Modelling the intracellular posttranslational protein targeting

Problem presented by: Ben Abell (Sheffield Hallam).

Study group contributors: John Ward (Loughborough), John Fozard, Anthony Hamstead, John King, Daniele Muraro (Nottingham), Verena Kriechbaumer (Sheffield Hallam) and Irene Lavagi (Warwick).

1 Introduction

Cellular organisation relies on the targeting of proteins to specific compartments; mislocalisation results in cellular disruption and is associated with diseases such as Alzheimer's and Parkinson's. The ability to control protein localisation is also crucial for genetic engineering, exemplified by recent success in the production of an artemisinin precursor in yeast for the treatment of malaria [18]. Hence, a quantitative understanding of protein targeting is expected to provide tools for biotechnology.

A crucial classification of protein targeting is whether it occurs during protein synthesis (translation) in the ribosome. The cotranslational mode is predominant in the targeting of endoplasmic reticulum (ER) proteins bearing an N-terminal signal sequence. However, proteins generally appear to be delivered to other intracellular organelles in the posttranslational mode, and their delivery to the correct intracellular organelle must be determined by factors that interact after ribosomal release. Many of these factors have been identified in the cytosol and at organellar membranes, but we have little understanding of how these factors compete with each other to generate reliable targeting outcomes. To study the crosstalk between targeting pathways we are utilising a group of membrane proteins termed tail-anchored (TA) proteins that possess a signal sequence at their C-termini, which commits them to follow posttranslational targeting pathways [10]. These TA proteins have a single transmembrane domain (TMD) at their C-terminus, which doubles as the signal sequence and membrane tether [16]. TA proteins are found in all cellular membranes, and can be delivered directly from the cytosol to ER, mitochondria, chloroplasts, and peroxisomes. Furthermore, TA proteins have been shown to interact with multiple cytosolic targeting factors, and therefore represent a model system for initiating a quantitative understanding of how signal sequences and targeting factors generate accurate targeting to specific organelles Figure 1.

Although some cytosolic binding factors that promote posttranslational targeting have been identified, the steps at which targeting specificity are generated are poorly understood. For example, the molecular chaperone Hsp70 has been shown to play an important role in protein targeting to ER [2, 14], mitochondria [3], and chloroplasts [20], yet the role of Hsp70 in regulating localisation is unclear. Targeting specificity is best understood for two independent pathways to the ER membrane, namely signal recognition particle (SRP) [1] and Asna1 [21]. Both of these targeting factors selectively bind tail-anchors that are highly hydrophobic [17], and cognate receptors for SRP and Asna1 are found at the ER membrane, thereby providing a conceptually simple mechanism for the specificity of ER targeting. However, some TA proteins can also be targeted to the ER by Hsp70 alone, which does not provide a unique delivery route to the ER. One possibility is that the tail-anchor is recognised directly by membrane-anchored receptors, and therefore it will be important to determine their contribution. Receptors for Hsp70 and another molecular chaperone, Hsp90, have been identified at mitochondrial [24] and chloroplast membranes [15], but their contribution to the specificity of protein targeting is not known. We have also identified an additional receptor for Hsp70 at the chloroplast membrane (unpublished). The other targeting factor known to be involved in TA protein targeting is AKR2A, which has recently been shown to promote targeting of Toc34 to the chloroplast, and

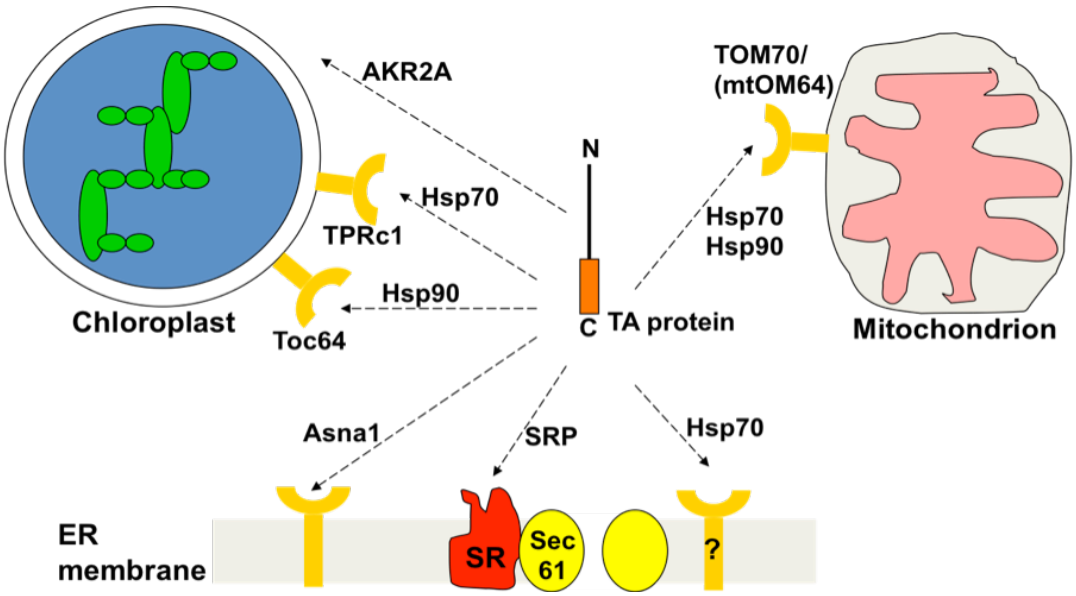

Figure 1: **Targeting pathways of TA proteins.** Targeting is promoted by several targeting factors that typically bind a receptor at the organelle surface

it also promotes the targeting of other classes of protein [4]. A receptor for AKR2A has not been identified to date.

1.1 Modelling approaches

We intend to apply modelling approaches to understand how protein targeting is controlled. The initial assumption is that binding to the signal sequence is competed by different targeting factors, and that this directs the protein to a specific organelle based on the localisation of receptors for the targeting factors. Although this mechanism may be sufficient in some cases, such as ER targeting mediated by SRP and Asna1, other pathways such as those mediated by Hsp70 may depend on direct recognition of the signal sequence by organellar receptors. A simple model could account for one or two organelles and a limited presence of targeting factors. Important considerations are the surface area of organelles and the distance from the site of protein synthesis to the organelle.

The generated models will be tested using cell free targeting assays, which are well established at Sheffield Hallam University, and can be manipulated to test a wide variety of models. Published data is available to quantify the binding affinities for some of the key targeting factors Therefore, we would be in a strong position to determine the importance of different events in protein targeting for generating specificity. We are currently acquiring quantitative data on the nature of interaction between targeting factors and their receptors using ellipsometry analysis, which could help to fix some of the unknown parameters. Taken together, we believe that there is potential to achieve fundamental insights into protein targeting by combining a modelling approach with cell free targeting experiments.

