Romanian Academy

Institute of Biochemistry

Structural and functional studies on the small GTPase Ypt7p and its interacting partners

A doctoral thesis

Ph.D. Student:
Alexandru Tudor Constantinescu

Ph.D. Supervisor:
Dr. Cecilia Motaş

Bucureşti, October 2001

"I asked the Great Sage 'How can I solve the secret of life?' He replied, 'The secret of life lies in the structure of proteins, and X-ray crystallography is the only way to solve it'. The 'Sage' was John Desmond Bernal (…) who had been the first to discover that protein crystals give detailed X-ray diffraction patterns (…)."

Max Perutz in "How the Secret of Life Was Discovered"

Acknowledgements

First of all, I owe my thanks to Dr. Kirill Alexandrov, who started this project. Without his constant help and encouragement during the whole project, this work would not have been possible. I have also to thank Dr. Axel J. Scheidig for introducing me to the field of crystallography and all his help on the crystallographic part of this project. Dr. Nicolas H. Thomä supervised many of the kinetic experiments described in this work and clarified to me the basics of the kinetics measurements; Nico was not only a good teacher, but also a good friend. I feel obliged to thank Prof. Roger S. Goody not only for being such a good boss, but, also, for the time and patience spent while explaining how to solve the problems I was encountering.

My Bucharest supervisors, Dr. Cecilia Motaş and Dr. Ştefana Petrescu ought to be acknowledged for the understanding and help they showed during all these years.

I have to thank Dr. Ilme Schlichting, Dr. Ingrid Vetter, Dr. Balaji Prakash and Dr. Louis Renault for the many hours they spent helping me on the crystallographic part of the work. The collaboration with Dr. Alexey Rak was extremely important for the construction and crystallization of Ypt7pΔC26. I must thank also Dr. Liviu Daia for his "remote technical support" which helped me to deal with the different Unix problems I was stumbling across.

My stay in Germany would not have been possible without the financial support provided by the Boehringer Ingelheim Fonds. They proved themselves not "yet another funding agency", but gave me the sentiment that they really care about their protégés and facilitated my participation to courses, seminars and meetings.

Outside the scientific work, I have to thank Emilia and Alice for all the discussions and exchange of opinions I had with them. Also, the members of the "Romanian community" in Dortmund had a major role in making life abroad so pleasant. Thank you Tudor, Betty, Cristina, Gina, Anca, Alexandra, Valeriu, Stefan!

Last but not at least, I have to thank my family for the education I received during my whole life and the moral support they gave me during these three years. Also, I have to thank my fiancée, Raluca, for the understanding and loving care she proved all these years far away from each other.

Dortmund, October 2001

CONTENTS

Crystallographic analysis of Ypt7p

Kinetic analysis of interaction between Mrs6p and different Ypt proteins

List of abbreviations

CHCA ..... 3,5-Dimethoxy-4-hydroxy cinnamic acid

Fpp ..... farnesyl pyrophosphate

FTase ..... farnesyl transferase

GGpp ..... geranylgeranyl pyrophosphate

GGTase I ..... Geranylgeranyl transferase

GGTaseII ..... Rab geranylgeranyl transferase

GppNHp ..... 5'-[β,γ-imido]-guanosintriphosphate (GppNHp)

MALDI-TOF ..... matrix-assisted laser desorption/ionization time-of-flight

MD ..... molecular dynamics

NMR ..... nuclear magnetic resonance

NSF ..... N-ethylmaleimide-sensitive factor

rmsd ..... root mean square deviation

SinA ..... sinapinic acid

SNAP ..... soluble NSF attachment protein

SNARE ..... SNAP receptor

Chapter I.

Introduction

I.1. Objectives of this work

Eukaryotic cells rely heavily on the compartmentalization of their cytoplasm; it is the only way they can perform multiple, sometimes almost contrary, reactions at the same time. This solution comes at a cost: the eukaryotic cell has to ensure that the transport of material is done in the correct way, that the vesicular traffic does not lead to problems. To this end, a whole subfamily of Ras–related small GTPases, termed Ypt in yeast and Rab in higher eukaryotes has emerged; in the yeast genome, it accounts for a third of the small GTPases.

