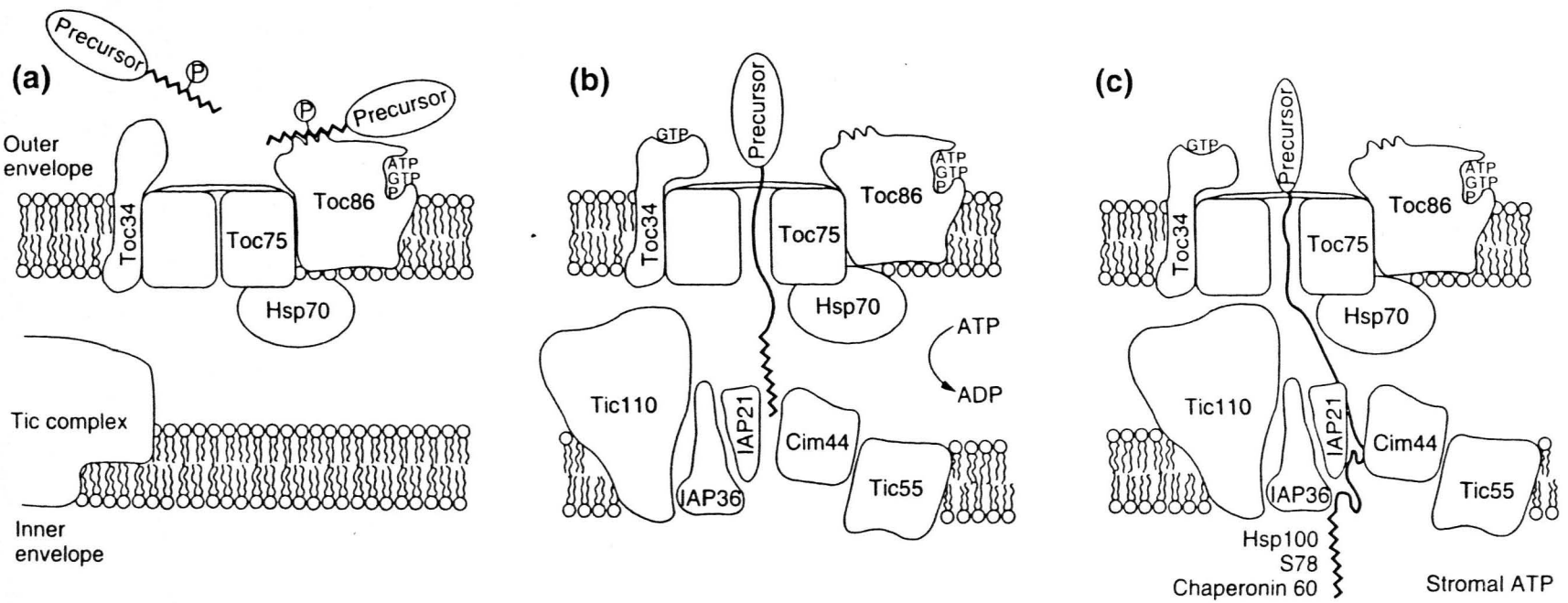
1.10.1 *Components of the outer envelope membrane*

Three major components of the outer envelope import apparatus have been identified, based on their stable association with the precursor protein trapped at an early step of translocation (Waegemann and Soll, 1991; Hirsch *et al.*, 1994; Schnell *et al.*, 1994; Tranel *et al.*, 1995). Toc34, Toc75 and Toc86 behave as integral membrane proteins, forming the core precursor translocon of the general-import pathway in the outer envelope membrane (**figure 9**). It can be assumed that Toc86 is a receptor for the recognition of the precursor protein. First, when the precursor binds in an initial, energy-independent step to the chloroplast surface, it is in intimate contact with Toc86 (demonstrated by label-transfer-chemical crosslinking). Moreover, further down the import pathway, in the presence of low ATP concentrations (<100μM), Toc75 is the major crosslinked product (Perry and Keegstra, 1994). Second, anti-Toc86 antibodies block the binding of precursor proteins to the envelope membrane (Hirsch *et al.*, 1994). Toc86 is a major chloroplast phosphoprotein, and has separate GTP- and ATP-binding sites (Kessler *et al.*, 1994; Seedorf *et al.*, 1995). ATP/GTP binding or hydrolysis could play an important role in a regulatory circuit of precursor handling. Toc86 has high homology to Toc34 in its cytosolically exposed N-terminus, and is in very close

Figure 9 *A model of protein import into chloroplasts across the outer and inner envelope membranes.*

(A) In an initial step, the phosphorylated precursor protein binds to the receptor Toc86 ('translocon at the outer envelope of chloroplasts'), in a high-affinity mode independent of ATP hydrolysis. The Toc86 activity might be influenced by GTP, ATP or phosphorylation. The Tic ('translocon at the inner envelope of chloroplasts') complex seems not to be involved at this step of precursor recognition. **(B)** Further translocation of the precursor protein to the translocation pore, Toc75, necessitates the dephosphorylation of the precursor protein and ATP hydrolysis in the intermembrane space. The conformation of Toc34 is probably altered by binding GTP, and as a consequence Toc34 possibly affects the gating properties of Toc75. Components of the Tic complex exposed to the intermembrane space probably participate for the first time at this point in the translocation. **(C)** Finally, joint translocation of precursor protein occurs simultaneously across the outer and inner envelope membranes. Stromal ATP is necessary to allow the passage of the precursor protein through the inner envelope membrane. In this model, components of the Toc and Tic translocon that are associated peripherally are not included (e.g. Com70 at the cytosolic leaflet, or Hsp100, S78 and chaperonin 60 at the stromal side). Furthermore, the arrangement of known components of the Tic translocon does not reflect the situation in vivo; especially, a Tic component functioning as a translocation pore has not yet been discovered. The names Cim and Com derive from 'chloroplast inner/outer envelope membrane protein'; IAP derives from 'import-intermediate-associated protein'.

Diagram taken from Heins *et al.* (1998).


(a)
Precursor
P
P
Precursor
Outer
envelope
ATP
GTP
P
Toc86
Toc34
Toc75
Hsp70
Tic complex
Inner
envelope
(b)
GTP
Precursor
Toc34
Toc75
Toc86
ATP
GTP
P
Hsp70
ATP
ADP
Tic110
IAP21
Cim44
IAP36
Tic55
(c)
GTP
Precursor
Toc34
Toc75
Toc86
ATP
GTP
P
Hsp70
Tic110
IAP21
Cim44
IAP36
Tic55
Hsp100
S78
Chaperonin 60
Stromal ATP

proximity to Toc75, as demonstrated by the formation of a covalent disulphide bridge *in vitro* (Seedorf *et al.*, 1995). Toc75 is largely buried in the outer envelope membrane, adopting a β-barrel conformation similar to that of bacterial pore proteins (Schnell *et al.*, 1994; Tranel *et al.*, 1995; Hinnah *et al.*, 1998). The observation that Toc75 specifically interacts with the transit sequence of precursor proteins, by the identification of crosslinked products and after reconstitution of Toc75 in liposomes, substantiates a function in the general import pathway (Ma *et al.*, 1996; Hinnah *et al.*, 1998). Patch-clamp studies demonstrate that Toc75 forms a water-filled channel *in vitro* (Hinnah *et al.*, 1998). Like Toc86, Toc34 is a new type of GTP-binding protein and is in close proximity to Toc75 (again demonstrated by the formation of a covalent disulphide bridge *in vitro* (Seedorf *et al.*, 1995)). Toc34 shows the classical GTP-binding motif, which resides in the N-terminal portion of the protein protruding into the cytosol. No further sequence homology to other eukaryotic G-proteins, except to Toc86, is known. A direct interaction between Toc34 and a precursor protein has not been demonstrated. Toc34 is in close proximity to Toc75, and the recombinant protein exhibits the predicted GTPase activity (Seedorf *et al.*, 1995). A current hypothesis is that hydrolysis of nucleotides and subsequent alteration of the conformation of Toc34 could influence the gating properties of Toc75 and its affinity for precursor proteins (Soll, 1998).

In addition, the precursor translocon of the outer envelope contains two associated molecular chaperones of the hsp70 family - Hsp70 and Com70 (a chloroplast outer envelope membrane protein) (Waegemann and Soll, 1991; Kourtz and Ko, 1997). Com70 generates a crosslinked product with a precursor protein at an early stage of import at the cytosolic leaflet of the outer envelope (Kourtz and Ko, 1997), but Hsp70 is probably localised on the other side at the intermembrane space side (Waegemann and Soll, 1991).

1.10.2 *Components of the inner envelope precursor translocon*

Five integral protein subunits of the inner envelope membrane are known: Tic110, Tic55, Cim44 (a chloroplast inner envelope membrane protein), IAP36 and IAP21 (import intermediate-associated proteins) (Wu *et al.*, 1994; Lübeck *et al.*, 1996; Ma *et al.*, 1996; Kessler and Blobel, 1996; Caliebe *et al.*, 1998). Tic110 was co-immuno-

precipitated with a precursor protein that was trapped in joint translocation sites during membrane passage. Because Tic110 is largely exposed to the intermembrane space, it might be involved in the formation of joint translocation sites and/or in the guidance of the presequences across the soluble lumen space to the inner envelope translocon (Lübeck *et al.*, 1996). Tic55 associates with an inner envelope membrane complex comprising Tic110 and the Toc translocon in the presence of a precursor protein. A Rieske-type, iron-sulphur cluster and a mononuclear iron-binding site are also characteristic features of Tic55. Diethylpyrocarbonate, a Rieske-type, protein modifying reagent, inhibits the translocation of precursor protein across the inner envelope membrane, whereas binding of the precursor to the outer envelope membrane is still possible (Caliebe *et al.*, 1998). Cim44 and IAP21 form crosslinked products at distinct stages of import and may be involved in precursor translocation across the inner envelope membrane (Wu *et al.*, 1994; Ma *et al.*, 1996;). Three different chaperones appear to associate peripherally with the stromal side of the Tic complex: Hsp100, Hsp70 and chaperonin 60 (Kessler and Blobel, 1996; Nielsen *et al.*, 1997). Recently, a stromal Hsp100 was found in stable association with the joint Toc and Tic translocation sites (Nielsen *et al.*, 1997). Another stromal Hsp70 homologue, S78, is probably associated with precursor translocation, but a stable association could not be demonstrated (Nielsen *et al.*, 1997). Whether Hsp100 and Hsp70 co-operate in 'pulling' the precursors into chloroplasts in a mechanism analogous to that in mitochondria is unknown. In a later stage of precursor translocation, chaperonin 60 co-immunoprecipitates with Tic110 (Kessler and Blobel, 1996), but the co-ordinated sequence of interactions of these molecular chaperones remains to be established.

1.10.3 *Translocation across the outer envelope membrane*

Patch-clap analysis shows that recombinant, purified Toc75 is a voltage-gated ion channel. The specific interaction of reconstituted Toc75 with precursor proteins was demonstrated by partial closing of the channels in the presence of precursor proteins (Hinnah *et al.*, 1998). The comparatively low ionic conductance of reconstituted Toc75 is compatible with a narrow aqueous pore of 8-9 Å diameter, which suggests that protein translocation into chloroplasts requires almost complete unfolding of the precursor protein. Recent evidence demonstrates that the pore size of the translocon

in the endoplasmic reticulum could be up to 60 Å diameter (Hamman *et al.*, 1997). However, the membrane of the endoplasmic reticulum remains sealed by a gating mechanism that probably involves the ribosomes, which are attached to the surface of the organelle during co-translational transport of precursor proteins (Siegel, 1997). In post-translational protein translocation there is no such seal, and there might be a need for a much narrower pore. The notion that proteins may need to be at least partially unfolded to make them competent for translocation is corroborated by findings that demonstrate that chloroplasts have a strong unfolding activity. When bound to the chloroplast surface, the A chain of ricin, fused to a chloroplast presequence, was highly accessible to protease treatment - indicating an unfolded conformation of the protein (Walker *et al.*, 1996). Furthermore, the precursor of SSU becomes increasingly protease-sensitive from an early stage of translocation up until it reaches a final protease-resistant location inside the chloroplasts (Waegemann and Soll, 1991). Two Hsp70 homologues in the outer envelope membrane (Waegemann and Soll, 1991; and Kourtz and Ko, 1997) could be involved in the unfolding mechanism. However, the translocation of a globular, chimeric protein consisting of pSSU and bovine-pancreatic-trypsin inhibitor, which is tightly folded and should provide an insurmountable obstacle to the Toc-complex, was observed (Clark and Theg, 1997). The tightly folded bovine-pancreatic-trypsin inhibitor domain has a diameter of about 18Å and a length of about 30Å. Assuming that the chimeric protein was translocated faithfully by all the components of the Toc and Tic complex, this implies that the Toc75 pore would be quite flexible and larger *in situ* than the diameter calculated by the electrophysiological measurements *in vitro*.

1.11 *Stromal events and proteolytic processing of stromal and thylakoid proteins*

Once an imported precursor protein has crossed the chloroplast envelope, or possibly during its passage across the inner envelope membrane, the transit peptide is cleaved by SPP (**section 1.11.1**). Should the protein be further sorted to the thylakoid lumen or inserted into the thylakoid membrane, then the thylakoid transfer signal is removed by the membrane-bound TPP (**section 1.11.2**). In the stroma (and in all probability, the thylakoid lumen too), a newly-imported protein is faced with the task of folding

into its correct conformation, and then, possibly assembly into a larger protein complex. On any given, unfolded polypeptide, there are likely to be numerous 'interactive surfaces' which can induce the formation of many non-productive tertiary or quaternary structures (Gatenby and Ellis, 1990). In order to prevent incorrect folding, there exists a family of proteins (ubiquitous throughout nature), whose function in mediating attainment of correct tertiary or quaternary structures has led to them being termed 'molecular chaperones' (Ellis, 1987). These chaperones probably do not dictate the correct conformation to the subject polypeptide, rather they appear to inhibit incorrect folding pathways by binding to the interactive surface, and only releasing it when the proper structure has been formed. They are involved in the assembly of rubisco in the stroma, and have been shown to associate with the Rieske iron-sulphur protein, FNR and LHCP after import to the chloroplast (Lubben *et al.*, 1989; Madueño *et al.*, 1993; Tsugeki and Nishimura, 1993). The two chaperones which assist in the assembly of rubisco are part of a distinct group, called the chaperonins: cpn60 and cpn10. Cpn60 is active as an oligomer of two stacked toroids, each containing seven subunits and a central cavity which is most important for chaperone function. The oligomers hydrolyse MgATP, and dissociate reversibly to smaller units. Cpn10 forms a ring of seven subunits, which binds cpn60 in the presence of MgATP. All this information was derived from studies on bacterial chaperonins (reviewed by Gatenby and Ellis, 1990). In brief, the chaperonins are believed to function by binding the substrate in the central cavity, in a 'molten globule'-like state where the secondary structure is complete, but stable tertiary interactions are not yet formed. The tertiary structure is reached in a protected environment, away from other folding polypeptides. Final release of the folded polypeptide may occur after many binding and release cycles, corresponding to cycles of ATP binding, hydrolysis and release in which cpn10 is intimately involved (reviewed by Hartl *et al.*, 1994). It is not yet clear to what extent cpn60 is involved in assembly of the rubisco monomers to form the hexadecamer.

Immunoprecipitation techniques, with and without native gel electrophoresis, enabled Madueño *et al.* (1993) and Tsugeki and Nishimura (1993) to identify interactions of imported Rieske iron-sulphur protein (a thylakoid membrane protein) and FNR (a stromal enzyme) with both cpn60 and hsp70 (another type of molecular

chaperone, well characterised in mitochondrial import). Interestingly, time-course analyses showed that the Rieske protein formed a complex first with cpn60, then with hsp70, whereas the temporal sequence was reversed for FNR. On the other hand, it was not possible to detect any high molecular weight complex containing the stromal form of imported 23K (a thylakoid lumen protein) by gel filtration (Creighton *et al.*, 1995), so chaperoning new arrivals in the stroma may not be universally necessary. Alternatively, the interaction may have been missed, due to a long import incubation time (fifteen minutes).

1.11.1 *Stromal processing peptidase*

SPP was purified approximately 350-fold by Robinson and Ellis (1984a) from pea chloroplasts, around 1000-fold from similar starting material by Dr. D. C. Bassham (Ph.D. thesis, 1993) and to homogeneity (as a pair of immunologically-related polypeptides of 143 and 145 kDa) by Oblong and Lamppa (1992b). Since then, a gene encoding a chloroplast processing enzyme has been cloned using antibodies raised against the 143/145 kDa polypeptide pair (Vander Vere *et al.*, 1995). The partially purified protease was shown to be highly specific for chloroplast precursor proteins (Robinson and Ellis, 1984a), and is believed to remove the envelope transit peptides of both stromal proteins and those destined to the thylakoid membrane and lumen.

SPP has a pH optimum of 8-9, is not ATP-dependent, is inhibited by EDTA and 1,10-phenanthroline but not by serine- or thiol-protease inhibitors, commensurate with the peptidase being a metalloprotease. In line with these findings, the enzyme identified by Vander Vere *et al.* (1995) contains a zinc-binding motif which is conserved in a family of metalloendopeptidases, including the β subunit of mitochondrial processing peptidase (MPP) and *E. coli* protease III. Interestingly, active 143/145 kDa peptidase is found in etioplasts and root plastids, which suggests both an essential role of the enzyme in green and non-photosynthetic tissues alike, and that the enzyme may act as a general stromal processing peptidase in all plastids.

Bassham *et al.* (1991) delineated the envelope transit peptides and thylakoid transfer signals of three thylakoid lumen proteins by carrying out Edman degradation of the

intermediate-sized precursor proteins (formed after their incubation with SPP) to find the SPP cleavage site. A comparison of thirty-two envelope transit peptides from a variety of precursors yielded a very loosely-conserved consensus sequence around the processing site of (Val/Ile)-X-(Ala/Cys)↓Ala, with the arrow marking the site of cleavage (Gavel and von Heijne, 1990). Introduction of mutations at residues -1 and +1 relative to the cleavage site (of the 33K protein) drastically reduces cleavage efficiency, but does not alter the fidelity of the reaction (Bassham *et al.*, 1994). Taken with the weakly-conserved consensus sequence, these results suggest that the precise site of cleavage is determined by other factors than the adjacent amino-acid side chains.

1.11.2 *Thylakoidal processing peptidase*

The first report of TPP activity emerged in Hageman *et al.* (1986), when it was shown that pre-PC was processed only to an intermediate form by SPP; complete maturation was found to require a thylakoidal processing peptidase which was active in detergent-solubilised thylakoids. Further studies (Kirwin *et al.* 1987; Kirwin *et al.*, 1988; James *et al.*, 1989) showed that TPP is involved in the maturation of other proteins synthesised with bipartite presequences, and is highly specific for thylakoid signal peptides. Unlike SPP, however, TPP is highly hydrophobic and can only be assayed in the presence of detergents (Kirwin *et al.*, 1988). The enzyme is present in non-appressed thylakoid membranes which are in contact with the stroma (not surprising, since proteins are presumably transported into and across these membranes) and the active site of the enzyme is on the lumenal face of the thylakoid membrane (Kirwin *et al.*, 1988). Pea TPP has been partially purified from Triton X-100-solubilised thylakoids (Kirwin *et al.*, 1987) and recently an *Arabidopsis* cDNA encoding TPP has been cloned and overexpressed (Chaal *et al.*, 1998).

The most interesting studies of the TPP reaction mechanism were prompted by comparisons of the thylakoid transfer signals of lumenal proteins. In each case these contain a hydrophobic core domain and short-chain residues at the -3 and -1 positions, relative to the TPP cleavage site - these are two of the diagnostic features of signal sequences which direct transport across the bacterial plasma membrane and endoplasmic reticulum (von Heijne *et al.*, 1989). Halpin *et al.* (1989) compared the

reaction specificities of purified *E. coli* leader (or signal) peptidase and partially purified TPP, and it was found that they were indistinguishable. The bacterial enzyme was shown to cleave lumenal protein precursors efficiently and accurately, and it was concluded that TPP is indeed a signal-type peptidase and belongs to the family of Type I leader peptidases (Dalbey and von Heijne, 1992). Members of this family include the leader peptidases from *E. coli* and the cyanobacterium *Phormidium laminosum*, where the enzyme also processes polypeptides of the cyanobacterial thylakoid lumen (Wolfe *et al.*, 1983; and Packer *et al.*, 1995). An inhibitor of the *E. coli* enzyme (Allsop *et al.*, 1995) has recently been shown to inhibit the chloroplast and cyanobacterial enzymes (Barbrook *et al.*, 1996). The structures of the bacterial leader peptides and the thylakoid-transfer domain have similar features. These include a positively charged N-terminal region, a hydrophobic central core and a polar C-terminal region in which the -3 and -1 residues relative to the cleavage site are short uncharged amino acid residues (Shackleton and Robinson, 1991).

Following these observations, and the demonstration that some lumenal protein precursors can be exported by *E. coli* (Seidler and Michel, 1990; Meadows and Robinson, 1991), it was widely believed that the thylakoid protein transport/maturation apparatus evolved from a prokaryotic-type apparatus in an ancestral cyanobacterial progenitor of the chloroplast. In keeping with this suggestion, it has been found that cyanobacterial thylakoid lumen proteins are synthesised with presequences which resemble both typical signal sequences and the thylakoid transfer signal of the higher plant counterparts (Kuwabara *et al.*, 1987).

Comparison of amino acid sequences among the Type I leader peptidases shows the presence of a few highly conserved regions, and site-directed mutagenesis of the *E. coli* leader peptidase shows particular serine and lysine residues to be involved in the catalytic activity (Black *et al.*, 1992; Black, 1993; Tshantz *et al.*, 1993; Dalbey *et al.*, 1997). Other prokaryotic and mitochondrial leader peptidases are believed to contain such a catalytic dyad, although the homologous subunits of the endoplasmic reticulum signal peptidase contain a histidine residue at the equivalent position to the lysine (reviewed in Dalbey *et al.*, 1997).

1.11.2a *Catalytic mechanism*

Since the discovery of signal peptidases, it was proposed that these enzymes are mechanistically novel peptidases as they are resistant to inhibitors of the classical serine, cysteine, aspartic acid, or metallo classes of peptidases. Site-directed mutagenesis approaches have been used to define the amino acids directly involved in catalysis by the bacterial enzymes. These studies revealed that the conserved serine in domain B of the family is essential for catalytic activity and may act as the active site nucleophile in signal peptide cleavage. Mutation of this residue to alanine in the *E. coli* (Sung and Dalbey, 1992) and *B. subtilis* (SipS; van Dijl *et al.*, 1995) signal peptidases, as well as the Imp2p subunit of the mitochondrial signal peptidase (Nunnari *et al.*, 1993) yielded inactive enzymes. Substitution of cysteine for serine-90 in the *E. coli* and the corresponding serine in *B. subtilis* (SipS) signal peptidases produced active peptidases. In the case of *E. coli*, the cysteine-90 signal peptidase can be inactivated with N-ethyl maleimide, a cysteine-specific reagent that does not inactivate the wild-type enzyme (Tschantz *et al.*, 1993).

In contrast to classical serine peptidases, the *E. coli* signal peptidase does not employ a histidine as the proton donor and proton acceptor. Each of the three histidine residues in this enzyme can be mutated to alanine without any observable effect on enzymatic activity (Sung and Dalbey, 1992; Black *et al.*, 1992). No histidine residues are conserved in amino acid sequence alignments of the prokaryotic signal peptidase family. lysine-145 in domain D of *E. coli* leader peptidase is present in all bacterial, mitochondrial signal peptidases described to date (Dalbey *et al.*, 1997) and is also present in the only TPP clone described (*Arabidopsis*; Chaal *et al.*, 1998). This lysine-145 is a strong candidate to play the role of general base that strips the proton from the nucleophilic serine during peptide bond cleavage. However, this lysine is notably absent in the ER signal peptidases (Dalbey *et al.*, 1997). Substitution of lysine-145 of the *E. coli* signal peptidase by methionine (Black, 1993) or by alanine, histidine or asparagine (Tschantz *et al.*, 1993) abolishes catalytic activity. Similar results were observed with the *B. subtilis* signal peptidase SipS (van Dijl *et al.*, 1995). These experiments established that the critical lysine residue is required for a

functional enzyme but the data do not directly show that this amino acid is involved in the proton transfer at the active site.

At this stage, the signal peptidases of bacteria, mitochondria and at least in *Arabidopsis* thylakoids appear to be unconventional serine peptidases that employ a lysine as a general base. The hydroxyl group of the serine side chain acts as the nucleophile that attacks the carbonyl carbon of the scissile peptide bond of the precursor protein cleavage site. The unprotonated form of the lysine ε-amino group serves to activate the hydroxyl group of the serine. There is a precedent for a mechanism involving a serine/lysine dyad for a peptidase. The LexA protein, which is involved in the SOS response in *E. coli*, undergoes a self-cleavage reaction that inactivates the protein. This self-processing event employs a serine as the nucleophile that attacks the peptide bond (Roland and Little, 1990) and a lysine that is deprotonated (Lin and Little, 1989; Little, 1993). Moreover, X-ray crystallographic analysis has shown that the serine/lysine dyad is at the active site of the UmuD protein, a member of the LexA peptidase family (Peat *et al.*, 1996).

1.11.2b *TPP has more stringent substrate requirements than other signal peptidases*

The first comparisons (Halpin *et al.*, 1989) of TPP with bacterial and ER signal peptidases suggested that the reaction specificities of the three enzymes were very similar; the *E. coli* enzyme was shown to cleave lumenal protein precursors, and TPP was shown to cleave one bacterial signal sequence and also that of yeast α-factor. More detailed studies (Shackleton and Robinson, 1991) on the processing of pre-33K by pea TPP involving mutagenesis of the terminal Ala-X-Ala motif, showed that TPP is in fact more selective than other known signal peptidases. In general, signal peptidases are known to require short chain residues to be present at the -3 and -1 positions in substrate precursors, and the *E. coli* and ER enzymes can usually tolerate alanine, serine or glycine at the -1 position and somewhat longer-chain amino acids (often including leucine) at the -3 position. However, efficient cleavage of pre-33K by TPP was completely dependent on the presence of alanine at -1, even serine or glycine at this position was detrimental to processing. At the -3 position the

substitution of alanine by valine had no effect, but the presence of leucine inhibited cleavage to a marked degree (Shackleton and Robinson, 1991). These data have been supported by sequence data and a total of 12 lumenal proteins have been cloned either in higher plants or *Chlamydomonas reinhardtii*, and all contain alanine at the -1 position. Alanine is also usually found at the -3 position although valine is found in one instance.

In addition to the -1 and -3 processing determinants, the helix-breaker, typically found at the -4 to -6 position of the signal peptide, can also be important for processing (Shen *et al.*, 1991; Nothwehr and Gordon, 1989). For example, the proline at the -6 position of the M13 procoat signal peptide is essential for efficient processing *in vivo*, although *in vitro* most -6 mutants can be cleaved by high levels of leader peptidase. This suggests that the presence of a helix breaker allows the pre-protein to bind with high affinity. A 'minimum' substrate for cleavage by *E. coli* leader peptidase *in vitro* is ALA↓KI (Dev *et al.*, 1990) (where ↓ is the cleavage site), but more efficient cleavage is observed for FSASALA↓KI, which includes a part of the hydrophobic region.

Finally, the hydrophobic region of the signal peptide can play a role in cleavage. Nothwehr and Gordon (1990) found that the amino terminus of the hydrophobic region of the human pre(Δpro)apolipoprotein A-II signal peptide influences the accuracy of processing. This is most likely to be an indirect effect resulting from a shift of the mutant signal peptide within the membrane such that the processing site is moved relative to the leader peptidase active site. When the hydrophobic region becomes very long, processing is also prevented, and the protein becomes permanently anchored to the membrane *via* its signal peptide (Chou and Kendall, 1990). Again, it is likely that the processing site is moved out of reach of the signal peptidase.

1.12 *Translocation of nuclear-encoded thylakoid lumen proteins*

Whilst valuable information about transport of proteins to the thylakoid lumen could be obtained by intact chloroplast import assays, the real advance in this area came

when protein transport into isolated thylakoid vesicles was demonstrated (Kirwin *et al.*, 1989; Bauerle and Keegstra, 1991; Cline *et al.*, 1992). The advantage of using isolated thylakoid membranes to study protein transport is that the translocation process may be examined independently of envelope transport, and further, that experimental conditions at the thylakoid membrane may be easily manipulated (for example, ATP concentration or presence/absence of stromal proteins). Kirwin *et al.* (1989) achieved reasonably efficient import of *in vitro*-translated, radiolabelled pre-33K by mixing the translation products with pea thylakoids, stromal extract and ATP. Later, Mould *et al.* (1991) greatly improved the efficiency of the assay by illuminating the thylakoids instead of including exogenous MgATP. The substrates for the assay were pre-23K and pre-33K, and a striking finding of the study was that the two precursors differed in their transport requirements. Efficient translocation of pre-33K was dependent upon the presence of stromal extract, but the opposite was true for pre-23K. A further point to emerge from this observation was that *in vitro* at least, processing of pre-23K to i23K by SPP was unnecessary for transport across the thylakoid membrane, a phenomenon also observed for pre-PC by Bauerle and Keegstra (1991). Cline *et al.* (1992) examined the transport characteristics of pre-PC, pre-33K, pre-23K and pre-16K and found them to be protein-specific; pre-23K and pre-16K could be translocated in the complete absence of ATP, but a transthylakoidal ΔpH was absolutely required. Klösgen *et al.* (1992) obtained similar results with pre-16K, and also showed that the full precursor protein could be transported across the thylakoid membrane without prior processing by SPP.

As discussed in **1.7.2**, thylakoid transfer signals markedly resemble prokaryotic (and eukaryotic) signal peptides. The endosymbiont hypothesis for the origin of chloroplasts from cyanobacterial progenitors is now widely accepted, so it was a natural assumption that the mechanism for protein translocation across the thylakoid membrane would be conserved from prokaryotes. This would involve a soluble chaperone such as SecB, and an ATP-dependent translocation factor, like SecA - both components of the bacterial translocation apparatus, discussed in **1.19**. As it turned out, the real situation is more complicated - there appear to be at least two pathways across the thylakoid membrane and into the lumen. One of these is almost certainly conserved from the ancestral prokaryote, and is mediated by a SecA

homologue, referred to as the Sec-type pathway. Until recently, the other was originally believed to be unique in the realm of protein translocation.

1.12.1 *Identification of two distinct pathways across the thylakoid membrane*

A variety of biochemical, and recently, genetic, lines of evidence supports the postulate that there are (at least) two mechanisms by which thylakoid lumen proteins reach their destination. Firstly, over-expressed precursor proteins have been used to conduct competition assays using both intact chloroplasts and isolated thylakoids (Cline *et al.*, 1993). The import of pre-PC, pre-33K, pre-23K and pre-16K into intact chloroplasts was inhibited by saturating concentrations of *E. coli*-over-expressed pre-23K, indicating that all four precursor proteins share a common component of the envelope translocation machinery. However, once in the chloroplast, *en route* to the thylakoid lumen, i23K and i16K compete at the thylakoid membrane, but they do not compete with i33K or iPC. The latter two intermediates compete with each other however, thus indicating that the four precursor proteins fall into two groups. Similar results were obtained using isolated thylakoid import assays: over-expressed i23K competed with pre-23K, i23K and pre-16K, but did not affect the transport of pre-33K, i33K or pre-PC. These results imply that there are two distinct translocases, each of which interacts with only one group of precursor proteins. That the pathways may converge after these separate interactions cannot be excluded.

Hulford *et al.* (1994) carried out a detailed dissection of the translocation requirements for pre-23K and pre-33K across the thylakoid membrane. A stromal protein factor was found to be necessary for 33K import, but not 23K import. SPP was eliminated as a candidate for the stromal factor by the use of artificial intermediates, i23K and i33K, which obviated any requirement for removal of the envelope transit peptide during the isolated thylakoid import assay. The 33K intermediate still required the presence of stromal extract, whereas i23K could be efficiently imported in the presence and absence of stroma. Nucleoside triphosphates (NTPs) were essential for the transport of 33K, but not 23K, into isolated thylakoids. This study was closely followed by further evidence for the existence of two distinct

translocation systems, involving chimaeric precursor proteins (Robinson *et al.*, 1994). Wild-type PC was demonstrated to require stromal extract and NTPs, just like 33K. A chimaera, consisting of the presequence of 23K, attached to the mature PC protein, was synthesised and imported into isolated thylakoids. The import of this chimaera required only a transthylakoidal ΔpH, and was abolished by the proton ionophore, nigericin. Import also took place in the absence of stromal extract and NTPs, showing that the chimaera behaved exactly like 23K, and unlike PC. The presequence of 23K had succeeded in diverting PC from its "natural" pathway to that of the 23K protein. A second chimaera, consisting of the presequence of 16K attached to mature PC was also transported across the thylakoid membrane in a ΔpH-dependent manner. An important conclusion from these findings is that all the information required for targeting a protein *via* a particular pathway is contained in the presequence, not the mature protein. Henry *et al.* (1994) also showed that the thylakoid transfer signal contains the information necessary for discriminating between two separate translocation pathways. They synthesised chimaeric precursors: 23K presequence/PC mature protein and 33K presequence/16K mature domain. These fusion proteins behaved in a manner dictated by the origin of the presequence when imported to intact chloroplasts and isolated thylakoids, although the 33K/16K chimaera was too poor a substrate in the thylakoid import assay to yield satisfactory results. A fifth protein, PSII-T (photosystem II), was shown to have identical import requirements to 23K and 16K.

Finally, two nuclear mutations in maize have been discovered, which disrupt the two thylakoid transport pathways discussed above (Voelker and Barkan, 1995). One mutation, named *tha*1, interferes with the targeting of PC, 33K, PSI-F and cyt*f* (a chloroplast-encoded protein). The other mutation, hcf106, interferes with the targeting of 16K and 23K. These mutations are extremely interesting, for two reasons. The first is because they constitute the only *in vivo* evidence of the existence of two distinct translocation mechanisms for thylakoid lumen proteins. The second reason is because a chloroplast-encoded protein, cyt *f*, is shown to share common translocation machinery with nuclear-encoded proteins. The authors provide evidence that the mutations interfere with targeting itself, rather than with TPP or SPP activity *per se*, do not disrupt the integrity of the thylakoid membrane and do not disrupt

global chloroplast gene expression. The *hcf*106 gene product has been localised to the thylakoid membrane, and is in the process of being characterised.

1.12.2 *The Sec-type pathway*

The first experimental evidence that the stromal protein factor necessary for the transport of 33K and PC across the thylakoid membrane is a SecA homologue came when Knott and Robinson (1994), showed that 33K and PC import into intact chloroplasts was markedly inhibited by the presence of azide, and completely inhibited when these precursors were imported into intact thylakoids. The translocation of precursors from the ΔpH-dependent pathway was unaffected by azide. Two groups were able to purify a SecA homologue from pea chloroplasts: Yuan *et al.* (1994) and Nakai *et al.* (1994). The former group of workers raised an antibody against a peptide conserved among bacterial SecA proteins, and those deduced from algal *secA* sequences. This antibody aided isolation of the SecA homologue from stromal extract, using conventional and high-performance chromatography, and a 250-fold purification was achieved. The purified protein (a homodimer of 200-250 kDa) could substitute for stromal extract in isolated thylakoid import assays for the uptake of PC and 33K, but could not replace stromal extract in assays for LHCP integration (a process which requires stromal protein, but which is not azide-sensitive). Nakai *et al.* (1994) used an antibody raised against cyanobacterial SecA to identify a 110 kDa protein in pea chloroplasts. They then used polymerase chain reaction (PCR) techniques to amplify a portion of the pea secA cDNA, which was over-expressed in *E. coli* and used to make antibodies against the pea SecA homologue. In agreement with the expected function of the homologous protein, and its distribution in *E. coli*, SecA was found mainly in the stroma and partly in the thylakoid membrane. The anti-pea SecA IgG inhibited the import of 33K, but not 23K into isolated thylakoid vesicles, agreeing with the results of Yuan *et al.* (1994).

Since then, Nohara *et al.* (1995) and Berghöfer *et al.* (1995) have sequenced cDNAs encoding pea and spinach chloroplast SecA homologues respectively. The pea cDNA encodes a polypeptide consisting of 1011 amino acids, with a calculated relative molecular mass of 114 130. It is 58% identical to the *Synechococcus* PCC7942 SecA

protein, and shares 48% identity with the *E. coli* protein. The pea SecA homologue is synthesised with a N-terminal presequence of around 60 amino acids, which serves as an envelope transit peptide. The spinach SecA homologue, on the other hand, is made as a polypeptide of 1036 amino acids (117 kDa) of which around 80 form a putative envelope transit peptide. Intact chloroplast import assays with an *in vitro*-translated, radiolabelled spinach SecA substrate demonstrated that it accumulates predominately in the stromal fraction, with some found in the thylakoids (in agreement with Nohara *et al.*, 1995). The spinach protein displays slightly higher identity to both *Synechococcus* PCC7942 (62%) and *E. coli* (49%) SecA proteins. Intriguingly, Berghöfer *et al.* (1995) present evidence that the Sec-dependent translocation pathway in spinach chloroplasts is azide-resistant, suggesting that the spinach SecA homologue differs structurally from other known SecA proteins - yet, the two residues conferring azide sensitivity in *E. coli* are conserved in spinach.

A full-length cDNA for a SecY homologue in *Arabidopsis thaliana* has been identified (Laidler *et al.*, 1995), and encodes a protein predicted to be 551 amino acids long, including a putative envelope transit peptide of approximately 120 amino acids. The deduced sequence of the mature protein shares 41% identity with Synechococcus and 33% with *E. coli*. Transcription and translation of the cDNA *in vitro* produces a 58 kDa precursor protein, which is imported into isolated pea chloroplasts, processed to a 46 kDa product and targeted to the thylakoid network. The identification of SecA and SecY homologues in plant chloroplasts is significant, since it represents the first identifiable components of the thylakoid protein translocation machinery.

1.12.3 *The ΔpH pathway*

As previously mentioned (section 1.12), identification by Voelker and Barkan (1995) of two mutants (*hcf*106 and *tha*1) allowed the only *in vivo* evidence for the two distinct thylakoid lumen protein translocation pathways. The *hcf*106 mutant was found to be affected in the ΔpH-dependent pathway and as a consequence the stromal intermediates of 23K and 16K were observed in large quantities. The *hcf*106 gene has recently been cloned (Settles *et al.*, 1997) and the encoded protein is, as expected, located in the thylakoid membrane. The predicted structure is consistent with a single

transmembrane span together with a large globular domain exposed to the stroma, and this protein may serve as the receptor for incoming substrates (although other proteins may well be involved in this translocation process).

Proteins that follow this pathway i.e. 23K and 16K, can be efficiently transported across the thylakoid membrane in the complete absence of stromal extract, which implies that the precursors do not need a molecular chaperone to maintain them in a translocation-competent form. That such a chaperone may be provided by the wheat germ extract used to synthesise the precursors is, however, possible, although Creighton *et al.* (1995) presented evidence that 23K is monomeric and tightly folded in the stroma. More recently, a study by Roffey and Theg (1996) suggested that a folded structure might even be required for translocation; these authors found that minor C-terminal deletions from one of the substrates affected the translocation by the ΔpH-dependent route. And further conformation of this came when Clark and Theg (1997) demonstrated that a chimaeric 16K (termed OE17 by the authors) bovine pancreatic trypsin inhibitor fusion could be translocated across the thylakoid membrane when the passenger protein was fully folded. However, the passenger protein in this case has a relatively small molecular size (6.5 kDa), and in its fully folded state the fusion protein could be transported across the chloroplast envelope membranes.

Further experiments by Creighton *et al.* (1995) indicate that surface-exposed proteins in the thylakoid membrane are necessary for i23K translocation, for at certain concentrations, trypsin treatment abolishes import without interfering with the action of the chloroplast ATPase (which can be driven in the reverse direction to supply a ΔpH in the dark).

1.13 *Targeting and insertion of nuclear-encoded thylakoid membrane proteins*

Membrane-spanning proteins are the major components of the photosynthetic complexes, yet relatively few have been characterised in terms of their insertion mechanism. Of these, the best studied example is the light-harvesting chlorophyll

a/b-binding protein (LHCP). In terms of targeting signal, this protein differs fundamentally from lumenal proteins because it is synthesised with a presequence that functions solely to target the protein across the envelope membranes. Thylakoid targeting and integration information is contained in the mature protein (Viitanen *et al.*, 1988; Lamppa, 1988), which contains three transmembrane spans (discussed in **section 1.13.3**). Indeed, most multispanning thylakoid membrane proteins are synthesised with this type of presequence, and cleavable thylakoid-targeting peptides appear to be used primarily, if not exclusively, for the targeting of lumenal proteins and a subset of membrane proteins that contain a single membrane span (discussed in **section 1.13.4**).

1.13.1 *Membrane topology*

The fact that integral membrane proteins are, by definition, at least partially buried in the lipid bilayer appears to limit the types of structures that they can adopt, due to the physiochemical constraints that the hydrophobic environment imposes. So far, the protein structures that have been seen in membranes are the α-helices (20-25 amino acid residues in length), composed mainly of hydrophobic amino acids flanked by clusters of polar residues, and the β-barrels, in which hydrophobic residues face outward, towards the lipid bilayer.