1.2 Aims

Protein localisation underpins cellular architecture and biochemical networks. Whilst we understand the steady state condition and many of the protein targeting processes, we lack a global model of protein segregation. In the simplest model, proteins possess a signal sequence that is bound by a targeting factor, which ensures delivery to a specific organelle. Although

this may be specific and robust in some cases, it is possible for proteins to become mistargeted in a wide range of conditions. Furthermore, some proteins are targeted to multiple locations to facilitate their normal functions [5]. Our hypothesis is that proteins are segregated by competitive binding of targeting factors, and by specific recognition at organellar receptors. The aim of this work is to devise a systematic framework for understanding how protein segregation occurs reliably and flexibly.

2 Mathematical modelling

Figure 2 shows a schematic of the process described by the model below. The proposed modelling is intended to be applicable to protein transport in plant cells and in *in vitro* experiments, in which the relevant cell components are extracted and set in a suitable matrix. We expect the general approach will be equally applicable to protein transport in eukaryotic cells in animals and yeast. When a protein is formed, or introduced in experiments, they are bound to ribosomes (concentration given by P_r); here, "concentration" can be taken to be the mass per unit volume or molar concentration. There, they are either released to the cell's cytosol (or external matrix) as free protein (cytosolic concentration P) or are picked up by the SRP targeting factor (SRP TF) to form a Protein-SRP complex (L_s). The free proteins can bind reversibly to a variety of TF classes, including SRP, to form a range of protein-TF complexes; the concentrations of which are denoted by L_m, where "m" is the index value of each of the non-SRP TFs being considered. The free protein and complexes diffuse around in the cytosol and at a sufficiently close proximity to an organelle's outer membrane receptor, the complex can bind reversibly to it. The surface density and affinity of these receptors will be dependent on the organelle and the TF; the surface densities of free and bound receptors on organelle "n" are denoted by $R_{n,*}$ and $B_{n,*}$, respectively, (where $* = \{m, s\}$). Once the complex has bound to the membrane receptor, special membrane proteins, known as translocons, cleave the protein from the receptor and TF and accepted either on the membrane or inside the organelle. The TF will also be removed from the receptor by this process, making it available for further complex formation.

The system can be viewed as a pathway of reactions that can be modelled in the usual way using mass action laws. We will make the following additional assumptions

- The system consists of a total of N types of organelles and M types of targeting factors.
- Protein conversion to an incompetent form, e.g. by misfolding or aggregation, is described by a simple decay term.
- TFs are assumed to be evenly distributed throughout the cytosol. The level of protein is typically measured to be in or introduced at concentrations of the order 1 pM, whilst the concentrations of targeting factors in the system are typically a 10-1000 or more times greater than this. Consequently, the concentration of free TF will be relatively unchanged by the reactions with the protein and will be thus assumed constant; the concentration of the TFs are accounted for within the reaction rate constants K_s, k_s and k_m (see Figure 2).
- Translocons on the organelle membranes are assumed to be evenly distributed.
- Explicit spatial considerations will be ignored. The location and number of ribosomes and the organelles can vary greatly across cells and any choice of distribution in the modelling or in simulations will be somewhat arbitrary. Given the previous assumptions, the diffusion of protein-TF complexes from the ribosomes to the organelle receptors is the

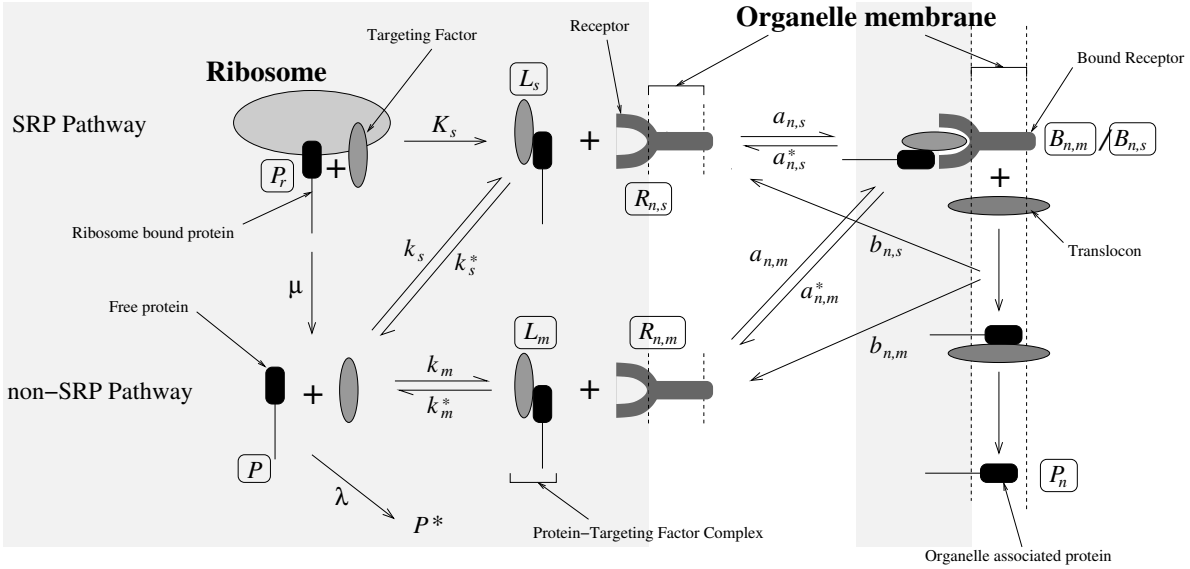

Figure 2: Schematic of the transit pathways of protein from ribosome to a target organelle for the SRP and HSP pathways. The pale shaded area is indicates the cytosol region, in which protein and complex diffusion occurs. The key components of the pathway are labelled, including the relevant model variables (names in boxes, see Table 1) and the rate constants.

only factor in which space is of any concern. The rate constants $a_{n,m}$ and $a_{n,s}$ (see Figure 2) will account for both the complex-receptor affinities and diffusion to the reaction site. This is discussed further in Section 2.1.

- Over the timescale of interest, the total surface density of receptors on the organelle membrane are assumed constant with a density denoted by $R^0_{n,*}$ for $* = \{m, s\}$.