Current data shows us that Ypt proteins play a major role in the catalysis/specificity of membrane traffic. Unfortunately, it does not (yet) tell us exactly in which substeps these proteins are implicated. Also, we do not know much about **how** the specificity is achieved and which are the molecular mechanisms that control the very complex endomembraneous system within the cell. Neither do we know what makes these Ypt proteins, structurally so much alike, the key players in a specificity process which requires the differentiation between different compartments. Another question which was unanswered at the beginning of this work was how big are the structural differences between the GTP- and the GDP- bound forms of the members of the family and if defining the "switch" regions based solely on the homology with Ras was a correct decision or not.

This work was done with the points outlined above in mind and with the hope of gaining some insight into the molecular basis of this specificity. To answer these questions, crystallographic analysis was combined with extensive kinetic characterization of different Ypt members.

The crystallographic analysis was performed on Ypt7p. This protein was truncated, in order to facilitate its crystallization, crystallized in complex with a non-hydrolyzable GTP analogue and its structure solved. Based on the semi-refined model, a second, more severe, truncation was performed. This second construct could be crystallized both in complex with GDP and a non-hydrolyzable GTP analogue and very high resolution structures were determined. Analysis of the models obtained, in comparison with other Rab/Ypt structures

emerged meanwhile, helped us to increase our knowledge on the specificity structural determinants and the structural changes that occur after the hydrolysis of GTP to GDP.

The purpose of the kinetic analysis was to gain a better understanding of the way prenylation is controlled. To this end, pre steady state kinetics methods, as well as equilibrium titrations were employed. By making use of a large array of techniques, the characterization of the interaction between the escort protein Mrs6p and a variety of members of the *Saccharomyces cerevisiae* Ypt family was achieved. The results obtained helped us to suggest a mechanism by which cells differentiate between the more and the less important members of the family.

In a trial to understand at molecular level the interaction between the Mrs6p and Ypt proteins, and since the complex did not seem amenable to crystallization, I tried to employ a new method of characterizing the interaction interface, using proton/deuteron exchange and mass spectrometry. Although it did not give information at high resolution and accuracy, the approach permitted pinpointing some regions involved in the interaction.

The experiments described in this work were performed between July 1998 and September 2001 in the Max-Planck-Institute Dortmund, Germany.

I.2. Ypt/Rab Proteins — Ras-like proteins involved in membrane fusion control

Eukariotic cells contain multiple lipidic membrane-bound compartments that are essential for their functioning. Exactly this compartmentalization gave them the evolutionary advantage over the prokaryotic systems. Each compartment is endowed with a specific proteic and lipidic composition and has a certain, well-defined role. A resulting problem is that the compartments, by themselves, are impermeable for the majority of macromolecules in the cell, such as proteins, glucides and ions and yet, such macromolecules have to be transported between compartments. This transport is achieved by what is termed "transport vesicles", which *bud* from the donor compartment, travel through the cytoplasm and *fuse* with the acceptor compartment, delivering there both their contents ("cargo") and lipid membrane components (Palade, 1975).

The budding process itself involves loading the vesicles with their cargo (Kirchhausen *et al.*, 1997) and the assembly of coat protein complexes (COP I and II, chlatrin etc.) (Wieland & Harter, 1999; Wendland *et al.*, 1998). After that, the vesicle is transported, with the help of cytoskeletal structures, to the destination (Sheetz, 1999), where it fuses with the target membrane in a process catalyzed by a very complex machinery, involving the Rab/Ypt small GTPases and the SNARE family proteins, to name just a few of them (Wickner & Haas, 2000; Jahn & Sudhof, 1999).

Cell's sustainable survival depends on the fidelity of cargo vesicles fusion with the right target membrane. First, because the surface aria of the different compartments (or that of the plasma membrane) has to be kept constant, despite the vesicular traffic (which comprises membrane transport as well as cargo) in which they are involved; moreover, there is the risk of non-specific fusion, a risk which has to be avoided in order to guarantee the survival of the cell.

According to the current view, the role of assuring a control on the specificity of the fusion event is assigned to small GTPases from the Rab/Ypt family (Wickner & Haas, 2000; Lazar *et al.*, 1997; Gotte *et al.*, 2000) (yeast proteins are termed *Ypt*, with the exception of Sec4p, while their counterparts from higher organisms are termed *Rab*).