The α-helix-type proteins are most abundant and can be made up of a single helix or of multiple helices packed together in bundles. The orientation (topology) of helices in the membrane of many membrane systems (such as the bacterial inner membrane, the plasma membrane of eukaryotic cells, the inner membrane of mitochondria and the thylakoid membrane of chloroplasts) is determined primarily by the distribution of positively charged amino acids flanking the transmembrane region according to the 'positive-inside' rule (von Heijne, 1994). By altering the distribution of the positively charged amino acids in loops between transmembrane segments, a proteins topology in the membrane can be altered (Nilsson and von Heijne, 1990).

1.13.2 *Identification of two distinct pathways for the insertion of thylakoid membrane proteins*

1.13.3 *The SRP-type mechanism*

The vast majority of research in this area has centred on the major light-harvesting chlorophyll *a/b* binding protein of PSII (LHCP). This protein is synthesised in the cytosol and imported post-translationally where it is cleaved, often resulting in multiple products (for example, 25 kDa and 26 kDa for wheat LHCP; Clark and Lamppa, 1991). The presequence of LHCP has been shown to be an envelope transit peptide only, with all the information for thylakoid membrane insertion residing in the mature protein sequence (Lamppa, 1988; Viitanen *et al.*, 1988). The exact location of this targeting information is unclear, but it may be collectively contained in all three membrane-spanning helices of the protein, based on studies by Huang *et al.* (1992). The authors reported both the synthesis of deletion mutants of LHCP, where the protein was deleted after the second and first membrane-spanning helices, and also the construction of three fusion proteins, where each helix and its flanking hydrophilic amino acids was fused to the LHCP envelope transit peptide. In every case, the mutant proteins were imported into chloroplasts, but were found primarily in the stromal fraction, and any association with the thylakoid membrane was found to be non-productive, because the polypeptides were sensitive to added proteases.

The mechanism by which LHCP integrates into thylakoids has been studied in detail, initiated by Cline (1986), who was able to reconstitute integration of pre-LHCP in chloroplast lysates. He found that the precursor protein integrated specifically into thylakoid, but not envelope membranes, and that ATP was required in the mixture. Soluble (stromal) protein extract was also necessary. The full precursor was integrated, suggesting that processing can take place after integration. This finding was in agreement with Chitnis *et al.* (1986) who found that pre-LHCP could be imported to barley etioplasts, then assembled into the light-harvesting complex of PSII as either the precursor or processed form. It was also found that a combination of nigericin and valinomycin (ionophores) inhibited integration of LHCP by up to 70%, demonstrating the importance of an energised thylakoid membrane (Cline *et al.*, 1989).

Chitnis *et al.* (1987) observed that the insertion of LHCP into the thylakoids of young barley plastids was dependent upon stromal factor, which was developmentally regulated. This finding was extended by Payan and Cline (1991) who demonstrated that a stromal protein factor binds imported LHCP to produce a 120 kDa complex, which prevents aggregation and maintains insertion competence. They emphasised however, that this complex still requires the presence of stromal extract for integration to thylakoid membranes. Due to the chaperone-like function of this stromal factor, it was suggested that hsp70 was the candidate (Yalovsky *et al.*, 1992), but this was refuted by Yuan *et al.* (1993) on the grounds that immunoprecipitation techniques did not reveal hsp70 in the complex, and stromal extract depleted of hsp70 could still support both transit complex formation and membrane integration. Similarly, cpn60 cannot be immunoprecipitated from the soluble transit complex (Payan and Cline, 1991). The mystery appears to have been at least partly solved by the identification and characterisation of a chloroplast homologue of the 54 kDa subunit of signal recognition particle (SRP), a soluble sorting factor involved in targeting proteins to the ER membrane, and implicated in prokaryotic protein export (Franklin and Hoffman, 1993; Li *et al.*, 1995). Franklin and Hoffman (1994) re-examined the NTP requirement for LHCP integration, and found that GTP, not ATP was needed, and that it was hydrolysed (non-hydrolysable analogues of GTP inhibit LHCP integration). Protein transport across (or integration to) the ER membrane depends on a mechanism mediated by GTP and SRP, so it was a natural progression to look for involvement of the chloroplast SRP homologue in the integration of LHCP. Li *et al.* (1995) used monospecific antibodies to the *A. thaliana* chloroplast SRP homologue (54CP) to immunoprecipitate a transit complex from stromal extract, containing both 54CP and LHCP. The association of LHCP with 54CP was confirmed by chemical cross-linking, followed by immunoprecipitation with anti-54CP antibodies. The cross-linked product was about 82 kDa, consistent with a stoichiometry of 1 molecule of 54CP: 1 molecule of LHCP. While *in vitro*-translated 54CP could not be made to form a transit complex when mixed with stroma, or with purified pre-LHCP, it did so after it had been imported into isolated, intact chloroplasts. Immunodepletion of 54CP from stromal extract abolished the ability of the extract to support integration of LHCP; a parallel depletion of hsp70 failed to exert any significant effect. Importantly, 54CP does not interact stably with either

pre-33K or pre-23K, which are representative proteins from the two mechanisms responsible for targeting to the thylakoid lumen. This is an interesting observation, because both proteins contain signal peptide motifs within their presequences which should be recognised by mammalian SRP. Significantly, this is the first demonstration of SRP supporting post-translational translocation in the absence of ribosomes (Wickner, 1995).

It is likely that 54CP does not function as a monomer - the stromal factor/LHCP complex of Payan and Cline (1991) had a molecular weight of around 120 kDa. Li *et al.* (1995) suggest that, because 54CP had to be imported to chloroplasts in order to form a transit complex, it may need the chloroplast protein synthesis apparatus for assembly into a multisubunit complex. Mammalian SRP is a ribonucleoprotein, containing six polypeptides and an RNA molecule. The authors believe that 54CP pilots LHCP to the thylakoid membrane, because only the integration process itself and not transit complex formation, requires GTP hydrolysis, and 54CP had GTPase activity.

1.13.4 *The direct insertion pathway*

The emergence of a fourth targeting pathway has made the biogenesis of thylakoid proteins even more complex. The first evidence for this pathway was with subunit II of the CF_0 component of the chloroplast ATP synthase (CF_0II). The gene encoding it (*atp*G) is found in the nucleus, but appears to be a duplication of *atp*F, the plastid gene encoding CF_0I (Herrmann *et al.*, 1993). CF_0II is synthesised in the cytosol as a precursor with a bipartite presequence typical of thylakoid lumen proteins. Unlike 16K, 23K, 33K and PC however, the precursor is not cleaved to an intermediate form by SPP in the chloroplast stroma (Michl *et al.*, 1994). Assays for integration of CF_0II to the thylakoid membranes of intact chloroplasts and isolated thylakoids, show that this protein differs from all previously studied precursors in having virtually no detectable requirements for membrane insertion. The thylakoid ΔpH does not appear to play an important role in membrane integration, for nigericin has only a very slight inhibitory effect (Michl *et al.*, 1994). Furthermore, prior treatment of thylakoids with proteases was found to block import by the Sec-, SRP- and ΔpH-dependent mechanisms, whereas the insertion of CF_0II was unaffected (Michl *et al.*, 1994). It

was concluded, therefore, that this protein is targeted to the thylakoid membrane by a novel pathway.

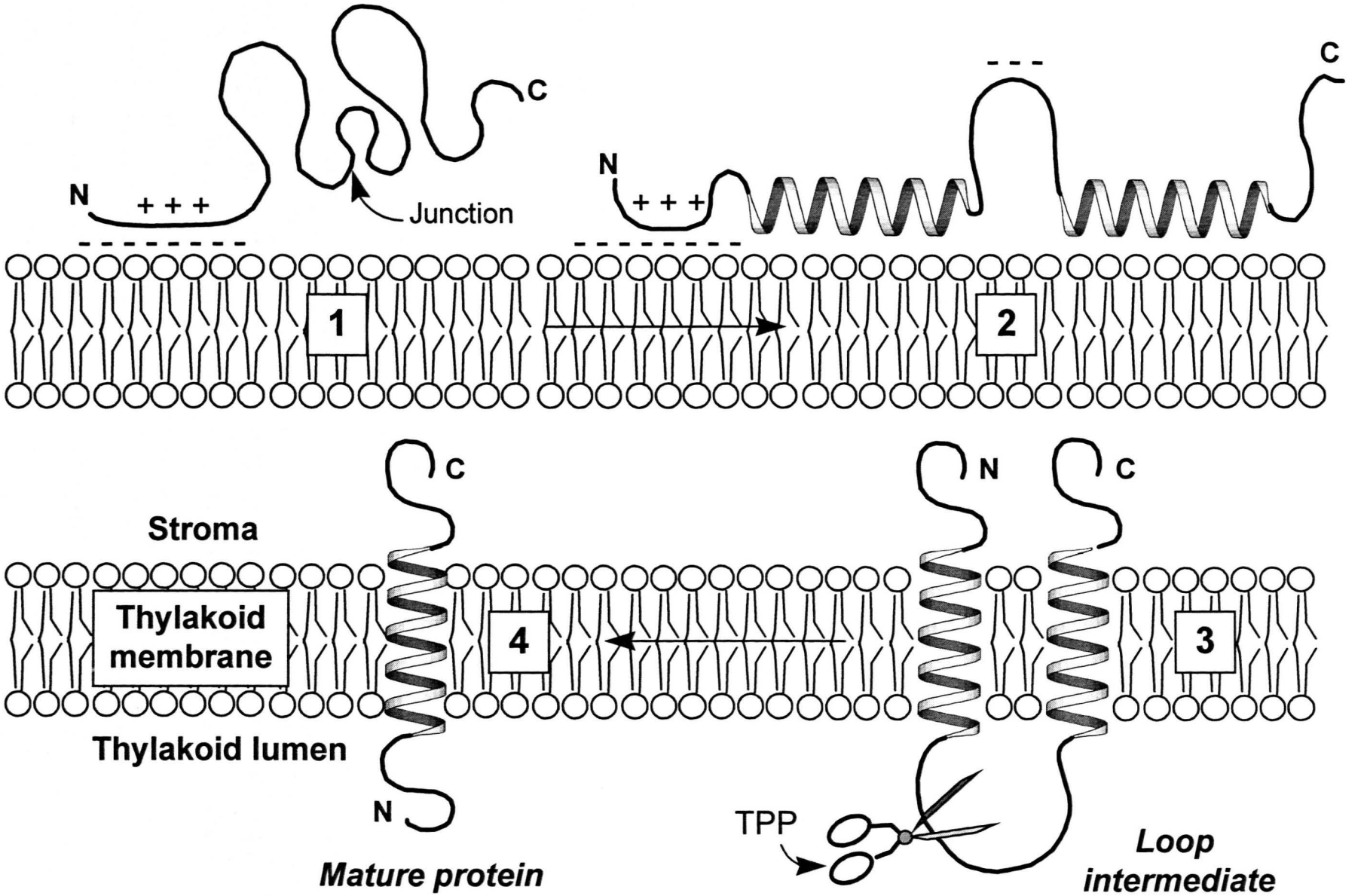
Two further proteins have been found to follow a similar route: subunits X and W of PSII (Lorokovic *et al.*, 1995; Kim *et al.*, 1996). Like CF_0II, each of these proteins spans the thylakoid membrane once, with the N-terminus in the lumen. And remarkably, these proteins insert into thylakoids in the absence of stromal factors, NTPs or trypsin-sensitive apparatus (Kim *et al*, 1998), and typical signal peptides can clearly be discerned in the presequence.

A direct insertion model has been proposed (**figure 10**) that takes into account structural similarities between these proteins and a phage protein, M13 procoat, that is likewise synthesised with a signal peptide and inserted by a Sec-independent mechanism in *E. coli* (Kuhn, 1995). In this model, the cleavable signal peptide does not interact with translocation machinery at all; its role is simply to provide an additional (and temporary) second hydrophobic domain. The two hydrophobic regions then synergistically partition into the thylakoid membrane and flip the intervening (negatively charged) sequence across, after which cleavage by TPP on the *trans* side of the membrane yields the mature protein in a transmembrane form.

There are also indications that a fourth protein may follow the same route but in this case the structure is rather different, at least in the initial stages of the import process. A cDNA clone encoding a novel thylakoid membrane protein has recently been found from *Arabidopsis thaliana* and shown to be homologous to the *ycf*32 unassigned reading frames present in cyanobacteria and the plastid genomes of several eukaryotic algae (Mant and Robinson, 1998). However, whereas the putative Ycf32 proteins are all small and predicted to contain a single transmembrane span, the *Arabidopsis* Ycf32 protein (termed PsbY in this thesis) is far more complex. The primary translation product includes an apparent bipartite presequence and the 'mature' protein contains three transmembrane spans (as shown by the authors hydrophobicity profile) of which two (the first and the third) show high homology to algal/bacterial Ycf32 proteins. The authors proposed that this is in fact a polyprotein that undergoes further maturation once in the thylakoids to yield two separate Ycf32-

Figure 10 *The direct insertion model for nuclear-encoded thylakoid membrane proteins.*

This is a working model for the direct insertion pathway for thylakoid membrane proteins. CF_0II, PSII-W and PSII-X (PS, photosystem) are believed to integrate by a direct mechanism that does not rely on protein transport apparatus in the thylakoid membrane. In this model, (**1**) the precursor protein may first bind to the thylakoid membrane by electrostatic interactions between the basic N-terminal region of the signal peptide and negatively charged membrane surface. At this point (**2**), possibly before, the hydrophobic regions in the signal peptide and mature protein are believed to form -helices, after which (**3**) the simultaneous partitioning of these helices into the membrane drives translocation of the intervening acidic region (the N-terminus of the mature protein) across the thylakoid membrane. Cleavage by TPP ('thylakoidal processing peptidase') leaves (**4**) the mature protein in a transmembrane form.


N
+ + +
Junction
C
N
+ + +
C
1
2
Stroma
Thylakoid membrane
Thylakoid lumen
C
N
4
N
C
3
TPP
Mature protein
Loop intermediate

like proteins, in which case each may be preceded by a hydrophobic signal-type peptide. The import and maturation of Ycf32 was found to be unaffected by nigericin or azide, suggesting that this protein(s) is imported by the 'spontaneous' pathway.

Recent sequencing projects have made it possible to trace the likely evolutionary development of the various signal-type peptides used in the targeting mechanisms described in **section 1.12**. Lumenal substrates for the Sec machinery are synthesised with signal peptides in cyanobacteria and the chloroplast genomes of several eukaryotic algae (for example, *Odontella sinesis* and *Cyanophora paradoxa*). It seems likely (and understandable) that the transfer of genes to the plant cell nucleus has resulted in the acquisition of an envelope transfer signal to get the protein into the organelle, with the subsequent thylakoid transfer step remaining almost unchanged over the intervening billion or so years. In contrast, PSII-X and CF_0II are not encoded with signal peptides in either cyanobacteria or the algal chloroplast genomes sequenced to date. The implication is that these proteins could insert directly into the membrane - perhaps co-translationally - when synthesised in the vicinity of the membrane, but that this mechanism could not be utilised when the protein was imported from the cytosol. Instead, the organisms have made good use of the TPP activity resident in the thylakoid, and the proteins have acquired signal peptides to assist in the membrane insertion process. Thus, this class of signal peptide differs fundamentally from true Sec substrate signal peptides in terms of both function and evolution.

The unprecedented simplicity of this integration mechanism (in terms of absence of a requirement for stromal factors, NTPs and ΔpH) indeed raises the possibility that these proteins insert into the thylakoid membrane spontaneously, and that the only thylakoid protein involved in the overall insertion process is TPP. However, it can not be ruled out that the incoming proteins may interact with alternative, as yet unidentified thylakoid membrane proteins.

1.14 *Targeting of proteins to the outer envelope membrane*

Among the five or so proteins whose targeting to the outer envelope membrane has been investigated, a variety of different mechanisms has been uncovered. Three

proteins: one of 6.7 kDa (Salomon *et al.*, 1990), one of 14 kDa (Li *et al.*, 1991) and an hsp70 homologue (Ko *et al.*, 1992) are not synthesised with a cleavable presequence, and their targeting requires neither ATP, nor a proteinaceous receptor. Li *et al.* (1991) showed that synthetic peptide analogues of the SUU presequence failed to inhibit specific integration of the 14 kDa protein to the outer envelope membrane, which strongly suggests that the protein does not use the usual import receptor.

Fischer *et al.* (1994) cloned and characterised an outer envelope polypeptide, named the 24 kDa protein (although this is much larger than the real molecular mass of 16.3 kDa). The protein is unusual for several reasons. Despite being acidic, with a low hydrophobicity and a high proline content, it is deeply embedded in the outer membrane. Unlike the three outer membrane proteins mentioned above, its integration is stimulated by ATP. The authors suggest that the ATP may be utilised by cytosolic components which aid interaction of the 24 kDa protein with the outer membrane.

The 75 kDa protein (Toc75) of the chloroplast import apparatus is targeted to the outer membrane by a completely different pathway, as shown by Tranel *et al.* (1995). Toc75 is synthesised as a precursor, and is processed to the mature size *via* two intermediates observed during chloroplast import assays. If the two intermediates really are productive, then Tranel *et al.* (1995) propose that one portion of the presequence directs the precursor to the chloroplast, and additional domains then target the protein to the outer membrane. This hypothesis implies that Toc75 utilises the general import apparatus, which is supported by the finding that surface-exposed, proteinaceous components are required, and that pre-Toc75 competes with pre-SSU for one or more of these components. A low concentration of ATP (*ca.* 50 μM) was sufficient to allow both binding and processing to the mature-size protein, but only supported binding of pre-SSU. It is possible that a complex system exists for targeting components of the import apparatus in order to ensure that their substrates are not subsequently mis-targeted. It is known that proteins from the mitochondrial import apparatus interact directly with one another in order to achieve correct targeting. In line with this, the 86 kDa chloroplast "import receptor" (Toc86) does

not compete with pre-SSU for import (Hirsch *et al.*, 1994), and may therefore fulfil an analogous role to Tom20, the mitochondrial "master receptor".

1.15 *Targeting of proteins to the inner envelope membrane*

So far, three proteins of the inner envelope membrane have been the subject of investigation: maize brittle-1 protein (Bt-1), which is a metabolite translocator located in the amyloplast membrane (studied by Li *et al.*, 1992), the triosephosphate-3-phosphoglycerate phosphate translocator (TPT) and the 37 kDa protein of unknown function (both studied by Brink *et al.*, 1994 and Brink *et al.*, 1995). In each case, the protein in question is synthesised as a precursor with an envelope transit peptide, and the information for integration resides in the mature protein domain. Interestingly, TPT and the 37 kDa protein both possess unusual envelope transit peptides, which display the amphipathic, α-helical characteristics more usually found in mitochondrial precursor proteins. As predicted by Brink *et al.* (1994), isolated plant and *Neurospora crassa* mitochondria were able to import these precursors, a process which was shown in the latter organism to be dependent on the main import receptor protein, Tom20 (**section 1.18.1**). Brink *et al.* (1995) synthesised chimaeric precursor proteins, linking TPT or 37 kDa presequences to the mature domains of LHCP and CF_0II (both thylakoid membrane proteins), and also linking the CF_0II presequence (containing signals for routing to the thylakoid membrane) to TPT and 37 kDa mature proteins. The inner membrane protein precursors directed the chimaeras into the stroma, whereupon the passenger protein (LHCP) reached its correct location in the thylakoid membrane. The bipartite presequence of CF_0II was unable to divert the mature TPT or 37 kDa proteins to the thylakoid membrane; they integrated correctly to the inner envelope, thus demonstrating the sufficiency of the targeting information contained in the mature protein domains of TPT and 37 kDa proteins. Deletion of a putative membrane anchor at the C-terminus of the 37 kDa protein resulted in the protein being found in thylakoid, envelope and stromal fractions, with envelope membrane integration being incomplete (as judged by sodium carbonate extraction). This implicated the putative anchor in both intraorganellar sorting, and anchoring of the protein in the membrane.

Similar results, using chimaeras consisting of the presequence of SSU fused to the mature domain of Bt-1 confirmed that Bt-1 integrates by means of information contained within the mature protein (Li *et al.*, 1992). It is not clear whether any of the proteins studied are sorted conservatively (targeted to the stroma, then re-directed to the inner envelope membrane), or whether they are inserted directly. Brink *et al.* (1995) favour the latter hypothesis, at least for the 37 kDa protein, based on structural resemblance to subunit Va of mitochondrial cytochrome *c* oxidase, which is thought to be inserted directly.

1.16 *Targeting of proteins to the intermembrane space*

At the moment, it is not possible to isolate components of the intermembrane space and distinguish them from stromal proteins, so there is no information available on targeting to this compartment. Neither Soll and Bennet (1988) nor Salvucci *et al.*, (1990) were able to make unequivocal identifications of proteins from the intermembrane space.

1.17 *Targeting of chloroplast-encoded proteins*

Some progress has been made in the investigation of two chloroplast-encoded proteins: cytochrome *f* (cyt *f*) from the b_6f complex and D1, one of the core polypeptides from photosystem II. Both these proteins are synthesised on ribosomes attached to the thylakoid membrane, as larger precursors. Cyt *f* (which is a transmembrane protein, discussed by Prince and George, 1995) is synthesised with an N-terminal signal sequence, similar to prokaryotic secretory signal sequences and thylakoid transfer signals. D1, on the other hand, has a C-terminal extension, which is cleaved (in *Synechocystis* PCC6803 at least), by the *ctpA* gene product (Anbudurai *et al.*, 1994).

Over ten years ago, SecA was implicated in the targeting of cyt *f* to the thylakoid membrane, based on experiments where pea cyt *f* gene was fused to the *E. coli* β-galactosidase gene to express chimaeric proteins in *E. coli* (Rothstein *et al.*, 1985). Chimaeras with the cyt *f* signal sequence intact were efficiently targeted to the bacterial cytoplasmic membrane, where secretion was initiated, but did not proceed

to completion. Membrane insertion was shown to require host SecA. More recently, cytochrome *f* was fused to the envelope transit peptide of SSU in order to assay integration to the thylakoid membrane in isolation from the translation process (Mould *et al.*, 1997). The chimaeric protein was imported to pea chloroplasts and integrated into the thylakoid membrane. A truncated chimaera (missing the stromal, C-terminal domain of cyt *f*) was imported to the thylakoid lumen instead. Integration of the full-length chimaera did not depend upon the presence of a Δp, but was inhibited by sodium azide, further suggesting the involvement of SecA in the integration process.

There is no evidence, however, for SecA taking part in the D1 integration process. D1 is one of the most strongly labelled proteins when chloroplast protein synthesis is monitored in the presence of radioactive amino acids, due to its high rate of turnover. Van Wijk *et al.*, (1995a) optimised procedures for analysing D1 replacement in both intact chloroplasts, and isolated thylakoids with attached ribosomes (a homologous ribosome run-off system). They showed that stromal factors are required for proper replacement of D1 into fully-assembled PSII complexes, and that the synthesis of D1 appears to be regulated in line with the availability of PSII complexes in which to insert. The stromal factors may not include SecA, however, as there is some evidence that sodium azide, and ionophores, fail to inhibit D1 insertion in intact chloroplasts (van Wijk *et al.*, 1995b). And to date, there is still no evidence favouring any particular stromal factor in the insertion of D1 protein.

1.18 *Comparison of chloroplast protein transport with that of other eukaryotic protein translocation systems*

1.18.1 *Mitochondria*

Around ten proteins required by this organelle are encoded by the mitochondrial genome; the rest (several hundred) must be imported from the cytosol (Pfanner *et al.*, 1994; Lill and Neupert, 1996). Most of the imported proteins are synthesised with N-terminal presequences, unless they are destined for the outer membrane (OM). Other exceptions include some of the proteins of the intermembrane space (IMS) and inner membrane (IM), and a small number of matrix proteins. Von Heijne *et al.*, (1989)

compared and contrasted mitochondrial presequences with those of chloroplasts. Like chloroplast envelope transit peptides, mitochondrial presequences do not contain stretches of highly conserved sequence. They are, however, divided into two sub-domains: an N-terminal region with potential to form an amphiphilic α helix, and a C-terminal domain, often containing arginine at positions -2 and -10, relative to the cleavage site. Chloroplast envelope transit peptides, on the other hand, display three sub-domains (see **section 1.7.1**). Statistical analyses showed that 90% of the time, mitochondrial presequences could be distinguished from chloroplast presequences purely by amino acid content. Mitochondrial presequences are enriched for arginine, serine and alanine, whereas threonine and especially serine are enriched in chloroplast envelope transit peptides. The intraorganellar targeting signals of both mitochondrial and chloroplast precursor proteins resemble signal peptides, although the final mitochondrial cleavage site does not conform to the -3, -1 pattern adhered to by TPP (chloroplasts), signal peptidase (eukaryotes) and leader peptidase (prokaryotes).

Very little is known about the mechanism by which cytosolic polypeptides are targeted to the surface of the chloroplast, but recently, knowledge about the mitochondrial system has been extended (reviewed by Ryan and Jensen, 1995; Mihara and Omura, 1996). Several cytosolic factors have been purified, which stimulate import into isolated mitochondria, but the most promising one so far is mitochondrial import stimulation factor (MSF), purified from rat liver cytosol. The reason MSF is so interesting is that it recognises mitochondrial protein presequences, and targets them to the import receptor. Unlike other cytosolic chaperones implicated in mitochondrial protein import, MSF is specific to mitochondria. It is a heterodimer of 32- and 30 kDa subunits, which has both precursor stabilising, and ATP-dependent unfolding activities. Importantly, MSF stimulates *in vitro* import of all mitochondrial precursors tested so far, even those without cleavable presequences, and proteins of the OM. Observations of MSF interactions with a variety of mitochondrial precursor proteins suggest that MSF does not just recognise the mitochondrial presequence (although this is an important feature), but also senses the conformational state of the mature domain. An alternative route to the surface of the mitochondrion is likely to be mediated by hsp70, acting in concert with another

factor conferring mitochondrial specificity. This pathway does not require ATP, unlike the MSF pathway. Mihara and Omura (1996) propose a model where precursor proteins which are able to maintain a relatively unfolded conformation associate with hsp70, are targeted to the OM receptors, Tom20 and Tom22 (independent of cytosolic ATP), and are imported. Other precursors are bound by MSF, which targets them to the receptors, Tom37 and Tom70, where they are transferred to Tom20 and Tom22 while MSF is released (ATP dependent). Their model is supported by *in vitro* import experiments using antibodies to OM receptors in the presence and absence of precursor proteins, ATP, non-hydrolysable ATP analogues, MSF and hsp70.

There are two, independent, mitochondrial translocases - one in the OM (reviewed by Lithgow *et al.*, 1995; Lill and Neupert, 1996) and one in the IM (reviewed by Horst *et al.*, 1993; Pfanner *et al.*, 1994). Many of the protein components have been characterised by different laboratories, so researchers in the field have agreed to unify the nomenclature, such that translocase components of the OM are given the prefix "Tom", followed by the relevant molecular mass, and components of the IM translocase are prefixed with "Tim". Corresponding genes are given the same name, but in capitals, and italicised (Pfanner *et al.*, 1996).

The two translocases (TOM and TIM complexes) may be independent, but they are co-ordinated both spatially and temporally to form transient contact sites for the translocation of precursor proteins across the two mitochondrial membranes. The TOM complex fulfils several functions. Firstly, it acts as a receptor, and then it passes the presequence across the OM - complete import requires the action of the TIM complex. The TOM complex is also responsible for the import of some proteins into the intermembrane space, and the insertion of others directly into the OM. **Figure 11** shows the identified components of the translocation machinery, as discovered in *N. crassa* and *S. cerevisiae*. The Tom proteins fall into two distinct groups: those with cytosolic domains, which act as receptors (Tom20, Tom70, Tom71, Tom37 and Tom22), and those which are embedded in the OM, and probably form some sort of translocation channel (Tom40, Tom6, Tom7 and Tom8). Tom70 and Tom71 are highly homologous, and contain seven 34-mer repeated

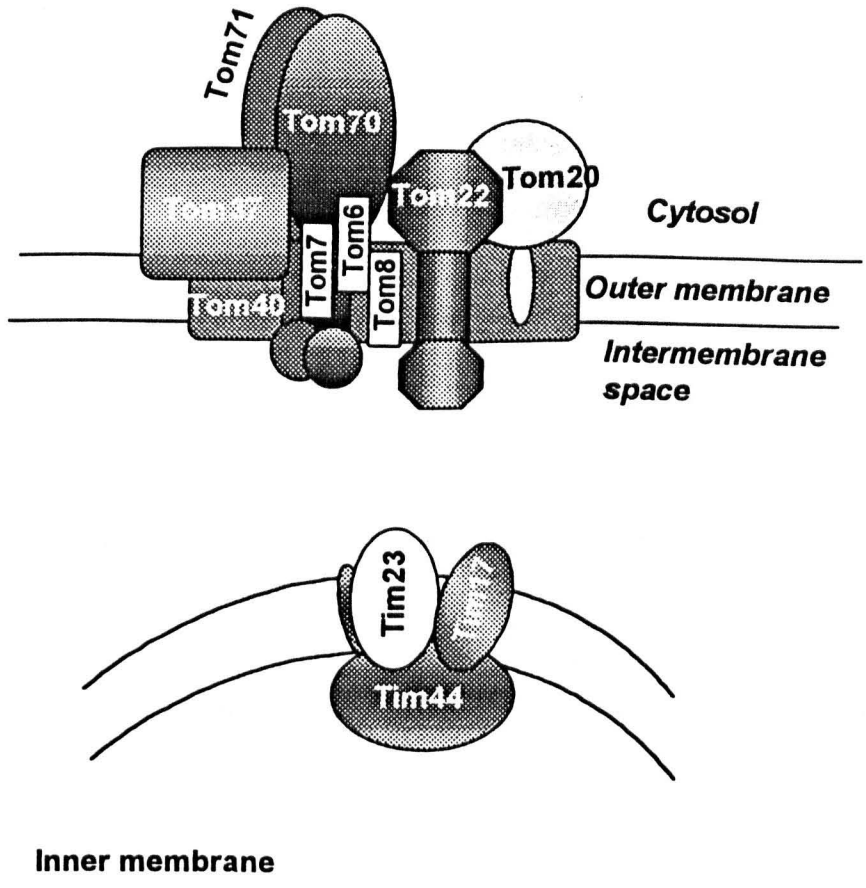


Figure 11 *The mitochondrial TOM and TIM complexes*

Adapted from Neupert (1997)

Tom70-Tom71-Tom37, and Tom20-Tom22 are the receptors. The remaining Tom proteins probably form the translocation channel. Tim17, Tim23 and Tim44 are the only components of the inner membrane identified so far. Soluble chaperones, and several processing peptidases are also involved in protein import

motifs, believed to function in protein-protein interactions. They act with Tom37 to form a functional unit which recognises a subset of mitochondrial precursor proteins (most likely those targeted by MSF), including alcohol dehydrogenase and cytochrome c_1. Other precursor proteins are directly recognised by Tom20-Tom22, which forms the second receptor, and also receives precursor proteins targeted initially to Tom70-Tom71-Tom37. The receptor is essential for the growth of *S. cerevisiae*. Although, Tom20 and Tom22 act together to bind precursor proteins *via* electrostatic interactions, Tom22 will substitute for Tom20, when over-expressed in cells lacking *TOM20*. Lill and Neupert (1996) propose that Tom20-Tom22 funnels the precursor protein into the translocation pore, after which the presequence associates with a specific binding site (termed the *trans* site), on the other side of the membrane (because matrix processing peptidase introduced to the lumen of OM vesicles can cleave the presequence of proteins attached to this binding site). This *trans* site has not been characterised, but a likely candidate is Tom40, because it contains some regions exposed to the IMS. The significance of the *trans* site is that it may mediate early unfolding of the precursor protein before matrix hsp70 can be involved - precursor proteins must be in an extended conformation to cross the membrane.

Three proteins of the TIM complex have been identified: Tim17, Tim23 and Tim44. The first two appear to form a translocation channel, while the latter has an interesting function, which is to co-operate with hsp70 in pulling proteins into the matrix. All four of these proteins are essential for the viability of *S. cerevisiae*. Tim17 and Tim23 are predicted to possess four membrane-spanning regions each, and show similarities in sequence and hydrophobicity, except in the hydrophilic N-terminus of Tim23, which is absent in Tim17. Both polypeptides can be cross-linked to precursor proteins crossing the IM. Tim44 is believed to be a peripheral membrane protein, with the majority exposed in the matrix. Around 10-20% of mitochondrial hsp70 is reversibly associated with Tim44, and both proteins can be cross-linked to translocating precursors. Two energy sources are required for the completion of the import process: matrix ATP and the inner membrane $\Delta\Psi$ (negative inside). Protein import into chloroplasts requires only ATP, as the chloroplast inner envelope, unlike the mitochondrial IM, is not the energy-transducing membrane. Pfanner and Meijer

(1995) discuss the mechanisms driving the translocation of mitochondrial precursor proteins across the IM. The contributions of the energy sources can be dissected by inactivating the mitochondrial hsp70 of a temperature-sensitive *S. cerevisiae* mutant. In the absence of hsp70, $\Delta\Psi$ drives the initial portion of the presequence across the IM - this transferral is stable if there is a high content of basic amino acids in the presequence, implying the mechanism is electrophoretic. Further directional import is absolutely dependent on the presence of hsp70, otherwise translocating precursor proteins were observed to oscillate in the membrane channels. There are two models for the function of mitochondrial hsp70 in protein import, both of which may be accurate explanations for the events taking place (described by Pfanner and Meijer, 1995). The first is the Brownian ratchet mechanism, whereby directional movement of the precursor protein is achieved by the binding of hsp70 to the N-terminus of the presequence, trapping it in the matrix. Brownian motion occurs in all directions, but if reverse and sideways movement are prevented (by hsp70 and the walls of the channel), net movement would occur in a forward direction. ATP hydrolysis drives the binding-release cycle of hsp70 in this case. The second model invokes concerted pulling of the precursor protein (which also functions to unfold it on the cytoplasmic side of OM) by hsp70 and Tim44. It is envisaged that the emerging polypeptide binds first to Tim44, and is then handed over to hsp70 (also bound to Tim44). Next, hsp70 undergoes a change in conformation, which pulls the polypeptide forward, and then hsp70 dissociates from Tim44, still clutching a segment of the polypeptide. The conformational change and release consume ATP. Only in partnership with Tim44, is hsp70 capable of unravelling tightly folded precursor proteins.

Like imported chloroplast proteins, new arrivals in the mitochondrial matrix interact with molecular chaperones to achieve their native conformations. Firstly, however, matrix processing peptidase cleaves the presequence to yield either the mature protein, or an intermediate ready for re-targeting to the intermembrane space (matured there by inner membrane protease I, related to chloroplast TPP). Two *S. cerevisiae* matrix chaperones homologous to bacterial ones have been identified (discussed by Stuart *et al.*, 1994). These are MDJ1 (resembling DnaJ) which is associated with the inner face of the IM, and MGE (resembling GrpE). The former is non-essential, but promotes refolding of denatured proteins at elevated temperatures,

while disruption of the gene for MGE is lethal in yeast. It is suggested that MDJ1 and MGE act after hsp70 to assist folding directly, or else they act together with hsp60 (utilising energy from ATP hydrolysis) and hsp10.

1.18.2 *Endoplasmic reticulum*

The best studied system is that of mammals, based on work carried out with canine pancreatic microsomes (the pancreatic cells have an important secretory function, and are therefore enriched with endoplasmic reticulum). Most secretory proteins in eukaryotes are transported across the ER by a co-translational mechanism.

The ER system is the first one where the environment in which precursor proteins cross the membrane has been elucidated. Use of fluorescent probes attached to the side-chain of the lysine moiety in Lys-tRNA, and thence incorporated into the nascent polypeptide, has enabled 'sensing' of the environment experienced by the signal peptide and the mature-sized protein (reviewed by Johnson, 1993; Matlack and Walter, 1993). The outcome of these experiments is that the signal sequence is in an aqueous pore, rather than inserting to the apolar core of the ER membrane, as was originally supposed. Furthermore, this pore is a tunnel, gated at both ends (reviewed by Gilmore, 1993). Similar experiments with fluorescent probes have shown that, although the signal peptide is in an aqueous environment early on in translocation, it is not accessible from the cytoplasm. Rather, the ribosome forms a tight seal around the circumference of the ribosome/membrane interface. Because of this, no special mechanisms have to be invoked for movement of the polypeptide chain through the pore. Simple diffusion down a polypeptide concentration gradient will result in transport through the aqueous pore, and the tight ribosomal seal would prevent backward diffusion into the cytosol. In the ER lumen, there are three major proteins which act as molecular chaperones: protein disulphide isomerase, binding protein (BiP, an hsp70 homologue) and GRP94 (an hsp90 homologue). The yeast homologue of BiP can be cross-linked to translocation intermediates in *in vitro* assays, which suggests that the chaperone assists in translocation as the polypeptide emerges from the membrane pore. This may take place mediated by cycles of BiP binding to, and dissociation from (which is dependent of ATP hydrolysis), the polypeptide (discussed by Dierks *et al.*, 1993; High and Stirling, 1993).

Possession of a signal peptide, with the properties described in **section 1.19.1**, is necessary for proteins to be targeted to the ER. Most secretory proteins are targeted to the membrane by signal recognition particle (SRP), which in mammals is a ribonucleoprotein, comprising a 7S RNA surrounded by six polypeptides of 9, 14, 19, 54, 68 and 72 kDa. SRP particles, some of which contain only the RNA and a 54 kDa homologue, have been identified in diverse organisms, including *E. coli*, *A. thaliana*, *Lycopersicon esculentum*, *S. cerevisiae* and *Canis familiaris*, which emphasises the central role of SRP complexes in protein targeting. In mammals, SRP will only recognise the signal peptides of nascent polypeptide chains as they emerge from the ribosome, although it now appears that NAC (nascent polypeptide associated complex) competes with SRP for early binding, and may regulate the association of SRP with the signal peptide (reviewed by Wickner, 1995; Rassow and Pfanner, 1996). It is the 54 kDa subunit of SRP which is sufficient for signal sequence binding (Lütcke *et al.*, 1992), and which also possesses GTPase activity. Binding of SRP to a signal peptide retards further elongation of the polypeptide until the sequence is released at the ER. The interaction of SRP with a signal peptide nearly always requires the simultaneous presence of the ribosome; therefore the ribosome is also targeted to the ER with the nascent polypeptide. At the membrane, GTP-dependent docking occurs at the SRP receptor, which is a heterodimeric protein, composed of α and β subunits, each possessing GTPase activity. In the presence of GTP, SRP detaches from the ribosome, binds tightly to the SRP receptor, and the signal sequence is released, to interact with integral membrane proteins while the rest of the polypeptide is translated. GTP hydrolysis allows dissociation of SRP from its receptor and a return to the cytosol. The steps in the SRP cycle are reviewed by Lütcke (1995).

Some proteins can be targeted to the ER independently of SRP, and are usually those which are translocated post-translationally. Hsp70 proteins have been implicated in this process, presumably maintaining the precursor proteins in a conformation competent for interaction with the membrane apparatus, and translocation (discussed by High and Stirling, 1993). Very little is known about SRP-independent targeting to the ER.

The core translocating protein, and the ribosome receptor, is the Sec61 complex, also found in *S. cerevisiae*, which is essential for insertion of proteins to the membrane (Oliver *et al.*, 1995). In mammals, it is composed of α, β and γ subunits, of which the α subunit is homologous to SecY, and therefore also to the chloroplast SecY homologue. Proteins trapped artificially in the membrane during translocation are predominantly cross-linked to Sec61α, which has ten membrane-spanning regions. Another component of the mammalian apparatus is the translocating chain associating membrane (TRAM) protein, which also has multiple membrane-spanning regions, and can be cross-linked to a translocating nascent chain. This protein is of variable importance, which appears to depend on the structure of the signal peptide, and therefore a proposed function is in determining the orientation of translocation (discussed by Jungnickel *et al.*, 1994). Signal peptidase, which removes the signal sequence from translocated proteins, is resident in the ER membrane, with its active site on the lumenal face. It is a member of the Type I signal peptidases that also contains leader peptidase, thylakoidal processing peptidase and mitochondrial inner membrane protease I, and has a similar reaction specificity to these enzymes.

1.18.3 *Peroxisomes*

Targeting and assembly of peroxisomal proteins have been recently reviewed by Olsen (1998). Emphasised are the ways in which peroxisomal protein targeting differs from all other systems studied to date. To begin with, peroxisomal targeting signals (PTS) come in more than one variety: N-terminal, C-terminal, both the aforementioned, or none at all.

The first PTS identified was a C-terminal signal, called PTS1, that consisted of a tripeptide with a consensus Ser-Lys-Leu (SKL), but several substitutions can be made within this without disrupting targeting. Unlike chloroplast envelope transit peptides, the motif is necessary, but not always sufficient for targeting (for example, the passenger chloramphenicol acetyltransferase only requires PTS1, but catalase A from *S. cerevisiae* needs other upstream sequences as well). The C-terminal location of the PTS1 is critical for its function. When a single amino acid is added to the C-terminus

of firefly luciferase, the protein is no longer targeted to peroxisomes (de Hoop and Ab, 1992).