Using these assumptions and adopting mass action kinetics we can write down the evolution of the variables listed in Table 1 in the form of differential equations as follows

$$\frac{dP_r}{dt} = -\mu P_r - K_s P_r, \tag{1}$$

$$\frac{dP}{dt} = \mu P_r - \sum_{m=1}^{M} k_m P + \sum_{m=1}^{M} k^*_m L_m - k_s P + k^*_s L_s - \lambda P, \tag{2}$$

$$\frac{dL_m}{dt} = k_m P - k^*_m L_m - \sum_{n=1}^{N} \beta_n a_{n,m} L_m R_{n,m} + \sum_{n=1}^{N} \beta_n a^*_{n,m} B_{n,m}, \tag{3}$$

$$\frac{dL_s}{dt} = K_s P_r + k_s P - k^*_s L_s - \beta_n a_{n,s} L_s R_{n,s} + \beta_n a^*_{n,s} B_{n,s}, \tag{4}$$

$$\frac{dB_{n,m}}{dt} = a_{n,m} L_m R_{n,m} - a^*_{n,m} B_{n,m} - b_{n,m} B_{n,m}; \tag{5}$$

$$\frac{dB_{n,s}}{dt} = a_{n,s} L_s R_{n,s} - a^*_{n,s} B_{n,s} - b_{n,s} B_{n,s}, \tag{6}$$

$$\frac{dP_n}{dt} = \beta_n b_{n,s} B_{n,s} + \beta_n \sum_{m=1}^{M} b_{n,m} B_{n,m}, \tag{7}$$

where from the final assumption listed above we can write for the free receptors

$$R_{n,m} = R^0_{n,m} - B_{n,m},$$
$$R_{n,s} = R^0_{n,s} - B_{n,s},$$

P_r	Ribosome associated protein concentration
P	Free protein concentration
P_n	Organelle "n" associated protein concentration
L_m	Protein–TF "m" complex concentration
L_s	Protein–SRP complex concentration
$R_{n,m}$	Surface density of free TF "m" receptors on organelle "n"
$R_{n,s}$	Surface density of free SRP TF receptors on organelle "n"
$B_{n,m}$	Surface density of bound TF "m" complex on organelle "n"
$B_{n,s}$	Surface density of bound SRP complex on organelle "n"
P^0	Initial concentration of ribosome associated protein.
$R^0_{n,m}$	Total surface density of rate of TF "m" complex receptor.
$R^0_{n,s}$	Total surface density of rate of SRP complex receptor.
μ	Ribosome associated protein release rate
λ	Free protein decay rate
K_s	Ribosome associated protein-SRP TF reaction rate constant
k_m	Free protein-TF "m" reaction rate constant
k_s	Free protein-SRP TF reaction rate constant
k^*_m	Protein-TF "m" complex dissociation rate constant
k^*_s	Protein-SRP TF complex dissociation rate constant
$a_{n,m}$	Protein-TF "m" complex – organelle "n" receptor binding rate constant*
$a_{n,s}$	Protein-SRP complex – organelle "n" receptor binding rate constant*
$a^*_{n,m}$	Organelle "n" TF "m" complex bound receptor dissociation rate constant
$a^*_{n,s}$	Organelle "n" SRP complex bound receptor dissociation rate constant
$b_{n,m}$	Protein uptake rate by organelle "n" via TF "m" complex bound receptor
$b_{n,s}$	Protein uptake rate by organelle "n" via SRP complex bound receptor
β_n	Surface area to volume ratio of organelle "n" membrane

Table 1: Table of the model variables and parameters.

where $n = 1, .., N$ and $m = 1, .., M$. The factor β_n in equations (14) and (15) is the organelle surface area to volume ratio in the system, thus a more wrinkled membrane will lead to a higher value of β_n. In equation (16), the quantity of organelle associated protein P_n is treated as a concentration; treating P_n as a surface density would require the introduction of the factor β_n as in equations (14) and (15). In total, the system (10)-(16) consist of $1 + (N + 1)(M + 2)$ nonlinear differential equations, with the nonlinear terms being of quadratic form.

The initial conditions in the experimental set up can be easily imposed on the model, and in what follows we assume

$$t = 0 : \qquad P_r = P^0, \quad P_n = P = L_m = L_s = B_{n,m} = B_{n,s} = 0,$$

i.e. at the start we introduce a quantity of ribosome associated protein into a virgin system of no bound targeting factors or receptors.

It is not possible to find closed form solutions for the full system and will require numerical treatment for their study (Section 3). However, by making biologically reasonable assumptions on the model parameters (see next section) we can analyse simplified forms of the model to gain deeper insights into the underlying dynamics of the system; an example of this is discussed in Section 4.

2.1 Parametrisation

The model contains a number of parameters and for many of these there is currently no data that can lead directly to a value. In this section, we discuss parameters estimation, highlighting those that were obtained from existing data and those that were estimated for this study.

2.1.1 Initial conditions

A proposed initial ribosome-associated protein concentration is about $P^0 = 1\,\text{pM}$, though this is likely to be an underestimate; nevertheless, the concentration will be significantly less than that of the associated TFs present. The only other non-zero initial conditions are the densities of organelle associated receptors, $R^0_{n,*}$, which are discussed in Section 2.1.4.

2.1.2 Ribosome-associated protein – SRP interaction

Although SRP concentrations in solution are of the order of 10 nM, SRP is thought to bind to the ribosome, and this results in the effective concentration being much higher.

The equilibrium dissociation constant for SRP binding to a translating ribosome with an exposed signalling sequence has been measured to be of the order of 0.05 nM [7]. However, there does not appear to be a great amount known about the kinetics of this process. The interaction is modelled as an irreversible first-order reaction, and we choose the rate constant K_s to fix the timescale, namely $K_s = 0.01\,\text{s}^{-1}$.

2.1.3 Protein – TF interaction

Data indicates that

	Typical concentration $[TF]$	Dissociation constant K_D
SRP	10 nM	0.05 nM [7]
HSP	2 μM	0.08 nM [19]

though these may vary markedly depending on the signal sequence emerging from the ribosome [7]. At equilibrium, $K_D = [P][TF]/[L] = k^*_m/k_m[TF]$, allowing us to estimate the ratio of the two reaction rates. We chose $k_s = k_m = 0.1\,\text{s}^{-1}$, so that the free protein concentration decays on a timescale of seconds. The backwards rates are then given by $k^*_i = K_D k_i/[TF]$, so we have $k^*_s = 5 \times 10^{-4}\,\text{s}^{-1}$ (i.e. for SRP) and $k^*_m = 4 \times 10^{-6}\,\text{s}^{-1}$ for HSP.

2.1.4 Complex–receptor interaction

The concentration of the TFs Hsp70 and Hsp90 used in cell free assays is about 2μM, which is considerably more concentrated than the ribosome-associated protein. The complex-receptor dissociation data suggests

	Dissociation constant K_D
SRP	15 nM [6]
HSP	100 nM [6]

The protein-TF complexes are distributed throughout the reaction medium, but the receptors are located on the organelle surfaces. In order for the ligand complex to bind to a receptor, it must first diffuse to the organelle. It may also be necessary for the ligand to be in the correct orientation (or overcome a potential barrier). The details of these short-range interactions are far beyond the scope of this study. Instead we are interested in a rough estimate of the influence of receptor surface density and organelle size (and geometry) on the reaction rates.