GTPases are considered to be the cell's molecular switches (Bourne *et al.*, 1990); they shuttle between two states, a GTP–bound, active, and a GDP–bound, inactive; the timer of the switch is represented by the rate the GTPase is able to hydrolyze its substrate (alone or with the catalytic help of a GTPase Activating Protein, termed generic GAP). Currently, there are more than 700 genes known for small (≈ 21..23kDa), monomeric GTPases of the Ras superfamily; the yeast *Saccharomyces cerevisiae* contains 30 open reading frames (ORFs) coding such small GTPases (Garcia-Ranea & Valencia, 1998). Out of them, four belong to the Ras family, six to Rho, two belong to Ran, seven to the Arf/Sar family and eleven, i.e. more than a third, to the Ypt/Sec4 family. This alone gives a hint on the importance of the Ypt family. The general characteristics of a Ras-related protein are: a GTP-binding domain and (in most of the cases) a prenylation signal at the C-terminus. For the special case of Rab/Ypt proteins, there is not an obvious consensus C-terminal prenylation signal; instead, there are different motifs, like xxxCC, xxCCx, xCCxx, CCxxx or xxCxC. Moreover, these proteins exhibit a hypervariable C-terminal tail, which is believed to contain the information for subcellular localization.

I.2.1. Generic structure of a Ras-like protein and highlights on Ypt proteins

Generally speaking, all Ras-related GTPases (and Ypt proteins as well) have a very similar fold, owing to their high sequence homology (see **Figure I-1** for a sequence alignment). They consist of a six-stranded beta sheet, surrounded by five alpha helices. Two important secondary structure elements are the helices α2, part of the so-called "switch II region", and the C-terminal helix, α5, at the end of which isoprenoid groups are attached. Also, three loops, L1, L2 and L4 play important roles in GTP binding, hydrolysis and interaction with binding partners (see **Figure I-2** for structural element naming and position).