Although the C-terminal tripeptide PTS1 is used by the vast majority of peroxisomal matrix proteins, other signals serve in a similar capacity. An N-terminal region on some peroxisomal matrix proteins constitutes the second type of PTS, termed PTS2. The PTS2 is a nonapeptide with a consensus of (R/K)(L/I/Q/V)X_5(H/Q)(L/A). The arginine residue and both leucines are critical for PTS2 function in *S. cerevisiae* (Glover *et al.*, 1994). In higher eukaryotes, including plants and mammals, the N-terminal presequence is proteolytically processed following import. A 110 kDa metalloprotease from mammalian peroxisomes was identified that degrades a synthetic presequence of thiolase, but does not process full-length *in vitro* synthesised thiolase (Authier *et al.*, 1995).

Putative receptors for PST1 and PST2 have been isolated by complementing transport-defective yeast peroxisomes. These are Pex5p and Pex7p, respectively, homologues of which have been discovered in humans too. Unusually, the receptors have been detected in the cytosol, peroxisomal membrane and matrix of different cell types, suggesting that the receptors may shuttle between the cytosol and the peroxisome. Furthermore, the discovery of yeast and mammalian mutants blocked in either or both of the import pathways has greatly facilitated the cloning and characterisation of several other proteins involved in the translocation of PTS1- and PST2-bearing precursors.

It is generally held that proteins must be largely unfolded as they cross membranes, but in peroxisomes, there is abundant evidence supporting, but not yet proving, that precursor proteins can be translocated as oligomers. For example, mature, octameric alcohol oxidase was microinjected to animal cells where it was rapidly taken up by organelles including the peroxisomes (judged by immunofluorescence). Chemically cross-linked proteins, whole antibodies and even colloidal gold particles coated with PTS1 are imported by peroxisomes when microinjected to animal cells. McNew and Goodman (1996) propose two models for oligomeric import: endocytosis or

regulated opening of pores. Either of these models should incorporate the observed need for ATP hydrolysis.

1.19 *Transport of proteins across prokaryotic membranes*

In both Gram-positive and Gram-negative bacteria, there is a vast exchange of molecules between the cell and environment. The cytoplasm of Gram-positive bacteria is separated from the extracellular medium by the cytoplasmic membrane, a cell wall and a protein layer. The barrier in Gram-negative bacteria is more complex however, consisting of the cytoplasmic membrane, periplasmic space, cell wall, outer membrane, lipopolysaccharide layer and protein layer. Extracytoplasmic proteins in either type of bacterium are referred to as "exported" proteins, but only those which are entirely extracellular (i.e. released anywhere outside the cytoplasmic membrane for Gram-positive, and anywhere outside the outer membrane for Gram-negative cells), are referred to as being "secreted". In *E. coli* (Gram-negative), the initial step of protein export across the cytoplasmic membrane occurs *via* the signal peptide-dependent general export pathway (GEP), in nearly all cases. This pathway operates using the *sec* gene products and signal peptidases, and the rest of the discussion will centre upon these proteins. Protein secretion across the Gram-negative outer membrane is reviewed by Wandersman (1992), while the topics of protein export and secretion in Gram-positive bacteria are reviewed comprehensively by Simonen and Palva (1993).

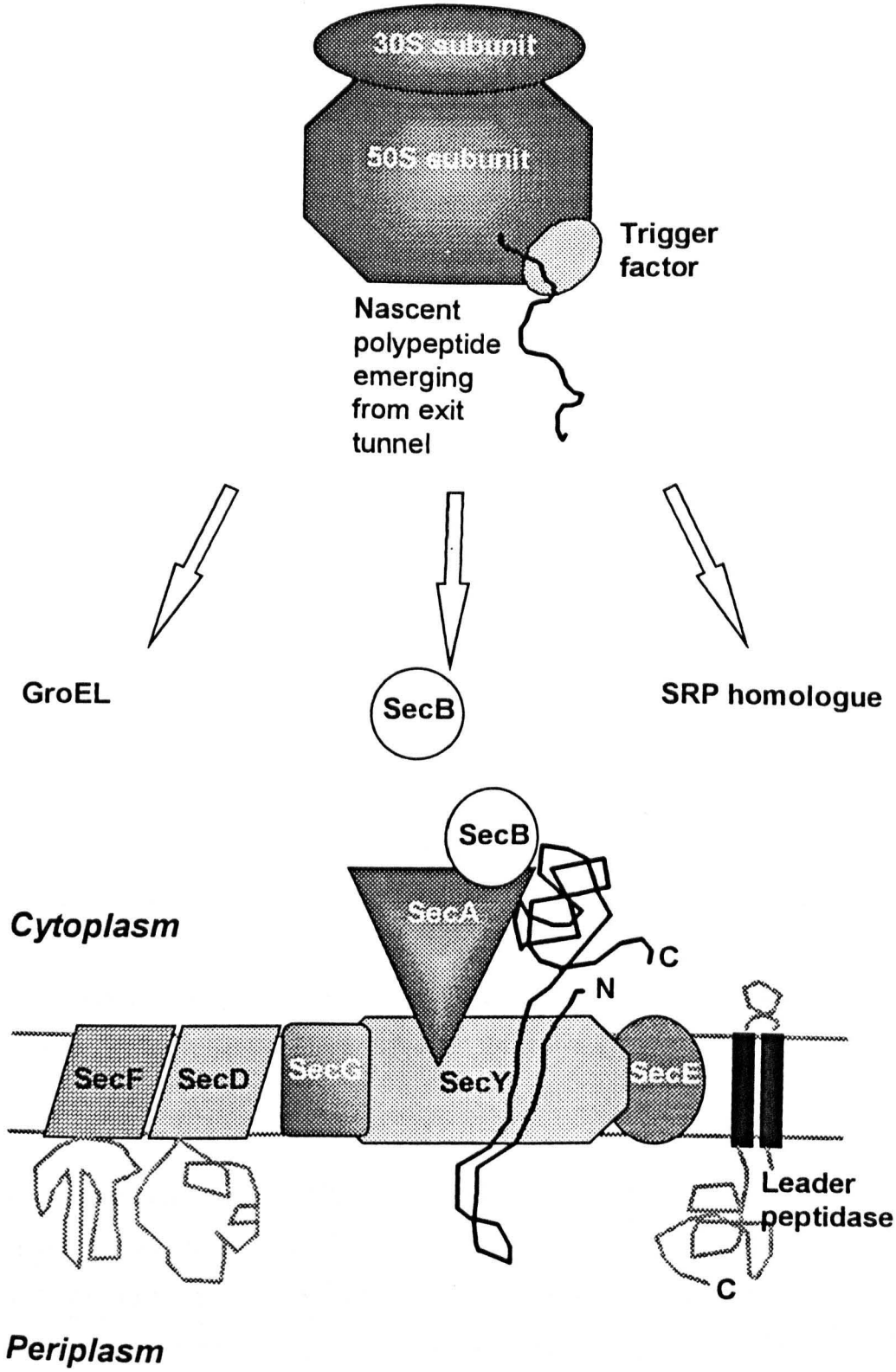
1.19.1 *The Sec proteins of the general export pathway (GEP)*

Major features of the GEP are the signal peptides of proteins to be exported, which assist in targeting them to the cytoplasmic membrane; a molecular chaperone, SecB; a translocation ATPase, SecA; integral membrane proteins, SecY, E, and G; SecD and SecF, which appear to be required for maintenance of Δp; and signal peptidases, required to remove the signal peptides from exported proteins. Based on copy numbers of the various Sec proteins, there are an estimated 500 translocases for every *E. coli* cell, but an alternative estimate obtained by jamming the GEP is much higher - 2000 translocation sites per bacterium (Pugsley, 1993).

Figure 12 shows the main components of the Sec apparatus. SecB is a molecular chaperone which binds some, but not all proteins destined for export (for example, pre-maltose-binding protein, pro-outer membrane protein A and pre-lambda receptor, but not pre-ribose-binding protein, pre-alkaline phosphatase or pre-β-lactamase). Purified SecB elutes from a gel filtration column with an apparent molecular mass of 64 kDa, which represents a tetramer of identical subunits, each of mass 17.3 kDa (Weiss *et al.*, 1988). There is some dispute as to whether SecB binds merely to unfolded proteins over their mature domains, or whether it binds signal sequences, thus rendering it a specific signal-recognition factor. In the former view (discussed by Randall and Hardy, 1995), the signal sequence acts indirectly to retard folding of the precursor in the cytosol, such that SecB may bind the denatured protein. This view is supported by the high affinity with which SecB binds four unrelated, denatured proteins, yet shows no detectable binding when they are in the native states. The model proposed by the authors invokes two types of binding site on SecB, each interacting with a separate region of the polypeptide ligand. The first type of site would be filled with one or more stretches of flexible polypeptide chain (abundant in an unfolded protein). Multiple interactions would release enough binding energy to make the likelihood of dissociation very low, and would induce a conformational change in SecB, exposing a region designed to bind hydrophobic portions of the ligand. Evidence to support the existence of both types of site has been supplied by mutational analysis of SecB/pre-maltose-binding protein (pre-MBP) complex formation (Kimsey *et al.*, 1995). The contrasting view holds that recognition of signal sequences by SecB, and their subsequent targeting to the membrane would be analogous to the function carried out by eukaryotic SRP (Watanabe and Blobel, 1995). These authors have assayed the binding of chemically pure SecB with pre-MBP and MBP, both forms of the ligand being truncated at the C-terminus to prevent protein folding. They found that SecB bound pre-MBP with one hundred-fold higher affinity compared to MBP, and since, in the denatured state, the only difference between the two ligands was the presence of a signal peptide, it was concluded that SecB interacts specifically with the signal peptide. Consistent with this explanation, a wild-type, but not a mutant, export-incompetent signal peptide competed with pre-MBP for SecB binding. Regardless of the specificity or otherwise by which SecB recognises and binds the ligand, its specific affinity for SecA is not in dispute.

Figure 12 *Proteinaceous components of the Sec export apparatus, and other relevant proteins.*

The highly schematic diagram (partially adapted from Rassow and Pfanner, 1996), shows trigger factor binding the nascent polypeptide chain as it emerges from the ribosome. Trigger factor can compete with SRP for binding signal sequences, but the former appears to bind first. GroEL (a prokaryotic chaperonin) promotes protein folding in the cytosol, and may bind trigger factor. SecB is a chaperone specific for exported and secreted proteins. SecB specifically binds SecA, which is known as the translocation ATPase. SecY, E and G are integral membrane proteins responsible for physically moving the protein across the cytoplasmic membrane. SecD and F are believed to be responsible for maintaining Δp. Leader peptidase is the major signal peptidase.


30S subunit
50S subunit
Trigger factor
Nascent polypeptide emerging from exit tunnel
GroEL
SecB
SRP homologue
SecB
SecA
Cytoplasm
C
N
SecF
SecD
SecG
SecY
SecE
Leader peptidase
C
Periplasm

SecA is an essential element of the GEP, and as such, a variety of mutant phenotypes exist that map to the *sec*A gene. These include *azi* mutants, which confer resistance to sodium azide (Oliver *et al.*, 1990; Fortin *et al.*, 1990), and temperature-sensitive mutants, which result in the accumulation of unsecreted protein in the cytosol (Oliver and Beckwith, 1981). If protein export is blocked or saturated, the synthesis of SecA increases (the only component of the GEP to do so), which suggests that SecA can compensate for temporary defects in protein export. Consistent with this is the *in vitro* observation that SecA alone can catalyse translocation across membranes under certain conditions (Cunningham *et al.*, 1989; Watanabe and Blobel, 1993).

SecA is 901 amino acids long (101 kDa), and acts as a homodimer (Driessen, 1993). It interacts with SecB, signal peptides, anionic phospholipids of the inner membrane, the integral membrane proteins SecY, E and G, and SecD and F. These multiple interactions underline the central role of SecA in protein translocation (reviewed by Oliver, 1993). The protein is found in both the cytosol and cytoplasmic membrane, and in both locations it will interact with precursor proteins. One line of evidence implying signal peptide interaction with SecA is provided by the *prlD* alleles of *secA* (Fikes and Bassford, 1989). These are mutations suppressing defects in the signal peptide, which would normally render the precursor protein incompetent for export. More directly, soluble SecA has been cross-linked to a signal peptide, with an efficiency that increased with increasing charge on the N-terminus of the peptide (Akita *et al.*, 1990). It has also been suggested that SecA possesses two binding sites: one for the signal peptide and one for the mature protein domain, because a combination of signal peptide and urea-denatured mature MBP or OmpA was required to stimulate SecA ATPase activity in liposomes (Lill *et al.*, 1990). The region of SecA which binds the signal peptide seems adjacent to a site where ATP is bound with high-affinity, and the two sites appear to act co-operatively, suggesting that the binding of ATP to SecA induces a conformational change leading to signal peptide binding (Shinkai *et al.*, 1991).

Hartl *et al.* (1990) demonstrated a SecA/SecB interaction in solution. As SecB and pro-OmpA bound SecA non-competitively, it was concluded that the two binding sites are separate. If SecB remained bound in a ternary complex with SecA and the

precursor protein, it could be envisaged that the chaperone could stabilise the precursor protein in a suitable conformation for SecA interaction (Oliver, 1993).

Anionic phospholipids play an important part, because without them, SecA will not bind the membrane, and defects in protein secretion result (the role of anionic phospholipids is reviewed by de Kruijff, 1994). Economou *et al.* (1995) have shown that SecA penetrates the cytoplasmic membrane by the insertion of a 30 kDa domain, along with a 20-30 aminoacyl stretch of the precursor protein. They also demonstrated that SecA inserts at the SecYEG complex, a result which is supported by the correspondence of the number of high-affinity SecA binding sites with the number of SecY molecules in membrane vesicles (Hartl *et al.*, 1990). SecD and SecF stabilise membrane-inserted SecA, a function separate from the one of maintaining Δp. Economou *et al.* (1995) propose that SecD and SecF regulate the withdrawal of SecA from the membrane, perhaps by means of their large periplasmic domains (see **figure 12**) which are near the membrane-inserted portion of SecA.

It is not yet obvious how many ATP binding sites SecA provides (Oliver, 1993). Photo-labelling with 8-azido-ATP suggests there are three sites, two of which are easily labelled and are necessary for translocation ATPase activity, and one which resists labelling and is the site of a basal ATPase activity (Lill *et al.*, 1989). The analyses carried out by Economou *et al.* (1995) utilised two SecA variants, one carrying a mutation (at position 209) in a high affinity nucleotide binding domain (NBD1), and the other, a mutation (position 509) in NBD2, the low affinity site. The first mutant (D209N) binds nucleotide, but does not possess translocation ATPase activity, because it cannot withdraw from the membrane. The other mutant (R509K) can insert and withdraw normally, driven by ATP hydrolysis at NBD1, but the process appears to be a futile cycle, because ATP hydrolysis at the NBD2 site is needed to couple protein translocation to SecA cycling in and out of the membrane (discussed further in **section 1.19.2**).

SecY is an integral membrane protein (molecular mass 49 kDa), which is very hydrophobic, due to the presence of ten membrane-spanning regions between the cytoplasmically-orientated N- and C-termini (Akiyama and Ito, 1987). This

arrangement in the membrane is similar to that of solute transport proteins, such as lactose permease, and hence it is believed that SecY either forms, or participates in the formation of a hydrophilic pore across the membrane. This is supported by the cross-linking of a translocating polypeptide chain to SecY (Joly and Wickner, 1993), and the significant structural homology displayed to SEC61 in the ER membrane, which also interacts with proteins during translocation (Görlich *et al.*, 1992). The cytoplasmic loops, and the N- and C-termini are thought to bind the precursor protein, but it is not known whether this involves specific recognition of the signal peptide (discussed in a review by Müller and MacFarlane, 1994). SecY is essential for the translocation process, as shown by reconstitution of protein translocation into liposomes containing highly-purified SecA, SecYEG complex, or SecY and SecE purified independently (Brundage *et al.*, 1990; Akimaru *et al.*, 1991). Only in an *in vitro* system has SecY been dispensed with, and in this case, over-expressed SecA and ATP were absolute necessities (Yamada *et al.*, 1989b).

SecE almost certainly forms a complex with SecY, because the two proteins can be isolated as a complex (Brundage *et al.*, 1990) and over-expression of SecY is enabled by simultaneous over-expression of SecE (Matsuyama *et al.*, 1990). The 13.6 kDa SecE protein has three transmembrane spans (with the N-terminus in the cytosol), but activity of SecE has been localised to just the third region, and its preceding cytoplasmic domain (Schatz *et al.*, 1991). There are opposing lines of evidence indicating the order in which SecY and SecE function. The first, based on the use of *prl* suppressor alleles, suggests SecE acts before SecY, because arrested translocation intermediates are not proteolytically processed in the presence of a mutant SecE, but are when SecY is mutated. The converse view is supported by the inability of SecE to function as the primary membrane receptor for precursor proteins when SecY is inactivated (Müller and MacFarlane, 1994). A further protein, SecG, has been identified as part of the SecYE complex, as it co-purifies with SecYE using either chromatography or immunoprecipitation methods. It became clear that SecG greatly stimulates translocation when it was possible to obtain pure preparations of SecA, SecY and SecE. When these three were reconstituted into proteoliposomes, their activity was much lower than in the membrane vesicles which had also contained SecG (Nishiyama *et al.*, 1993).

Although *in vitro* studies suggested SecD (67 kDa) and SecF (39 kDa) were not essential components of the GEP, mutations in the *secD* and *secF* genes *in vivo* told a different story. Null mutants for the two genes show severe defects in protein export, and are cold-sensitive for growth (Pogliano and Beckwith, 1994). The two proteins are not present in stoichiometric amounts: there is much less SecF than any other component of the GEP, suggesting a regulatory role of the protein. SecD and SecF show similar membrane topology, each possessing six transmembrane stretches, with large (45 kDa and 11 kDa respectively) periplasmic domains between the first two membrane-spanning regions (reviewed by Pugsley, 1993). Matsuyama *et al.* (1993) incubated spheroplasts with anti-SecD IgGs, which inhibited the secretion of pro-OmpA and pre-MBP. Anti-SecE and anti-SecY IgGs (to periplasmic epitopes) did not inhibit secretion, therefore the researchers concluded that SecD is required for precursor protein release from the Sec machinery. Economou *et al.* (1995) present evidence that SecD and SecF may regulate withdrawal of SecA from the membrane, it seems possible that prevention of this function by anti-SecD IgGs could account for the observed accumulation of un-transported precursor proteins by Matsuyama *et al.* (1993). Arkowitz and Wickner (1994) showed that SecD and SecF are also necessary to maintain Δp in inverted, inner membrane vesicles.

Leader peptidase cleaves the signal peptide from the N-terminus of newly-transported membrane and secretory proteins. The peptidase is located in the cytoplasmic membrane, with its active site facing the periplasm, and has been used as a model system to study the topogenesis of membrane proteins (reviewed by Dalbey *et al.*, 1995). The peptidase is insensitive to most types of protease-inhibitor, but as discussed in **section 1.11.2a** it belongs to a class of serine proteases which do not need a histidine residue to act as a proton donor and acceptor (Sung and Dalbey, 1992). The purified protein cleaves precursor proteins from many different sources, including plants, yeast, honeybees, humans and bacteria; all these share a similar pattern of small side-chain residues at the -1 and -3 positions relative to the cleavage site, and a helix-breaking residue between positions -4 to -6.

The properties of signal peptides are described in an excellent review by Izard and Kendall (1994). Although there is little in the way of primary structure homology, typical *E. coli* signal peptides possess an N-terminal, positively charged region of about five or six amino acids (usually two of these are charged), a hydrophobic core of about twelve residues' length and a C-terminal region about six amino acids long, of which the -1 and -3 residues are neutral, with small side chains. A helix-breaking amino acid, such as glycine or proline, may also be in this region.

The type of positively-charged residue in the N-terminal domain does not matter, but the position is important, because the introduction of positive residues to other regions of the signal peptide, or to the N-terminus of the mature protein disrupts protein export. One explanation of the requirement for a positively-charged N-terminus, is the need for an initial interaction with anionic phospholipids in the inner membrane. After the initial interaction, the hydrophobic core could insert to the lipid bilayer, or a part of the translocation apparatus. As mentioned above, the positive charges in the signal peptide are also implicated in SecA binding (Akita *et al.*, 1990). The physical properties of the hydrophobic core are more important than the identity of the residues from which it is composed, for a polyleucine core will replace a natural signal peptide with no decrease in the efficiency of processing. Also, within certain length parameters, the "hydrophobic density" is a critical determinant for translocation efficiency. For example, increasing the length of a synthetic polyvaline core region will partially compensate for the lower hydrophobicity of this amino acid, compared to leucine. Precursor proteins with mutations in the hydrophobic core are associated with the membrane, but not translocated. There is some indirect evidence that the hydrophobic core may interact with Sec proteins, as well as with the membrane, for *prl* mutations in SecY which suppress mutations in the hydrophobic core are clustered in discrete domains of the protein, suggesting the existence of a binding site. Also, sodium azide, which inhibits SecA, has less effect upon the processing of mutant precursor proteins when the core hydrophobicity is increased. This phenomenon may represent an increased affinity of the signal peptide for SecA, enabling the mutant precursor proteins to bind the remaining target sites unaffected by azide.

The analysis of the role of secondary structure in signal peptide function is complicated by the problems of measuring the parameters in a physiological environment. For example, signal peptides display random coil in aqueous solutions, but in SDS micelles, or phospholipid vesicles, they usually take up an α-helical conformation, which is not just confined to the hydrophobic core. Such helical arrangements enhance the hydrophobicity of the signal peptide, because polar groups are shielded by the hydrogen bonding which occurs in the structure. Helix-breaking amino acids are defined by their properties in globular proteins, and the term is not necessarily accurate when applied to the signal peptide context. Therefore, the role of these amino acids is unknown. It has become apparent, through studies on artificial signal peptides, and mutations which suppress defects in other parts of the signal peptide, that sub-optimal properties in one region may be compensated by an increase in the 'quality' in another region. In fact, Izard and Kendall (1994) propose that signal peptides contain "natural flaws and wobbles" to tune the structure for interactions with a number of components, rather than too tight in an association with just one of them, which is an intriguing idea.

1.19.2 *A model for protein export by the Sec apparatus*

The components of the GEP have been studied from almost every possible angle, in an attempt to construct a detailed picture of the mechanism by which the Sec proteins mediate protein export across the cytoplasmic membrane. Over the years, pieces of information on protein-protein interactions, protein-lipid interactions and energetic requirements have been steadily assembled until the jigsaw is in its final stages of completion. Of all the experimental work published to date, the papers that best delineate a model for the export process are: Schiebel *et al.*, 1991 ("$\Delta\mu_{H+}$ and ATP function at different steps of the catalytic cycle of preprotein translocase"); Driessen, 1992 ("Precursor protein translocation by the *Escherichia coli* translocase is directed by the proton motive force") and Economou *et al.*, 1995 ("SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF"). In the following paragraphs, an attempt is made to summarise the model as succinctly as possible.

Initially, nascent polypeptide chains associate with a cytosolic factor, such as SecB, or the SRP homologue (see **section 1.19.3**). Binding to SecB may or may not be mediated by specific recognition of the signal peptide. Not all precursor proteins associate with SecB. The SecB/precursor protein complex binds membrane-bound SecA *via* the affinities of the translocation ATPase for both SecB and the precursor protein. SecA itself binds to the cytoplasmic membrane due to its affinity for anionic phospholipids and the integral membrane complex, SecYEG. The binding of a model precursor protein, pro-OmpA, encourages the release of ADP from a high-affinity binding site on the SecA protein, which activates an ATPase activity. When this happens, ATP binding causes changes in the conformation of SecA, leading to the insertion of a 30 kDa domain of the ATPase in the membrane at SecYEG, accompanied by the simultaneous "stuffing" of approximately twenty residues of the precursor protein. These two insertion events are thought to be mechanistically coupled, because there is evidence that the 30 kDa SecA domain contains both NBD1 and a binding site for the precursor protein. Hydrolysis of ATP at sites NBD1 and NBD2 leads to the release of the precursor protein, and the withdrawal of SecA from the membrane, where it exchanges with cytosolic SecA. At this stage, it is the Δp which takes over in translocating the "stuffed" segment of precursor protein, but meanwhile, the cycle of precursor binding to SecA, followed by ATP-dependent insertion to the membrane may continue on the cytoplasmic side of the membrane. While SecYEG is essential for the cycling activity of SecA, SecD and SecF are required for optimum rates of cycling and translocation. They are believed to regulate the withdrawal of SecA from the membrane.

ATP hydrolysis is not coupled to net movement of the precursor polypeptide. This is achieved by the Δp, although in the absence of an energised membrane, multiple rounds of SecA/precursor binding and membrane insertion can result in complete translocation of the precursor protein. By reconstituting protein translocation in inverted inner membrane vesicles, containing SecYE, and either bacteriorhodopsin or cytochrome *c* oxidase (in order to obtain Δp of opposite polarities), Driessen (1992) showed that SecA-dependent translocation of pro-OmpA is stimulated by Δp of physiological orientation. A Δp of reversed polarity inhibits translocation, and

actually causes backward movement of a membrane-spanning intermediate. Either component of the Δp: ΔΨor ΔpH, is effective in promoting or inhibiting translocation. In the presence of a Δp, far less ATP molecules per molecule of pro-OmpA are hydrolysed than in its absence. Conversely, a reverse Δp causes a far greater number of ATP hydrolysis events, indicating that a physiologically-oriented Δp effectively eases the task of SecA, and also drives translocation in a forward direction.

The exact mechanism by which a Δp stimulates translocation is not known. Electrophoretic movement of negatively-charged regions of the precursor protein to the relatively positive periplasm is excluded because Δp even stimulates the translocation of an uncharged protein. Similarly, the effect is not due to the removal of protons from basic amino acids before translocation. By comparison with solute transport mechanisms, one would expect the movement of proteins to be coupled to a return of flux of protons. In the presence of a full Δp, inverted inner membrane vesicles will translocate pro-OmpA containing a disulphide bridge, without cleaving it (Tani *et al.*, 1990). This begs the question of whether exported proteins must be unfolded during translocation, as is believed to be the case in most protein transport systems. Arkowitz *et al.* (1993) demonstrated that the translocation process itself can cause the unfolding of a precursor protein domain, by analysing the transport of a pro-OmpA/DHFR fusion protein in the presence and absence of methotrexate and NADPH. A section of the DHFR domain was unfolded and translocated in the absence of SecA, ATP and Δp. In this case, therefore, the protein is translocated in an unfolded conformation, accomplished by the "intrinsic energy of translocation", which is the balance of free energy obtained by summing the various energies of folding and unfolding either side of the membrane, and the changing environments to which the precursor protein is exposed.

1.19.3 *The Escherichia coli signal recognition particle*

After cloning of the gene for mammalian SRP54 protein, database searches, and later, immunoprecipitation experiments showed that *E. coli* contains a homologue of this protein, termed Ffh, which is complexed with a 4.5S RNA (Poritz *et al.*, 1990;

Ribes *et al.*, 1990). This provided evidence for SRP-dependent targeting in *E. coli*, reinforced by the discovery that a gene formerly implicated in cell division, *ftsY*, encodes a protein homologous to the α subunit of the mammalian SRP receptor (Römisch *et al.*, 1990). Although at first, genetic screens designed to identify secretion mutants failed to implicate the Ffh/4.5S ribonucleoprotein or FtsY in protein targeting, recent evidence does support such a role. Luirink *et al.* (1992) identified Ffh as the major protein cross-linked to signal peptides in lysates of *E. coli* cells. Importantly, this interaction was only observed for signal peptides which were part of the nascent chain complex with the ribosome. Ffh will also assemble into mammalian SRP, to form a partially functional chimaeric particle (Bernstein *et al.*, 1993). A cytoplasmic membrane location has been demonstrated for FtsY, and depletion of this protein from cells resulted in the accumulation of β-lactamase and ribose-binding protein, but not MBP, correlating with the lack of SecB binding exhibited by the former two exported proteins (Luirink *et al.*, 1994). MacFarlane and Müller (1995) proposed that the major function of the bacterial SRP homologue is in the targeting and integration of hydrophobic membrane proteins, by-passing SecA. This proposal is based on their disruption of the Ffh/4.5S ribonucleoprotein, which prevented membrane integration of lactose permease. Several possibilities exist. Firstly, SRP and SecB/SecA may both converge at the SecYEG complex in the cytoplasmic membrane. Alternatively, SRP may mediate co-translational insertion of a subset of proteins which do not utilise the Sec machinery at all, due to their high hydrophobicity and tendency to aggregate immediately after translation. Another possibility, which does not exclude the second suggestion, is that a range of exported proteins utilise both the Sec machinery and Ffh/4.5S/FtsY, the choice of which depends upon unknown regulatory factors.

1.19.4 *Sec-independent export system*

Recently, a protein has been found in plants that is critical for the ΔpH-dependent pathway. This protein, Hcf106, was cloned by Settles *et al.* (1997), and has provided major insights into the origins and likely mechanism of this system. Homologues of *hcf*106 have been found in a wide range of bacterial genomes, providing clear indications that a similar system might operate in prokaryotes. At about the same

time, Berks (1996) showed that periplasmic proteins bearing any of a range of complex redox factors (e.g. FeS centres or molybdopterin cofactors) are invariably synthesised with signal peptides containing a twin-Arg motif, promoting speculation that they might be recognised by a system akin to the thylakoid ΔpH system. Furthermore, many of these cofactors are believed to be inserted in the cytosol, therefore requiring folding of the precursor to a large extent. Also, there is a widely held belief that the Sec system is incapable of exporting highly folded proteins (Randall and Hardy, 1986), however, recent data on the ΔpH-dependent system in thylakoids (Clark and Theg, 1997; Hynds *et al.*, 1998) demonstrated that fully folded proteins (using fusion constructs) can be transported by this translocase. The consequence of this data implies that a translocation mechanism in the plasma membrane is required, that allows transport of fully folded proteins and these proteins have signal peptides with twin-Arg motifs. Neoteric data has shown this scenario to be correct and the first details of a Sec-independent bacterial export system have emerged. The *E. coli* genome contains two *hcf*106 homologues, one of which appears to be monocistronic (*ybe*C) whereas the other forms part of an apparent operon (*yig*T). The original *yig*T gene was later shown to comprise two separate genes and these were followed by two further genes, designated *yig*U (renamed *tat*C) and *yig*W, in the same operon.

Sargent *et al.* (1998) probed the function of these genes using in-frame deletions and found that disrupting either *yig*T or *ybe*C caused a marked inhibition of the export of at least five periplasmic proteins synthesised with twin-Arg containing signal peptides. A double deletion strain was totally defective in export, confirming a central, and perhaps overlapping role for these gene products. Sec-dependent export was unaffected in these deletion strains confirming the operation of a separate export pathway. The genes have now been designated *tat* genes (for twin-arginine-translocation), and the *hcf*106 homologues are *tat*A and *tat*E. A separate study focused on the operon by Weiner *et al.* (1998). An *E. coli* mutant defective in this export pathway was isolated and the mutation believed to lie in the *yig*T region. However it now appears the mutation was instead in the second gene of the operon, which is unrelated to *hcf*106. Effectively, they found a critical requirement for the *tat*B gene product. A fourth component of this export pathway has now been

identified; Bogsch *et al.* (1998) disrupted the *tat*C gene which led to a complete block in the export of a range of cofactor-containing proteins bearing twin-Arg motif signal peptides.

There are still many questions to be asked about this pathway: for example, is there a mechanistic similarity between the bacterial and chloroplast Sec-independent transport pathways? Wexler *et al.* (1998) investigated this by assessing the ability of bacterial twin-Arg containing signal peptides to direct thylakoid import. They fused the presequences of trimethylamine *N*-oxide reductase (TorA) and the catalytic subunit of formate dehydrogenase-*N* (FdnG) with the mature protein of 23K. Both TorA-23K and FdnG-23K were imported and processed to a smaller form by thylakoids by a ΔpH-dependent, ATP-independent mechanism, demonstrating that bacterial signal peptides are capable of directing efficient translocation of proteins across the plant thylakoid membrane in a Sec-independent manner.

1.19.5 *M13 procoat*

The coat protein of the filamentous phage M13 is the classical example of a relatively small protein which inserts into a membrane independently of the translocation machinery (Kuhn, 1995). The M13 coat protein is synthesised with a 23 amino acid signal sequence in the cytoplasm of *E. coli* and is then inserted into the plasma membrane. The membrane insertion process of M13 procoat protein has been extensively studied genetically and biochemically (reviewed by Kuhn and Troschel, 1992). There is experimental evidence that indicates that the binding of the protein to the negatively charged membrane occurs *via* the positively charged amino acid residues at the N- and C-termini of the procoat protein (Gallusser and Kuhn, 1990). The subsequent partitioning of the two hydrophobic regions into the membrane results in a loop configuration. The topology of this intermediate has been investigated using procoat protein with either N- or C-terminal extensions (Kuhn *et al.*, 1986; Kuhn 1987), and it was found that the N- and C-termini do not leave the cytoplasm during the insertion process forming a loop intermediate prior to cleavage by leader peptidase (Kuhn, 1987). Only the negatively charged central region traverses the membrane in accordance with the "positive-inside" rule (von Heijne, 1994).

The *in vivo* insertion process requires a transmembrane potential for the wild-type procoat protein (Kuhn and Wickner, 1985). Furthermore, *in vitro* translocation efficiency has been linked to the ability of the signal peptide to induce the formation of non-bilayer lipid structures (Killian *et al.*, 1990); proposed to allow the partitioning of the precursor protein across the *E. coli* plasma membrane (Gounaris *et al.*, 1983). Biophysical investigations have shown that at least part of the predicted transmembrane regions adopt an α-helical conformation (Thiaudière *et al.*, 1993). After membrane insertion, the signal sequence is cleaved by the *E. coli* leader peptidase which has its enzymatic activity at the periplasmic face of the cytoplasmic membrane (Dalbey and von Heijne, 1992).

1.20 *Endosymbiotic theory and organelle evolution*

The discovery of organellar DNA enabled the testing of two hypotheses for the evolution of chloroplasts and mitochondria: autogenous origins (from within the eukaryotic cell) and xenogenous origins (from outside the eukaryotic cell). Analysis of organellar genomes has supported the most popular xenogenous theory; that of endosymbiosis. According to this theory, chloroplasts and mitochondria originated as eubacterial endosymbionts in a progenitor nucleus-containing cell (the urkaryote). During the establishment of an endosymbiotic relationship, genes were transferred to the nucleus, resulting in the present situation (which is unlikely to be static) whereby most of the organellar proteins are encoded by the nucleus, and hence must be imported into the organelle post-translationally. Targeting signals for translocation into chloroplasts or mitochondria would therefore have evolved at the same time as gene transfer to the nucleus. The retention of prokaryotic signal peptides would allow subsequent 'conservative sorting' of proteins to sub-cellular locations.

Gray (1989) discusses several forms of proof for the endosymbiotic hypothesis, including: a) a demonstrably different evolutionary history of nuclear and organellar genomes; b) a demonstration that chloroplasts and mitochondria arose from different lineages of bacteria; c) proof that either type of organelle is polyphyletic. Data for these sorts of analyses are best obtained from rRNA sequences, because of the ubiquitous nature of large and small ribosomal subunit genes. The degree of

relatedness of an organellar gene to either its partner nuclear gene, or to a bacterial homologue is decided by sequence alignments and comparisons. Unfortunately, a basic assumption, that the history of a single gene must reflect the history of the whole genome, does not always hold true for the mitochondrial genome, due to its horizontal acquisition of 'promiscuous' DNA sequences from chloroplasts (Ellis, 1982). Despite this, such analyses have confirmed that chloroplasts arose from early cyanobacteria, and mitochondria from purple photosynthetic bacteria (although the best evidence for the latter comes from biochemical comparisons of respiratory chain enzymes). Furthermore, it appears that chloroplasts in different photosynthetic groups of eukaryotes may have arisen independently from different bacterial phyla (first suggested by Raven, 1970). There is also a possibility that mitochondria are polyphyletic, for metaphyte mitochondrial rRNA genes appear to have originated later than those from other eukaryotic mitochondria.

1.21 *Aims of project*

At the commencement of this project, the direct insertion pathway was beginning to be termed as the fourth 'mainstream' pathway in thylakoid import. Since the insertion mechanism was known only by its lack of insertion requirements, a mechanistic analysis was required to try and determine whether precursor proteins assigned to this pathway did in fact follow the same pathway. Furthermore, to determine whether structures known to be involved in M13 procoat insertion (the only other protein known to insert in a Sec-, ΔpH-, and SRP-independent manner, requiring a signal peptide) were similarly required. Therefore, this project had the following aims:

1. To determine whether TPP cleaves proteins that follow this novel pathway.
2. To further investigate if the two precursor proteins pre-PSII-W and pre-PSII-X do in fact follow the same insertion pathway and to determine whether the precursor proteins pre-PSII-W and pre-PSII-X form loop intermediates during the insertion into the thylakoid membrane.
3. To resolve the question of whether a new precursor protein, pre-PsbY is a polyprotein.

Chapter 2

Materials and Methods

Chapter 2

Materials and Methods

Solutions and media were sterilised by autoclaving which was carried out at 15psi for 25 min or filter-sterilised by passing through a 2 micron filter. Sucrose solutions were autoclaved at 10psi for 20 min or filter-sterilised. The microfuge used during the protocols outlined below was an MSE microcentaur. Sambrook *et al.* (1989) is a laboratory manual and not necessarily the originator of the protocols described below.

2.1 *Growth and Maintenance of Escherichia coli strains.*

2.1.1 *Escherichia coli strains.*

Strain	References
TG1	T. J. Gibson, 1984
JM109	Yanisch-Perron *et al.*, 1985

2.1.2 *Short-term storage of E. coli.*

Stock strains of *Escherichia coli* were maintained by streaking out individual colonies onto LB medium plates followed by incubation at 37°C overnight and storage at 4°C. This process was repeated monthly.

All components of L-agar medium plates were made up with sterile distilled water (SDW), autoclaved or filter-sterilised separately and cooled before mixing aseptically.

2.1.3 *Long-term storage E. coli.*

A single bacterial colony was grown overnight at 37°C in 10ml L-broth in an orbital shaker at 250rpm. 850µl of overnight culture was transferred into a vial containing 150µl of sterile glycerol and mixed. Glycerinated cultures were stored at -70°C.

2.2 *Transformation of E. coli.*

2.2.1 *Preparation of Competent Cells.*

5 ml of LB-medium were inoculated with a single bacterial colony, from an overnight plate, this was incubated overnight at 37°C. 0.2 ml was used to inoculate 20 ml of LB-medium. These cells were grown until the OD_{600} reached 0.3-0.5 (approximately 2-3 hours). The cells were then centrifuged at 10 000 rpm for 5 min at 4°C. The supernatant was then decanted and the cells resuspended in 1ml of ice-cold solution A (10 mM MOPS, pH7.0 and 10 mM rubidium chloride). The volume was brought up to 20ml with solution A and the sample was centrifuged at 10 000rpm for 5 min at 4°C. Again the supernatant was decanted, and the cell pellet was resuspended in 1ml of ice-cold solution B (100 mM MOPS, pH 6.5, 10 mM rubidium chloride and 50 mM calcium chloride) . Solution B was next added to bring the final volume up to 10 ml. The cells were incubated on ice for 30 min and then centrifuged at 10 000 rpm for 5 min at 4°C. Again the supernatant was removed and the tube thoroughly drained. The cells were then gently resuspended in 2 ml solution B. These freshly prepared competent cells were then stored on ice for a minimum of 24 hours. For long-term storage glycerol was added (10% of the final cell volume) to the cell solution and 200 µl aliquots stored at -70°C.

2.2.2 *Transformation of E. coli with plasmid DNA.*

Approximately 1 ng of DNA or an aliquot of ligation mix (see **section 2.4.2**) was gently mixed with 100 µl competent cells in pre-chilled Eppendorfs and incubated on ice for 15 min. Cells were heat shocked by incubation at 37°C for 45 seconds then diluted with 1ml L-broth and incubated at 37°C for 30 min. The transformed cells were then pelleted in a microfuge by centrifuging for 30 seconds at high speed, then resuspended in 100 µl L-broth prior to spreading on L-agar plates, containing ampicillin if required.

<u>Ampicillin</u>

Ampicillin was prepared as a 100 mg/ml stock solution in SDW and stored at -20°C. This stock solution was added to liquid media or cooled agar-containing

media just before pouring, to give a final concentration of 100 μg/ml. The following control transformations were always carried out:

-A transformation with no added DNA to check for contamination

-A transformation containing 1 ng ds DNA to monitor cell competence.

2.3 *Preparation of DNA - small scale isolation of plasmid DNA from E. coli.*

Plasmid mini-preparations were carried out using the Qiagen QIAprep-spin Plasmid Kit, which produces very pure plasmid samples, capable of being used for *in vitro* transcription and translation. The manufacturer's protocols were followed. The basic principle is to bind plasmid DNA (released by alkaline lysis of the host bacteria) on a silica matrix in the presence of high salt concentrations. RNA, protein and other cellular contaminants are not adsorbed by the matrix.