The ligand diffuses with diffusion coefficient D_s or D_m, whilst the receptors are restricted to the surfaces of the organelles. This affects the on and off rates for ligand-receptor binding: the on-rate is reduced by the need for the ligand to diffuse to the organelle surface (and competition for the free ligand between nearby receptors), and the off-rate is reduced by dissociating ligands re-binding to free receptors. Goldstein and Dembo [8] propose the following effective rate coefficients (see also [11] §4.2)

$$a_{n,i} = \frac{\alpha_{n,*}}{1 + (R_{n,i} r_n \alpha_{n,i}/D_i)} \qquad a^*_{n,i} = \frac{\alpha^*_{n,i}}{1 + (R_{n,i} r_n \alpha_{n,i}/D_i)} \tag{8}$$

where $\alpha_{n,i}$ and $\alpha^*_{n,i}$ are the intrinsic on- and off-rates for binding between the ligand and the receptor (e.g. for reactions in which the receptor is diffusing freely) and r_n is the radius of the (assumed spherical) organelle. Note that these rates depend on the surface density of free receptors, $R_{n,i}$. When the surface density is low ($R_{n,i} \ll D_i/r_n\alpha_{n,i}$), the reaction is second-order: $a_{n,i} \to \alpha_{n,i}$ as $R_{n,i} \to 0$ (the reaction itself is the rate-limiting step and diffusion is unimportant). When the surface density is high, the rate of receptor-ligand binding is limited by diffusion to the organelles, and the reaction becomes first-order : $a_{n,i} \sim D_i/(r_n R_{n,i})$ as $R_{n,i} \to \infty$ (so the total reaction rate per unit volume $a_{n,i} R_{n,i} L_i \beta_n \to (D_i/r_n) L_i \beta_n$). The estimate (8) is derived from the approximate solution to the reaction-diffusion equation outside a cell, with a boundary condition that relates the net diffusive flux into the cell to the rate of change of bound ligand ([8], equations (9) and (10)).

SRP, combined with the protein, is roughly 400 kDa in size [23], whilst the other TFs, such as HSPs are roughly 100 kDa (e.g. Hsp60, Hsp70 and Hsp90 are 60, 70 and 90 kilodaltons respectively [12]) ; similar-sized proteins have been measured [22] to have diffusion coefficients of about $D_s = 3.5 \times 10^{-7}\,\mathrm{cm^2\,s^{-1}}$ (for SRP) and $D_m = 5 \times 10^{-7}\,\mathrm{cm^2\,s^{-1}}$ (for the other TFs).

Chloroplasts and mitochondria are roughly cylindrical in shape and we estimate that they are $r_n = 2.5\,\mu\mathrm{m}$ **([9])** and $r = 0.5\,\mu\mathrm{m}$, respectively. The geometry of the endoplasmic reticulum is far from spherical, and the current approximation is not likely to be accurate. We propose an estimate of a 1000 receptors per chloroplast, and as the surface area of a chloroplast is estimated to be $30\,\mu\mathrm{m}^2$ [9] we have that $R^0_{n,i} = 1000/30\,\mu\mathrm{m}^2 = 30\,\mathrm{molecules}\,\mu\mathrm{m}^{-2}$.

For SRP-receptor binding on the endoplasmic reticulum, [13] measured the intrinsic on and off rates to be $\alpha_{n,s} = 1 \times 10^6\,\mathrm{M^{-1}\,s^{-1}}$ and $\alpha^*_{n,s} = 7 \times 10^{-3}\,\mathrm{s^{-1}}$ (note that the on rate was 20 times slower than this at a physiological ionic strength); these are consistent with typical on and off rates for ligand-receptor binding [11]. These values are roughly consistent with the measurement of $K_D = 15\,\mathrm{nM}$ for SRP-receptor binding [6]. For the other TFs, we take the intrinsic on-rate $\alpha_{n,m}$ to be the same as for SRP, and the estimate $K_D = 100\,\mathrm{nM}$ gives the off-rate $\alpha^*_{n,m} = K_D \alpha_{n,m} = 10^{-1}\,\mathrm{s^{-1}}$.

We also need to estimate the organelle surface area per unit volume, β_n, which we will obtain using the concentration of chloroplasts; a number of estimates were found during the study-group week. In vivo, using an estimate of 100 chloroplasts in a cell of volume $10^6\,\mu\mathrm{m}^3$, gives a chloroplast concentration of $[C] = 100/10^6\,\mu\mathrm{m}^3 = 10^{-4}\,\mu\mathrm{m}^{-3}$. With a chloroplast surface area of $30\,\mu\mathrm{m}^2$, this gives $\beta_n = 30\,\mu\mathrm{m} \cdot [C] = 3 \times 10^{-3}\,\mu\mathrm{m}^{-1}$. In vitro, there are two estimates. One appears to be that we have a a chloroplast concentration of $[C] = 5 \times 10^8\,\mathrm{ml^{-1}} = 5 \times 10^{-4}\,\mu\mathrm{m}^3$ (where we have made the assumption that the cellular volume consists of 5% chloroplasts, using a chloroplast volume estimate of $10^{-10}\,\mathrm{ml} = 100\,\mu\mathrm{m}^3$), corresponding to $\beta_n = 1.5 \times 10^{-2}\,\mu\mathrm{m}^{-1}$. The other is an estimate of 1000 chloroplasts in $60\mu l$, which gives $[C] = 1000/60\,\mu\mathrm{l} = 2 \times 10^5\,\mathrm{ml^{-1}} = 2 \times 10^{-7}\,\mu\mathrm{m}^{-3}$, and $\beta_n = 6 \times 10^{-6}\,\mu\mathrm{m}^{-1}$ (significantly smaller than the other estimates).

2.1.5 Other reaction rates

The other processes are modelled as first-order reactions. We do not presently have data with which to estimate their rates, and instead choose these such that the reaction timescales are sensible compared with other processes in the system. We take the rate of release of the protein from the ribosome to be $\lambda = 0.1\,\mathrm{s}^{-1}$. The rate at which free protein became incompetent was taken to be $\mu = 1\,\mathrm{s}^{-1}$. The rate at which protein bound to a surface receptor is incorporated into the organelles by translocases was chosen to be $b_{n,i} = 1\,\mathrm{s}^{-1}$; amongst the limited related data is that there are approximately 7 translocons per receptor [25].

2.2 Non-dimensionalisation

We rescale using

$$t = \frac{1}{\mu}\,\hat{t}, \quad \{P_r, P, L_i, P_n\} = P^0\,\{\hat{P}_r, \hat{P}, \hat{L}_i, \hat{P}_n\}, \quad \{B_{n,i}, R_{n,i}\} = R^0_{n,i}\,\{\hat{B}_{n,i}, \hat{R}_{n,i}\}, \tag{9}$$

where $i = \{m, s\}$. Here, time has been scaled with the fastest timescale, which is of $O(1$ second). The non-dimensional parameters are then

$$\{\hat{\lambda}, \hat{K}_s, \hat{k}_i, \hat{k}^*_i, \hat{a}^*_{n,i}, \hat{b}_{n,i}\} = \frac{1}{\mu}\{\lambda, K_s, k_i, k^*_i, a^*_{n,i}, b_{n,i}\},$$
$$\hat{\beta}_{n,i} = \frac{R^0_{n,i}\beta_n}{P^0}, \quad \hat{\alpha}_{n,i} = \frac{P^0}{\mu}\alpha_{n,i}, \quad \hat{D}_{n,i} = \frac{P^0 D_i}{R^0_{n,i} r_n \mu}.$$