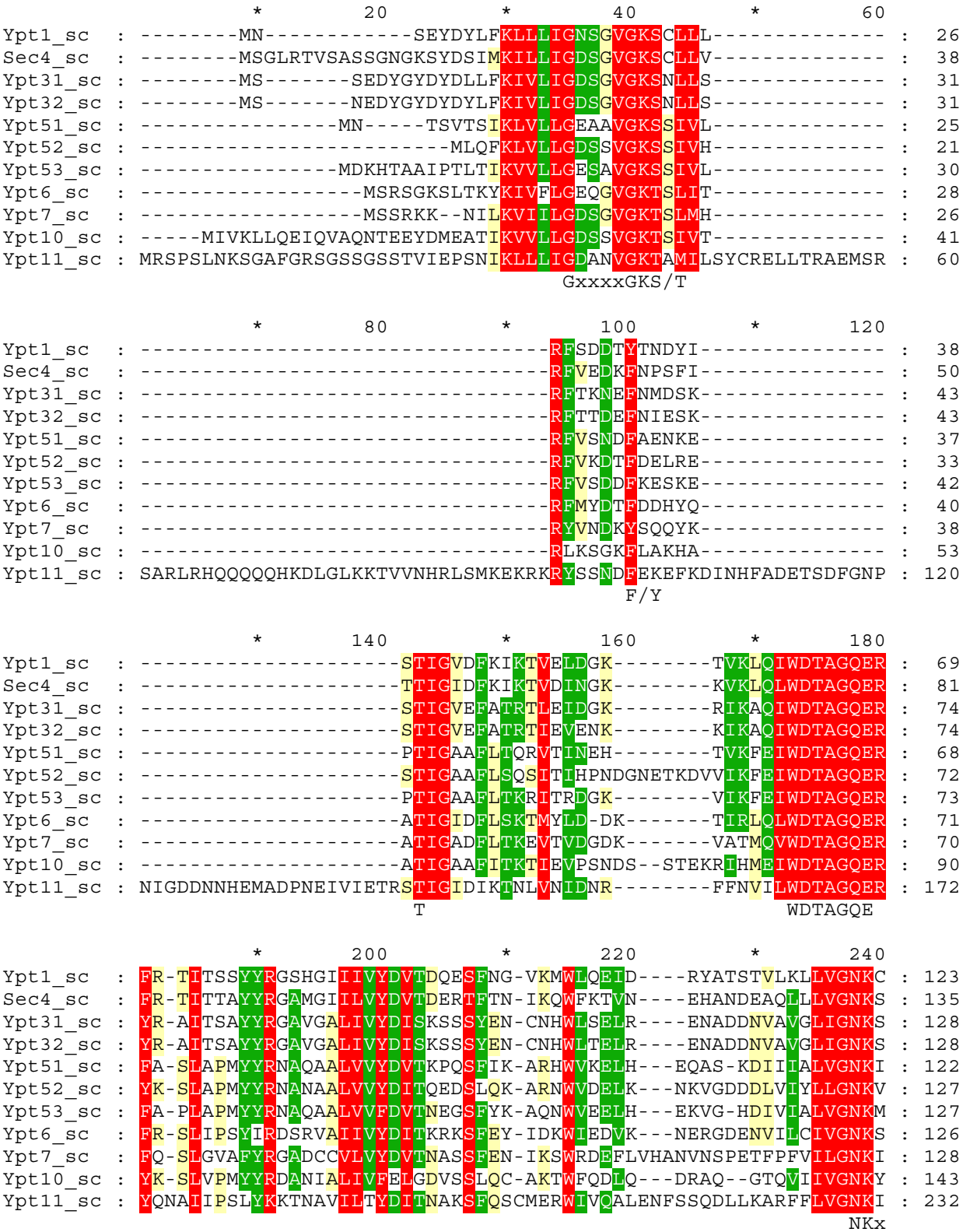


Figure I-1: Structural alignment of *Saccharomyces cerevisiae* Ypt proteins. Consensus motifs described below are marked on the bottom row of the alignment. It is worth noting that, except Ypt11p, the other members align very well. (continued on the next page)