2.4 *DNA Manipulation and Analysis.*

2.4.1 *Digestion of DNA with restriction Endonucleases.*

The following were pipetted into a clean Eppendorf tube:

DNA - 0.1 μg to 4 μg in SDW or TE (10 mM Tris-HCl (pH 7.4) 1 mM EDTA (pH 8.0))

2 μl 10x Restriction buffer

SDW (to make the final volume 20 μl)

Restriction endonuclease was added at 1 to 5 units per μg DNA and the reaction incubated at the temperature recommended by the enzyme manufacturers (usually 37°C, but 30°C for *Sma*I) for 1 to 4 hours.

The volume of restriction enzyme(s) added was less than 1/10 the reaction volume. Appropriate restriction buffers were supplied with the enzyme by the manufacturers. In principle, 1 unit of restriction endonuclease completely digests 1 μg of purified DNA in 60 min, using the recommended assay conditions.

If digestion by two restriction endonucleases was required, a restriction buffer was chosen in which both enzymes are relatively active. If the buffers of the enzyme were too dissimilar, the DNA was first digested by the enzyme active at the lower salt concentration, then sufficient 1 M NaCl (1 to 3 µl for a 20 µl reaction) so that the final concentration was appropriate for digestion by the second enzyme which was then added to the incubation.

2.4.2 *Ligations*

T4 ligase and its buffer came from Boehringer Mannheim and Gibco Life Technologies; reaction mixtures were set up according to the manufacturers' instructions. In nearly every case, ligation reactions were carried out using vector and insert cut with two restriction enzymes, so ligation could only occur in one orientation. The small fragment of polycloning site derived from the vector was removed by agarose gel electrophoresis (**section 2.4.3**) to reduce the background of nonrecombinant clones.

2.4.3 *Preparation and Running of Agarose Gels.*

A 1% (w/v) agarose gel was prepared by adding 1 g agarose to 100 ml 0.5 x TBE and heating in a microwave oven with intermittent mixing until the agarose was completely dissolved. The solution was allowed to cool to approximately 50°C and then ethidium bromide added (from a stock solution of 10 mg/ml in SDW) to give a final concentration of 0.5 g/ml. Gels were cast in a BRL "Horizon 58" horizontal slab gel apparatus with the required well-former inserted. When the gel was completely set, the well-former and end-sealers were removed and the gel immersed in 0.5 x TBE, to cover the gel to a depth of 1 mm. DNA samples and appropriate DNA molecular weight markers, were mixed with at least 1/6 volume of 6 x DNA sample buffer, loaded into the wells and the gel run at a constant voltage of 1-6 V/cm. Since DNA is negatively charged, it migrates towards the anode. DNA in the gel was visualised with an ultra-violet transilluminator, since ethidium bromide bound to DNA displays an increased fluorescent yield compared to that free in solution.

10 x TBE	6 x DNA sample buffer
0.89M Tris Base	30% (w/v) glycerol
0.89M Boric Acid	0.25% (w/v) Bromophenol blue
0.02M EDTA (pH 8.0)	0.25% (w/v) Xylene cyanol

2.4.4 *Isolation of DNA from agarose gels.*

Bands were cut from 1% agarose TBE gels (**section 2.4.3**) and the DNA they contained was prepared using the Qiagen gel extraction kit. Protocols supplied with the kit were followed throughout. The basic principle is to dissolve the agarose/DNA containing band using solutions supplied, then addition of the solution to the matrix. This silica matrix binds DNA in the presence of high salt concentrations. Contaminants are not adsorbed by the matrix.

2.4.5 *Blunt Ending of 5' or 3' Overhangs.*

(Sambrook *et al.*, 1989)

2 µg of DNA was digested to completion with the appropriate restriction enzyme(s) in a 20 µl reaction volume. 1 µl of a solution containing dATP, dCTP, dGTP and dTTP each at 2 mM in SDW was then added.

To blunt-end 5' overhangs, the recessed 3' termini were filled using the Klenow fragment of *E. coli* DNA polymerase 1. 1 unit of Klenow fragment was added for each microgram of DNA in the reaction and the sample incubated at room temperature for 15 min.

To blunt-end 3' overhangs, the protruding 3' termini were removed by utilising the 3' to 5' exonuclease activity of bacteriophage T4 DNA polymerase. 2 units of this enzyme were added for each microgram of DNA in the reaction and the sample left at room temperature for 30 min.

After the blunt-ending reaction, enzymes in the sample were inactivated by heating to 75°C for 10 min.

2.4.6 *Dideoxy Sequencing with modified T7 DNA Polymerase.*

Sequencing was carried out by the dideoxy chain termination method (Sanger *et al.*, 1977) using a Sequenase^{TM}II DNA sequencing kit from United States Biochemical Corporation/Amersham.

The chain termination method for sequencing involves synthesis of a DNA strand by *in vitro* using a single-stranded DNA template. A synthetic oligonucleotide, complementary to a portion of the DNA template, is annealed to the template and the DNA polymerase initiates synthesis from this complimentary strand. The primer is extended using dCTP, dGTP, dTTP and radioactively labelled [α-^{35}S] dATP. The synthesis reaction is terminated by random incorporation of dideoxynucleotide triphosphates (ddNTPs) which are nucleotide analogues that lack the 3'-OH required for chain elongation.
Sequenase^{TM}II is a modified form of bacteriophage T7 DNA polymerase which has low 3' to 5' exonuclease activity and, unlike Klenow fragment, has high processivity therefore DNA strands elongating from the template terminate only when a dideoxynucleotide is incorporated.

2.4.6.a *Preparation of double stranded plasmid DNA templates for sequencing.*

To prepare single stranded DNA 3 µg of plasmid DNA had 0.25 volumes of fresh 1 M NaOH, 1 mM Na_2EDTA added. This was incubated at 37°C for 30 min. Then 0.1 volumes of 3 M Sodium acetate, pH 4.8 followed by 2.5 volumes of absolute ethanol. This was allowed to precipitate at -20°C for 30 min and the sample was centrifuged at 15 000 g 4°C for 15 min. The pellet was washed with 70% ethanol and then centrifuged for 5 min at 15 000 g. The pellet was then allowed to dry. The pellet was then resuspended in 7 µl of water. The sequencing reaction was then carried out as in **section 2.4.6.b**.

2.4.6.b *Sequencing Reaction.*

1. Annealing Template and Primer - Annealing of the primer to the template was carried out by combining 1 µl primer (0.5 pmol/µl), 2 µl 5 x sequencing buffer (supplied in the kit)and 7 µl ssDNA (~1-2 µg).

2. Labelling Reaction - To the annealed template-primer the following was added: 1 µl 0.1 M DTT, 2 µl labelling mix, 0.5 µl [α-^{35}S]- dATP (10 Ci/µl) and 2 µl Sequenase™ II. This was mixed thoroughly and incubated for 15 min at room temperature.

3. Termination Reaction - 2.5 µl of each dideoxy termination mixture (ddATP, ddCTP, ddGTP, ddTTP) was pipetted into 4 separate wells of a microtiter plate. When the labelling reaction was complete, 3.5 µl of the labelling reaction was placed into each of the four wells. After 10 min incubation at 37°C, 4 µl of stop solution was added to each termination reaction and the samples were stored at 4°C.

2.4.7 *Preparation and Running of Sequencing Gels.*

Sequencing gels are high-resolution polyacrylamide gels which are used to fractionate radiolabelled single-stranded oligonucleotides on the basis of size. To reduce the effects of DNA secondary structure on electrophoretic mobility, gels contain 7 M urea and are run at sufficient voltage to heat them to approximately 65°C.

Sequencing gel plates (49.8 cm x 19.8 cm x 0.5 cm) were cleaned thoroughly, wiped with ethanol and then acetone. The cuspid plate had Repelcote™ applied according to the manufacturers instructions, then plates were assembled using 0.4 mm spacers. A 6% sequencing gel was prepared by mixing 42 ml of 6% acrylamide mix with 252 µl 10% (w/v) ammonium persulphate and 11 µl TEMED. The gel was poured immediately by sucking the mixture into a 60 ml syringe then slowly squirting it between the two plates, avoiding trapping air bubbles. A shark-toothed comb 0.4 mm thick, was inserted upside down 1 cm into the top of the gel to create a well, and the gel left to polymerise at room temperature for 1 hour. Once polymerised, the comb and bottom spacer were removed and the exposed surfaces rinsed with distilled water. The gel was secured into the sequencing gel tank (BRL). The top and bottom reservoirs was

filled with 1 x TBE. The buffer was squirted between the plates at the bottom of the gel to remove any air bubbles. If the gel was being run for less than two hours the gel was pre-warmed by running at 45 V/cm until hot to the touch. Just before loading the gel, the comb-well was rinsed by squirting buffer into it then the shark-tooth comb was inserted the correct way up with the points just sticking into the gel. Prior to running samples were heated to 80°C for 2 min. After the samples were loaded, the gel was run at 45 V/cm. When a sequence of over 200 bases needed to be read, half the sample was loaded and the gel run for approximately 3 hours (until the bromophenol blue dye in the sequencing stop solution was at the bottom of the gel) then the other half loaded and the gel run for a further three hours.

After running, the plate treated with Repelcote™ was removed. A piece of Whatman 3MM paper, slightly larger than the gel, was laid on top of the gel then gently peeled off with the gel attached. The gel was dried onto the paper, under vacuum for 1 hour at 80°C. When dry, the gel was exposed to X-ray film (Fuji 35.6 cm x 43.2 cm Medical) overnight. The film was developed in a dark-room using LX24 developer (Kodak) and fixed with Unifix™ (Kodak) according to the manufacturers instructions.

6% Acrylamide Mix	40% Acrylamide Mix
43 g Urea	38% (w/v) Acrylamide
5 ml 10xTBE	2% (w/v) Bis-acrylamide
15 ml 40% Acrylamide mix	(made up with SDW)
Made up to 100ml with SDW.	

2.4.8 *Oligonucleotide-directed Mutagenesis of Cloned DNA.*

2.4.8a *Preparation of Oligonucleotides.*

Oligonucleotides were synthesised on an Applied Biosystems 380B DNA synthesiser and purified by high performance liquid chromatography.

Oligonucleotides were 5'-phosphorylated as follows:

The oligonucleotide was diluted to approximately 20 nmol/ml and 2.5 µl transferred to a clean Eppendorf. The sample was gently mixed with 25 µl of SDW, 3 µl of 10 x kinase buffer (1 M Tris pH8.0/ 100 mM $MgCl_2$/ 70 mM DTT/ 10 mM ATP, filter-sterilised) and 2 units of bacteriophage T4 polynucleotide kinase then incubated at 37°C for 15 min. The reaction was stopped by heating to 70°C for 10 min and the sample stored at -20°C until required.

2.4.8b *Oligonucleotide-directed Mutagenesis using the iPCR Method.*

Hemsley *et al.*, (1989)

In this procedure, a sequence change in a plasmid is directed by one of two opposing primers which anneal "back-to-back". PCR results in amplification of the entire plasmid with the change now incorporated, and this product can then be purified and self-ligated to regenerate a circular vector. The procedure is rapid and 50-100% of resulting colonies harbour the mutant sequence.

iPCR can be used to engineer limited base substitutions and insertions, and also deletion mutants.

The following are alterations to the Hemsley *et al* (1989) protocol, used for the production of mutations in this thesis:

(a) The original protocol can produce deletions at the ligation point. Therefore, kinased oligos were used which prevents this and furthermore provides the necessary phosphate groups for self-ligation of the iPCR-product.

(b) Often the original plasmid will co-purify with the iPCR-product on a gel. A good yield of mutants, therefore, depends on amplifying from a low concentration of starting DNA (1-10 f mol).

(c) Using a polymerase with proof-reading qualities (e.g. Pwo-polymerase from Boehringer Mannheim) prevents the addition of extra nucleotides (usually A at the 3' end - as found by Hemsley *et al.*, (1989) using Taq polymerase) and enables a blunt-end self-ligation of the iPCR-product immediately after purifying the product through agarose gel electrophoresis.

The oligonucleotides used throughout this thesis were as follows:

cDNA mutated	Position of Mutation	amino acid altered
PSII-W (single mutation)	AA 78	Ala to Thr
PSII-X (single mutation)	AA 74	Ala to Thr
PsbY (single mutation)	AA 66 (PsbY/1)	Ala to Thr
	AA 66 (PsbY/1Leu)	Ala to Leu
	AA 78 (PsbY/1*)	Ala to Thr
	AA 143 (PsbY/2)	Ala to Thr
	AA 150 (PsbY/2*)	Ala to Thr
PsbY (double mutation)	AA 66 (PsbY/1, 2)	Ala to Thr
	and AA 143	Ala to Thr

N.B. AA refers to the amino acid which was mutated, numbers are relative to the initiation codon.

(1) For Ala to Thr mutations the first nucleotide of the codon was altered from G to A (region in primer underlined).

(2) For the Ala to Leu mutation the first two nucleotides were altered from GC to CU (region in primer underlined).

All mutant primers prepared were forward primers and 35mers. Reverse primers were prepared as 35mers and contained no alterations to the wild-type sequence.

The following primers were used for this thesis -

(1) Pre-PSII-W primers (alanine to threonine)

Forward (mutant primer):

TgAgcAATcccgcgATgAcTTTggTTgATgAgAgg

Reverse:

TcAccgccgTcAAAgcAgcTgTTgcgcAgcTgAA

(2) Pre-PSII-X primers (alanine to threonine)

Forward (mutant primer):

TTccggAgATAgcTgAAAcTgcTggTTcTggAATc

Reverse:

TcAccATcgAAgcAgTAAgAgcAgccgcggAgATA

(3) Pre-PsbY primers (1 alanine to threonine)

Forward (mutant primer):

AcTcTgAgccAgcTTTAAccATccAAcAgATcgcT

Reverse:

AgcTgAgggAAgAgAAgAcggcTccggcTAgggcg

(4) Pre-PsbY primers (1* alanine to threonine)

Forward (mutant primer):

AgcTcgcTgcAgcAAAcAccAgTAgcgAcAAccgT

Reverse:

gAgcgATcTgTTggTTggcTAAAgcTggcTcTcAgA

(5) Pre-PsbY primers (2 alanine to threonine)

Forward (mutant primer):

cAccTccggAggcTTATAcTgcggcAgAAgcggc

Reverse:

gAgTcAAAAgcccTgAcgcTgcAAgAccAccA

(6) Pre-PsbY primers (2* alanine to threonine)

Forward (mutant primer):

cggcAgAAgcggcggcTAcTAgTTcAgAcAgcAgA

Reverse:

cAgcATAAgccTccggAggTggAgTcAAAAgcccT

(7) Pre-PsbY primers (1leu alanine to leucine)

Forward (mutant primer):

AcTcTgAgccAgcTTTAcUcATccAAcAgATcgcT

Reverse:

AgcTgAgggAAgAgAAgAcggcTccggcTAgggcg

N.B. for the double mutation (PsbY/1, 2) primer set 3 and primer set 5 were used consecutively.

iPCR mix -

5.0 µl	10 X PCR buffer (1.5 mM Mg^{2+})
2.5 µl	template (10 ng/l = 10 fmol)
1.0 µl	primer 1 (kinased, 50 pmol/µl)
1.0 µl	primer 2 (kinased, 50 pmol/µl)
4.0 µl	dNTP mix (2.5 mM each)
36.0 µl	H_2O
0.5 µl	Pwo-polymerase (5 units/µl)
50.0 µl	

All these components were mixed on ice and the polymerase was added last. To prevent starting the reaction before the template had been completely denatured a hot start PCR program was used.

Program -

1. first denaturing	94°C	2 min
2. Denaturing	94°C	1 min
3. Annealing (Tm of primer -8°C)		1 min
4. Extension	72°C	6 min
5. Last extension	72°C	10 min
25 cycles of steps 2-4 were used		

Self-ligation -

14 µl	iPCR-product (gel purified)
4 µl	5 x ligation buffer
2 µl	T4 ligase
made up to 20 µl with SDW	

10 µl of the ligation mix is used for transforming 100 µl competent cells. The average transformation rate is 500-1000 colonies/plate. Much lower rates gave an indication of incorrect PCR products - usually multiple identical insertions of 20-30 bp following the primers.

2.5 *Transcription, translation and protein analysis.*

2.5.1 *In vitro transcription of DNA in pGEM™-4Z and -5Z.*

Full length cDNA inserts containing Pre-PSII-X and Pre-PSII-W (Kim *et al.*, 1996) were cloned into *Eco*R1-*Bam*H1 sites of the vector pGEM-4Z (Promega), in the orientation for transcription from the T7 promoter (kindly provided by S. J. Kim, University of Warwick). cDNA encoding for Pre-PsbZ was cloned into the pGem-5Z vector, in the orientation for transcription from the T7 promoter (kindly provided by A. Mant, University of Warwick).

To 2 µl of vector and insert DNA (1 µg/µl) the following were added:
15.5 µl transcription premix (40 mM Tris-Cl pH 7.5/ 6 mM $MgCl_2$/ 2 mM

spermidine/ 10 mM DTT/ 500 μM rATP/ 500 μM rCTP/ 500 μM rGTP/ 500 μM rUTP/ 100 μM per ml BSA)

20 units RNasin (Promega. 40 units/μl)

500 μm monomethyl cap (m7G(5')ppp(5')G) (Pharmacia)

15 U SP6 or T7 RNA polymerase.

In a total volume of 20μl

SP6 RNA polymerase (GibCo BRL. 15 units/μl) or T7 RNA polymerase (GibCo BRL. 50 units/μl).

The sample was incubated for 30 min at 37°C, then rGTP was added to a final concentration of 600 μm. The mixture was further incubated at 37°C for 30 min and the process was terminated by placing the Eppendorf tube on dry ice. The transcription reaction product was stored at -70°C until required for translation.

2.5.2 *In vitro translation using a wheat germ lysate system.*

All precursor proteins used in this project were produced by *in vitro* translation using the Promega Wheat Germ Extract system.

In a 50μl wheat germ translation reaction, the components are:

wheat germ extract	25 μl
1mM amino acid mixture (minus met or leu)	4 μl
RNA substrate from **2.5.1**	2 μl
RNasin at 40 U μl^{-1}	1 μl
[^{35}S] methionine / [^{3}H] leucine	70 / 50 Ci
nuclease-free water	to final volume of 50 μl
Incubation 60 min at 27°C.	

The translation reaction product was stored at -70°C until required. Wheat germ translation products were separated by SDS-PAGE and visualised by fluorography followed by autoradiography as outlined below.

2.5.3 *SDS-Polyacrylamide gel electrophoresis (SDS-PAGE).*

2.5.3a *Preparation and running of SDS-Polyacrylamide tricine gels*

(Schägger and von Jagow, 1987)

Gel plates (15.5cm x 12cm x 0.5cm) were cleaned and assembled using 1mm spacers on the vertical edges. The bottom of the gel was sealed with an agarose plug by placing the assembled plates into a plastic trough containing melted 1% (w/v) agarose. Once the plug had set, resolving gel mix was poured between the plates, leaving enough space at the top of the plates for insertion of the well-forming comb. The gel was immediately overlaid with water saturated butan-1-ol. Once the resolving gel had polymerised, the butan-1-ol was rinsed away with distilled water and the gel surface carefully dried by inserting a piece of Whatman paper between the plates. Stacking gel mix was then poured into the remaining space between the plates and the required well-forming comb inserted. Once the stacking gel had polymerised, the comb was gently removed and the wells rinsed with distilled water to remove any unpolymerised acrylamide. The gel was then removed from the trough, taking care not to dislodge the agarose plug, and assembled in the gel tank (Scie-Plas). The top gel tank reservoir was filled with 1 x cathode buffer, sample wells were rinsed out with this buffer using a Hamilton syringe, and the lower gel tank reservoir was filled with 1 x anode buffer. Samples were heated to 95°C for 4 min with at least an equal volume of 2 x protein sample buffer before loading with a Hamilton syringe. Initially, electrophoresis was carried out at 90 volts until samples were in the stacking portion of the gel and then the gel was run at a current of 30 mA on constant voltage until the bromophenol blue, in the sample buffer, ran into the agarose plug (usually about 16 hours).

2 x Protein Sample Buffer	10 x Anode (+ve) Buffer
50 mM Tris / HCl pH 6.8	2.0 M Tris / HCl pH 8.9
0.4% (w/v) SDS	
1.2% (v/v) Glycerol	10 x Cathode (-ve) Buffer
0.2% (v/v) 2-β mercaptoethanol	1M Tris
2.4% (w/v) Urea	1M Tricine
Sufficient Bromophenol Blue	1% SDS
Stain protein gel fix	Destain protein gel fix
0.25% (w/v) Coomassie Brilliant blue R250	40% (v/v) Methanol
	7% (v/v) acetic acid
45% (v/v) H_2O	53% (v/v) H_2O
45% (v/v) Methanol	
10% (v/v) acetic acid	

Resolving Gel Mix
10 ml 48% Acrylamide / 1.5% bis-acrylamide (w/v)
10 ml 3M Tris / HCl pH 8.45
9.9 ml water
300 µl 10%SDS
100 µl APS
40 µl TEMED

Stacking Gel Mix

1 ml 48% Acrylamide / 1.5% bis-acrylamide (w/v)

3.1 ml 3M Tris / HCl pH 8.45

8.3 ml water

93 µl 10% SDS

100 µl APS

30 µl TEMED

For both resolving and stacking gels, the ammonium persulphate and TEMED solutions were added and mixed with the other components just before pouring.

2.5.3b *Staining, drying and autoradiography of SDS-polyacrylamide gels.*

The gel was removed from the tank and the plates prised apart. The gel was carefully removed from the plates and placed in stain protein gel fix for at least 30 min, with gentle shaking. The gel was then placed in destaining protein gel fix for 1-2 hours, with gentle shaking. The gel was then placed in Amplify™ (Amersham) for 20 to 30 min then placed on a piece of Whatman 3MM paper before drying under vacuum at 80°C for 1.5 hours. The gel was then placed in a light-proof cassette and exposed to X-ray film (Fuji 18 cm x 24 cm Medical). The film was developed in a dark-room using LX24 developer (Kodak) and fixed with Unifix (Kodak) according to the manufacturer's instructions.

2.6 *Preparation of Isolated Chloroplasts, Thylakoid Vesicles and Import Assays.*

2.6.1 *Growth of Pea Plants.*

Seedlings of *Pisum sativum*, var. Feltham First were grown for 9 days, under a 12 h photoperiod, of light intensity 50 µmol photons $M^{-2}s^{-1}$.

The growth room temperature was maintained between 18°C and 22°C, as far as possible.

2.6.2 *Isolation of intact chloroplasts and thylakoid vesicles.*

Isolation of pea chloroplasts was as described by Brock *et al.*, 1993.

2.6.3 *Determination of chlorophyll concentration.*

Chlorophyll concentrations were measured as described by Hipkins and Baker (1986) after the method of Arnon (1949). Chloroplast numbers were counted using a heamocytometer, and were found to vary directly with chlorophyll concentration. Given the described growth conditions, ages and cultivars of plants, the following figures were used as an approximate guide to chloroplast numbers.

Pea seedlings: 25 μg chlorophyll corresponds to 5×10^7 chloroplasts.
(20×10^6 per μg chlorophyll).

2.6.4 *Solutions used for the isolation of chloroplasts and import assays.*

5 x Sorbitol resuspention Media (5 x SRM)	SRM-EDTA (SRME)
1.65 M sorbitol 250 mM Hepes-KOH pH 8.0	1 x SRM 50 mM EDTA, pH 8.0
Hepes-magnesium (HM)	Hepes-magnesium-EDTA (HME)
10mM Hepes-KOH pH 8.0 5mM $MgCl_2$,	HM buffer containing 10mM EDTA (for ending thermolysin treatments)
Percoll pads	
7 ml 100% Percoll 9 ml distilled water 4 ml 5 x SRM (percoll concentration 35%)	

2.6.5 *Assay for Protein Import into Intact Chloroplasts.*

Import incubations contained:

Chloroplasts equivalent to 50 μg chlorophyll

8 mM MgATP (60 mM stock solution pH 7.0, dissolved in SRM)

SRM upto a volume of 270 μl

These components were mixed (using wide-aperture, yellow, pipette tips) and preincubated in the light bath (intensity 150 mol photons $M^{-2}s^{-1}$ at 26°C) for 10 min. The import assay was initiated by addition of the following mixture:

12.5 μl unlabelled methionine/leucine (60 mM stocks in 2 x SRM)

12.5 μl radiolabelled, *in vitro*-translated precursor protein.

Eppendorf tubes were returned to the light bath for a 20 minute duration. Incubations were terminated by the addition of 1ml ice-cold SRM, placing the tubes upon ice. Procedures for fractionation of the chloroplasts are described in the following numbered paragraphs.

1. Recovery of the chloroplasts: Diluted chloroplasts were centrifuged at 6500 rpm for 1 min (using a microcentrifuge). After removing the radioactive supernatant, the pellet was carefully resuspended in 120 μl ice-cold SRM (wide-aperture, yellow pipette tips).

2. Total chloroplast sample (non-protease-treated chloroplasts): 30 μl chloroplasts were removed to a fresh Eppendorf tube (sample C-).

3. Protease-treatment of chloroplasts: to the remaining 90 μl chloroplasts were added 2.5 μl $CaCl_2$ (such that the final concentration in the Eppendorf tube was 2.5 mM Ca^{2+}) and 9 μl thermolysin (2 mg/ml). Protease-treatment was carried out for 40 min upon ice, and was terminated by the addition of 2 μl 0.5 M EDTA, followed by centrifugation for 1 min at 6500 rpm in a microcentrifuge. The supernatant was removed and the pellet resuspended in 90 μl HME. 30 μl was removed to a fresh Eppendorf tube (sample C+).

4. Lysis of the chloroplasts to obtain stroma and thylakoids: the remaining chloroplasts were allowed to lyse for 10 min on ice, then they were centrifuged at 15 000 rpm for 1 min at 4°C. The supernatant (stromal extract) was removed into a fresh Eppendorf tube (sample S). The pellet was resuspended in 60 µl HM and the sample centrifuged for 1 min at 15 000 rpm in a microcentrifuge.

5. Non-protease-treated thylakoids sample: the pellet was resuspended in 60 µl HM and 1.67 µl $CaCl_2$ (such that the final concentration in the Eppendorf tube was 2.5 mM Ca^{2+}). 30 µl was removed to a fresh Eppendorf tube (sample T-).

6. Protease-treated thylakoids sample: to the remaining 30 µl of thylakoids were added 3 µl thermolysin. Protease-treatment was for 30 min on ice, and was terminated by the addition of 200 µl HME. The mixture was centrifuged at 15 000 rpm for 1 min, and then resuspended in 30 µl HME (sample T+).

7. Preparation of samples for SDS-PAGE: all samples were kept on dry ice until the end of the experiment, whereupon they were boiled for 5 min and centrifuged to remix condensed moisture from the lids of the tubes. See **section 2.5.3**.

2.6.6 *Topological Probing of Proteins Imported into Isolated Chloroplasts using Proteases.*

All protease treatments were performed on untreated thylakoid fractions from a chloroplast import. All the proteases used for topological analysis were prepared in 1 x HM. The incubations used for each protease are indicated below. The following proteases were used:

Protease	Cleaves at	Incubation time
Endoproteinase Glu-C (V8 protease)	Carboxylic side of Glu and Asp.	60 min
Thermolysin	Hydrolyses peptide bonds involving hydrophobic amino acids e.g. Ile, Leu, Met, Phe, Trp and Val. Cleaves from C-terminus.	30 min
Trypsin	Hydrolyses proteins at carboxylic side of basic amino acids Arg and Lys.	30 min

2.6.7 *Determination of protein stability in the thylakoid membrane.*

2.6.7.a *Carbonate washing*

Michl *et al.* (1994) protocol followed throughout.

2.6.7.b *Urea washing*

Breyton *et al.* (1994) protocol followed throughout.

2.6.8 *Inhibitor studies with thylakoids*

All protease treatments were performed on untreated thylakoid fractions from a chloroplast import. All the proteases used for topological analysis were prepared in 1 x HM. The incubations used for each protease are indicated below. The following proteases were used:

Protease Inhibitor	Inhibits	Incubation time
E64	cysteine proteases	30 min
EDTA	metalloproteases	30 min
1, 10-Phenanthroline	metalloproteases	30 min
Phenylmethylsulphonyl fluoride (PMSF)	Serine proteases	30 min
N-ethyl maleimide	Cysteine proteases	30 min

2.7 *Reagent Suppliers.*

Reagents and materials were analytical grade when available, and were obtained from the following suppliers:

Aldrich Chemical Co.	Tween 20	
	Tricine	
Amersham International plc.	All radiochemicals	
	restriction enzymes	
	Amplify™ fluorographic reagent	
	Sequenase™ version 2.0 sequencing kit	
BDH Merck Ltd	acetone	ammonium acetate
	ammonium sulphate	APS
	boric acid	bromophenol blue
	butan-1-ol	chloroform

	dithiothreitol	DMSO
	ethanol	glycerol
	HEPES	isopropanol
	magnesium sulphate	
	manganese chloride	
	methanol	PEG 6000
	potassium acetate	sodium chloride
	sodium hydroxide	SDS
	thermolysin	Tris
	Triton X-100	
Bethesda Research Laboratories	agarose	
Bio-Rad Laboratories Ltd.		
Boehringer Mannheim Ltd.	DNAaseI	Pwo DNA polymerase
	T4 ligase	
Calbiochem Ltd.		
Difco	agar	tryptone
Eastman Kodak Co.	bis-acrylamide	LX24-devolper
	Unifix	
Fisons Scientific Equipment	acetic acid	acrylamide
	ammonium sulphate	caesium chloride
	calcium chloride	
	disodium hydrogen orthophosphate	
	EDTA	glucose
	glycerine	magnesium chloride
	2-mercaptoethanol	phenol
	potassium chloride	
	potassium hydroxide	
	potassium dihydrogen orthophosphate	
	sodium acetate	sucrose
	urea	
Fuji Photo Film Co. Ltd.	RX medical X-ray film	
Gibco Life Technologies Ltd.	T3 RNA polymerase	

	T4 ligase
	T4 polynucleotide kinase
	restriction enzymes
ICN Pharmaceuticals, Inc.	
New England Biolabs	restriction enzymes
Pharmacia Biotech	cap (m7G(5')ppp(5')G)
	Percoll
Promega corporation	wheatgerm extract
	RNasin
	SP6 and T7 RNA polymerases
	T7 primer
	amino acid mixtures for *in vitro* translation
	rATP, rCTP, rGTP, rUTP
	pGEM-4Z, pGEM-5Z
	ultra-pure dithiothreitol for transcription
QIAGEN Ltd.	QIAprep-spin plasmid kit
	QIAEX gel extraction kit
Sigma Chemical Co.	ampicillin apyrase
	ATP (NA and Mg salts)
	Coomassie Brilliant Blue
	cysteine dimethyldichlorosilane
	ethidium bromide IPTG
	leucine lysozyme
	Mes methionine
	nigericin RNaseA
	sodium azide sorbitol
	TEMED tetracycline
	X-gal
Stratagene Ltd.	pBluescript
Unipath Ltd.	Yeast extract
Whatman Paper Ltd.	Whatman filter paper

Chapter 3

Mechanistic analysis of pre-PSII-W and pre-PSII-X into the thylakoid membrane

Chapter 3

Results 1 - Mechanistic analysis of pre-PSII-W and pre-PSII-X insertion into the thylakoid membrane

3.1 *Introduction*

A large proportion of bacterial, chloroplast, and mitochondrial membrane proteins are inserted post-translationally, and interest has centered on the mechanisms by which these proteins are initially maintained in the soluble form, and how the hydrophobic regions are then transferred from an aqueous phase into the hydrophobic environment of the membrane bilayer. Certain elements of the insertion events have been characterised in detail, one example being the influence of positive charges on overall topology, since these residues have a strong tendency to remain on the *cis* side of the membrane ('positive inside rule'; von Heijne, 1994). It has also been possible to recognise two broad categories of insertion mechanism, which can be regarded as "assisted" and "unassisted" according to whether protein translocation machinery is relied upon. In bacteria, for example, the secretory (Sec) apparatus used for the translocation of periplasmic proteins is also required for the insertion of some membrane proteins (Wolfe *et al.*, 1985; Gebert *et al.*, 1988; Sääf *et al.*, 1995; Traxler *et al.*, 1996). In addition, another element of the export machinery, the signal recognition particle (SRP), is involved in the insertion of a range of membrane proteins (De Gier *et al.*, 1997; Ulbrandt *et al.*, 1997; High *et al.*, 1997). However, other proteins are integrated by Sec-independent mechanisms and these may well insert spontaneously into the plasma membrane (Andersson and von Heijne, 1993; Kuhn, 1995; Brassilana and Gwidzek, 1996; Kiefer *et al.*, 1997). Even though the overall requirements have been detailed for several membrane proteins, in most cases the actual insertion mechanisms are poorly understood. Relatively few proteins have been examined in genuine mechanistic detail, apart from the coat proteins of the M13 and Pf3 phages which are notable exceptions that have been shown to integrate into the *E. coli* plasma membrane by Sec-independent, possibly spontaneous mechanisms (reviewed in Kuhn, 1995).

At present, only two main pathways have been identified for integral thylakoid membrane proteins. In the first case, the vast majority of research has focused on the light-harvesting chlorophyll *a/b* binding protein (LHCP). This is a multispanning membrane protein, which only has an envelope transit signal. This protein integrates into the thylakoid membrane by means of information contained in the mature protein (Lamppa, 1988; Viitanen *et al.*, 1988). The insertion process requires GTP and a stromal homologue of the 54-kDa protein of signal recognition particles (Li *et al.*, 1995), and is thus likely to be similar in certain respects to the SRP-dependent pathway identified in bacteria. For the second pathway, a very different mechanism has been demonstrated for three thylakoid proteins: subunit II of the integral CF_0 component of the thylakoid ATP synthase (CF_0II) and subunits X and W of photosystem II (PSII-X and PSII-W). These proteins are synthesised with bipartite presequences that strongly resemble typical Sec-type signal peptides (Hermann *et al.*, 1993; Lorkovic *et al.*, 1995; Kim *et al.*, 1996), yet they integrate into thylakoid membranes in the absence of stromal factors, nucleoside triphosphates, or ΔpH (Michl *et al.*, 1994; Kim *et al.*, 1998). Furthermore, mild proteolysis of thylakoids blocks the Sec-, SRP- and ΔpH-dependent mechanisms but has no effect on the insertion of CF_0II, PSII-W, and PSII-X (Robinson *et al.*, 1996; Kim *et al.*, 1998). It has therefore been proposed that they insert spontaneously into the thylakoid membrane. This particular insertion mechanism is highly unusual in that no other mainstream insertion process studied to date involves the use of cleavable signal peptides for "unaided" membrane insertion, and M13 procoat is in fact the only other protein known to use this type of mechanism. In this chapter, I have analysed this integration pathway in greater detail in order to elucidate a possible role for TPP in the insertion mechanism, and to identify mechanistic details of the membrane-integration events. A priority in this work was to determine whether these proteins insert by means of critical loop intermediates that are central to the insertion process.

3.2 *Mutagenesis and sequencing of pre-PSII-W and Pre-PSII-X*

Both cDNA templates were supplied by Soo Jung Kim (Warwick) and consisted of *Arabidopsis* pre-PSII-W and pre-PSII-X cloned into pGEM-4Z. The method of mutagenesis used was the iPCR method (see chapter 2). The PCR products were run

on 1% agarose gels, and the bands of appropriate size (in comparison to the construct restricted at a single site) were gel excised and purified. These purified bands were ligated overnight and transformed into competent *E. coli*. Colonies were taken at random and grown overnight in LB broth. Plasmid DNA was extracted from these cultures and sequenced to determine the presence of the correct mutation. On finding the mutation in the plasmid, the total mutant cDNA was fully sequenced. All *in vitro* translations used in this chapter were labelled with [^{3}H]-leucine unless otherwise mentioned.

3.3 *Cleavage by TPP is not required for the insertion of PSII-W*

So far, TPP is the only known proteinaceous component involved in this insertion mechanism. This leads to a hypothesis that TPP cleavage might behave in an analogous role to that of a receptor for insertion into the thylakoid membrane, perhaps rendering the insertion process unidirectional. Therefore, the first aim was to determine whether the removal of the signal peptide was a requirement for stable insertion into the thylakoid membrane.

To address this question, I inhibited the cleavage step catalysed by TPP (whose active site is in the thylakoid lumen, Kirwin *et al.*, 1988), in order to probe the location and topology of this protein immediately prior to the final maturation step. **Figure 14** shows the structures of the relevant sections of both precursors, including the signal peptide and the mature region of the precursor protein. TPP is known to depend on the presence of short-chain residues at the -3 and -1 positions in the substrate, relative to the processing site, and alanine at the -1 position is essential with even serine and glycine unable to support efficient cleavage (Shackleton *et al.*, 1991). A threonine was therefore introduced at the position in pre-PSII-W, in the expectation that processing would be affected without substantially altering the characteristics of the translocated region.

Figure 13. *The Amino Acid sequence of wild-type Pre-PSII-W and Pre-PSII-X. Top* line of both sequences is the chloroplast envelope transit signal. The *Middle* line is the thylakoid transit peptide ending in *Ala-X-Ala.* The *bottom* line is the mature protein. *Underlined* amino acids are the hydrophobic regions, and the alanine (-1, in relation to TPP cleavage) denoted by the **bold** *italic* is the amino acid mutated in both pre-PSII-W and pre-PSII-X to threonine.

(A) Pre-PSII-W

MASFTASASTVSAARPALLLKPTVAISAPVLGLPPMGK

KKGGVRCSMETKQGNVSVMGAGVSAAATAALTAVMSNPAM***A***

LVDERMSTEGTGLPFGLSNNLLGWILFGVFGLIWTFFFVYTSSLEEDEESGLSL

TPP Cleavage

(B) Pre-PSII-X

MASTSAMSLVTPLNQTRSSPFLKPLPLKPSKALVATGGRAQ

RLQVKALKMDKALTGISAAALTASMVIPEIAE***A***

AGSGISPSLKNFLLSIASGGLVLTVIIGVVVGVSNFDPVKRT

TPP Cleavage

Figure 14. *Structural features of CF_0II (C), PSII-X (X), and PSII-W (W), and M13 procoat (Pc). A*, the figure illustrates the full sequence of M13 procoat, spinach CF_0II, and PSII-W and PSII-X from *A.thaliana*; the thylakoid proteins start with the signal-type peptides (the N-terminal envelope transit peptides are omitted). The extreme C-terminal region of CF_0II has also been omitted. Hydrophobic domains are *underlined*, acidic residues are shown italicised and the TPP cleavage site denoted by an *asterisk*. The figure also shows the locations of the individual sections within iPSII-W (membrane, lumen or stroma exposed) deduced in this study. *B*, alignment of PSII-X sequences from *Arabidopsis* (*A. thal*), *O. sinensis* (*O. sinen*), *C. paradoxa* (*C. para*), and *P. purpurea* (*P. purp*). The *Arabidopsis* sequence includes the signal peptide of the presequence region; the remaining sequences are given in full. Identical residues are denoted by *asterisks*; conserved residues by *dots*. Adapted from Thompson *et al.*, (1998).