On dropping hats, the dimensionless equations become

$$\frac{dP_r}{dt} = -P_r - K_s P_r, \tag{10}$$
$$\frac{dP}{dt} = P_r - \sum_{m=1}^{M} k_m P + \sum_{m=1}^{M} k^*_m L_m - k_s P + k^*_s L_s - \lambda P, \tag{11}$$
$$\frac{dL_m}{dt} = k_m P - k^*_m L_m - \sum_{n=1}^{N} \beta_{n,m} a_{n,m} L_m (1 - B_{n,m}) + \sum_{n=1}^{N} \beta_{n,m} a^*_{n,m} B_{n,m}, \tag{12}$$
$$\frac{dL_s}{dt} = K_s P_r + k_s P - k^*_s L_s - \beta_{n,s} a_{n,s} L_s (1 - B_{n,s}) + \beta_{n,s} a^*_{n,s} B_{n,s}, \tag{13}$$
$$\frac{dB_{n,m}}{dt} = a_{n,m} L_m (1 - B_{n,m}) - a^*_{n,m} B_{n,m} - b_{n,m} B_{n,m}, \tag{14}$$
$$\frac{dB_{n,s}}{dt} = a_{n,s} L_s (1 - B_{n,s}) - a^*_{n,s} B_{n,s} - b_{n,s} B_{n,s}, \tag{15}$$
$$\frac{dP_n}{dt} = \beta_{n,s} b_{n,s} B_{n,s} + \beta_{n,m} \sum_{m=1}^{M} b_{n,m} B_{n,m}, \tag{16}$$

where

$$a_{n,i} = \frac{\alpha_{n,i} D_{n,i}}{D_{n,i} + \alpha_{n,i} R}, \qquad a^*_{n,i} = \frac{\alpha^*_{n,i} D_{n,i}}{D_{n,i} + \alpha_{n,i} R},$$

and conservation of receptors yield

$$B_{n,i} + R_{n,i} = 1. \tag{17}$$

Table 2 shows the values of the dimensionless parameters using the data in Section 2.1. Of most interest is the comparison between the protein-TF complex diffusion rate and complex-receptor

K_s	0.01	k_i	0.1	k_s^*	5×10^{-4}	k_m^*	4×10^{-6}
$\alpha_{n,i}$	1×10^{-6}	$\alpha_{n,s}^*$	7×10^{-3}	$\alpha_{n,m}^*$	0.1	$b_{n,i}$	1
λ	0.1	$\beta_{n,i}$	0.3 - 750	$D_{n,s}$	$3 - 14 \times 10^{-4}$	$D_{n,m}$	$4 - 20 \times 10^{-4}$

Table 2: Table of dimensionless parameter values. The ranges for $\beta_{n,i}$ and $D_{n,s}/D_{n,m}$ result from $\beta_n = 6 - 500 \times 10^{-6} \mu\text{m}^{-1}$ and $r_n = 0.5 - 2.5\mu\text{m}$.

binding rate. Here, $\alpha_{n,i}/D_{n,i} = O(0.001 - 0.0001)$, so that the binding reaction is occurring at a rate significantly slower than diffusion transfer, which is counter to what was expected; this means that $a_{n,i} \approx \alpha_{n,i}$ and $a_{n,i}^* \approx \alpha_{n,i}^*$. Further investigation is required to review our estimates and to confirm whether or not the binding process really is the rate limiting step in this process.

3 Results

Figures 3 and 4 shows simulation results for the evolution of all the variables following the introduction of SRP and HSP, respectively, in the presence of single organelle. The parameters used are those in listed Table 2 with $\beta_{n,i} = 750$, $D_{n,s} = 3 \times 10^{-4}$ and $D_{n,m} = 4 \times 10^{-4}$. In both simulations shown, the unbound protein binds very quickly with TFs to form ligands; this is happening so rapidly that the drop of P_r and P from the initial values to near zero is indistinguishable to the $t = 0$ axis. As P and P_r drops rapidly we see that the ligand concentration rises rapidly to $L_i \approx 0.5$; this is in fact in agreement with that predicted by the analysis of Section 4. In this simulation, approximately 50% of the protein is becomes incompetent, "IP" in the figures. Following this initial phase, the ligands concentration decay relatively slowly, due to the relatively low ligand-receptor affinity. Consequently, the bounded receptor density $B_{1,1}$ is almost negligible, reaching a maximum near $t = 0$ of about 5×10^{-7}, due to a relatively rapid dissociation to unbound receptor and organelle-associated proteins P_1.

The results presented here seem to possess a number of anomalies that suggest the parameter values obtained require further scrutiny. Currently, there seems to be a paucity of data for which serious quantitative comparisons with the model's results. However, based on experimental observation we expect the organelle associated proteins P_1 to increase, but decelerate, in time towards a maximal limit. The lower middle plot in both figures seems to agree well with this qualitative description, and thus provides a basis for future work. An analysis similar to that described in Section 4 would suggest that P_1 will evolve approximately in the manner of $P_1 \sim P_{max}(1 - e^{-ct})$, i.e. P_1 will decay exponentially towards a final concentration; knowledge of such underlying behaviour could help establish more accurate estimates for the parameter values given appropriate experimental data.

4 Two organelle competition for single protein

In this section we analyse the model by making simplifications based on the estimated parameter values discussed in Section 2.1. The approach is applicable to any number of proteins, targeting factors and organelle combinations, but as a demonstration of the value of the analysis we focus on the case of two organelles competing for a single protein and targeting factor combination; though all types of targeting factor will be considered. We thus focus on the case $n = 2$ and one targeting factor ($m = 0$ if SRP is considered, else $m = 1$). The analysis will give a simple expression that can be validated experimentally to test the hypothesis, based on the discussion of parameter $a_{n,m}$ in Section 2.1, that the diffusion of the protein-TF complex makes