A. Procoat and thylakoid membrane proteins

```
(Pc)        MKKSLVLKASVAVATLVPMLSFA*AEGDDPAKAAFNSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSKAS
(C) -HLNLSVLKSAAITATPLTLSFLLPYPSLA*EEIEKASLFDFNLTLPIIMAEFLFLMFALIYYTPLGDFMDKRDASIKEQ---
(X) -KALKMDKALTGISAAALTASMVIPEIAEA*AGSGISPSLKNFLLSIASGGLVLTVIIGVVVGVSNFDPVKRT

(W)  -GNVSVMGAGVSAAATAALTAVMSNPAMA*LVDERMSTEGTGLPFGLSNNLLGWILFGVFGLIWTFFFVYTSSLEEDEESGLSL
```

stroma	membrane	lumen	membrane	stroma

B. PSII-X sequences

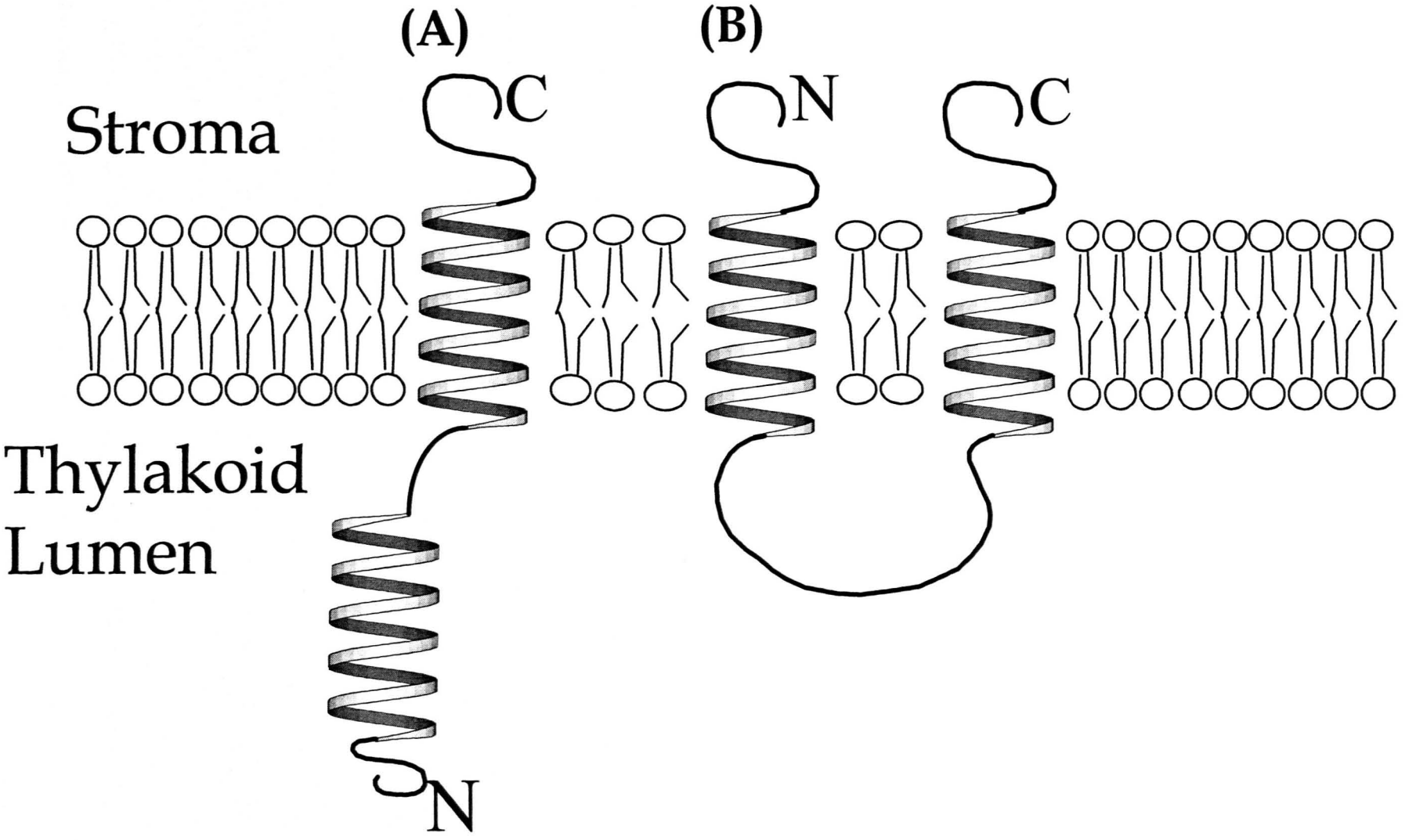
```
A.thal.      --DKALTGISAAALTASMVIPEIAEA*AGSGISPSLKNFLLSIASGGLVLTV-IIGVVVGVSNFDPVKRT
O.sinen.                                MTPSLANFISSLTAGGLVVLT-IAVALIVISRTDRVTRF
C.para.                                 MTSSLSAFINSLLWGAIVVVIPLSAALLFISQKDKIQRN
P.purp.                                 MTPSLSSFLNSLILGAVIVVVPITLALLFVSQKDRTIRS
                                        .. **  *. *.  *....   .   .. .*  *   *
```

The import characteristics of the PSII-W mutant (W/A78T) are shown in **figure 16**. The wild-type precursor protein is imported into the chloroplasts, processed to the mature size (6.1kDa), and fractionation of the chloroplasts confirm that this polypeptide is located exclusively in the membrane fraction (lane T). Protease treatment of the thylakoid membranes (lane T+) results in digestion of the exposed C-terminal region (as found in previous studies on authentic PSII-W, Irrgang *et al.*, 1995), and the production of a slightly smaller degradation product. The mutant protein, on the other hand, is converted to a mixture of mature- and intermediate-size proteins (6.1 and 10.3 kDa respectively), the latter form presumably resulting from the action of SPP, which cleaves the envelope transit domains from most bipartite presequences. Importantly, this iPSII-W form is also found only in the thylakoid fraction and further tests show that it is completely resistant to carbonate extraction, because after this treatment both the intermediate and mature bands are recovered only in the pellet fraction (lane P). Since this procedure effectively removes extrinsic proteins from the thylakoid membrane (Michl *et al.*, 1994), it can exclude the possibility that the action of TPP is required for the insertion of this protein. Thermolysin treatment of the iPSII-W results in a mobility shift that is comparable to that of mature PSII-W, confirming that at least one section of the intermediate is exposed on the stromal face of the thylakoid membrane.

3.4 *A loop intermediate in the insertion of PSII-W*

The next step was to determine whether the cleavable signal-type peptides of the thylakoid proteins are used for a similar purpose as that of M13 procoat: the formation of a loop-intermediate in which the hydrophobic region (H-domain) spans the membrane. This is probably the most important role of the M13 signal peptide and the priority was to determine whether such an intermediate is a core feature in the biogenesis of PSII-W and/or PSII-X. Alternatively, from Shackleton *et al.* (1991) it was shown that inhibiting TPP cleavage by site-directed mutagenesis on a lumenal protein (pre-33K), firstly demonstrated that mutations at these positions (-1 in relation to TPP cleavage) did not affect import across the thylakoid membrane shown with protease protection assays; and secondly, has an increased association with the lumenal face of the thylakoid membrane. This presents two possible arrangements of

Figure 15. *Possible arrangements of the signal peptide and mature protein of pre-PSII-W in the thylakoid membrane.* The data available shows two possible topologies of the signal peptide and the mature protein in the thylakoid membrane. Previous work using TPP cleavage site mutations (Shackleton *et al.*, 1991), demonstrated that inhibition of TPP cleavage by introduction of various amino acids at the -1 position (in relation to the TPP cleavage site) enhanced the ability, compared with wild-type, to adhere to the lumenal face of the thylakoid membrane. Therefore giving rise to arrangement *A*. The second possible topology, arrangement *B*, is derived from work on M13 procoat. Previous studies have found a loop intermediate to be an absolute requirement for insertion.


(A)
(B)
Stroma
C
N
C
Thylakoid
Lumen
N

the signal peptide and mature protein in relation to the thylakoid membrane (see **figure 15*a*** and **15*b***)

The topology of the iPSII-W polypeptide was mapped by determining the accessibility of trypsin and thermolysin cleavage sites. **Figure 13** shows that the PSII-W mature protein contains only a single basic residue which is located in the lumenal N-terminal region (arginine at position 5); the presequence contains many basic residues but these are all located prior to the H-domain in the signal peptide. In a loop structure these sites are predicted to be on the lumenal and stromal sides of the membrane, respectively. **Figure 17** shows that trypsin cleaves iPSII-W to a smaller form that remains significantly larger than mature-size PSII-W; low concentrations of trypsin (e.g. 10 μg/ml) generate a mixture of iPSII-W and a smaller degradation product, whereas 25 μg/ml is sufficient to fully convert the iPSII-W to the degradation product, without cleavage of the mature protein. Trypsin therefore cleaves the N-terminus of the iPSII-W signal peptide on the stromal surface of the membrane. A loop intermediate would also contain the C-terminus of the mature protein on this face of the membrane, and **figure 18** shows that this is the case. The C-terminus of mature PSII-W is exposed on the stromal face (as shown by Irrgang *et al.*, 1995) and the data in **figure 16** showed that it is cleaved by thermolysin; **figure 18** shows that thermolysin also cleaves a small fragment from iPSII-W, generating a degradation product similar in mobility to that produced by trypsin (compare lanes Th and Tp). The important point is that incubation of iPSII-W with both thermolysin and trypsin generates a much smaller degradation product (lane Th/Tp) indicating that trypsin and thermolysin cleave at different ends of the molecule (the N and C termini, respectively). Lane Tp/son of **figure 18** provides final confirmation of a loop structure: thylakoids containing iPSII-W were incubated with trypsin and the thylakoids were sonicated to allow access of the protease to the lumenal side of the membrane. Both the mature PSII-W and the iPSII-W are almost quantitatively converted to a polypeptide that is smaller than mature PSII-W, indicating that cleavage has taken place after the arginine at position 5 in the mature protein, which must therefore be exposed on the lumenal side of the membrane.

Figure 16. *Stable insertion of PSII-W occurs in the absence of cleavage by thylakoidal processing peptidase.* Pre-PSII-W (*left panel*) or pre-PSII-W containing an Ala78Thr mutation (W/A78T) were imported into intact chloroplasts. After import, samples were analysed of the total chloroplast fraction (*lane C*), protease-treated chloroplasts (*lane C+*), and of the stromal (*S*) and thylakoid (*T*) fractions after lysis. *Lane T+*, protease-treated thylakoids (200 μg/ml thermolysin for 30 min on ice). *Lanes Tr*, translation products. In the case of the W/A78T import, further samples of the thylakoids were subjected to carbonate washing (Michl *et al.*, 1994) and samples were analysed of the pellet and supernatant samples (P, S). iPSII-W, intermediate-size form of PSII-W; *DG1* and *DG2* denote proteolytic degradation products.

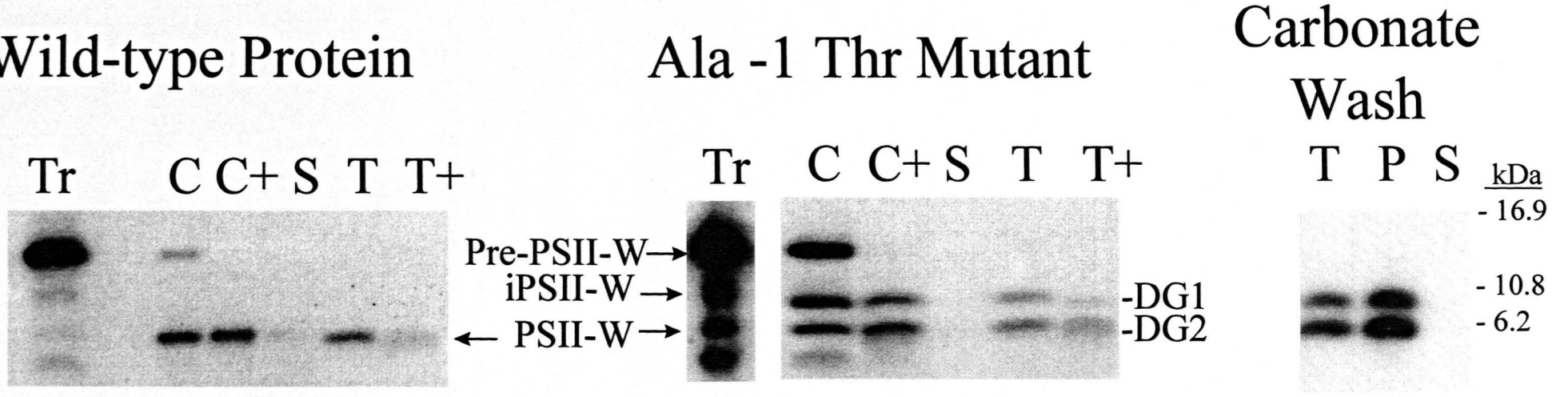
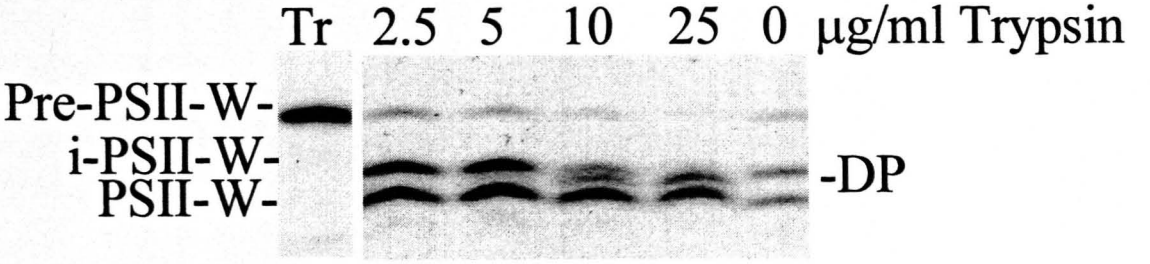

Wild-type Protein
Ala -1 Thr Mutant
Carbonate Wash
Tr
C C+ S T T+
Tr
C C+ S T T+
T P S
kDa
- 16.9
- 10.8
- 6.2
Pre-PSII-W
iPSII-W
PSII-W
-DG1
-DG2

Figure 17. *Trypsin cleaves the intermediate but not the mature pre-PSII-W.* The W\A78T mutant described in **figure 16** was imported into chloroplasts and the thylakoid fractions isolated. *A*, thylakoids were incubated with trypsin (at the concentrations indicated *above* the lanes) for 30 min on ice. *DP* denotes degradation product generated from the iPSII-W polypeptide. *B*, a diagrammatic representation of the proposed signal peptide and mature protein arrangement in the thylakoid membrane. Included on the model is the TPP cleavage site and the relative positions of the proposed trypsin cleavage sites.


Tr 2.5 5 10 25 0 µg/ml Trypsin
Pre-PSII-W-
i-PSII-W-
PSII-W-
-DP

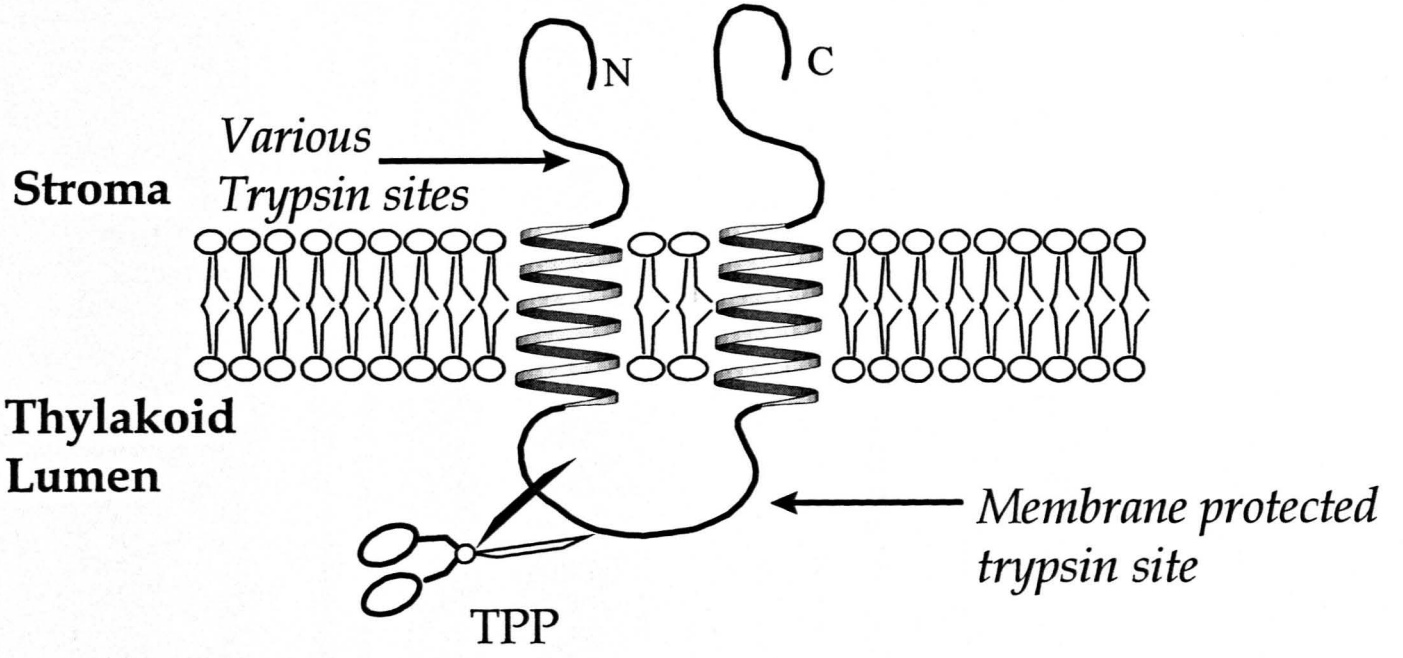
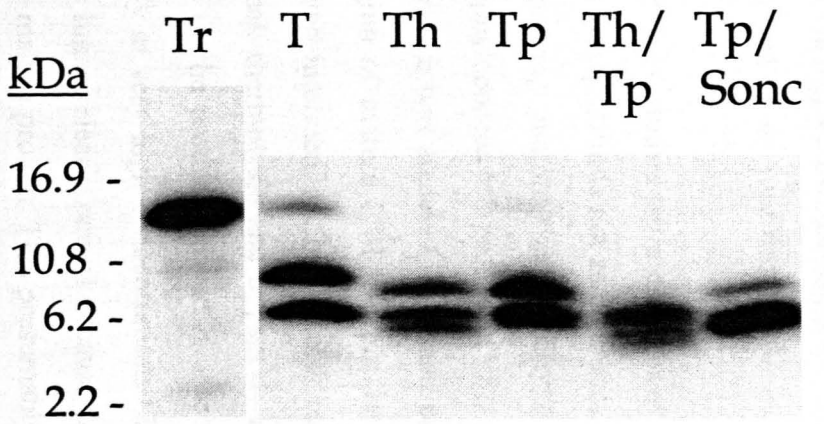
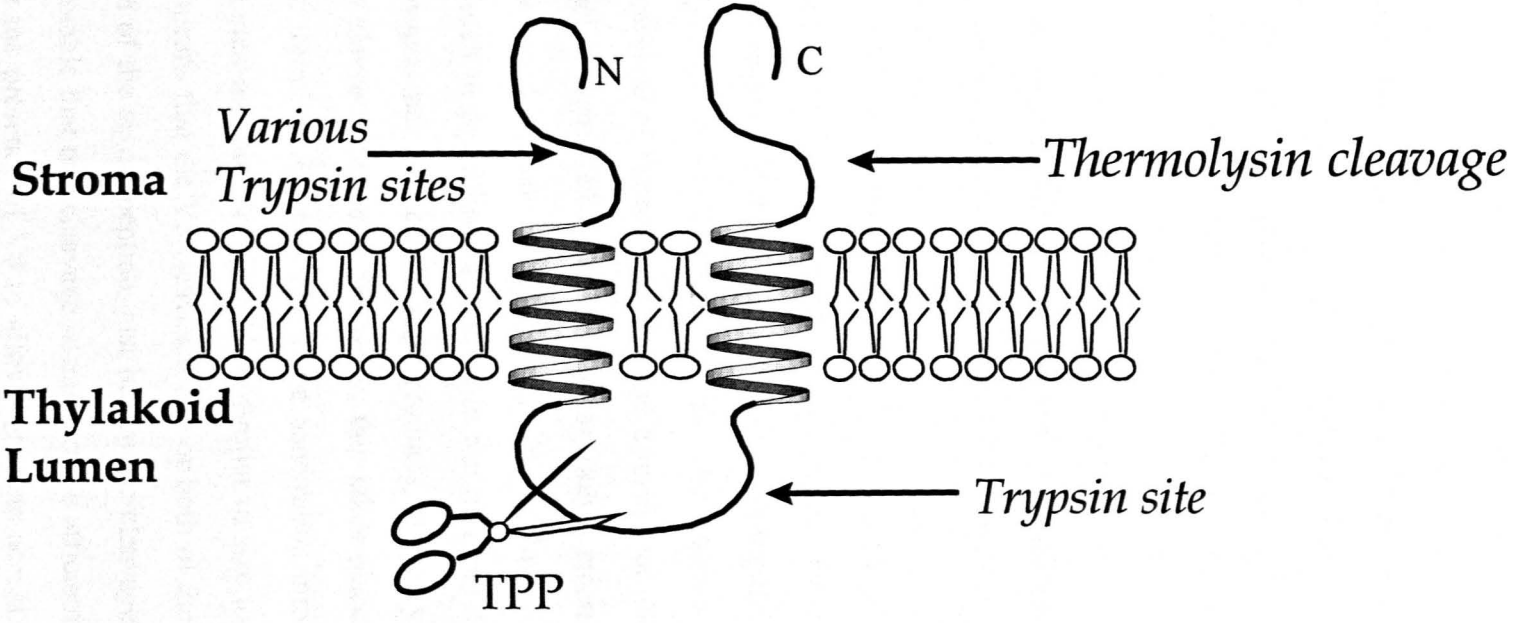

N
C
Various
Trypsin sites
Stroma
Thylakoid
Lumen
Membrane protected
trypsin site
TPP

Figure 18. *A loop intermediate in the PSII-W insertion process.* The W\A78T mutant described in **figure 16** was imported into chloroplasts and the thylakoid fractions isolated. *A*, the thylakoids were incubated with 200 μg/ml thermolysin (*Th*), 25 μg/ml trypsin (*Tp*), or 200 μg/ml thermolysin followed by 25 μg/ml trypsin (*Th/Tp*). In the latter case, the thylakoids were washed once in 1ml of HME between protease treatments. A further sample was incubated with 25 μg/ml trypsin and sonicated for 10 min (*Tp/son*). All samples were incubated for 30 min on ice. *B*, a diagrammatic representation of the proposed signal peptide and mature protein arrangement in the thylakoid membrane. Included on the model is the TPP cleavage site, the relative positions of the proposed trypsin and thermolysin cleavage sites.


Tr
T
Th
Tp
Th/ Tp
Tp/ Sonc
kDa
16.9 -
10.8 -
6.2 -
2.2 -


N
C
Various Trypsin sites
Stroma
Thermolysin cleavage
Thylakoid Lumen
Trypsin site
TPP

3.5 *Evidence favouring a similar insertion process for PSII-X*

It is still unclear whether the three thylakoid proteins currently viewed as a group (CF_0II, PSII-X, and PSII-W) actually insert by a similar mechanism, therefore the TPP cleavage site mutation approach was taken to investigate the insertion of PSII-X. The cleavage site of pre-PSII-X was similarly mutated by the introduction of a threonine at the -1 position and the effects examined as for PSII-W. **Figure 19** shows that this mutation also inhibits processing by TPP although the effects are in fact more drastic; analysis of the thylakoid fraction following an import reaction shows hardly any mature PSII-X is detectable, indicating that the action of TPP is almost totally blocked. Once again, an intermediate- size form (int) accumulates which is presumed to result from the action of stromal processing peptidase. This intermediate form is located exclusively in the thylakoid membrane showing that TPP action is not required for the insertion of PSII-X.

Topology studies were carried out on the iPSII-X polypeptide but the results show that this protein is not as amenable to analysis as PSII-W. Thylakoids containing the imported, mature-size wild-type protein were incubated with thermolysin, Staphylococcus V8 protease (V8), a mixture of thermolysin/V8, and trypsin, but no cleavage was apparent (left-hand panel of **figure 20**). It would appear, only a small region of mature PSII-X protrudes into the stroma. Unfortunately, the membrane-bound intermediate form is also resistant to proteolysis as shown in the right-hand panel; thermolysin and trypsin do not generate defined cleavage products, and only V8 of the proteases tested is able to cleave the protein. However, this takes place only at very high concentrations (200 μg/ml for 60 min) and some penetration into the lumen clearly occurs because the cleavage product (lane V) is similar in size to the mature protein. This presumably means that the V8 cleaves at one or both of the Glu residues in the C-terminal region of the signal peptide, just before the cleavage site (see **figure 13**). However, it is notable that this cleavage is much more efficient when the thylakoids are sonicated in the presence of V8 to allow complete access. Under these conditions the intermediate is quantitatively converted to the degradation product (lane Vs) providing evidence, albeit circumstantial, that this region is in the lumen. This would imply the formation of a loop intermediate as with PSII-W, but further tests will be required to confirm this point.

Figure 19. *Cleavage by TPP is not required for the insertion of PSII-X.* Pre-PSII-X containing an Ala74Thr mutation (X/A74T; *lane Tr*) was imported into intact chloroplasts. After import, samples were analysed of the total chloroplast fraction (*lane C*), protease-treated chloroplasts (*lane C+*), and the stromal (*S*) and thylakoid (*T*) fractions after lysis. In the case of the X/A74T import, further samples of thylakoids were subjected to carbonate washing and samples were analysed of the pellet and supernatant samples (P,S). *Int*, intermediate-size form of PSII-X. *Pre* denotes precursor protein.

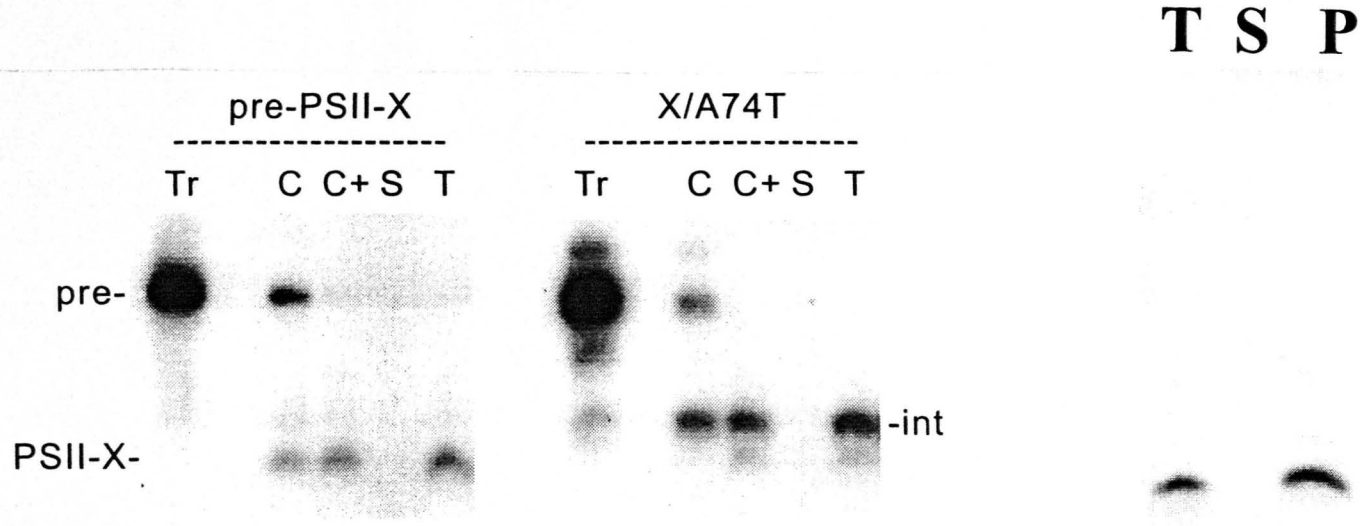
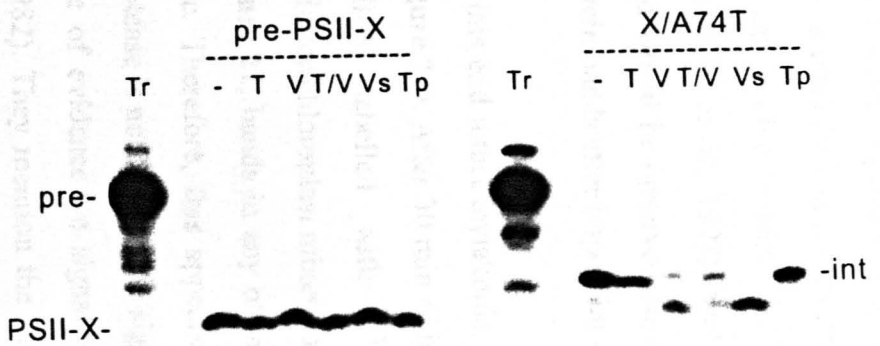
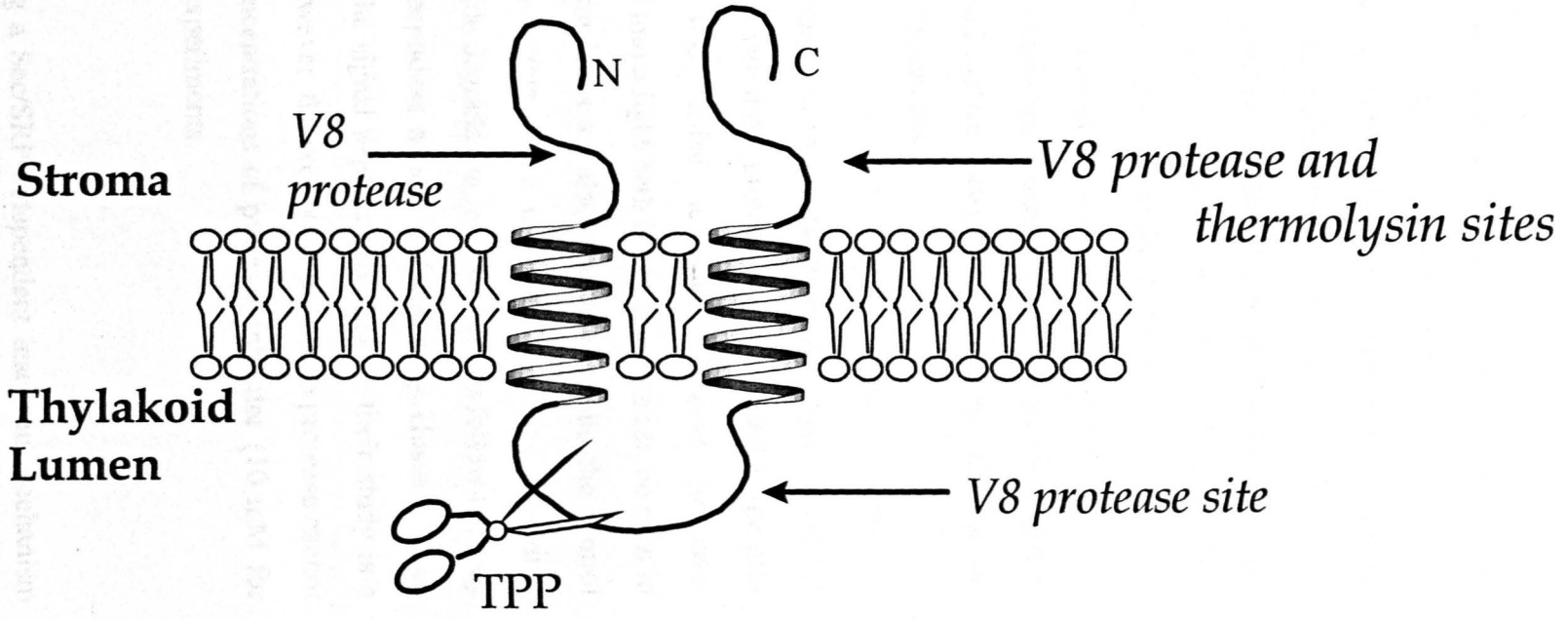

pre-PSII-X
X/A74T
Tr C C+S T
Tr C C+S T
pre-
PSII-X-
-int
T S P

Figure 20. *Topological analysis of iPSII-X.* PSII-X and X\A74T mutant were imported into chloroplasts as described in the legend to **Figure 16** and the thylakoid fractions (denoted -) incubated with proteases as follows: 200 μg/ml thermolysin (*T*), 200 μg/ml Staphylococcus V8 (*V*), thermolysin followed by V8 (*T/V*), or 25 μg/ml trypsin (*Tp*). *Lane Vs*, V8 was incubated with thylakoids and the vesicles sonicated to allow access of the protease into the lumen. All protease treatments were for 30 min on ice, except the V8 incubations that were for 60 min on ice. *Int*, intermediate form. a diagrammatic representation of the proposed signal peptide and mature protein arrangement in the thylakoid membrane. Included on the model is the TPP cleavage site, the relative positions of the proposed V8 protease and thermolysin cleavage sites.


pre-PSII-X
Tr
- T V T/V Vs Tp
X/A74T
Tr
- T V T/V Vs Tp
pre-
PSII-X-
-int


N
C
V8
protease
Stroma
V8 protease and
thermolysin sites
Thylakoid
Lumen
V8 protease site
TPP

3.6 *Signal peptide analysis with inhibitors*

Both pre-PSII-W and pre-PSII-X are unusual in that they contain methionines in their precursor form, but on cleavage of the signal peptide they lose all (PSII-X), or all but one (PSII-W), methionine residues. It is for this reason that experiments presented thus far, in this chapter, have been labelled with ^{3}H-leucine. However, this offers a unique possibility to attempt analysis of the degradation of the signal peptide. Little so far has been published in this area, specifically those signal peptides cleaved by type I signal peptidases e.g. TPP and leader peptidase in *E. coli*. Furthermore, the system used for analysis of proteins after import is a tricine SDS-PAGE system, and this can resolve proteins of 2 kDa. The signal peptide of PSII-W is approximately 4 kDa (Kim *et al.*, 1998), excluding the envelope transit peptide. Thus, this protein band would be observed on a tricine gel system between the precursor and mature protein bands after import into isolated thylakoids.

To this end intact thylakoids were incubated with several types of protease inhibitor (**figure 21**). After 30 min of incubation with these protease inhibitors, wild-type pre-PSII-W labelled with [^{35}S]-met was added to the thylakoid protease-inhibitor/chloroplast mixes and placed into a light-bath for 15 min. As can be seen in **figure 21**, bands in any of the inhibitor lanes are also contained within the control lane. Therefore, this appears to demonstrate that a cysteine-, metallo-, or serine-protease is not involved in signal peptide degradation, which does not follow the only line of evidence on signal peptide peptidase action, presented by Hussain *et al.* (1982). They mention the action of the signal peptide peptidase in their study is a either a cysteine or serine protease, however, the precise nature of the protease cannot be identified because of the high concentrations of protease inhibitor (10 mM for each protease) used throughout their experiments.

3.7 *Discussion*

The experimental evidence favouring a Sec/SRP-independent insertion mechanism (Michl *et al.*, 1994; Lorokovic *et al.*, 1995; Robinson *et al.*, 1996; Kim *et al.*, 1996; Kim *et al.*, 1998) is consistent with the unusual evolution of the signal peptides. Genes encoding PSII-X and CF_0II have been identified in cyanobacteria and also

within the plastid genomes of several eukaryotic algae/diatoms, such as *O. sinensis* and *C. paradoxa*. None of these cyanobacterial or plastid-encoded proteins is synthesised with a signal-type peptide, strongly suggesting that the signal peptides were acquired after the transfer of these genes to the nucleus in higher plants.

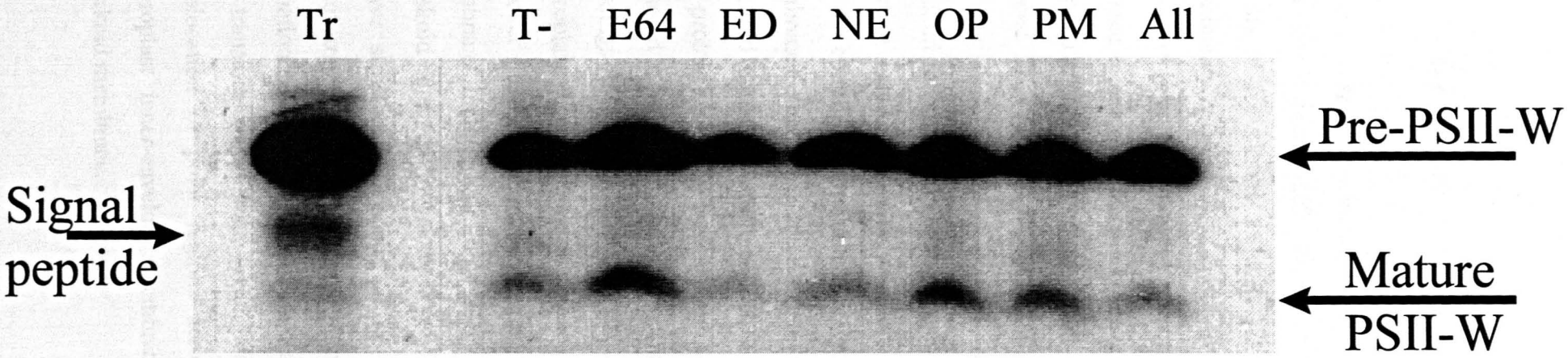


Figure 21. *Analysis of pre-PSII-W with protease inhibitors.*
Isolated thylakoids were incubated with either no protease inhibitors (*T-*), 10 mM E64 (*E64*), 10 mM EDTA (*ED*), 10 mM N-ethyl maleimide (*NE*), 10 mM 1, 10 phenanthroline (*OP*), 10 mM PMSF (*PM*), and all of these combined (*ALL*). All samples were incubated for 30 min on ice and then radiolabelled pre-PSII-W was added and a thylakoid import performed as in Brock *et al*. (1986). After import, samples were washed in HM containing the appropriate protease inhibitor and samples were then placed on to dry ice, prior to SDS-PAGE analysis. Denoted by *signal peptide* is the proposed size of the envelope transit peptide and thylakoid signal peptide (without separation by SPP) after cleavage by TPP.

3.7.1 *So how do proteins on the spontaneous insertion pathway insert into the appropriate membrane?*

The hypothesis that TPP may have a receptor-like role in the insertion mechanism has been shown to be incorrect. Therefore, what dictates the specific insertion into the thylakoid membrane? There are two plausible explanations: (1) Internal signals - analysis of the mature proteins of CF_0II, PSII-W and PSII-X, shows the only commonality between these mature proteins is their inherent hydrophobic character; it is therefore unlikely that the mature proteins contain a signal for insertion. Further evidence comes from Michl *et al.* (in press), in this paper they compare the integral thylakoid membrane proteins CF_0I and CF_0II, the former protein is chloroplast-encoded. They found that addition of either an envelope transit peptide or a bipartite targeting signal to CF_0I, allowed the correct insertion of this protein into the thylakoid membrane. However, in the case of CF_0II the removal of the thylakoid targeting signal, but not the envelope transit peptide, allowed import of the protein into the chloroplast stroma but the protein was unable to insert into the thylakoid membrane. What is perhaps more interesting was the fact that they added the CF_0II thylakoid signal peptide to several lumenal proteins (16K, 23K and 33K), and the only protein capable of translocation across the thylakoid membrane was 33K, which is known to follow the Sec-type mechanism. (2) Lipid composition of the membrane - findings reflect differences in lipid compositions of the various membranes of the chloroplast. The thylakoid membrane is most unusual in that galactolipids account for over 80% of membrane lipids, whereas phospholipids predominate in most membranes including those of *E. coli*. It has been proposed that one of the major thylakoid lipids, monogalactosyldiacylglycerol, has a pronounced ability to form non-bilayer structures (Gounaris *et al.*, 1983). Furthermore, Killian *et al.* (1990) demonstrated that the functional signal peptide from M13 procoat can induce the formation of these non-bilayer lipid structures, and this appears to be related to *in vivo* translocation efficiency. And thus, it is possible that this property may facilitate translocation of polar regions across the thylakoid membrane. However, the chloroplast inner-envelope membrane has a similar lipid composition as the thylakoid membrane.

3.7.2 *Conformation of a loop intermediate.*

The suggestion that the thylakoid transit peptide evolved after the transfer of the genes encoding these proteins to the nucleus contrasts starkly with "genuine" Sec-type signal peptides such as those found in the precursors of lumenal proteins, because the cyanobacterial/plastid-encoded counterparts of these proteins are invariably synthesised with 'classical' signal peptides. It is therefore believed that the signal peptides of these membrane proteins differ from Sec-type signal peptides in terms of both function and origin. However, there appear no large-scale structural differences among the mature proteins that might explain why nuclear-encoded, imported proteins are now synthesised with signal peptides. Mature PSII-X proteins, for example, are structurally very similar in higher plants and cyanobacteria (see **figure 14**). Given these similarities, it may be simplistic to suggest that the signal peptide functions simply by providing a critical, additional hydrophobic segment, because the mature proteins in cyanobacteria and the above mentioned algae can clearly insert despite possessing only a single hydrophobic region. The only obvious difference concerns the charge distributions in the N-terminal regions; these regions are neutral in the cyanobacterial and plastid-encoded proteins whereas that of *Arabidopsis* PSII-X is basic due to the lysine at position 10. This could conceivably render the translocation process more difficult to the extent that the additional hydrophobic region is required to aid insertion. However, it is also possible that this unusual insertion mechanism has been forced on these proteins as a consequence of their more complex translocation/insertion pathway; the proteins may insert co-translationally in cyanobacteria, in which case the protein may never be free in the aqueous phase and insertion may be more favourable. Another possibility, is that the signal peptide renders the precursor protein more stable in solution, perhaps by forming a "helical-hairpin" in which the signal peptide partially masks the hydrophobic character of the transmembrane segment. Structural studies on this type of precursor protein should prove instrumental in resolving these points.

CF_0II, PSII-W, and PSII-X represent an unusual group of membrane proteins in that no other membrane protein, with the exception of M13 procoat, has been found to rely on a cleavable signal peptide for membrane insertion by a Sec- and SRP-independent mechanism. As shown in **figure 14**, the three thylakoid proteins share

additional common features with procoat: each contains a single transmembrane span in the mature protein and the translocated regions are all negatively charged. It was thus unsurprising that initial models for the Sec-independent insertion of CF_0II (Michl *et al.*, 1994) were based on the procoat insertion model, and this study has confirmed that in one fundamental respect the insertion of PSII-W does indeed resemble that of procoat. Both proteins adopt a loop intermediate during the insertion process, strongly suggesting that the function of the PSII-W signal peptide is to provide a second hydrophobic domain which, in concert with the transmembrane segment in the mature protein, is then able to drive the translocation of the hydrophilic intervening region into the thylakoid lumen. Interestingly, signal-type peptides appear also to form loop structures when directing soluble proteins across membranes by means of protein translocation systems (Kuhn *et al.*, 1994; Fincher *et al.*, 1998).

Although the insertion of these thylakoid proteins is reminiscent of M13 procoat insertion, there are equally significant differences in the mechanisms used. The first concerns the initial events in the insertion process. Gallusser and Kuhn (1990) have shown that electrostatic interactions play a critical role in the insertion of procoat, in which basic residues in both the extreme N- and C-termini bind to the negatively charged membrane surface; removal of either set of basic residues renders procoat wholly insertion incompetent. Like procoat, pre-CF_0II contains positively charged N- and C-termini and a similar mechanism was considered possible for this protein (Michl *et al.*, 1994). However, the C-terminal region of pre-PSII-W is extremely negatively charged (a series of five acidic residues follows the transmembrane section), precluding any electrostatic binding between this region and the membrane surface. Further tests are required to determine whether the basic N-terminal region plays any role through an electrostatic binding to the thylakoid surface.

A second difference concerns the energetics of insertion. Most Sec-dependent bacterial proteins are at least partially dependent on the transmembrane potential for efficient insertion, but the best characterised Sec-independent proteins (M13 procoat and Pf3 coat) are entirely dependent on the $\Delta\mu^{H+}$ (Kuhn *et al.*, 1990). Nevertheless, Thompson *et al.* (1998) shows that the two systems have different tolerances and it

remains to be determined how such superficially similar translocated regions can be so potentially dependent on $\Delta\mu^{H+}$ in *E. coli*, but not in thylakoids.

3.7.3 *The fate of the signal peptide*

As shown in **figure 21**, cysteine-, serine-, and metallo-protease inhibitors have no apparent effect on the degradation of the signal peptide once it has been cleaved from the mature protein, even in the high concentrations (10 mM in all cases) presented. In *E. coli*, cleaved signal peptides are digested by signal peptide peptidases (Hussain *et al.*, 1982). This degradation is thought to be initiated by membrane-bound protease IV, which cleaves the peptide within the hydrophobic core (Novak and Dev, 1988). The resulting fragments can be released into the cytosol and further hydrolysed by Oligopeptidase A (Novak and Dev, 1988). Both signal peptide peptidases are endopeptidases and account for the majority of signal peptide degrading activity *in vitro* (Novak *et al.*, 1986).

Chapter 4

Generation and use of thylakoidal processing peptidase cleavage mutations in PsbY

Chapter 4

Results 2 - Generation and use of thylakoidal processing peptidase cleavage mutations in PsbY

4.1 *Introduction*

In the past years great progress has been made in understanding the mechanisms that directly target proteins to the thylakoid and especially the lumen. Multiple pathways are clearly distinguishable by different requirements for soluble factors, ATP and a ΔpH gradient across the thylakoid membrane, and the involvement of protease-sensitive components at the thylakoid surface. There are currently, four independent pathways that allow either insertion into or translocation across the thylakoid membrane: (1) a SecA-dependent pathway; (2) a ΔpH-dependent pathway; (3) a signal recognition particle (SRP)-dependent pathway; and (4) a spontaneous insertion mechanism. The details of each pathway have been reviewed recently by Robinson *et al.* (1998).

Although the majority of proteins targeted to the thylakoid follow one of the first three pathways mentioned above, several reports have shown that certain thylakoid membrane proteins integrate in the presence of azide (a specific inhibitor of SecA) and nigericin (an ionophore inhibiting ΔpH-dependent transport). These observations demonstrate that the mechanism of integration does not require ATP hydrolysis or a pH gradient as an energy source. Furthermore, proteolytic pre-treatment of the thylakoid fails to block import, suggesting that the process does not require protein components in or on the thylakoid membrane. So far, three membrane proteins have been shown to integrate into the thylakoid under such conditions: subunit II of CF_0 (CF_0II; Michl *et al.*, 1994), and two subunits of photosystem II W and X (PSII-W and PSII-X).

To date, cleavable signal peptides have been shown to assist in the SRP-/Sec-independent insertion of three proteins - CF_0II, PSII-W and PSII-X. All three are

simple single span proteins raising the possibility that this type of apparently spontaneous insertion mechanism is used only by such simple proteins. In this chapter, I have analysed the insertion and maturation of an unusual thylakoid membrane protein, PsbY. Mant and Robinson (1998) characterised an unusual *Arabidopsis thaliana* cDNA encoding a protein containing two separate regions bearing high homology to *ycf32* open reading frames encoded by several algal plastid genomes. Whereas *ycf32* genes encode small single-span proteins, the *Arabidopsis* protein was predicted to contain four hydrophobic regions and it was proposed that this was in effect a polyprotein of two separate Ycf32-related proteins, each of which was preceded by a signal-type peptide. It has now been shown that the two proteins are indeed found in thylakoids, associated with photosystem II, and the gene has been designated *psbY* (Gau *et al.*, in press). The individual polypeptides are designated PsbY-A1 and -A2. Interestingly, the two separate regions bearing high homology to Ycf32 proteins are preceded by additional hydrophobic regions that resemble signal peptides (see below). It was therefore considered possible that PsbY may uniquely insert through the use of two separate cleavable signal peptides and I have investigated this possibility through mutagenesis of potential TPP cleavage sites. The aims were two-fold: To determine whether PsbY is indeed synthesised as a polyprotein and to unravel the insertion/maturation pathway.

4.2 *Mutagenesis and sequencing of pre-PsbY*

The cDNA template was supplied by Dr. A. Mant (Warwick) and consisted of pre-PsbY cloned into pGem-5Z. The method of mutagenesis used was the iPCR method (see **chapter 2**). The PCR products were run on 1% agarose gels, and the bands of appropriate size (in comparison to the construct restricted at a single site) were gel excised and purified. These purified bands were ligated overnight and transformed into competent *E. coli*. Colonies were taken at random and grown overnight in LB broth. Plasmid DNA was extracted from these cultures and sequenced to determine the presence of the correct mutation. On finding the mutation in the plasmid, the total mutant cDNA was fully sequenced.

4.3 *Results - PsbY contains four hydrophobic regions and two regions of homology to algal Ycf32*

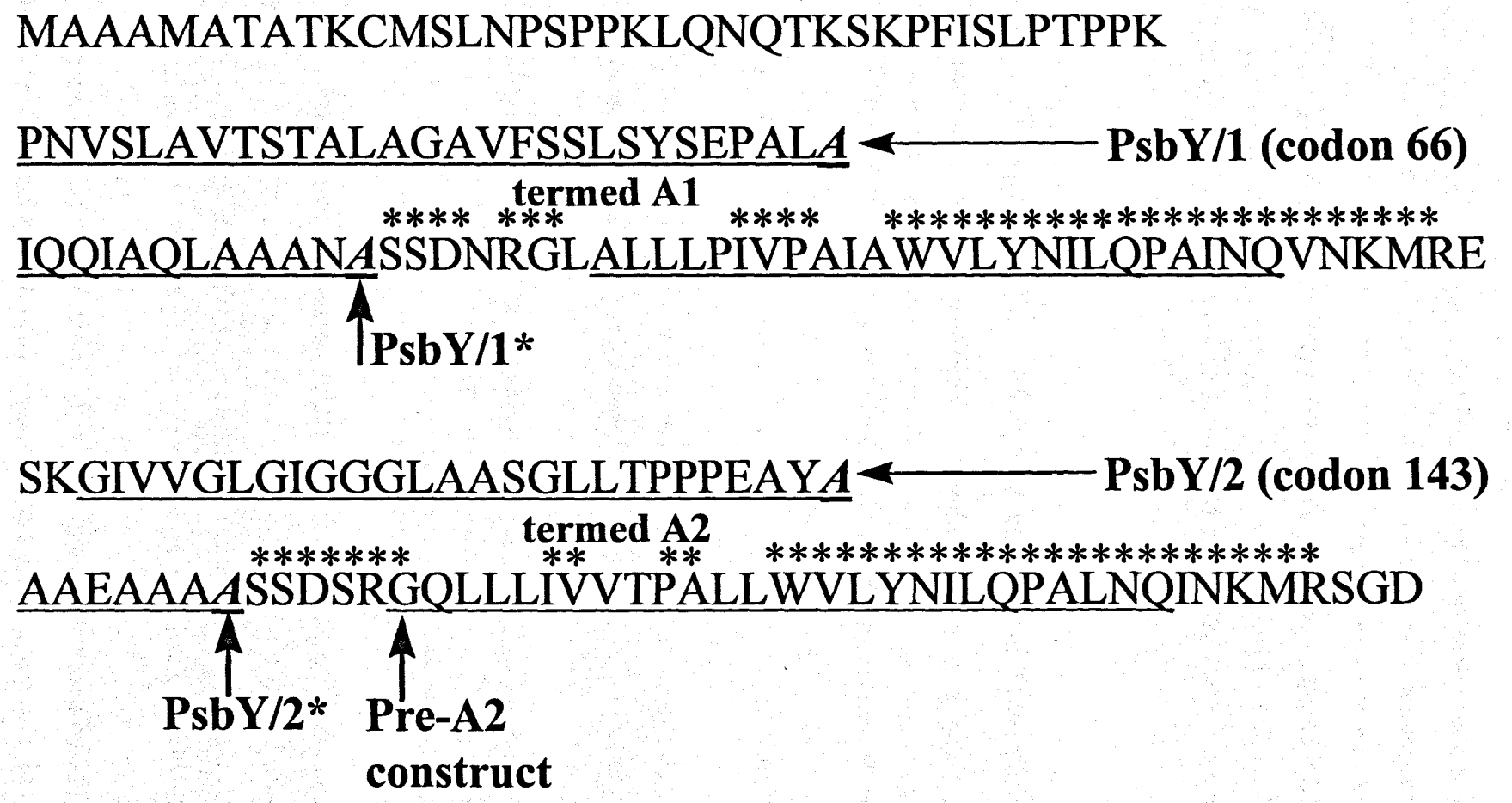
The initial aim of the study was to determine whether the PsbY gene product is indeed processed to two small single-span proteins through the use of dual signal peptides as proposed by Mant and Robinson (1998). **Figure 22** shows the overall primary structure of the full precursor protein (which is termed pre-PsbY) encoded by the *Arabidopsis* PsbY cDNA protein, in which the structure is divided into five domains. The predicted sequence reveals an apparently typical stroma-targeting envelope transit peptide that is basic, hydrophilic and enriched in hydroxylated residues. This is followed by the initial 'mature' protein (PsbY) which contains four hydrophobic regions : (i) a predicted signal peptide, (ii) a region that is closely homologous to algal ycf32 open reading frames (protein A1), (iii) a second possible signal peptide, and (iv) a second region homologous to single-span Ycf32 proteins (protein A2). Thylakoid signal peptides are cleaved by a membrane-bound, lumen-facing TPP activity which cleaves after short chain residues at the -3 and -1 positions in the substrate relative to the processing site, and the presence of alanine at -1 is essential for efficient cleavage (Shackleton and Robinson, 1991). Alanine is also usually found at the -3 position. Hydrophobic regions (i) and (iii) above were considered to be possible signal peptides on the basis that these regions are followed by potential Ala-X-Ala TPP cleavage sites.

4.4 *The PsbY translation product contains two signal peptides*

The sequence information strongly suggests that pre-PsbY contains a total of four hydrophobic regions including two cleavable signal peptides. Because TPP is known to be active on the lumenal side of the thylakoid membrane, this protein offers attractive possibilities in terms of identifying the topology of the polypeptide chain during membrane insertion. In order to address this topic more directly advantage was taken of the highly precise nature of the TPP reaction , substituting threonine residues at possible -1 positions. This markedly inhibits the TPP processing reaction (Shackleton and Robinson, 1991; Thompson *et al.*, 1998) and it would also lead to the identification of defined intermediates on the PsbY biogenesis pathway. Any observed inhibition would furthermore help to define these cleavable peptides as

Figure 22. *Primary structure and proposed domain arrangement organisation of pre-PsbY.*

The figure shows the full predicted sequence of Arabidopsis pre-PsbY (previously designated Ycf32 by Mant and Robinson, 1998). The N-terminal region contains an envelope transit peptide that is believed to be removed after import by the stromal processing peptidase (note that this processing site has not been identified). Homology with the Ycf32 open reading frame from plastid genome of Porphyra purpurea is indicated with * (from Mant and Robinson, 1998). Proteins A1 and A2 are indicated (see text), each of which is preceded by an apparent signal peptide. The exact position of the junction between protein A1 and signal peptide A2 has not been determined. Hydrophobic regions in the signal peptides and mature proteins are underlined. Candidate -1 alanine residues of TPP cleavage sites are shown bold, italicised and underlined; these were altered to Thr residues in four single mutants whose designations are shown. Codon 66, indicated with an arrow and in bold was also changed to leucine. Construct pre-A2 was synthesised by amplification of the coding region for the C-terminal region of PsbY and the introduction of a start codon in place of the glycine indicated by an arrow (this construct was kindly provided Dr. A. Mant, Warwick).


MAAAMATATKCMSLNPSPPKLQNQTKSKPFISLPTPPK
PNVSLAVTSTALAGAVFSSLSYSEPALA
PsbY/1 (codon 66)
termed A1
IQQIAQLAAANASSDNRGLALLLPIVPAIAWVLYNILQPAINQVNKMRE
PsbY/1*
SKGIVVGLGIGGGLAASGLLTPPPEAYA
PsbY/2 (codon 143)
termed A2
AAEAAAASSDSRGQLLLIVVTPALLWVLYNILQPALNQINKMRSGD
PsbY/2*
Pre-A2
construct

substrates for TPP since only minor, single substitutions were made at each site and identical substitutions had no effect on the insertion of loop structures in pre-PSII-W or pre-PSII-X (**Chapter 3**). Both terminal alanine residues were therefore altered to Thr by site-specific mutagenesis of the cDNA clone (the relevant residues are shown italicised, bold and underlined in **figure 22**). These mutants were designated PsbY/1 and PsbY/2 according to whether the first or second cleavage site was affected. Two further potential TPP cleavage sites were also identified and mutated; these mutants were designated PsbY/1* and PsbY/2* as shown in **figure 22**.

Analyses of the four individual mutants are shown in **figures 23-24**. All of the mutant proteins are imported by isolated chloroplasts and found exclusively in the thylakoid fraction, lane T. The import and processing profiles of mutants PsbY/1* and PsbY/2* (**figure 23**) were found to be exactly as those found for the wild-type protein (**figure 23**) giving a persuasive argument that these mutations do not lie at TPP cleavage sites. However, the remaining mutants exhibit severe defects in processing. PsbY/2 is imported and converted to a 10.8 kDa product as shown in **figure 24**. Clearly, the threonine residue prevents cleavage at the second site by TPP and a larger polypeptide accumulates. There is, however, good evidence that the first signal peptide has been cleaved because the mobility of the protein is consistent with a three-span protein rather than a protein containing all four hydrophobic regions (see below).

PsbY/1 is imported and converted to two polypeptides, one of apparently mature size (7 kDa) together with a processing intermediate (denoted 'int') which migrates as 10.2 kDa. The 7 kDa protein co-migrates with protein A2 from chloroplast import assays with the wild-type protein and therefore the assumption is that the A2 protein is correctly removed from the PsbY/1 polyprotein upon insertion. Almost no A1 is formed, however, indicating that the presence of the Thr residue has a drastic effect on the release of this protein and providing very strong evidence that the first hydrophobic region is in fact a cleavable signal peptide that is recognised by TPP. A larger polypeptide (indicated by an arrow) is also apparent in this experiment but the identity of this band is unclear since it is usually present in very low quantities and is

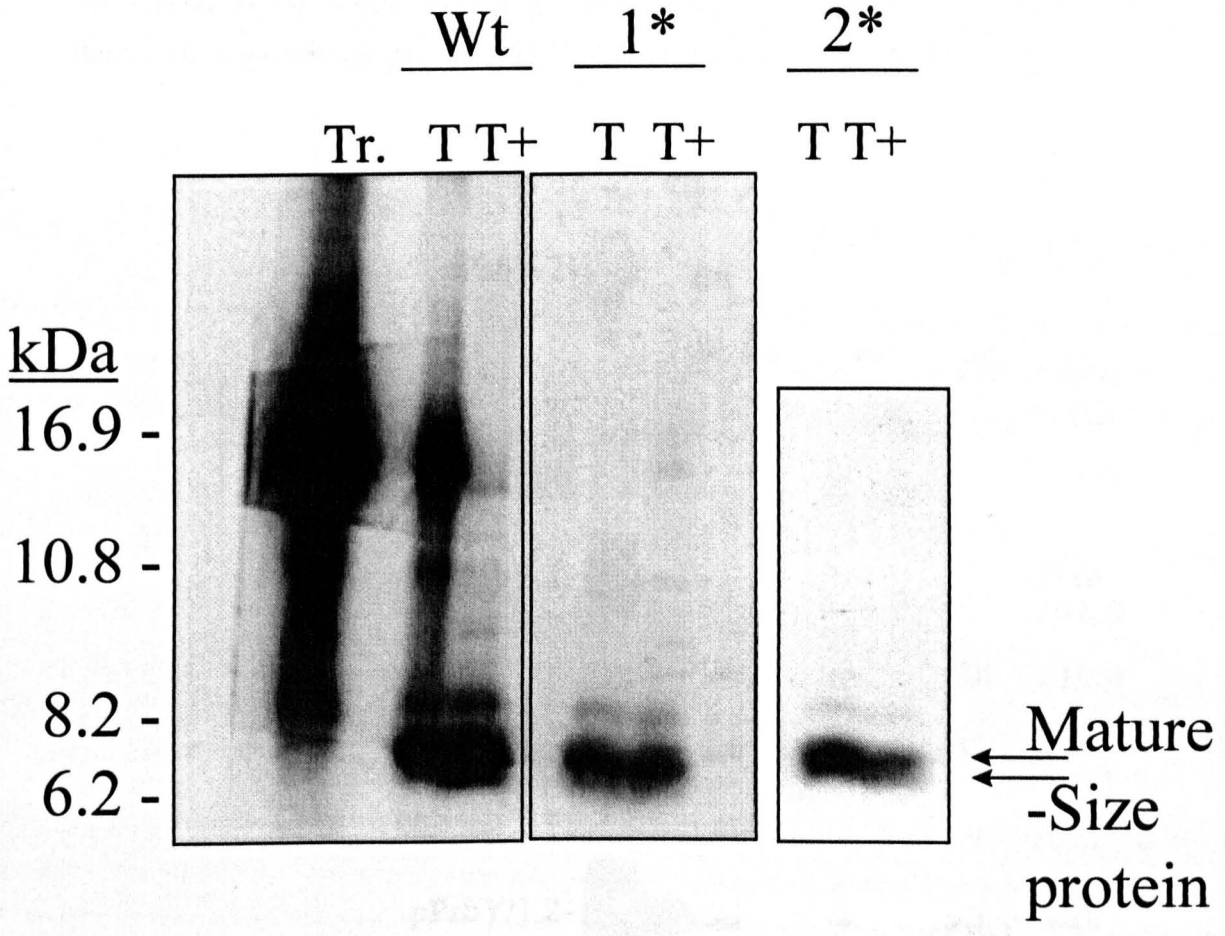
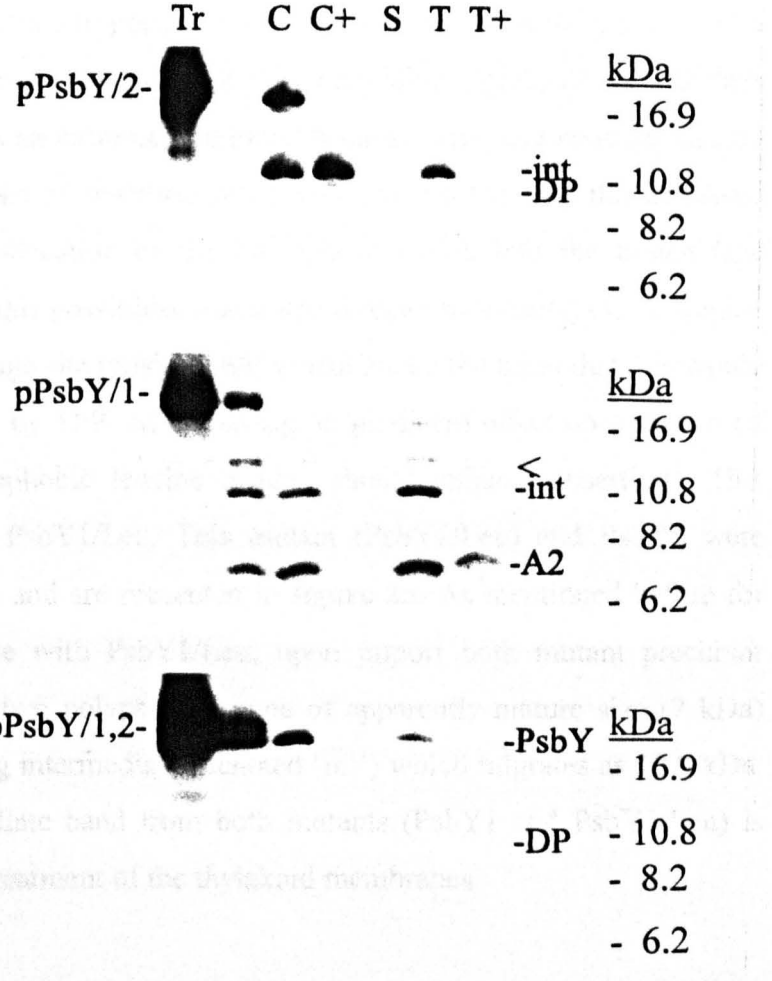


Figure 23. *A comparison of Mutant PsbY/1*, PsbY/2* and wild-type PsbY.* Wild-type PsbY, mutant PsbY/1* and mutant PsbY/2* were imported into isolated, intact chloroplasts and the thylakoid fraction isolated (T). A fraction of these thylakoids was further treated with thermolysin (lane T+).

Figure 24. *Accumulation of processing intermediates in the Ala>Thr mutants.* Mutants PsbY/1, PsbY/2 and the PsbY/1, 2 double mutant (lanes Tr) were imported into intact chloroplasts and samples analysed of the chloroplasts (lanes C) and thermolysin-treated chloroplasts (lanes C+). Other aliquots of chloroplasts were pelleted after protease treatment and lysed, after which centrifugation yielded samples of stroma (S) and thylakoids (T). Lane T+: thermolysin-treated thylakoids. Processing intermediates (int) and degradation products (DP) are denoted. Arrow denotes larger cleavage product of unknown significance (see text).



sometimes virtually absent. In contrast, the 10.2 kDa and 7 kDa proteins are always observed in similar quantities.

4.5 *Substitution of a threonine residue in PsbY1 does not alter membrane stability in thylakoid membranes*

The PsbY/1 mutant shows some interesting characteristics and the appearance of the larger intermediate suggests that TPP cleavage has been inhibited at the first site. However, after thermolysin treatment of the thylakoid fractions the band is degraded and therefore, it was considered important to test the possibility that the presence of a Thr residue might affect the translocation of the intervening regions as well as their removal by TPP. This was an extremely unlikely because threonine residues had no detectable effect on the rate of insertion of PSII-W and PSII-X into the thylakoid membrane or on the translocation of the hydrophilic region into the lumen (see **chapter 3**). Nevertheless, this possibility was tested directly by altering the -1 alanine at the first predicted cleavage site (position 66) to leucine on the basis that this would similarly prevent cleavage by TPP, while having no predicted effect on insertion (if anything, the more hydrophobic leucine residue should enhance insertion). This mutation was designated PsbY1/Leu. This mutant (PsbY1/Leu) and PsbY1 were imported into chloroplasts and are presented in **figure 25**. As mentioned before for the PsbY/1, and this time with PsbY1/Leu, upon import both mutant precursor proteins are converted to two polypeptides, one of apparently mature size (7 kDa) and an apparent processing intermediate (denoted 'int') which migrates as 10.2 kDa. Furthermore, the intermediate band from both mutants (PsbY1 and PsbY1/Leu) is degraded by thermolysin treatment of the thylakoid membranes

4.6 *Inhibition of single TPP sites in PsbY gives intermediate forms that contain three hydrophobic regions*

A further mutant was constructed in which both sites were disrupted (PsbY/1,2). The chloroplast import experiment using the PsbY/1,2 double mutant is shown in **figure 24**. This protein is imported and cleaved to a larger product that migrates as a 17.5 kDa protein, and which is thus only slightly smaller than the full precursor protein.

Therefore indicating the probable removal of the envelope transit peptide (see below).

Thylakoid samples from chloroplast imports of the three processing mutants are shown in **figure 26**. Panel *A* shows a comparison of all the intermediate forms together with appropriate marker proteins. The result confirms that the imported PsbY/1,2 polypeptide is significantly larger than the intermediates generated during import of either PsbY/1 or PsbY/2, and that the latter intermediates are significantly larger than the pre-A2 translation product containing two hydrophobic regions. This result strongly suggests that the two single mutants are imported and processed to polypeptides containing three hydrophobic regions. The imported PsbY/1,2 mutant is only marginally smaller than the full precursor and clearly contains all four hydrophobic regions. **Figure 26** also confirms the point made in **section 4.4**, namely that the A2 protein is cleaved from mutant PsbY/1, since the smaller import product in lane '1' co-migrates precisely with the A2 protein generated in a thylakoid assay (adjacent lane T). Therefore, neither of the single mutations prevents cleavage at the unmutated TPP site. The only difference in the import profiles of PsbY/1 and PsbY/2 is that a lower molecular mass cleaved product is visible in the former but not in the latter chloroplast import. This reflects the nature of the smaller cleaved species. Protein A2 is released from PsbY/1 and this protein is stable under these conditions, whereas with PsbY/2 cleavage at the first processing site leads to release of the signal peptide. Possible loss of either methionine residues and hence labelling under these assay conditions has been shown not to be responsible because labelling with ^{3}H-leucine still leaves the signal peptide undetectable (data not shown). After cleavage, these polypeptides are clearly turned over very rapidly and attempts to visualise them have failed to date (see **section 3.6**). Similar findings have been made with the single-span proteins PSII-W and PSII-X, which are also synthesised with cleavable signal peptides. Both proteins insert into thylakoids yet the cleaved signal peptides are completely undetectable despite being almost as large as the mature proteins (Kim *et al.*, 1998).

Figure 25. *Accumulation of similar size processing intermediates with the Ala>Thr and Ala>Leu mutants.*

Mutants PsbY/1, and PsbY/1 Leu (lanes Tr) were imported into intact chloroplasts and samples analysed of the chloroplasts (lanes C) and thermolysin-treated chloroplasts (lanes C+). Other aliquots of chloroplasts were pelleted after protease treatment and lysed, after which centrifugation yielded samples of stroma (S) and thylakoids (T). Lane T+: thermolysin-treated thylakoids. Processing intermediates (int) are denoted. Arrow denotes larger cleavage product of unknown significance (see text).

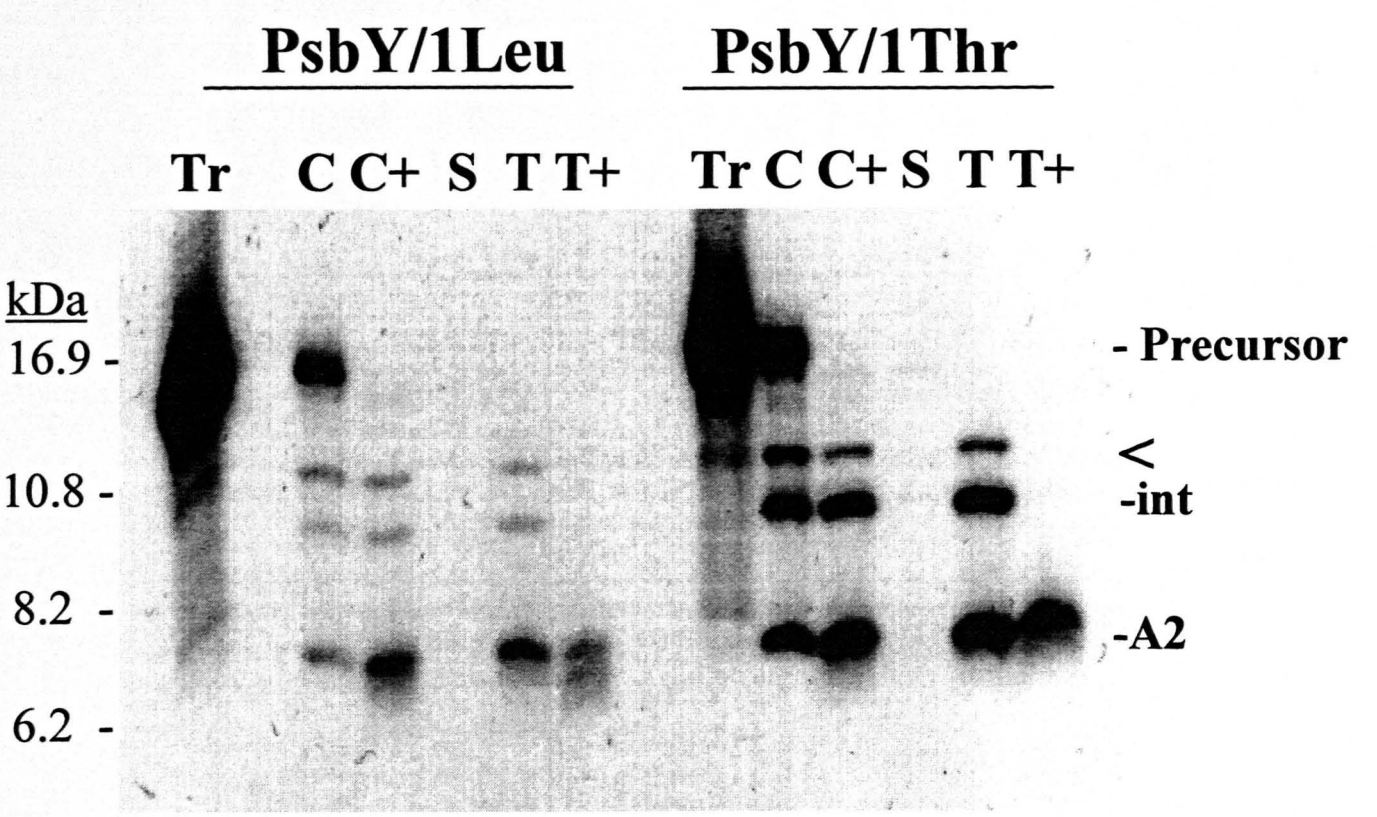
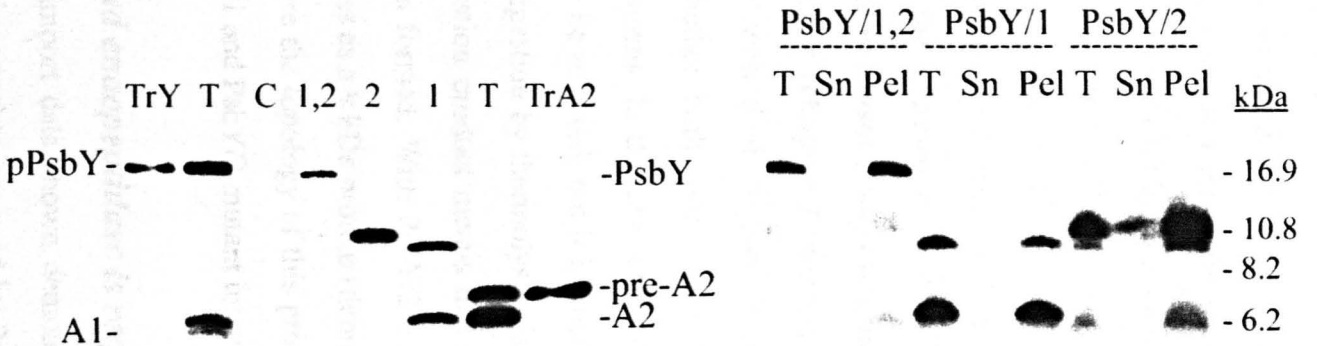

PsbY/1Leu
PsbY/1Thr
Tr C C+ S T T+
Tr C C+ S T T+
kDa
16.9 -
10.8 -
8.2 -
6.2 -
- Precursor
<
-int
-A2

Figure 26. *PsbY processing intermediates contain three hydrophobic regions and are stably inserted into the thylakoid membrane. A*: size comparison of the imported proteins. The figure shows the mobilities of the PsbY/1, PsbY/2 and PsbY/1, 2 import products (lanes 1, 2 and 1, 2 respectively) together with a sample from a thylakoid import of pre-A2 and sample of pre-A2 translation products (lanes T and TrA2 on the right of panel) and a sample of the thylakoids from thylakoid- and chloroplast-imports of pPsbY (lanes T and C respectively, on left of panel). Mobilities of molecular mass markers are indicated on right hand side. *B*: the three mutants described in *A* were imported into chloroplasts and the thylakoid fraction prepared after lysis (lanes T). Samples of the membranes were then washed with urea and samples analysed of the supernatant fraction (Sn) and pellet fraction (Pel) containing the membranes.


PsbY/1,2
PsbY/1
PsbY/2
TrY T C 1,2 2 1 T TrA2
T Sn Pel T Sn Pel T Sn Pel
kDa
pPsbY-
-PsbY
-pre-A2
-A2
A1-
- 16.9
- 10.8
- 8.2
- 6.2

4.7 *PsbY inserts into thylakoids as a double loop structure*

All three of the mutant forms are stably inserted into the thylakoid membrane because each is resistant to extraction by urea washing. This procedure effectively removes extrinsic membrane proteins from thylakoids (Breyton *et al.*, 1994; Mant and Robinson, 1998) and **figure 26 panel *B*** shows that each of the intermediates is almost completely resistant to this extraction procedure. In all cases, the protein from the thylakoid fraction of a chloroplast import experiment (lanes T) is almost totally recovered in the pellet fraction containing the urea-washed membranes (lanes Pel) and very little protein is recovered in the supernatants (Sn). However, the precise topologies of the intermediates are difficult to determine. Studies on PSII-W have shown that the precursor protein inserts as a loop intermediate prior to cleavage by TPP in the lumen (**chapter 3**) and the two signal peptides in PsbY probably form similar loops with their cognate 'mature' proteins. It is, however, difficult to determine whether both loops have formed in the inserted PsbY/1 or PsbY/2 intermediate forms. In the case of PsbY/1, the second loop must have formed for protein A2 to be released, but it is notable that the remaining intermediate is highly sensitive to digestion by thermolysin (lane T+ in **figure 24**). The lack of any defined protease digestion product means that there can be no determination that the first loop has been formed. With PsbY/2 a thermolysin degradation product is apparent which migrates as a 9 kDa protein (denoted DP in **figure 24**), but further analyses are required before the topology of this protein can be determined. Hypothetical models for the PsbY/1 and PsbY/2 mutant intermediates are presented in **figures 27-28**.

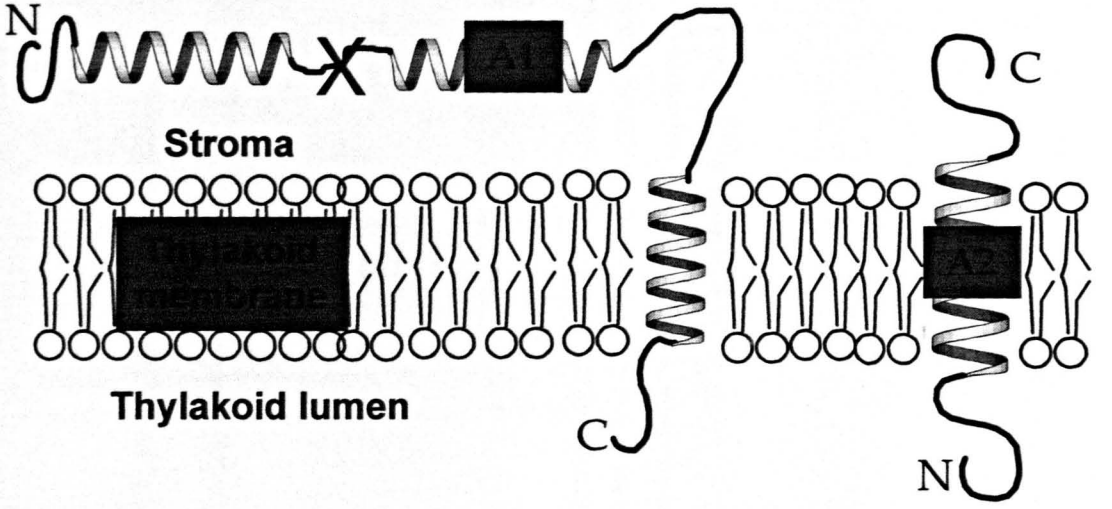
4.8 *A second endopeptidase is required for final maturation of PsbY*

Finally, the import data shown demonstrate another important implication for the overall insertion mechanism used by PsbY. Hydrophobic regions (i) and (iii) can be designated as cleavable signal peptides that are processed upon reaching the thylakoid lumen, and these peptides must therefore form loop structures with their partner A1 or A2 proteins. This means that an additional cleavage event must take place, between the A1 protein and the second signal peptide, and this event must

Figure 27. *A hypothetical topology model of the intermediates produced by the mutation PsbY/1.*

A: A model of the imported intermediates from PsbY/1 in the thylakoid membrane. *B*: Size comparison of the imported proteins. Wild-type pre-PsbY and pre-PsbY/1 were imported into chloroplasts. The figure shows the mobilities of the wild-type and PsbY/1 import products from thylakoid fractions of the chloroplast import assays. For both wild-type and mutant proteins untreated (T-) and thermolysin-treated (T+) thylakoids are presented.

X - Position of TPP mutation



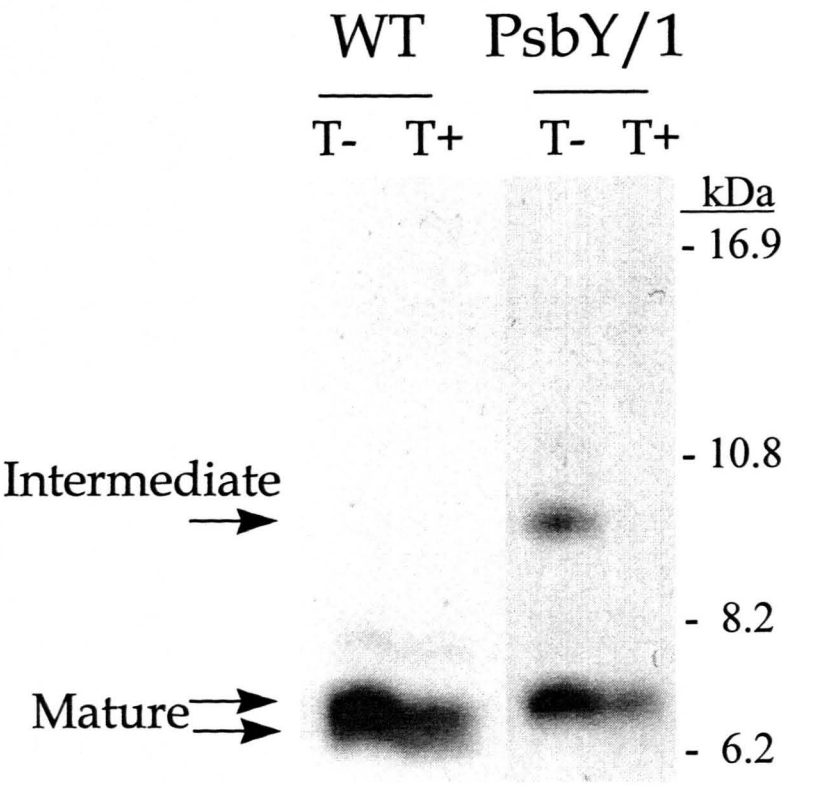
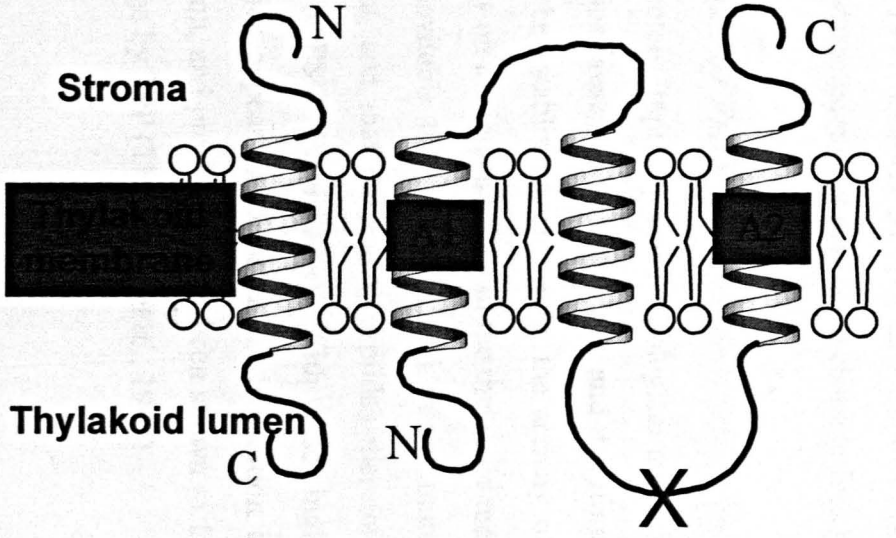
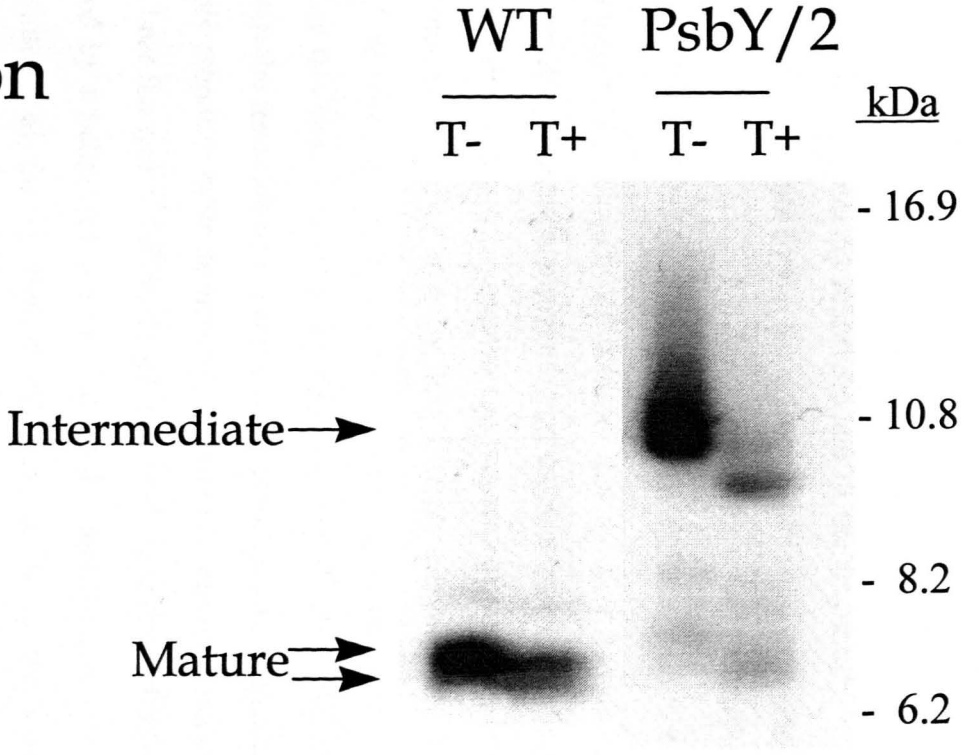


Figure 28. *A hypothetical topology model of the intermediates produced by the mutation PsbY/2.*

A: A model of the imported intermediates from PsbY/2 in the thylakoid membrane. *B*: Size comparison of the imported proteins. Size comparison of the imported proteins. Wild-type pre-PsbY and pre-PsbY/2 were imported into chloroplasts. The figure shows the mobilities of the wild-type and PsbY/2 import products from thylakoid fractions of the chloroplast import assays. For both wild-type and mutant proteins untreated (T-) and thermolysin-treated (T+) thylakoids are presented.