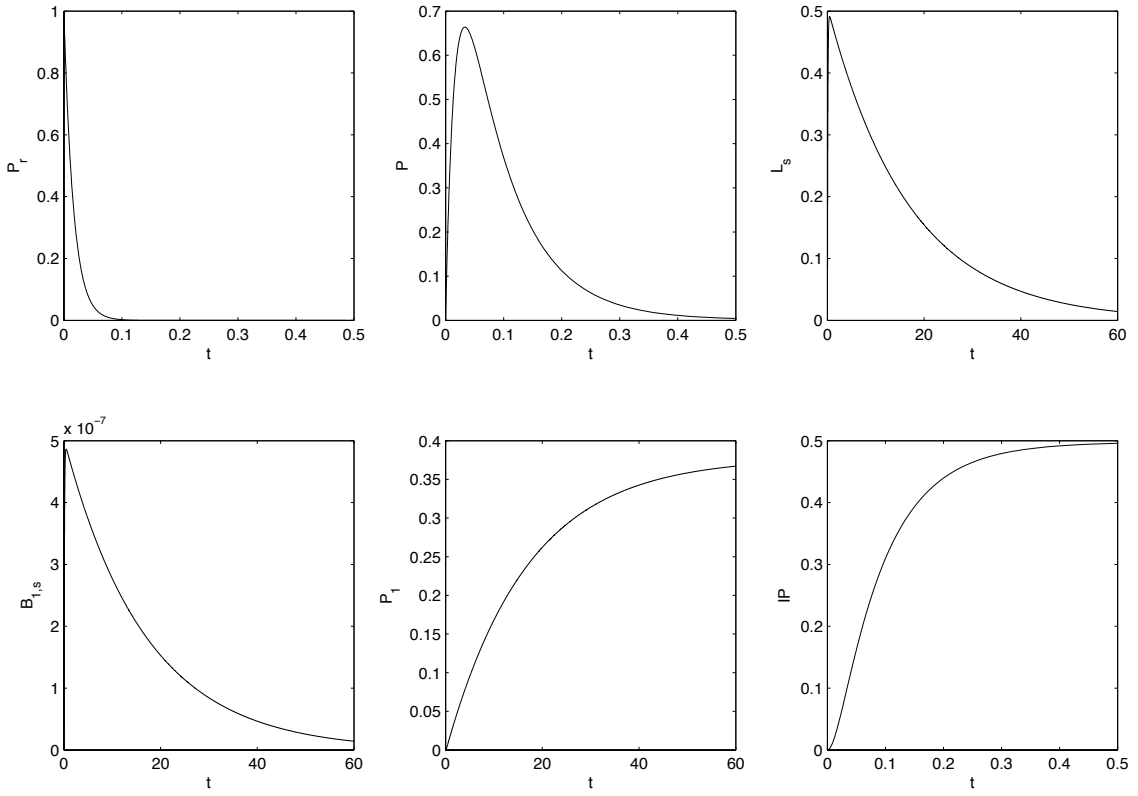

Figure 3: Simulated results for the SRP pathway and single organelle, with the time t shown in minutes.

a negligible contribution in the overall dynamics of the system. For notational convenience we will remove subscripts s and m, whereby the solutions for SRP or the other TFs correspond to the cases $K_s > 0$ or $K_s = 0$, respectively.

As discussed in Section 2.1, we expect the forward reactions to be occurring very rapidly (seconds) and the reverse ones very slowly O(10+ mins). The parameter values suggest that the timescale for diffusion is significantly less than that of the receptor binding and hence $\alpha_{n,m} \ll D_{n,m}$, so that $a_{n,m} \sim \alpha_{n,m}$. Furthermore, the parameters given suggest that $a_{n,m} \approx 10^{-5}$; however, this is likely to be a considerable under-estimate, as this suggests that the binding process operates on a timescale of days. Nevertheless, we will proceed on the assumption that $a_{n,m} \ll 1$. Dropping the m component of $a_{n,m}$ for notational convenience, we write $\epsilon = a_1 \ll 1$, and assume that $\hat{\lambda}, \hat{b}_1, \hat{b}_2, \hat{k}, \hat{K}_s = O(1)$, $\hat{a}_2 = O(\epsilon)$ and $\hat{a}_1^*, \hat{a}_2^*, k^* = o(\epsilon)$. Making the appropriate parameter rescalings in terms of ϵ, equations (10)-(16) up to $O(\epsilon)$ are

$$\frac{dP_r}{dt} = -P_r - K_s P_r, \tag{18}$$

$$\frac{dP}{dt} = P_r - (k+\lambda)P, \tag{19}$$

$$\frac{dL}{dt} = kP + K_s P_r - \epsilon\,\beta_1\,L\,(1-B_1) - \epsilon\,\beta_2\,a_2\,L\,(1-B_2), \tag{20}$$

$$\frac{dB_1}{dt} = \epsilon\,L(1-B_1) - b_1 B_1, \tag{21}$$

$$\frac{dB_2}{dt} = \epsilon\,a_2\,L(1-B_2) - b_2 B_2, \tag{22}$$

$$\frac{dP_1}{dt} = \beta_1 b_1 B_1, \tag{23}$$

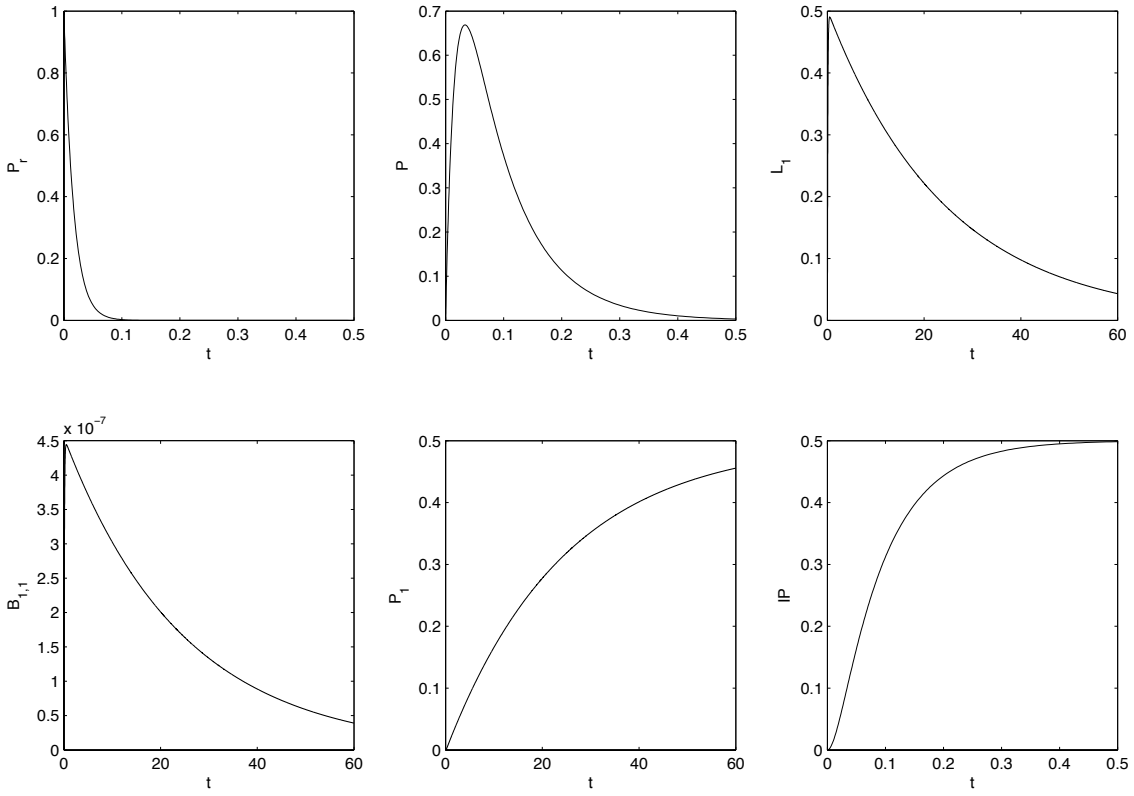

Figure 4: Simulated results for the HSP pathway and single organelle, with the time t shown in minutes.

$$\frac{dP_2}{dt} = \beta_2 b_2 B_2, \tag{24}$$

subject to

$$t = 0: \qquad P_r = 1,\ P = L = B_1 = B_2 = P_1 = P_2 = 0.$$

In the limit $\epsilon \to 0$, there are two timescales

1. $t = O(1)$: protein are released from the ribisome and combine with TFs to form complexes.
2. $t = O(1/\epsilon)$: the complexes diffuse in the media and harvested by organelle receptors, on which the protein is rapidly dissociated from the complex and taken in by the organelle.