X - Position of TPP mutation





furthermore take place on the stromal side of the membrane (see **section 4.9.3**). The data would imply that this occurs relatively late in the maturation process because none of the processing mutants shown has undergone this cleavage step.

4.9 *Discussion*

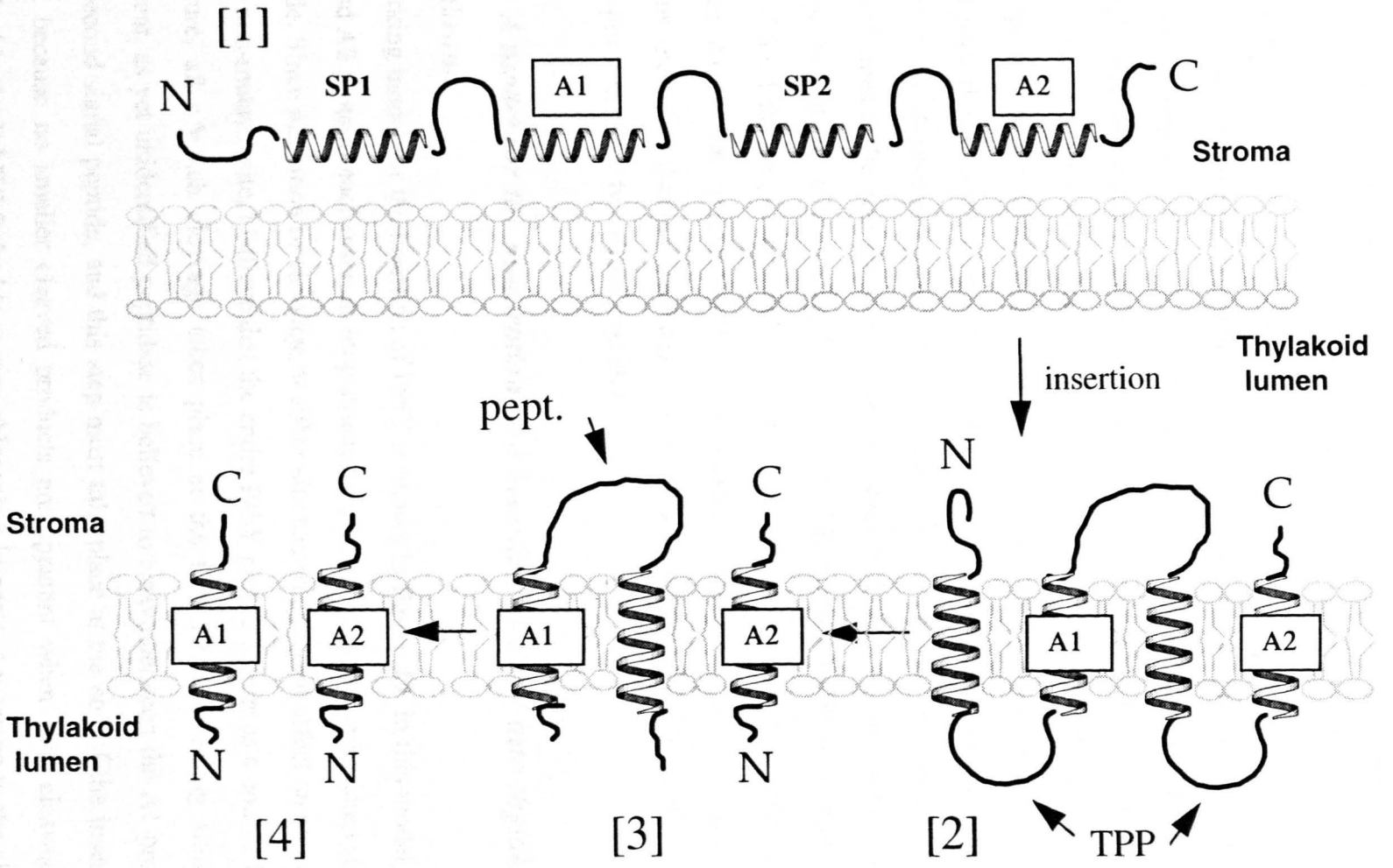
The data presented in this chapter demonstrate that the insertion and maturation of the PsbY A1 and A2 proteins take place by an unusual pathway involving multiple proteolytic processing steps. After import, the removal of the first, envelope transit peptide is presumably carried out by the stromal enzyme stromal processing peptidase (SPP) as with most other imported stromal/thylakoid proteins (Robinson *et al.*, 1998). After this step, it is proposed that two distinct cleavable signal peptides are used to assist the insertion of the A1 and A2 proteins. The evidence presented is in favour of this proposal: these sequences are certainly removed and then degraded, the sequences bear the typical hallmarks of thylakoid signal peptides (a hydrophobic region followed by a helix-breaking proline residue and then an Ala-X-Ala motif) and the substitution of the -1 alanine residues results in the almost complete inhibition of TPP activity as found in other studies using thylakoid signal peptides (Shackleton and Robinson, 1991; **chapter 3**). The data implies that the *psbY* gene encodes the first chloroplast-targeted polyprotein to be characterised in higher plants.

4.9.1 *Topological analysis of PsbY*

The very nature of a polyprotein implies precise topological analysis is much more complex compared to proteins presented in **chapter 3,** and therefore, the topologies of A1 and A2 have yet to be established. However, the activity of TPP has been demonstrated to reside solely on the lumenal face of the thylakoid membrane (Kirwin *et al.*, 1988), it appears inevitable that the N-termini of the mature proteins are located in this compartment, and that the signal peptides therefore function in a manner analogous to those of PSII-W and PSII-X. With these single-span proteins, the role of the signal peptide appears to be to assist integration by providing an additional hydrophobic region, and pre-PSII-W has been shown to form such a loop intermediate prior to cleavage by TPP (Thompson *et al.*, 1998).

Figure 29. *Model for the insertion of precursor PsbY.*

Stage [**1**]: stromal PsbY contains four hydrophobic regions (shown as helices) which include proteins A1 and A2 together with associated signal peptides (SP1, SP2). This protein inserts as a double loop structure (stage [**2**]) with the N-termini of the mature proteins located in the lumen. Cleavage at the TPP cleavage sites yields (stage [**3**]) the mature A2 protein together with protein A2 and the attached signal peptide 2. Cleavage of the latter intermediate by an unknown peptidase (pept.) on the stromal face of the membrane generates the mature A1 protein (stage [**4**]).


[1]
N
SP1
A1
SP2
A2
C
Stroma
insertion
Thylakoid lumen
pept.
Stroma
C
C
C
N
A1
A2
A1
A2
Thylakoid lumen
N
N
[4]
[3]
[2]
TPP

4.9.2 *Evolutionary considerations*

The requirement of the bipartite presequence appears only to be required if the protein is nuclear-encoded and thus required to translocate into the chloroplast. Genes encoding PSII-X and CF_0II have been identified in cyanobacteria and the plastid genomes of several eukaryotic algae, and in no case is the protein preceded by a signal-type peptide (Hermann *et al.*,1993; Kim *et al.*, 1996). This raises the possibility that the signal peptides have been acquired only after the transfer of the genes to nucleus since the initial endosymbiotic events involving a cyanobacterial-type organism, and that the more complex import pathway necessitates the presence of the second hydrophobic region for some reason. A basically similar situation applies to the *Arabidopsis psbY* gene: the homologous proteins encoded by open reading frames in the cyanobacterium Synechocystis PCC6803 and in several plastid genomes are devoid of signal peptides (Mant and Robinson, 1998) suggesting that both of the PsbY signal peptides were acquired after transfer of the gene to the nucleus. All known cyanobacterial/plastid-encoded ycf32 genes encode single-span proteins, indicating that the *Arabidopsis* gene arose by gene duplication together with the acquisition of the two signal peptides.

4.9.3 *A model for the maturation and insertion of PsbY into thylakoid membranes*

A working model for the insertion of PsbY is shown in **figure 29**. In this model, the A1 and A2 proteins each insert as loop structures together with their associated signal peptide. Since inhibition of cleavage at either site has no apparent effect on cleavage at the non-mutated site, I believe that the entire PsbY protein inserts as a double-loop structure, after which cleavage takes place at the two TPP processing sites. A different, as yet unidentified peptidase is believed to cleave between the A1 protein and second signal peptide, and this step must take place at the end of the insertion event because no smaller cleaved products are apparent when TPP cleavage is inhibited in the PsbY/1,2 double mutant chloroplast imports. According to the above model, this step has to take place on the stromal face of the thylakoid membrane; any other scenario would require transmembrane segments to reverse orientation and this would be extremely unusual. One interesting possibility follows on from suggestions

in **chapter 3** and further discussed in **chapter 5**, that the purpose of the signal peptide is not only to aid import into thylakoid membranes but also to stabilise the proteins in solution as first suggested by the helical hairpin hypothesis (Engelman and Steitz, 1981). In the case of PsbY the tertiary cleavage site is tightly constrained in the PsbY polyprotein (perhaps as a tight loop) but that it becomes accessible when the two TPP cleavage events release a smaller, more flexible structure.

Chapter 5

Final discussion

Chapter 5

Final discussion

5.1 *Introduction*

Nuclear-encoded thylakoid proteins are targeted into and across the thylakoid membrane by four distinct mechanisms. Precursors of lumenal proteins are translocated by either a Sec- or ΔpH-dependent mechanism, the choice of the pathway being determined by a specific sequence within the signal peptide. In both cases, thylakoidal processing peptidase (TPP) removes this signal peptide to release the mature protein. A structurally similar signal peptide is present in the fourth pathway used by a subset of integral membrane proteins including CF_0II, PSII-W, PSII-X and PsbY. In this case the integration process does not involve an identifiable energy source or known protein transport machinery. This Sec-independent direct insertion mechanism is unlike any other known, with the exception of a single protein, M13 procoat, which similarly inserts in the *Escherichia coli* plasma membrane by means of a signal peptide. This Sec-, ΔpH- and SRP-independent translocation mechanism shown with thylakoid proteins still left many questions to be answered prior to a possible conclusion that they all followed the same translocation mechanism.

5.2 *TPP cleavage is not required for insertion of proteins that conform to Sec-, ΔpH-, SRP-independent translocation criteria*

Both **chapter 3** and **4** show TPP cleavage mutant precursor proteins that follow the direct insertion pathway. Each of these mutant proteins was mutated at the same position relative to the TPP cleavage site (-1) and in each the terminal alanine of this motif was mutated to threonine. In every case presented in this thesis, stringent methods of extraction (carbonate and urea) failed to totally remove the intermediate forms of the precursor proteins, produced after chloroplast import, from the thylakoid membranes. The cases of pre-PSII-W and pre-PSII-X, demonstrate at least that although TPP is the only apparent protein to interact with these precursor proteins, cleavage is not required for stable insertion. However, in the case of pre-PsbY there

appears to be a logical requirement for a second protease to separate the two proteins from one another. But still, inhibition of TPP cleavage has no discernible effect on stability of the protein in the thylakoid membrane. So is there a requirement for TPP cleavage at all? One would have to say there is, cleavage by TPP in this and at least the Sec-dependent pathway logically has to be to remove the signal peptide. This is for either of two reasons: (a) with the signal peptide still attached to the mature protein it is unable to adopt the correct structure for insertion into the PSII complex; or secondly (b) the protein is unable to move from the site of import, stromal lamellae, to the granal lamellae for insertion into the PSII complex.

The production of alanine -1 threonine mutants presented in this thesis has further demonstrated that the inhibition of the signal peptidase has no effect on the translocation process itself.

5.3 *Loop intermediates - a "common theme" in translocation mechanisms?*

In **chapter 3**, there are conclusive arguments for a loop intermediate on at least the pre-PSII-W insertion pathway, this is demonstrated with the trypsin and thermolysin protease data (**figure 18**). Furthermore, the V8 protease data shown in **figure 20** combined with the sonication data, also presented, gives convincing, albeit circumstantial, evidence for a loop intermediate with pre-PSII-X. In **chapter 4** there are a variety of TPP cleavage mutants presented that evince a loop structure although further experiments are undoubtedly required. Because of the complex nature of pre-PsbY the most productive method for topological analysis of this protein is probably using various antibodies designed to the predicted loops.

Loop intermediates are not an uncommon two-dimensional structure in protein translocation. There have been reports of loop intermediates in proteins on the thylakoid ΔpH-dependent pathway (Fincher *et al.*, 1998), and the Sec-dependent export mechanism in *E. coli* (Kuhn *et al.*, 1994). However, in these cases the loop intermediates are caused by, firstly, the interaction of the signal peptide with the biological membrane prior to insertion (de Kruijff, 1994), followed by a "whip-like"

movement into the translocation machinery thus producing a 'loop intermediate' in the translocon, at which point the signal peptide is cleaved by the signal or leader peptidase. However, the difference demonstrated with the pre-PSII-W mutant presented in **chapter 3**, and most probably pre-PSII-X, was that this protein has been shown not to interact with proteinaceous machinery prior to insertion and therefore the insertion mechanism relies wholly on this loop structure for insertion as with the analogous bacteriophage protein M13 procoat (Kuhn, 1987). Furthermore, in the cases of these three proteins, the loop is probably an attempt to protect the highly hydrophobic nature of the mature protein from the aqueous phases these proteins pass through along the journey to their final destination, as was first proposed in the 'helical hairpin hypothesis' (Engelman and Steitz, 1981).

5.4 *Good housekeeping - signal peptide degradation*

The precursor protein PSII-W offers an opportunity to study the separate fate of the signal peptide from the mature protein. Of the many considerations those important for this work include: (1) the signal peptide of PSII-W contains five methionines and the mature protein only contains a single methionine; (2) relative to the size of pre-PSII-W, the signal peptide is very large (estimated to be about 4kDa; Kim *et al.*, 1998) and thus observable on the tricine SDS-PAGE system used throughout these studies. Therefore, labelling the precursor with [^{35}S]-methionine gives a powerful tool to study the signal peptide after cleavage by TPP. Unfortunately, the inhibitor study presented, and others omitted, have demonstrated no influence on the fate of the signal peptide. Perhaps, this is a conclusion in its own right, the fact hydrolysis of the signal peptide is an extremely rapid and efficient process. With previous studies, inhibition of the signal peptide peptidase causes translocation across the *E. coli* plasma membrane to be radically reduced (Chen and Tai, 1989), thus demonstrating a purpose for this efficiency.

Wickner and co-workers (Zwizinski and Wickner, 1980; Silver *et al.*, 1981) initially described a signal peptide hydrolase activity in the soluble fraction (S-100) of *E. coli* responsible for the degradation of the signal peptide cleaved from M13 procoat, although no further work has been published about this fraction. A membrane-bound signal peptidase activity that degrades the signal peptide (Hussain *et al.* 1982) has

also been shown. Hussein *et al.* (1982) found serine-thiol protease inhibitors affected the hydrolysis of the signal peptide from the *E. coli* major outer-membrane lipoprotein. Later, Protease IV, an integral membrane protease, was attributed to the initial signal peptide degradation (Ichihara *et al.*, 1984) and that subsequent hydrolysis of the signal peptide required a further cascade of proteases.

5.5 *Precursor PsbY is a polyprotein*

Chapter 4 shows two pieces of evidence to support the hypothesis proposed by Mant and Robinson (1998) that pre-PsbY (originally termed Ycf-32) is a polyprotein. Firstly, as previously mentioned, from four TPP mutations two mutants were found to cause the production of intermediate forms of pre-PsbY. Secondly, a construct (developed by Dr. A. Mant) is presented that consists of the 3rd and 4th hydrophobic regions of PsbY, that imports into isolated thylakoids as would be predicted.

Therefore, pre-PsbY is a polyprotein, but for what purpose? In cyanobacterial and algal genomes, Ycf32 (PsbY) is only encoded by a single cDNA, however, in *A. thaliana* there are two copies. There are several possibilities: (a) mature proteins A1 and A2 may be required in stoichiometric quantities in the PSII complex. But SecY and E are two *E. coli* membrane proteins known to be required in stoichiometric amounts, with excess, or more strictly uncomplexed, SecY being removed by the proteolytic function of FtsH (Kihara *et al.*, 1995); (b) or there is a requirement for the two PsbY proteins because they exist as heterodimeric structure, and in the cases of algal and cyanobacterial Ycf32 the requirement is for a homodimer.

Chapter 6

Bibliography

Chapter 6
Bibliography

A

Akimaru, J., S. Matsuyama, S., Tokuda, H., and Mizushima, S. (1991) Reconstitution of a protein translocation system containing purified Sec Y, SecE and SecA from *Escherichia coli. Proc Natl Acad. Sci. USA* **88** 6545-6549

Akita, M., Sasaki, S., Matsuyama, S. and Mizushima, S.(1990) SecA interacts with secretory proteins by reorganising the positive charge at the positive charge at the amino terminus of the signal peptide in *Escherichia coli. J. Biol. Chem.* **265** 8164-8169

Akiyama, Y. and Ito, K. (1987) Topology analysis of the SecY protein, and integral membrane protein involved in protein export in *Escherichia coli. EMBO J.* **6** 3465-3470

Alberts, B., Bray, D, Lewis, J., Raff, M., Roberts, K. and Watson, J.D. (1989) In *Molecular Biology of the Cell* 2nd Ed. Garland Publishing , Inc., New York

Allsop, A. E., Brooks, G., Bruton, G., Coulton, S., Edwards, P. D., Hatton, I. K., Kaura, A. C., McLean, S. D., Pearson, N. D., Smale, T. C. and Southgate, R. (1995) Penem inhibitors of bacterial signal peptidase. *Biorganic and Medicinal Chemistry Lett.* **5** 443-448

America, T., Hageman, J., Geura, A., Rook, F., Archer, K., Keegstra., K., and Weisbeek, P. (1994) Methotrexate does not block import of a DHFR fusion protein into chloroplast. *Plant Mol. Biol.* **24** 283-294

Anbudurai, P.R., Mor, T..S., Ohad, I., Shestakov, S.V. and Pakrasi, H.B. (1994) The *ctp*A gene encodes the C-terminal processing protease for the D1 protein of the photosystem II reaction centre complex. *Proc. Natl. Acad. Sci .USA* **91** 8082-8086

Andersen, B. and Scheller, H.V. (1993) In *Pigment-Protein Complexes in Plastids: Synthesis and Assembly* (Sundqvist, C. and Ryberg, M. Eds) Academic Press, Orlando, Florida pp383-418

Anderson, S. and Smith, S. M. (1986) Synthesis of the small subunit of ribulose-bisphosphate carboxylase from genes cloned into plasmids containing the SP6 promoter. *Biochem. J.* **240** 709-71

Andersson, H. and von Heijne, G. (1993) Sec dependent and sec independent assembly of E.coli inner membrane proteins: the topological rules depend on chain length. *EMBO J.* **12** 683-691

Andersson, H. and von Heijne, G. (1994) Positively charged residues influence the degree of SecA dependence in protein translocation across the E.coli inner membrane. *FEBS Lett.* **347** 169-172

Arkowitz, R. A., Joly, J.C. and Wickner, W. (1993) Translocation can drive the unfolding of a preprotein domain. *EMBO J.* **12** 243-253

Arkowitz, R. A. and Wickner, W. (1994) SecD and SecF are required for the protein electrochemical gradient stimulation of preprotein translocation. *EMBO J.* **13** 954-963

Arnon, D. I. (1949) Copper enzymes in isolated chloroplasts. Polyphenoloxidases in *Beta vulgaris*. *Plant Physiol.* **24** 1-15

Authier, F., Bergeron, J. J. M., Ou, W-J., Rachubinski, R. A., Postner, B. I. and Walton, P. A. (1995) Degradation of the cleaved leader peptide of thiolase by a peroxisomal proteinase. *Proc. Natl. Acad. Sci. USA* **92** 3859-3863

B

Bakker, E. P. and D Randall, L. L. (1984) The requirement for energy during export of β lactamase in *Escherichia coli* is fulfilled by the total protonmotive force. *EMBO J.* **3** 895-900

Barbrook, A. C., Packer, J. C. L. and Howe, C. J. (1996) Inhibition by Penems of processing peptidases from cyanobacteria and chloroplast thylakoids. *FEBS Lett.* **398** 198-200

Bassham, D. C., Bartling, D, Mould, R. M., Dunbar, B., Weisbeek, P., Herrmann, R.G. and Robinson, C. (1991) Delineation of envelope "transit" and thylakoid "transfer" signals within the pre-sequences of three imported thylakoid lumen proteins. *J. Biol. Chem.* **266** 23606-23610

Bassham, D. C., Creighton, A. M., Karnauchov, I., Herrmann, R. G., Klösgen, R. B. and Robinson, C. (1994) Mutations at the stromal processing peptidase cleavage site of a thylakoid lumen protein precursor affect the rate of processing but not fidelity. *J. Biol. Chem.* **269** 16062-16066

Bauerle, C. and Keegstra, K. (1991) Full length plastocyanin precursor is translocated across isolated thylakoid membranes. *J. Biol. Chem.* **266** 5876-5883

Berghöfer, J., Karnauchov, I., Herrmann, R.G. , Klösgen, R. B, (1995) Isolation and characterisation of a cDNA encoding the SecA protein from spinach chloroplasts. Evidence for azide resistance of Sec-dependent protein translocation across thylakoid membranes in spinach. *J. Biol. Chem.* **270** 18341-18346

Berks, B. C. (1996) A common pathway for proteins binding complex redox cofactors? *Mol. Microbiol.* **22** 393-404

Bernstein, H. D., Zopf, D., Freymann, D. M. and Walter, P. (1993) Functional substitution of the signal recognition particle 54-Kda subunit by its Escherichia coli homologue. *Proc. Natl. Acad. Sci. USA* **90** 5229-5233.

Betts, S. and Pichersky, E., (1992) Nucleotide sequence of cDNA encoding the precursor of the 23 kDa photosystem II protein of tomato. *Plant Mol. Biol.* **18** 995-996

Black, M. T., Munn, J. G. and Allsop, A. E. (1992) On the catalytic mechanism of prokaryotic leader peptidase I. *Biochem. J.* **282** 539-543

Black, M. J. (1993) Evidence that the catalytic activity of prokaryotic leader peptidase depends upon the operation of a serine-lysine catalytic dyad. *J. Bacteriol.* **175** 4957-4961

Blobel, G. and Dobberstein, B.(1975) Transfer of proteins across membranes. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* **67** 835-851

Blobel, G. and Dobberstein, B. (1975) Transfer of proteins across membranes. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* **67** 835-851

Bogsch, E. G., Brink, S. and Robinson, C. (1997) Pathway specificity for a delta pH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein. *EMBO J.* **16** 3851-3859

Bogsch, E. G., Sargent, F., Stanley, N. R., Berks, B. C., Robinson, C. and Palmer, T. (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem.* **273** 18003-18006

Brassilana, M. and Gwidzek, C. (1996) *In vivo* membrane assembly of the *E. coli* polytopic protein, melibiose permease, occurs via a Sec-independent process which requires the proton motive force. *EMBO J.* **15** 5202-5208

Breyton, C., de Vitry, C. and Popot, J.-L. (1994) Membrane association of cytochrome b_6f subunits. *J. Biol. Chem.* **269** 7597-7602

Briggs, L. M., Pecoraro, V. L. and McIntosh, L. (1990) Copper expression, cloning and regulatory studies of the plastocyanin gene from the cyanobacterium Synechocystis sp. PCC 6803 Plant Mol. Biol.. 15 633-642

Brink, S., Flügge, U.-I., Chaumont, F., Boutry M., Emmermann, M., Schmitz, U., Becker, K. and Pfanner, N. (1994) Preproteins of chloroplast envelope inner membrane contain targeting information for receptor-interdependent import into fungal mitochondria. *J. Biol. Chem.* **269** 16478-16485

Brink, S., Fischer, K., Klösgen, R. B. and Flügge, U.-I. (1995) Sorting of nuclear encoded chloroplast membrane proteins to the envelope and the thylakoid membrane. *J. Biol. Chem.* **270** 20808-20815.

Brock, I.W., Hazell, L., Michl, D., Nielson, V.S., Møller, B. L., Herrmann, R.G., Klösgen, R.B. and Robinson, C. (1993) Precursors of one integral and five lumenal thylakoid proteins are imported by isolated pea and barley thylakoids: optimisation of *in vitro* assays. *Plant Mol. Biol.* **23** 717-725

Brock, I.W., Mills, J.D., Robinson, D. and Robinson, C. (1995) The ΔpH-driven, ATP-independent protein translocation mechanism in the chloroplast thylakoid membrane. *J. Biol. Chem.* **270** 1657-1662

Bruce, B. D. (1998) The role of lipids in plastid protein transport. *Plant Mol. Biol.* **38** 223-246

Brundage, L., Hendrick, J. P., Schiebel, E., Driessen, A..J.M. and Wickner, W. (1990) The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* **62** 649-657

C

Cai, D., Herrmann, R.G. and Klösgen, R. B. (1993) the 20 kDa apoprotein of the CP24 complex of photosystem II: and alternative model to study import and intraorganellar routing of nuclear-encoded thylakoid proteins. *Plant J.* **3** 383-392

Caliebe, A., Grimm, R., Kaiser, G., Lübeck, J., Soll, J. and Heins, L. (1998) The chloroplastic protein import machinery contains a Rieske type iron-sulphur cluster and mononuclear iron binding protein. *EMBO J.* **16** 7342-7350

Cary., J. W., Lax, A.R. and Flurkey, W.H. (1992) Cloning and characterisation of cDNAs coding for *Vicia faba* polyphenol oxidase. *Plant Mol. Biol.* **20** 245-253

Chaal, B. and Howe, C. J. (1998) Characterisation of a cDNA encoding the thylakoidal processing peptidase from *Arabidopsis thaliana. J. Biol. Chem.* **273** 689-692

Chaddock, A. M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R. G., Klösgen, R. B. and Robinson, C. (1995) A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the ΔpH-dependent thylakoidal protein translocase, *EMBO J.* **14** 2715-2722

Chen, L. and Tai, P. C. (1989) Effects of inhibitors of membrane signal peptide peptidase on protein translocation into membrane vesicles. *Arch. Microbiol.* **153** 90-94

Chia, C. P. and Arntzen, C. J. (1986) Evidence for two-step processing of nuclear-encoded chloroplast proteins during membrane assembly. *J. Cell Biol.* **103** 725-731

Chitnis, P. R., Harel, E., Kohorn, B. D., Tobin, E. M. and Thornber, J. P. (1986) Assembly of the precursor and processed light-harvesting chlorophyll a/b protein of *Lemna* into the light-harvesting complex II or barley etiochloroplasts. *J. Cell. Biol.* **102** 982-988

Chitnis, P. R., Nechushtai, R. and Thornber, J. P. (1987) Insertion of the precursor of the light-harvesting chlorophyll a/b -protein into the thylakoids requires the presence of a developmentally regulated stromal factor. *Plant Mol. Biol.* **10** 3-11

Chou, M. M. and Kendall, D. A. (1990) Polymeric sequences reveal a functional interrelationship between hydrophobicity and length of signal peptides. *J. Biol. Chem.* **265** 2873-2880

Chua, N. -H. and Gilham, N. W. (1977) The sites of synthesis of the principle thylakoid membrane polypeptides in *Chlamydomonas reinhardtii. J. Cell Biol.* **74** 441-452

Chua, N. -H. and Schmidt, G. W. (1978) Post-translational transport into intact chloroplasts of a precursor to the small subunit of ribulose-1, 5-biphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* **75** 6110-6114

Clark, S. A. and Theg, S. M. (1997) A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol. Biol. Cell* **8** 923-934

Clark, S. E. and Lamppa, G. K. (1991) Determinants for cleavage of the chlorophyll a/b binding protein precursor: a requirement for a basic residue that is not universal for chloroplast imported proteins. *J. Cell Biol.* **114** 681-688

Cline, K. (1986) Import of proteins into chloroplasts. Membrane integration of a thylakoid precursor protein reconstituted in chloroplast lysates. *J. Biol. Chem.* **261** 14804-14810

Cline, K., Fulsom, D. R. and Viitanen, P. V. (1989) An imported thylakoid protein accumulates in the stroma when insertion into thylakoids is inhibited. *J. Biol. Chem.* **264** 14225-14232

Cline, K., Ettinger, W. F., and Theg, S. M. (1992) Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J. Biol. Chem.* **267** 2688-2696

Cline, K., Henry, R., Li, C. and Yuan, J. (1993) Multiple pathways for protein transport into or across the thylakoid membrane. *EMBO J.* **12** 4105-4114

Cline, K. and Henry, R. (1996) Importing and routing of nucleus-encoded chloroplast proteins. *Annu. Rev. Dev. Biol.* **12** 1-16

Creighton, A. M., Hulford, A., Mant, A., Robinson, D. and Robinson, C. (1995) A monomeric, tightly folded stromal intermediate on the pH-dependent thylakoid protein transport pathway. *J. Biol. Chem.* **270** 1663-1669

Cunningham, K., Lill, R., Crooke, E., Rice, M., Moore, K., Wickner, W. and Oliver, D. (1989) SecA protein, a peripheral protein of the *E. coli* plasma membrane, is essential for the functional binding and translocation of proOmpA. *EMBO J.* **8** 955-959

D

Dalbey, R. E. and von Heijne, G. (1992) Signal peptidases in prokaryotes and eukaryotes - a new protease family. *TIBS* **17** 474-478

Dalbey, R. E., Kuhn, A. and von Heijne, G. (1995) Directionality in protein translocation across membranes: the N-tail phenomenon. *Trends Cell Biol.* **5** 380-383

Dalbey, R. E., Lively, M. O. Bron, S. and Van Dijl, J. M. (1997) The chemistry and enzymology of the type I signal peptidases. *Protein Science* **6** 1129-1138

De Gier, J-W. L., Valent, Q. A., von Heijne, G. and Luirink, J. (1997) The *E. coli* SRP: preferences of a targeting factor. *FEBS Lett.* **408** 1-4

de Hoop, M. J. and Ab, G. (1992) Import of proteins into peroxisomes and other microbodies. *Biochem. J.* **286** 657-669

De Kruijff, B. (1994) Anionic phospholipids and protein translocation. *FEBS Lett.* **346** 78-82

Dev, I. K., Ray, P. H. and Novak, P. (1990) Minimum substrate sequence for signal peptidase I of *E. coli. J. Biol. Chem.* **265** 20069-20072

Dierks, T., Klappa, P., Wiech, H. and Zimmermann, R. (1993) The role of molecular chaperones in protein transport into the endoplasmic reticulum. *Phil. Trans. R. Soc. Lond.* B **339** 335-341

Douce, R and Joyard, J. (1990) Biochemistry and function of the plastid envelope. *Annu. Rev. Cell Biol.* **6** 173-216

Driessen, A. J. M. (1992) Precursor protein translocation by the *E. coli* translocase is directed by the proton motive force. *EMBO J.* **11** 847-853

Driessen, A. J. M. (1993) SecA, the peripheral subunit of the *E. coli* precursor protein translocase, is functional as a dimer. *Biochemistry* **32** 13190-13197

E

Economou, A., Pogliano, J. A., Beckwith, J., Oliver, D. B. and Wickner, W. (1995) SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell* **83** 1171-1181

Ellis, R. J. (1982) Promiscuous DNA - chloroplast genes inside plant mitochondria. *Nature* **299** 678-679

Ellis, R. J. (1987) Proteins as molecular chaperones. *Nature* **328** 378-379

Engelman, D. M. and Steitz, T. A. (1981) The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis. *Cell* **23** 411-422

F

Fikes, J. D. and Bassford Jr., P. J. (1989) Novel *secA* alleles improve export of maltose-binding protein synthesised with a defective signal peptide. *J. Bacteriol.* **171** 402-409

Filetici, P. (1996) Sequence of the *GLT1* gene from *S. cerevisiae* reveals the domain structure of yeast glutamate synthase. *Yeast* **12** 1359-1366

Fincher, V., McCaffery, M. and Cline, K. (1998) Evidence for a loop mechanism of protein transport by the thylakoidal delta pH pathway. *FEBS Lett.* **423** 66-70

Fischer, K., Weber, A., Arbinger, B., Brink, S., Eckerskorn, C. and Flügge, U. I. (1994) The 24 kDa outer envelope membrane protein from spinach chloroplasts: molecular cloning, *in vivo* expression and import pathway of a protein with unusual properties. *Plant Mol. Biol.* **25** 167-177

Flügge, U. I., Fischer, K., Gross, A., Sebald, W. Lottspeich, F. and Eckerskorn, C. (1989) The triose phosphate-3-phosphoglycerate-phosphate translocator from spinach chloroplasts: Nucleotide sequence of a full-length cDNA clone and import of the *in vitro* synthesised precursor proteins into chloroplasts. *EMBO J.* **8** 39-46

Franklin, A. E. and Hoffman, N. E. (1993) Characterisation of a chloroplast homologue of the 54 kDa subunit of the signal recognition particle. *J. Biol. Chem.* **268** 22175-22180

Fortin, Y., Phoenix, P. and Drapeau, G. R. (1990) Mutations conferring resistance to azide in *E. coli* occur primarily in the *SecA* gene. *J. Bacteriol.* **172** 6607-6610

G

Gallusser, A. and Kuhn, A. (1990) Initial steps in protein membrane insertion. Bacteriophage M13 procoat protein binds to the membrane surface by electrostatic interaction. *EMBO J.* **9** 2723-2729

Gau, A. E., Thole, H. H., Sokenko, A., Altschmied, L., Herman, R. G. and Pistorius, E. K. (1998) PsbY, a novel maganese-binding, low-molecular-mass protein associated with photosystem II. *Mol. Gen. Gene.* **in press**

Gavel, Y. and von Heijne, G. (1990) A conserved cleavage-site motif in chloroplast transit peptides. *FEBS Lett.* **261** 455-458

Gatenby, A. A. and Ellis, R. J. (1990) Chaperone function: the assembly of ribulose bisphosphate carboxylase-oxygenase. *Annu. Rev. Cell Biol.* **6** 125-149

Gebert, J. F., Overhoff, B., Manson, M. and Boos, W. (1988) The Tsr chemosensory transducer of *E. coli* assembles into the cytoplasmic membrane via a SecA-dependent process. *J. Biol. Chem.* **263** 16652-16660

Gibson, T.J. (1984) Studies on the Epstein-Barr virus genome. PhD thesis, University of Cambridge

Gilmore, R. (1993) Protein translocation across the endoplasmic reticulum: a tunnel with toll booths at entry and exit. *Cell* **75** 589-592

Glover, J. R., Andrews, D. W., Subramani, S. and Rachubinski, R. A.. (1994) Mutagenesis of the amino targeting signal of *Saccharomyces cerevisiae* 3-ketoacyl-CoA thiolase reveals conserved amino acids required for import into peroxisomes *in vivo*. *J. Biol. Chem.* **269** 7558-7563

Görlich, D., Prehn, S., Hartmann, E., Kalies, K. U. and Rapoport, T. A. (1992) A mammalian homologue of SEC61p and SECYp is associated with ribosomes and nascent polypeptides during translocation. *Cell* **71** 489-503

Gounaris, K., Mannock, D. A.., Sen, A., Brain, A. P. R. Williams, W. P. and Quinn, P. J. (1983) Polyunsaturated fatty acyl residues of galactolipids are involved in the control of bilayer/non-bilayer transitions in higher plant chloroplasts. *Biochim. Biophys. Acta.* **732** 229-242

Gray, M. C. (1989) The evolutionary origins of organelles. *Trends Genet.* **5** 294-299

H

Hageman, J., Robinson, C., Smeekens, S and Weisbeek, P. (1986) A thylakoid processing peptidase is required for complete maturation of the lumen protein plastocyanin. *Nature* **324** 567-569

Hageman, J., Baecke, C., Ebskamp, M., Pilon, R., Smeekens, S. and Weisbeek, P. (1990) Protein import and sorting inside the chloroplast are independent processes. *Plant Cell* **2** 479-494

Halpin, C., Elderfield, P. D. James, H. E., Zimmermann, R., Dunbar, B. and Robinson, C. (1989) The reaction specificities of the thylakoidal processing peptidase and Escherichia coli leader peptidase are identical. *EMBO J.* **9** 1743-1748

Hamman, B. D. Chen, J. C., Johnson, E. E. and Johnson, A. E. (1997) The aqueous pore through the translocon has a diameter of 40-60 Å during co-translational protein translocation at the ER membrane. *Cell* **89** 535-544

Harris, E. H., Boynton, J. E. and Gillham, N. W. (1994) Chloroplast ribosomes and protein synthesis. *Microbiol. Rev.* **58** 700-754

Hartl, F. -U., Lecker, S., Schiebel, E., Hendrick, J. P. and Wickner, W. (1990) The Binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* **63** 269-279

Hartl, F. -U., Hlodan, R. and Langer, T. (1994) Molecular chaperones in protein folding: the art of avoiding sticky situations. *Trends Biochem. Sci.* **19** 20-25

Heins, L., Collinson, I. and Soll, J. (1998) The protein translocation apparatus of chloroplast envelopes. *Trends Plant Sci.* **3** 56-61

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G. and Galas, D. J. (1989) A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Research* **17** 6545-6551

Henry, R., Kapazoglou, A., McCaffery, M. and Cline, K. (1994) Differences between lumen targeting domains of chloroplast transit peptides determine pathway specificity for thylakoid transport. *J. Biol. Chem.* **269** 10189-10192

Herrmann, R. G., Steppuhn, J., Herrmann, G. S. and Nelson, N. (1993) The nuclear-encoded polypeptide CF_0II from spinach is a real, ninth subunit of chloroplast ATP synthase. *FEBS Lett.* **326** 192-198

High, S. and Stirling, C. J. (1993) Protein translocation across membranes: common themes in divergent organisms. *Trends Cell Biol.* **3** 335-339

High, S., Henry, R., Mould, R. M., Valent, Q. A., Meacock, S., Cline, K., Gray, J. C. and Luirink, J. (1997) Chloroplast SRP54 interacts with a specific subset of thylakoid precursor proteins. *J. Biol. Chem.* **272** 11622-11628

Highfield, P. E., and Ellis, R. J. (1978) Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. *Nature* **271** 420-424

Hinnah, S. C., Hill, K., Wagner, R., Schlicher, T. and Soll, J. (1998) Reconstitution of a protein translocation channel of chloroplasts. *EMBO J.* **16** 7351-7360