In effect, it is the transfer of the complexes to the receptors that is the rate limiting process; this two timescale structure has already been demonstrated in the results shown in Figures 3 and 4. The model solutions in the two timescales are summarised below.

$t = O(1)$

For $t = O(1)$, the leading order system is obtained by substituting $\epsilon = 0$ into (18)-(24), which provides on solution

$$P_r \sim e^{-(1+K_s)t}, \quad P \sim \frac{1}{k+\lambda-1-K_s}\left(e^{-(1+K_s)t} - e^{-(k+\lambda)t}\right),$$
$$L \sim \frac{k}{k+\lambda-1-K_s}\left(\frac{1-e^{-(1+K_s)t}}{1+K_s} - \frac{1-e^{-(1+K_s)t}}{k+\lambda}\right) + \frac{K_s}{1+K_s}\left(1 - e^{-(k+\lambda)t}\right),$$
$$B_1 \sim 0, \quad B_2 \sim 0, \quad P_1 \sim 0, \quad P_2 \sim 0,$$

where we observe that the organelle associated protein levels are negligible in this initial period. In large time ribosome-associated and free protein will have all been picked up by TFs or been degrade and/or sequestered in the media. The overall fraction of protein that will ultimately form complexes with TF is given approximately by L_∞, where L_∞ is the limiting value of L as $t \to \infty$, namely

$$L_\infty \;=\; 1 \;-\; \frac{\lambda}{(1+K_s)(k+\lambda)}.$$

We note using the parameter values in Table 2 that $L_\infty \approx 0.5$, which agrees with the profiles for L_1 shown in Figures 3 and 4 in the vicinity of $t = 0$. The details regarding the correction terms will be omitted for brevity, the key conclusion being that these approximations will breakdown around $t = O(1/\epsilon)$ and that

$$B_1 \;\sim\; \epsilon\,\tfrac{1}{b_1} L_\infty, \quad B_2 \;\sim\; \epsilon\,\tfrac{a_2}{b_2} L_\infty, \quad P_1 \;\sim\; \epsilon\, L_\infty t, \quad P_2 \;\sim\; \epsilon\, a_2 L_\infty t,$$

as $t \to \infty$, which will be used to match with the solutions of the next timescale.

$t = O(1/\epsilon)$

We rescale time using $t = \tilde{t}/\epsilon$ and the matching conditions (25) suggest that we should write $B_1 = \epsilon \tilde{B}_1$ and $B_2 = \epsilon \tilde{B}_2$; these latter rescalings indicate that the occupation time of a complex bound-receptor is small in comparison to the diffusion timescale and thus bound receptors remain at a low density. Substitution of the new rescalings into (18)-(24) and solving provides at leading order

$$\begin{aligned} L \;&\sim\; L_\infty e^{-(1+a_2)\tilde{t}}, \quad \tilde{B}_1 \;\sim\; \tfrac{L_\infty}{b_1} e^{-(1+a_2)\tilde{t}}, \quad \tilde{B}_2 \;\sim\; \tfrac{a_2 L_\infty}{b_2} e^{-(1+a_2)\tilde{t}}, \\ P_1 \;&\sim\; \tfrac{\beta_1 L_\infty}{1+a_2}(1 - e^{-(1+a_2)\tilde{t}}), \quad P_2 \;\sim\; \tfrac{a_2\,\beta_2\, L_\infty}{1+a_2}(1 - e^{-(1+a_2)\tilde{t}}). \end{aligned}$$

The solutions of P_1 and P_2 are functions that decay exponentially to a constant, steady-state level, thus agreeing qualitatively the experimental results.

The aim of the analysis is to establish the outcome of the competition between two organelles for the same protein, and this can be assessed via the ratio of P_1 and P_2 as $t \to \infty$; in terms of the original dimensional parameters the ratio at steady-state is

$$\frac{P_1}{P_2} \;=\; \frac{\beta_1\,\alpha_1\,R_1^0}{\beta_2\,\alpha_2\,R_2^0}; \tag{25}$$

thus an organelle with greater surface area to volume ratio (β_n), and is more numerous or closer to the ribosomes (a_n) and has more receptors (R_n^0) will obtain a greater proportion of the protein. Though, this is not unexpected, the analysis leading to (25) provides a quantitative understanding of the effects the parameters have on the organelle-associated protein yield. We note that the ratio (25) in fact holds throughout this timescale and not just at steady-state.

Figure (5) shows three plots of the evolution of P_1 and P_2, in each case P_1 being the upper curve. In each simulation, the parameters are fixed, except, from top to bottom, $R_2^0 = R_1^0/2$, $\beta_2 = \beta_1/2$ and $a_2 = a_1/2$. In each case $P_2 \approx P_1/2$ as predicted by (25).

5 Discussion

During the week of the studygroup we derived and studied a mathematical model to describe the passage of proteins from ribosome to organelles (e.g. chloroplasts, mitochondria, endoplasmic reticulum) via a range of targeting factors. The modelling approach can be made to account

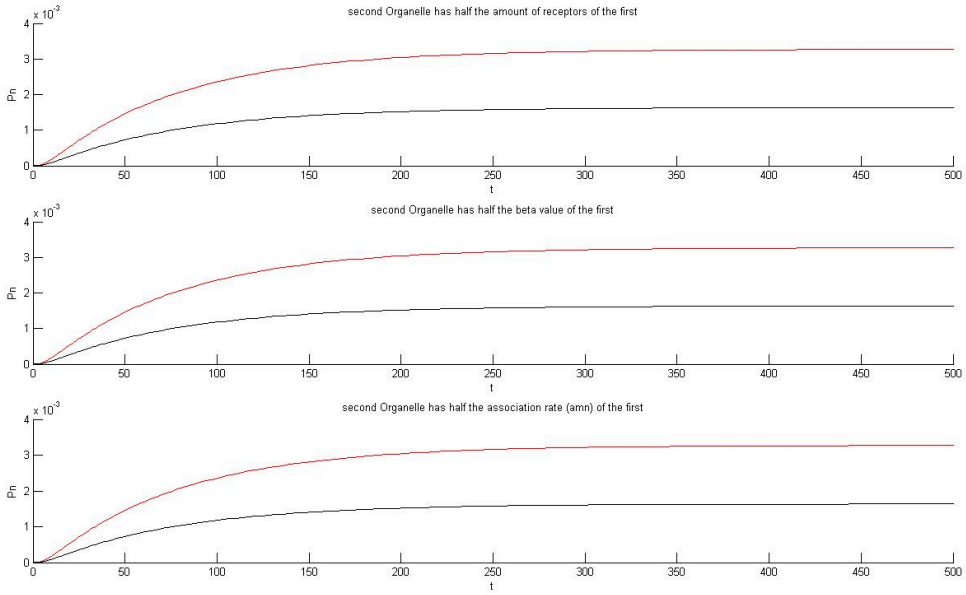

Figure 5: Plots comparing the evolution of the organelle associated protein yield P_1 and P_2; in each plot the upper curve is P_1. The parameters used here are the same except $R_2^0 = R_1^0$ (top), $\beta_2 = \beta_1/2$ (middle) and $a_2 = a_1/2$ (bottom).

for multiple organelles and targeting factors. We have attempted to establish reasonable estimates for the model's parameters, though some of these required significant extrapolation. In particular, it seemed at first reasonable to assume that the process' rate limiting step is that of diffusion from the ribosome to the organelle receptor site, however, our current estimates surprisingly suggest that it is the ligand-receptor binding kinetics that is time limiting. Refinement of the parameter values presented in this report is necessary and further experimentation would be needed to this end. Nevertheless, using the given parameter values, the simulations of the model do produce results that are qualitatively consistent with experimentally measured organelle associated protein yield; though, the timescale at which saturation levels are reached appears to be exaggerated by the model (about an hour as opposed to 5-10 minutes [2]). Nevertheless, the results so far are encouraging and suggests that further investigation is highly worthwhile. Furthermore, the model though complex, is amenable to mathematical analysis, from which model predictions can be expressed in terms of simple formula (i.e. the simulation of the full model system is not required); for example, the formula given by (25) for the comparative protein yield between two organelles.

References

[1] Abell, B. M., Pool, M. R., Schlenker, O., Sinning, I. and High, S. (2004). Signal recognition particle mediates post-translational targeting in eukaryotes. Embo. J. 23, 2755-64.

[2] Abell, B. M., Rabu, C., Leznicki, P., Young, J. C. and High, S. (2007). Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. J. Cell. Sci. 120, 1743-51.

[3] Artigues, A., Iriarte, A. and Martinez-Carrion, M. (2002). Binding to chaperones allows import of a purified mitochondrial precursor into mitochondria. J. Biol. Chem. 277, 25047-55.

[4] Bae, W., Lee, Y. J., Kim, D. H., Lee, J., Kim, S., Sohn, E. J. and Hwang, I. (2008). AKR2A-mediated import of chloroplast outer membrane proteins is essential for chloroplast biogenesis. Nat. Cell. Biol. 10, 220-7.

[5] Chew, O., Rudhe, C., Glaser, E. and Whelan, J. (2003). Characterization of the targeting signal of dual-targeted pea glutathione reductase. Plant Mol Biol 53, 341-56.

[6] Connolly, T. and Gilmore, R. (1993). GTP hydrolysis by complexes of the signal recognition particle and the signal recognition particle receptor. J. Cell Biol. 123,799-807.

[7] Flanagan, J.J. and Chen, J.C. and Miao, Y. and Shao, Y. and Lin, J. and Bock, P.E. and Johnson, A.E. (2003). Signal recognition particle binds to ribosome-bound signal sequences with fluorescence-detected subnanomolar affinity that does not diminish as the nascent chain lengthens. J. Biol. Chem. 278, 18628-18637.

[8] Goldstein, B. and Dembo, M. (1995). Approximating the effects of diffusion on reversible reactions at the cell surface: ligand-receptor kinetics. Biophys. J. 68, 1222-1230.

[9] Heber, U. and Heldt, H.W. (1981). The chloroplast envelope: structure, function, and role in leaf metabolism. Ann. Rev. Plant Physiol. 32, 139-168.

[10] Kriechbaumer, V., Shaw, R., Mukherjee, J., Bowsher, C. G., Harrison, A. M. and Abell, B. M. (2009). Subcellular Distribution of Tail-Anchored Proteins in Arabidopsis. Traffic.

[11] Lauffenburger, D.A. and Linderman, J.J. (1993). Receptors: models for binding, trafficking and signalling, OUP.

[12] Li, Z. and Srivastava, P. (2003). Heat-shock proteins, in Current protocols in immunology, Ed. J.E. Coligan. John Wiley and Sons.

[13] Mandon, E. C. and Jiang, Y. and Gilmore, R. (2003). Dual recognition of the ribosome and the signal recognition particle by the SRP receptor during protein targeting to the endoplasmic reticulum. J. Cell Biol. 162, 575-585.

[14] Ngosuwan, J., Wang, N. M., Fung, K. L. and Chirico, W. J. (2003). Roles of cytosolic Hsp70 and Hsp40 molecular chaperones in post-translational translocation of presecretory proteins into the endoplasmic reticulum. J Biol Chem 278, 7034-42.

[15] Qbadou, S., Becker, T., Mirus, O., Tews, I., Soll, J. and Schleiff, E. (2006). The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. EMBO J 25, 1836-47.

[16] Rabu, C., Schmid, V., Schwappach, B. and High, S. (2009). Biogenesis of tail-anchored proteins: the beginning for the end? J Cell Sci 122, 3605-12.

[17] Rabu, C., Wipf, P., Brodsky, J. L. and High, S. (2008). A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum. J Biol Chem 283, 27504-13.

[18] Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J. et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 940-3.

[19] Rüdiger, S., Schneider-Mergener, J. and Bukau, B. (2001). Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. EMBO J., 20, 1042

[20] Soll, J. (2002). Protein import into chloroplasts. Curr Opin Plant Biol 5, 529-35.

[21] Stefanovic, S. and Hegde, R. S. (2007). Identification of a targeting factor for posttranslational membrane protein insertion into the ER. Cell 128, 1147-59.

[22] Tyn, M. T. and Gusek, T. W. (1990). Prediction of diffusion coefficients of proteins. Biotechnol. Bioeng. 35, 327–338.

[23] Walter, P. and Johnson, A. E. (1994). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. Annu. Rev. Cell Biol. 10, 87-119.

[24] Young, J. C., Hoogenraad, N. J. and Hartl, F. U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell 112, 41-50.

[25] Zahedi, R.P., Völzing, C., Schmitt, A., Frien, M., Jung, M., Dudek, J., Wortelkamp, S., Sickmann, A., Zimmermann, R. (2009). Analysis of the membrane proteome of canine pancreatic rough microsomes identifies a novel Hsp40, termed ERj7. Proteomics, 9, 3463-3473.