Hipkins, M.F. and Baker, N.R. (1986) *Photosynthesis energy transduction: a practical approach.* IRL Press, Oxford, UK, Chapter 4

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C. -R., Meng, B. -Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K and Sugiura, M. (1989) The complete sequence of the rice (Oryza sativa) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217** 185-194

Hirel, B., Phillipson, B., Murchie, E., Suzuki, A., Kunz, C., Ferrario, S., Limani, A., Chaillou, S., Deleens, E., Brugière, N., Chaumont-Bonnet, M., Foyer, C. and

Morot-Gaudry, J-F. (1997) Manipulating the pathway of ammonia assimilation in transgenic non-legumes and legumes. *J. Plant Nutr. Soil Sci.* **160** 283-290

Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G. and Soll, J. (1994) A receptor component of the chloroplast protein translocation machinery. *Science* **266** 1989-1992

Horniak, L., Pilon, M., van't Hof, R. and de Kruijff, B. (1993) The secondary structure of the ferredoxin transit sequence is modulated by its interaction with negatively charged lipids. *FEBS Lett.* **334** 241-246

Horst, M., Kronidou, N. G. and Schatz, G. (1993) Through the mitochondrial inner membrane. *Curr. Biol.* **3** 175-177

Huang, J. Hack, E. Thornburg, R. W. and Myers, A. M. (1990) A yeast mitochondrial leader peptide functions *in vivo* as a dual targeting signal for both chloroplasts and mitochondria. *Plant Cell* **2** 1249-1260

Huang, L., Adam, Z. and Hoffman, N. E. (1992) Deletion mutants of chlorophyll a/b binding proteins are efficiently imported into chloroplasts but do not integrate into thylakoid membranes. *Plant Physiol.* **99** 247-255

Hulford, A., Hazell, L., Mould, R. M. and Robinson, C. (1994) Two distinct mechanisms for the translocation of proteins across the thylakoid membrane, one requiring the presence of a stromal protein factor and nucleotide triphosphates. *J. Biol. Chem.* **269** 3251-3256

Hurt, E. C., Soltanifar, N., Goldschmidt-Clermont, M., Rochaix, J. -D. and Schatz, G. (1986) The cleavable presequence of an imported chloroplast protein directs attached polypeptides into yeast mitochondria. *EMBO J.* **5** 1343-1350

Hussain, M., Ozawa, Y., Ichihara, S. and Mizushima, S. (1982) Signal peptide digestion in *E. coli*. Effect of protease inhibitors on hydrolysis of the cleaved signal peptide of the major outer-membrane lipoprotein. *Eur. J. Biochem.* **129** 233-239

Hynds, P. J., Robinson, D. and Robinson, C. (1998) Sec-independent twin-arginine system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J. Biol. Chem.* **273** 34868-34874

I

Ichihara, S., Beppu, N. and Mizushima, S. (1984) Protease IV, a cytoplasmic membrane protein of *E. coli*, has signal peptide peptidase activity. *J. Biol. Chem.* **259** 9853-9857

Ichihara, S., Suzuki, T., Suzuki, M. and Mizushima, S. (1986) Molecular cloning and sequencing of the sppA gene and characterisation of the encoded protease IV, a signal peptide peptidase, of *E. coli*. *J. Biol. Chem.* **261** 9405-9411

Irrgang, K.D., Shi, L.X., Funk, C., and Schröder, W.P. (1995) A nuclear-encoded subunit of the photosystem II reaction center. *J. Biol. Chem.* **270** 17588-17593

Izard, J. W. and Kendall, D. A. (1994) Signal peptides: exquisitely designed transport promoters. *Mol. Microbiol.* **13** 765-773

J

Jacobs, H. T. and Lonsdale, D. M. (1987) The selfish organelle. *Trends Genet.* **3** 337-341

James, H. E., Bartling, D., Musgrove, J. E., Kirwin, P. M. Herrmann, R. G. and Robinson, C. (1989) Transport of proteins into chloroplasts. Import and maturation of precursors to the 33, 23 and 16 kDa proteins of the oxygen-evolving complex. *J. Biol. Chem.* **264** 19573-19576

Johnson, A. E. (1993) Protein translocation across the ER membrane: a Fluorescent light at the end of the tunnel. *Trends Biochem. Sci.* **18** 456-458

Joly, J. C. and Wickner, W. (1993) The SecA and SecY subunits of translocase are the nearest neighbours of a translocating preprotein, shielding it from phospholipids. *EMBO J.* **12** 255-263

Jungnickel, B., Rapoport, T. and Hartmann, E. (1994) Protein translocation: common themes from bacteria to man. *FEBS Lett.* **346** 73-77

K

Karlin-Neumann, G. A. and Tobin, E. M. (1986) Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. *EMBO J.* **5** 9-13

Kerber, B. and Soll, J. (1992) Transfer of chloroplast-bound precursor protein into the translocation apparatus is impaired after phospholipase C. *FEBS Lett.* **306** 71-74

Kessler, F. Blobel, G., Patel, H. A., Schnell, D. J. (1994) Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science* **266** 1035-1039

Kessler, F. and Blobel, G. (1996) Interaction of the protein import and folding machineries in chloroplasts. *Proc. Natl. Acad. Sci. USA.* **93** 7684-7689

Kiefer, D., Hu, X., Dalbey, R. and Kuhn, A. (1997) Negatively charged amino acid residues play an active role in orienting the sec-independent Pf3 coat protein in the *E. coli* inner membrane. *EMBO J.* **16** 2197-2204

Kihara, A., Akiyama, Y. and Ito, K. (1995) FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit. *Proc. Natl. Acad. Sci. USA* **92** 4532-4536

Killian, J. A., de Jong, A. M. Ph., Bijvelt, J., Verkleij, A. J. and de Kruijff, B. (1990) Induction of non-bilayer lipid structures by functional signal peptides. *EMBO J.* **9** 815-819

Kim, S. J., Robinson, D., and Robinson, C. (1996) An Arabidopsis thaliana cDNA encoding PSII-X, a 4.1 kDa component of photosystem II: a bipartite presequence mediates SecA/pH-independent targeting into thylakoids. *FEBS Lett.* **390** 175-178

Kim, S. J., Robinson, C. and Mant, A. (1998) Sec/SRP-independent insertion of two thylakoid membrane proteins bearing cleavable signal peptides. *FEBS Lett.* **424** 105-108

Kimsey, H. H., Dagarag, M. D. and Kumamoto, C. A. (1995) Diverse effects of mutation on the activity of the *E. coli* export chaperone SecB. *J. Biol. Chem.* **262** 16386-16390

Kirwin, P. M., Elderfield, P. D. and Robinson, C. (1987) Transport of proteins into chloroplasts. Partial purification of a thylakoidal processing peptidase involved in plastocyanin biogenesis. *J. Biol. Chem.* **262** 16386-16390

Kirwin, P. M., Elderfield, P. D., Williams, R. S., and Robinson, C (1988) Transport of proteins into chloroplasts. Organisation, orientation, and lateral distribution of the plastocyanin processing peptidase in the thylakoid network. *J. Biol. Chem.* **263** 18128-18132

Kirwin, P. M., Meadows, J. W., Shackleton, J. B., Musgrove, J. E., Elderfield, P. D., Mould, R. M., Hay, N. A. and Robinson, C. (1989) ATP-dependent import of a lumenal protein by isolated thylakoid vesicles. *EMBO J.* **8** 2251-2255

Klösgen, R. B., Brock, I. W., Herrmann, R. G. and Robinson, C. (1992) Proton gradient-driven import of the 16 kDa oxygen-evolving complex protein as the full precursor protein by isolated thylakoids. *Plant Mol. Biol.* **18** 1031-1034

Knott, T. G. and Robinson, C. (1994) The SecA inhibitor, azide, reversibly blocks the translocation of a subset of proteins across the chloroplast thylakoid membrane. *J. Biol. Chem.* **269** 7843-7846

Ko, K. and Cashmore, A. R. (1989) Targeting of proteins to the thylakoid lumen by the bipartite transit peptide of the 33 kDa oxygen-evolving protein. *EMBO J.* **8** 3187-3194

Ko, K., Bornemisza, O., Kourtz, L., Ko, Z. W., Plaxton, W. C. and Cashmore, A. R. (1992) Isolation and characterisation of a cDNA clone encoding a cognate 70 kDa heat shock protein of the chloroplast envelope. *J. Biol. Chem.* **267** 2986-2993

Kourtz, L. and Ko, K. (1997) The early stages of chloroplast protein import involves Com70. *J. Biol. Chem.* **272** 32264-32271

Kuhn, A. and Wickner, W. (1985) Conserved residues of the leader peptide are essential for cleavage by leader peptidase. *J. Biol. Chem.* **260** 15914-15918

Kuhn, A., Kreil, G. and Wickner, W. (1986) Both domains of M13 procoat are required to initiate membrane insertion. *EMBO J.* **5** 3681-3685

Kuhn, A. (1987) Bacteriophage M13 procoat protein inserts into the plasma membrane as a loop structure. *Science* **238** 1413-1415

Kuhn, A., Zhu, H-Y. and Dalbey, R. E. (1990) Efficient translocation of positively charged residues of M13 procoat across the membrane excludes electrophoresis as the primary force for membrane insertion. *EMBO J.* **9** 2385-2389

Kuhn, A. and Troschel, D. (1992) In *Membrane Biogenesis and Protein targeting* (Neupert, W. & Lill, R., Eds) pp33-47, Elsevier, New York.

Kuhn, A., Kiefer, D., Khöne, C., Zhu, H-Y, Tschantz, W. R. and Dalbey, R. E. (1994) Evidence for a loop-like insertion mechanism of pro-Omp A into the inner membrane of *E. coli. Eur. J. Biochem.* **226** 891-897

Kuhn, A. (1995) Major coat proteins of bacteriophage PF3 and M13 as model systems for Sec-independent protein transport. *FEMS Micro. Rev.* **17** 185-190

Kuwabara, T., Reddy, K. J. and Sherman, L. A. (1987) Nucleotide sequence of the gene from the cyanobacterium *Anacystis nidulans* R2 encoding the Mn-stabilising protein involved in photosystem II water oxidation. *Proc. Natl. Acad. Sci. USA* **84** 8230-8235

L

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227** 680-685

Laidler, V., Chaddock, A. M., Knott, T. G., Walker, D. and Robinson, C. (1995) A SecY homologue in Arabidopsis thaliana. Sequence of a full-length cDNA clone and import of the precursor protein into chloroplasts. *J. Biol. Chem.* **270** 17664-17667

Lamppa, G. K. (1988) The chlorophyll *a/b* -binding protein inserts into the thylakoids independent of its cognate transit peptide. *J. Biol. Chem.* **270** 17664-17667

Lawlor. D. W. (1993) *Photosynthesis* 2nd Ed. Longman Scientific & Technical, U.K., Chapter 6

Li, H., Moore, T. and Keegstra, K. (1991) Targeting of proteins to the outer envelope membrane uses a different pathway than transport into chloroplasts. *Plant Cell* **3** 709-717

Li, H., Sullivan, T. D. and Keegstra, K. (1992) Information for targeting to the inner envelope membrane is contained in the mature region of the maize *Bt1*-encoded protein. *J. Biol. Chem.* **267** 18999-19004

Li, X., Henry, R., Yuan, J., Cline, K. and Hoffman, N. E. (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the post-translational integration of a protein into thylakoid membranes. *Proc. Natl. Acad. Sci. USA* **92** 3789-3793

Lill, R., Dowhan, W. and Wickner, W. (1990) The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. *Cell* **60** 271-280

Lill, R. and Neupert, W. (1996) Mechanisms of protein import across the mitochondrial outer membrane. *Trends Cell Biol.* **6** 56-61

Lin, L. L. and Little, J. W. (1989) Autodigestion and RecA-dependent cleavage of Ind-mutant LexA proteins. *J. Mol. Biol.* **210** 439-452

Lithgow, T., Glick, B. S. and Schatz, G. (1995) The protein import receptor of mitochondria. *Trends Biochem. Sci.* **20** 98-101

Little, J. W. (1993) LexA cleavage and other self-processing reactions. *J. Bacteriol.* **175** 4943-4950

Lorokovic, Z. J. Schröder, W. P., Pakrasi, H. B., Irrgang, K-D. and Herrmann, R. G. and Oelmüller, R. (1995) Molecular characterisation of PSII-W, the only nuclear-encoded component of the photosystem II reaction centre. *Proc. Natl. Acad. Sci. USA* **92** 8930-8934

Lubben, T. H., Donaldson, G. K., Viitanen, P. V. and Gatenby, A. A. (1989) Several proteins imported into chloroplasts form stable complexes with the GroEL-related chloroplast molecular chaperone. *Plant Cell* **1** 1223-1230

Lübeck, J., Soll, J. Akita, M., Nielsen, E. and Keegstra, K. (1996) Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. *EMBO J.* **15** 4230-4238

Luirink, J., High, S., Wood, H., Giner, A., Tollervey, D. and Dobberstein, B. (1992) Signal sequence recognition by an *E. coli* ribonucleoprotein particle. *Nature* **359** 741-743

Luirink, J., ten Hagen-Jongman, C. M., van der Weijden, C. C., Oudega, B., High, S., Dobberstein, B. and Kusters, R. (1994) An alternative protein-targeting pathway in *E. coli*: studies on the role of FtsY. *EMBO J.* **13** 2289-2296

Lütcke, H. (1995) Signal recognition particle (SRP), a ubiquitous initiator of protein translocation. *Eur. J. Biochem.* **228** 531-550

Lütcke, H., High, S., Römisch, K., Ashford, A. J. and Dobberstein, B. (1992) The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. *EMBO J.* **11** 1543-1551

M

Ma, Y., Kouranou, A., LaSala, S. E. and Schnell, D. J. (1996) Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. *J. Cell Biol.* **134** 315-327

MacFarlane, J. and Müller, M. (1995) The functional integration of a polytopic membrane protein of *E. coli* is dependent on the bacterial signal-recognition particle. *Eur. J. Biochem.* **233** 766-771

McNew, J. A. and Goodman, J. M. (1996) The targeting and assembly of peroxisomal proteins: some old rules do not apply. *Trends Biochem. Sci.* **21** 54-58

Madueño, F., Napier, J. A. and Gray, J. C. (1993) Newly imported Rieske iron-sulfur protein associates with both Cpn60 and Hsp70 in the chloroplast stroma. *Plant Cell* **5** 1865-1876

Maeda, H., Watanabe, T., Kobayashi, M., and Ikegami, I (1992) Presence of two chlorophyll *a*' molecules at the core of photosystem I. *Biochim. Biophys. Acta* **1099** 74-80

Mant A. and Robinson, C. (1998) An Arabidopsis cDNA encodes an apparent polyprotein of two non-identical thylakoid membrane proteins that are associated with photosystem II and homologous to algal *ycf*32 open reading frames. *FEBS Lett.* **423** 183-188

Matlack, K. E. S. and Walter, P. (1993) Shedding light on the translocation pore. *Curr. Biol.* **3** 677-679

Matsuyama, S. -I., Akimaru, J. and Mizushima, S. (1990) SecE-dependent overproduction of SecY in *E. coli*. Evidence for interaction between two components of the secretory machinery. *FEBS Lett.* **269** 96-100

Matsuyama, S. -I., Fujita, Y. and Mizushima, S. (1993) SecD is involved in the release of translocated secretory proteins from the cytoplasmic membrane of *E. coli*. *EMBO J.* **12** 265-270

Meadows, J. W. and Robinson, C. (1991) The full precursor of the 33 kDa oxygen-evolving complex protein of wheat is exported by *Escherichia coli* and processed to the mature size. *Plant Mol. Biol.* **17** 1241-1243

Michl, D., Robinson, C., Shackleton, J.B., Herrmann, R.G. and Klögen, R.B. (1994) Targeting of proteins to the thylakoids by bipartite presequences: CFoII is imported by a novel, third pathway. *EMBO J.* **13** 1310-1317

Michl, D. Karnauchov, I., Berghöfer, J., Herrmann, R. G. and Klösgen, R. B. (1998) Phylogenetic transfer of organelle genes to the nucleus can lead to new mechanisms of protein integration into membranes. *Plant J.* **in press**

Mihara, K. and Omura, T. (1996) Cytoplasmic chaperones in precursor targeting to mitochondria: the role of MSF and hsp70. *Trends Cell Biol.* **6** 104-108

Mishkind, M. L., Wessler, S. R. and Schmidt, G. W. (1985) Functional determinants in transit sequences: import and partial maturation by vascular plant chloroplasts of the ribulose-1, 5-bisphosphate carboxylase small subunit of *Chlamydomonas. J. Cell Biol.* **100** 226-234

Mould, R. M. and Robinson, C. (1991) A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. *J. Biol. Chem.* **266** 12189-12193

Mould, R. M., Shackleton, J. B., and Robinson, C. (1991) Requirements for the efficient import of two lumenal oxygen-evolving complex proteins into isolated thylakoids. *J. Biol. Chem.* **266** 17286-17289

Mould, R. M., Knight, J. S., Bogsch, E. and Gray, J. C. (1997) Azide-sensitive thylakoid membrane insertion of chimaeric cytochrome *f* constructs imported by isolated pea chloroplasts. *Plant J.* **11** 1051-1058

Müller, M. and MacFarlane, J. (1994) Membrane Assembly in Bacteria. In *Subcellular Biochemistry, Volume 22: Membrane Biogenesis* Maddy, A. H. and Harris, J. R. (eds) Plenum Press, New York pp327-359

N

Nakai, M., Goto, A., Nohara, T., Sugita, D. and Endo, T. (1994) Identification of the SecA protein homologue in pea chloroplasts and its possible involvement in thylakoidal protein transport. *J. Biol. Chem.* **269** 31338-31341

Neupert, W. (1997) Protein import into mitochondria. *Annu. Rev. Biochem.* **66** 863-917

Nielsen, E., Akita, M., Davila-Aponte, J. and Keegstra, K. (1997) Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO J.* **16** 935-946

Nilsson, I. and von Heijne, G. (1990) Fine-tuning the topology of a polytopic membrane protein: role of positively and negatively charged amino acids. *Cell* **62** 1135-1141

Nilsson, I. and von Heijne, G. (1992) A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a Sec-independent protein. *FEBS Lett.* **3** 243-246

Nishiyama, K., Mizushima, S. and Tokuda, H. (1993) A novel membrane protein involved in protein translocation across the cytoplasmic membrane of *E. coli*. *EMBO J.* **12** 3409-3415

Nohara, T., Nakai, M., Goto, A. and Endo, T. (1995) Isolation and characterisation of the cDNA for pea chloroplast SecA. Evolutionary conservation of the bacterial-type SecA-dependent protein transport within chloroplasts. *FEBS Lett.* **364** 305-308

Nothwehr, S. F. and Gordon, J. L. (1989) Eukaryotic signal peptide structure/function relationships. Identification of conformational features which influence the site and efficiency of co-translational proteolytic processing by site directed mutagenesis of human pre (Δ) apolipoprotein A-II *J. Biol. Chem.* **264** 3979-3987

Nothwehr, S. F. and Gordon, J. L. (1990) Structural features in the NH2-terminal region of a model eukaryotic signal peptide influence the site of the cleavage by signal peptidase. *J. Biol. Chem.* **265** 17202-17208

Novak, P., Ray, P. H. and Dev, I. K. (1986) Localisation and purification of two enzymes from *E. coli* capable of hydrolysing a signal peptide. *J. Biol. Chem.* **261** 420-427

Novak, P. and Dev, I. K. (1988) Degradation of a signal peptide by protease IV and oligopeptidase A. *J. Bacteriol.* **170** 5067-5075

Nunnari, J., Fox, T. D., and Walter, P. (1993) A mitochondrial protease with two catalytic subunits of non-overlapping specificities. *Science* **262** 1997-2004

O

Oblong, J. E. and Lamppa, G. K (1992b) Identification of two structurally related proteins involved in proteolytic processing of precursors targeted to the chloroplast. *EMBO J.* **11** 4401-4409

Ohyama, K. Fukuzawa, H. Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Complete nucleotide sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Plant Mol. Biol. Rep.* **4** 148-175

Oliver, D. B. and Beckworth, J. (1981) *E. coli* mutant pleiotropically defective in the export of secreted proteins. *Cell* **25** 765-772

Oliver, D. B., Cabelli, R. J., Dolan, K. M. and Jarosik, G. P. (1990) Azide-resistant mutants of *E. coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. *Proc. Natl. Acad. Sci. USA* **87** 8227-8231

Oliver, D. B. (1993) SecA protein: autoregulated ATPase catalysing preprotein insertion and translocation across the *E. coli* inner membrane. *Mol. Microbiol.* **7** 159-165

Oliver, J., Jungnickel, B., Görlich, D., Rapoport, T. and High, S. (1995) The Sec61 complex is essential for the insertion of proteins into the membrane of the endoplasmic reticulum. *FEBS Lett.* **362** 126-130

Olsen, L. J. (1998) The surprising complexity of peroxisome biogenesis. *Plant Molecular Biology* **38** 163-189

P

Packer, J. C. L. and André, D. and Howe, C. J. (1995) Cloning and sequence analysis of a signal peptidase I from the thermophilic cyanobacterium *Phormidium laminosum*. *Plant Molecular Biology* **27** 199-204

Payan, L. A. and Cline, K. (1991) A stromal protein factor maintains the solubility and insertion competence of an imported thylakoid membrane protein. *J. Cell. Biol.* **112** 603-613

Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R., Hendrickson, W. A. (1996) Structure of the UmuD' protein and its regulation in response to DNA damage. *Nature* **380** 727-730

Perry, S. E. and Keegstra, K. (1994) Envelope membrane proteins that interact with chloroplast precursor proteins. *Plant Cell* **6** 93-105

Pfanner, N., Craig, E. A. and Meijer, M. (1994) The protein import machinery of the mitochondrial inner membrane. *Trends Biochem. Sci.* **19** 368-372

Pfanner, N. and Meijer, M. (1995) Pulling in proteins. *Curr. Biol.* **5** 132-135

Pfanner, N., Douglas, M. G., Endo, T., Hoogenraad, N. J., Jensen, R. E., Meijer, M., Neupert, W., Schatz, G., Schmitz, U. and Shore, G. C. (1996) Uniform nomenclature for the protein transport machinery of the mitochondrial membranes. *Trends Biochem. Sci.* **21** 51-52

Pilon, M., de Kruijff, B. and Weisbeek, P. J. (1992) New insights into the import mechanism of the ferredoxin precursor into chloroplasts. *J. Biol. Chem.* **267** 2548-2556

Pilon, M., Wienk, H., Sips, W., de Swaarf, M., Talboom, I., Van't Hof, R., Weisbeek, P. J. and de Kruijff, B. (1995) Functional domains of the ferredoxin transit sequence involved in chloroplast import. *J. Biol. Chem.* **270** 3882-3893

Pinnaduwage, P. and Bruce, B. D (1996) *In vitro* interaction between a chloroplast transit peptide and chloroplast outer envelope lipids is sequence-specific and lipid class-dependent. *J. Biol. Chem.* **271** 32907-32915

Poritz, M. A., Bernstein, H. D., Strub, K., Zopf, D., Wilhelm, H. and Walter, P. (1990) An *E. coli* ribonucleoprotein containing 4.5S RNA resembles mammalian signal recognition particle. *Science* **250** 1111-1117

Pogliano, J. A. and Beckwith, J. (1994) SecD and SecF facilitate protein export in *E. coli*. *EMBO J.* **13** 554-561

Prince, R. C. and George, G. N. (1995) Cytochrome *f* revealed. *Trends Biochem. Sci.* **20** 217-218

Pugsley, A. P. (1993) The complete general secretory pathway in Gram-negative bacteria. *Microbiol. Rev.* **57** 50-108

Q

R

Randall, L. L. and Hardy, J. S. (1986) Correlation of competence for export with lack of tertiary structure of the mature species: A study *in vivo* of maltose binding protein in *E. coli*. *Cell* **46** 921-928

Randall, L. L. and Hardy, J. S. (1995) High selectivity with low specificity: how SecB has solved the paradox of chaperone binding. *Trends Biochem. Sci.* **20** 65-69

Rassow, J. and Pfanner, N. (1996) Protein biogenesis: chaperones for nascent polypeptides. *Curr. Biol.* **6** 115-118

Raven, P. H. (1970) A multiple origin for plastids and mitochondria. Many independent events may have been involved in the origin of these cellular organelles. *Science* **169** 641-646

Raven, P. H., Evert, R. F. and Eichhorn, S. E. (1986) In *Biology of Plants* 4th Ed. Worth Publishers, Inc. New York.

Ribes, V., Römisch, K., Giner, A., Dobberstein, B. and Tollervey, D. (1990) *E. coli* 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell* **63** 591-600

Robinson, C. and Ellis, R. J. (1984a) Transport of proteins into chloroplasts: partial purification of a chloroplast protease involved in the processing of imported precursor polypeptides. *Eur J. Biochem* **142** 337-342

Robinson, C., Cai, D., Hulford, A., Brock, I. W., Michl, D., Hazell, L., Schmidt, I. Herrmann, R. G. and Klösgen, R. B. (1994) The presequence of a chimeric construct dictates which of two mechanisms are utilised for translocation across the

thylakoid membrane: evidence for the existence of two distinct translocation systems. *EMBO J.* **13** 279-285

Robinson, D. Karnauchov, I., Herrmann, R. G., Klösgen, R. B. and Robinson, C. (1996) Protease-sensitive thylakoidal import machinery for the Sec-, ΔpH- and SRP-dependent protein targeting pathways, but not for CF_0II integration. *Plant J.* **10** 149-155

Robinson, C., Hynds, P. J., Robinson, D. and Mant, A. (1998) Multiple pathways for the targeting of thylakoid proteins in chloroplasts. *Plant Mol. Biol.* **38** 209-221

Roffey, R. A. and Theg, S. M. (1996) Analysis of the import of carboxy-terminal truncations of the 23-kilodalton subunit of the oxygen-evolving complex suggests that its structure is an important determinant for thylakoid transport. *Plant Physiol.* **111** 1329-1338

Roland, K. L. and Little, J. W. (1990) Reaction of LexA repressor with diisopropyl fluorophosphate. A test of the serine protease model. *J. Biol. Chem.* **265** 12828-12835

Römisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M. and Dobberstein, B. (1990) Homology of 54 K protein of signal-recognition particle, docking protein and two *E. coli* proteins with putative GTP-binding domains. *Nature* **340** 478-482

Rothstein, S. J., Gatenby, A. A., Willey, D. L. and Gray, J. C. (1985) Binding of pea cytochrome *f* to the inner membrane of *E. coli* requires the bacterial *SecA* gene product. *Proc. Natl. Acad. Sci. USA* **82** 7955-7959

Ryan, K. R. and Jensen, R. E. (1995) Protein translocation across mitochondrial membranes: what a long, strange trip it is. *Cell* **83** 517-519

S

Sääf, A., Andersson, H., Gafvelin, G. and von Heijne, G. (1995) SecA-dependence of the translocation of a large periplasmic loop in the *E. coli* MalF inner membrane protein is a function of sequence context. *Mol. Membr. Biol.* **12** 209-215

Salomon, M., Fischer, K., Flügge, U. I. And Soll, J. (1990) Sequence analysis and protein import studies of an outer chloroplast envelope polypeptide. *Proc. Natl. Acad. Sci. USA* **87** 5778-5782

Salvucci, M. E., Drake, R. R., Broadbent, K. P., Haley, B. E., Hanson, K. R. and Mchale, N. A. (1990) Identification of the 64 kDa chloroplast stromal phosphoprotein as phosphoglucomutase. *Plant Physiol.* **93** 105-109

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual. 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger, F., Niklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74** 5463-5467

Sargent, F., Bogsch, E. G., Stanley, N. R., Wexler, M., Robinson, C., Berks, B. C. and Palmer, T. (1998) Overlapping functions of components of a bacterial Sec-independent export system. *EMBO J.* **17** 3640-3650

Schägger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulphate-Polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166** 368-379

Schatz, P. J., Bieker, K. L., Ottemann, K. M., Silhavy, T. J. and Beckwith, J. (1991) One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the *E. coli* secretion machinery. *EMBO J.* **10** 1749-1757

Schiebel, E., Driessen, A. J. M., Hartl, F. -U. and Wickner, W. (1991) $\Delta\mu_{H+}$ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* **64** 927-939

Schmidt, G. W. and Mishkind, M. L. (1986) The transport of proteins into chloroplasts. *Annu. Rev. Biochem.* **55** 879-912

Schnell, D. J., Kessler, F. and Blobel, G. (1994) Isolation of components of the chloroplast protein import machinery. *Science* **266** 1007-1012

Seedorf, M., Waegemann, K. and Soll, J. (1995) A constituent of the chloroplast import complex represents a new type of GTP-binding protein. *Plant J.* **7** 401-411

Seidler, A. and Michel, H. (1990) Expression in *E. coli* of the PsbO gene encoding the 33 kDa of the oxygen-evolving complex from spinach. *Plant Mol. Biol.* **9** 1743-1748

Settles, M. A., Yonetani, A., Baron, A., Bush, D. R., Cline, K. and Martienssen, R. (1997) Sec-independent protein translocation the maize Hcf106 protein. *Science* **278** 1467-1470

Shackleton, J. B., and Robinson, C. (1991) The thylakoid processing peptidase is a signal-type peptidase with stringent substrate requirements at the -3 and -1 positions. *J. Biol. Chem.* **266** 12152-12156

Shen, L. M., Lee, J. I., Cheng, S. Y., Jutte, H., Kuhn, A., and Dalbey, R. E. (1991) Use of site-directed mutagenesis to define the limits of sequence variation tolerated for processing of the M13 procoat protein by the *E. coli* leader peptidase. *Biochemistry* **30** 11775-11781

Shinkai, A., Mei, L. H., Tokuda, H. and Mizushima, H. (1991) The confirmation of SecA, as revealed by its protease sensitivity, is altered upon interaction with ATP,

presecretory proteins, inverted membrane vesicles, and phospholipids. *J. Biol. Chem.* **266** 5827-5833

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N. Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, J., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organisation and expression. *EMBO J.* **5** 2043-2049

Siegel, V. (1997) Recognition of a transmembrane domain: another role for the ribosome? *Cell* **90** 5-8

Silver, P., Watts, C. and Wickner, W. (1981) Membrane assembly from purified components. I. Isolated M13 procoat does not require ribosomes or soluble proteins for processing by membranes. *Cell* **25** 341-345

Simonen, M. and Palva, I. (1993) Protein secretion in Bacillus species. *Microbiol. Rev.* **57** 109-137

Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K. and Weisbeek, P. (1986) Sequence of the precursor of the chloroplast thylakoid lumen protein plastocyanin *Nature* **317** 456-458

Smith, S. M. and Ellis, R. J. (1979) Processing of small subunit precursor of ribulose bisphosphate carboxylase and its assembly into whole enzyme are stromal events. *Nature* **278** 662-664

Soll, J. and Bennett, J. (1988) Localisation of a 64 kDa phosphoprotein in the lumen between the outer and inner envelopes of pea chloroplasts. *Eur. J. Biochem.* **175** 301-307

Soll, J. and Tien, R. (1998) Protein translocation into and across the chloroplast envelope membranes. *Plant Mol. Biol.* **38** 191-207

Stuart, R. A., Cyr, D. M., Craig, E.A. and Neupert, W. (1994) Mitochondrial molecular chaperones: their role in protein translocation. *Trends Biochem. Sci.* **19** 87-92

Subramanian, A. R. (1993) Molecular genetics of chloroplast ribosomal proteins. *Trends Biochem. Sci.* **18** 177-181

Sung, M. and Dalbey, R. E. (1992) Identification of potential active-site residues in the *E. coli* leader peptidase. *J. Biol. Chem.* **267** 13154-13159

Sutton, A., Sieburth, L. E. and Bennett, J. (1987) Light-dependent accumulation and localisation of photosystem II proteins in maize. *Eur. J. Biochem.* **164** 571-578

Suzuki, T., Itoh, A., Ichihara, S. and Mizushima, S. (1987) Characterisation of the *sppA* gene coding for protease IV, a signal peptide peptidase of *E. coli*. *J. Bacteriol.* **169** 2523-2528

T

Tani, K., Tokuda, H. and Mizushima, S. (1990) Translocation of proOmpA possessing an intramolecular disulphide bridge into membrane vesicles of *E. coli*. *J. Biol. Chem.* **265** 17341-17347

Theg, S. M. and Geske, F. J. (1992) Biophysical characterisation of a transit peptide directing chloroplast protein import. *Biochemistry* **31** 5053-5060

Thompson, S. J., Kim, S. J. and Robinson, C. (1998) Sec-independent insertion of thylakoid membrane proteins: analysis of insertion forces and identification of a loop intermediate. *J. Biol. Chem.* **273** 18979-18983

Thiaudière, E., Soekarjo, M., Kuchinka, E., Kuhn, A. and Vogel, H. (1993) Structural characterisation of membrane insertion of M13 procoat, M13 coat, and Pf3 coat proteins. *Biochemistry* **32** 12186-12196

Tranel, P. J., Froehlich, J., Goyal, A. and Keegstra, K. (1995) A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. *EMBO J.* **14** 2436-2446

Traxler, B. and Murphy, C. (1996) Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery. *J. Biol. Chem.* **271** 12394-12400

Tschantz, W. R., Sung, M., Delgado-Partin, V. M., and Dalbey, R. E. (1993) A serine and a lysine residue implicated in the catalytic mechanism of the *E. coli* leader peptidase. *J. Biol. Chem.* **268** 27349-27354

Tschantz, W. R., Paetzel, M., Cao, G. Suciu, D., Inouye, M. and Dalbey, R. E. (1995) Characterisation of a soluble, catalytically active form of *E. coli* leader peptidase: Requirement of detergent or phospholipid for optimal activity. *Biochemistry* **34** 3935-3941

Tsugeki, R. and Nishimura, M. (1993) Interaction of homologues of Hsp70 and Cpn60 with ferredoxin-NADP+ reductase upon its import into chloroplasts. *FEBS Lett.* **320** 198-202

U

Ulbrandt, N. D., Newitt, J. A. and Bernstein, H. D. (1997) The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* **88** 187-196

V

Van den Broeck, G., Timko, M. P. Kausch, A. P. Cashmore, A. R., Van Montagu, M. and Herrera-Estrella, L. (1985) Targeting of foreign proteins to chloroplasts by fusion to the transit peptide from the small subunit of ribulose-1, 5-bisphosphate carboxylase. *Nature* **313** 358-363

van Dijl, J. M., de Jong, A., Venema, G. and Bron, S. (1995) Identification of the potential active site of the signal peptidase SipS of Bacillus subtilis. Structural and functional similarities with LexA-like proteases. *J. Biol. Chem.* **270** 3611-3618

Vander Vere, P. S., Bennett, T. M., Oblong, J. E. and Lamppa, G. K. (1995) A chloroplast processing enzyme involved in precursor maturation shares a zinc-binding motif with a recently recognised family of metalloendopeptidases. *Proc. Natl. Acad. Sci. USA* **92** 7177-7181

Van t'Hoff, R., Demel, R. A., Keegstra, K. and de Kruijff, B. (1991) Lipid-peptide interactions between fragments of the transit peptide of ribulose 1, 5-bisphosphate carboxylase / oxygenase and chloroplast membrane lipids. *FEBS Lett.* **291** 350-354

Van t'Hoff, R., van Klompenburg, W., Pilon, M., Kozubek, A., de Korte-Kool, G., Demel, R. A., Weisbeek, P. J. and de Kruijff, B. (1993) The transit sequence mediates the specific interaction of the precursor of ferredoxin with chloroplast envelope membrane lipids. *J. Biol. Chem.* **268** 4037-4042

Van t'Hoff , R. and de Kruijff, B. (1995) Characterisation of the import process of a transit peptide into chloroplasts. *J. Biol. Chem.* **270** 22368-22373

Van Wijk, K. J., Bingsmark, S., Aro, E. -M. and Andersson, B. (1995a) *In vitro* synthesis and assembly of photosystem II core proteins. The D1 protein can be incorporated into photosystem II in isolated chloroplasts and thylakoids. *J. Biol. Chem.* **270** 25685-25695

Van Wijk, K. J., Knott, T. G. and Robinson, C. (1995b) Evidence for SecA- and ΔpH-dependent insertion of D1 into thylakoids. *FEBS Lett.* **368** 263-266

Viitanen, P. V., Doran, E. R. and Dunsmuir, P. (1988) What is the role of the transit peptide in thylakoid integration of the light-harvesting chlorophyll *a/b* protein? *J. Biol. Chem.* **263** 15000-15007

Voelker, R. and Barkan, A. (1995) Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J.* **14** 3905-3914

von Heijne, G., Steppuhn, J. and Herrmann, R. G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* **180** 535-545

von Heijne, G. and Nishikawa, K. (1991) Chloroplast transit peptides. The perfect random coil? *FEBS Lett.* **278** 1-3

von Heijne, G. (1994) Sec-independent protein insertion into the inner *E. coli* membrane. A phenomenon in search of an explanation. *FEBS Lett.* **346** 69-72

W

Waegemann, K., Paulsen, H. and Soll, J. (1990) Translocation of proteins into isolated chloroplasts requires cytosolic factors to obtain import competence. *FEBS Lett.* **261** 89-92

Waegemann, K. and Soll, J. (1991) Characterisation of the protein import apparatus in isolated outer envelopes of chloroplasts. *Plant. J.* **1** 149-158

Waegemann, K., and Soll, J. (1996) Phosphorylation of the transit sequence of chloroplast precursor proteins. *J. Biol. Chem.* **271** 6545-6554

Walker, D., Chaddock, A. M., Chaddock, J. A., Roberts, L. M., Lord, J. M. and Robinson, C. (1996) Ricin A chain fused to a chloroplast-targeting signal is unfolded on the chloroplast surface prior to import across the envelope membranes. *J. Biol. Chem.* **271** 4082-4085

Wandersman, C. (1992) Secretion across the bacterial outer membrane. *Trends Genet.* **8** 317-322

Watanabe, M. and Blobel, G. (1993) SecA protein is required for translocation of a model precursor protein into inverted vesicles of *E. coli* plasma membrane. *Proc. Natl. Acad. Sci. USA* **90** 9011-9015

Watanabe, M. and Blobel, G. (1995) High-affinity binding of *E. coli* SecB to the signal sequence region of a presecretory protein. *Proc. Natl. Acad. Sci. USA* **92** 10133-10136

Weiner, J. H., Bilous, P. T., Shaw, G. M., Lubitz, S. P., Frost, L., Thomas, G. H., Cole, J. A. and Turner, R. J. (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93** 93-100

Weiss, J. P., Ray, P. H. and Bassford Jr., P. J. (1988) Purified SecB protein of *E. coli* retards folding and promotes membrane translocation of the maltose-binding protein *in vitro*. *Proc. Natl. Acad. Sci. USA* **85** 8978-8982

Wexler, M., Bogsch, E. G., Klösgen, R. G., Palmer, T., Robinson, C. and Berks, B. C. (1998) Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the pH pathway. *FEBS Lett.* **431** 339-342

Whelan, J. and Glaser, E. (1997) Protein import into mitochondria. *Plant Mol. Biol.* **33** 771-789

Wickner, W. (1995) The nascent-polypeptide-associated complex: Having a "NAC" for fidelity in translocation. *Proc. Natl. Acad. Sci. USA* **92** 9433-9434

Wolfe, P. B., Wickner, W. and Goodman, J. M. (1983) Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope. *J. Biol. Chem.* **258** 12073-12080

Wolfe, P. B., Rice, M. and Wickner, W. (1985) Effects of two *sec* genes on protein assembly into the plasma membrane of *E. coli*. *J. Biol. Chem.* **260** 1836-1841

Wu, C., Seibert, F. S. and Ko, K. (1994) Identification of chloroplast envelope proteins in close proximity to a partially translocated chimaeric precursor protein. *J. Biol. Chem.* **269** 32264-32271

X

Y

Yalovsky, S., Paulsen, H., Michaeli, D. Chitnis, P. R. and Nechushtai, R. (1992) Involvement of a chloroplast HSP70 heat shock protein in the integration of a protein (light-harvesting complex precursor protein) into the thylakoid membrane. *Proc. Natl. Acad. Sci. USA* **90** 8552-8556

Yamada, H., Matsuyama, S., Tokuda, H. and Mizushima, S. (1989b) A high concentration of SecA allows proton motive force-independent translocation of a model secretory protein into *E. coli* membrane vesicles. *J. Biol. Chem.* **264** 18577-18581

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains. Nucleotide sequences of the M13mp8 and pUC19 vectors. *Gene* **33** 103-119

Yuan, J., Henry, R. and Cline, K. (1993) Stromal factor plays an essential role in protein integration into thylakoids that cannot be replaced by unfolding or by heat shock protein Hsp70. *Proc. Natl. Acad. Sci. USA* **90** 8552-8556

Yuan, J., Henry, R. McCaffery, M. and Cline, K. (1994) SecA homologue in protein transport within chloroplasts: evidence for endosymbiont-derived sorting. *Science* **266** 796-798

Z

Zwizinski, C. and Wickner, W. (1980) Purification and characterisation of leader (signal) peptidase from *E. coli*. *J. Biol. Chem.* **255** 7973-7977