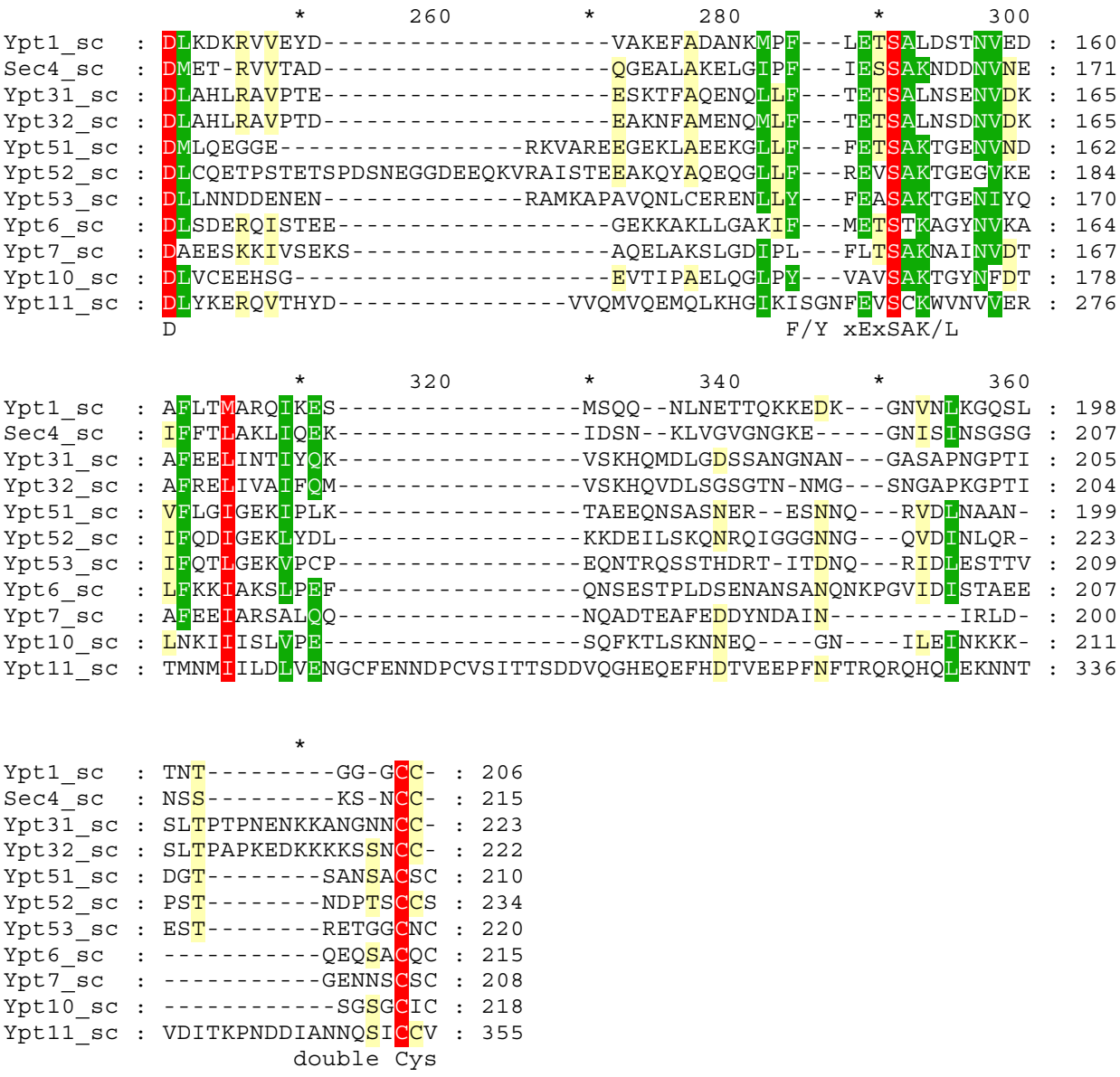


Figure I-1: (Continued)

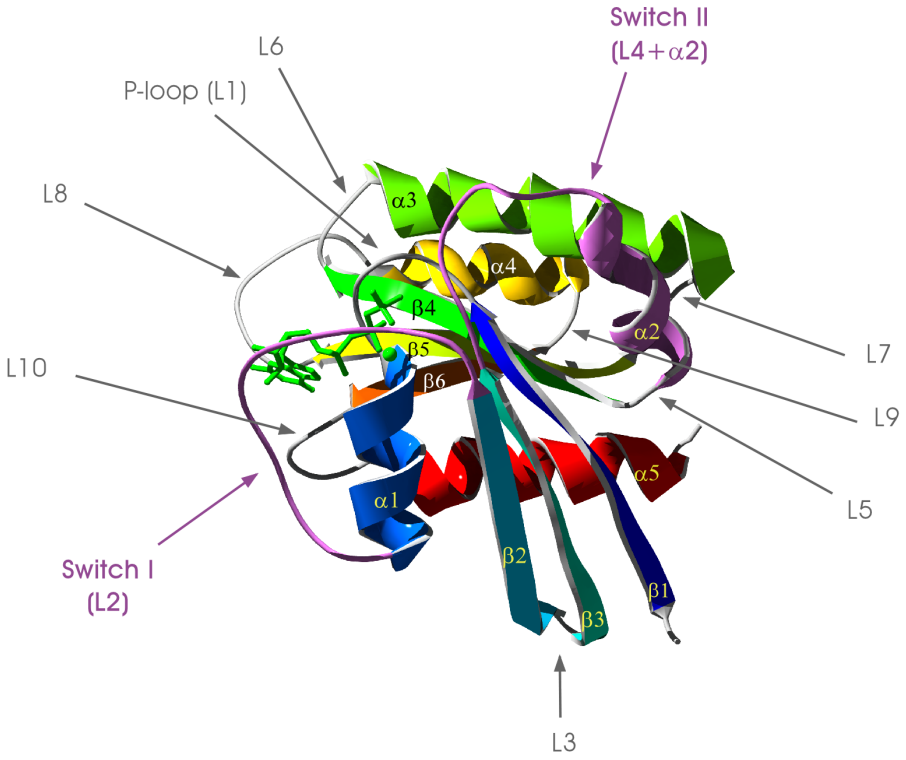


Figure I-2: The overall fold of H-Ras p21. The secondary structure elements of the protein are gradient-colored from dark-blue at the N-terminus to red at the C-terminus, the only exceptions being the loop L2 (Switch I) and the Switch II region (L4 and helix α2), depicted in pink.

Two major motifs define Ras-related protein and Ypt proteins are no exception to this rule as well. One is the nucleotide binding site and its associated "switch" regions, whereas the other one is the lipidation signal at the C-terminus. Specific for the Rab/Ypt proteins is the hyper-variability of their C-terminus, feature with important functional implications. These features will be described in the following subchapters.

I.2.1.a. The nucleotide binding site

comprises six conserved, small motifs. The main principle is that the binding pocket for the base should provide both hydrophobic and H-bonding interactions for the purine base and H-bonding and electrostatic interactions for the charged part (the phosphate moiety). Three motifs are involved in the binding of the phosphate groups and the Mg^{2+} ion (hence their P/M

name) while the remaining three are involved in the binding of the purine base (the G motifs) (Valencia *et al.*, 1991).

Motif PM1 (GxxxxGKS/T) is part of loop L1, called the P-loop (where P stands for *p*hosphate-binding). The name is justified by the observation that this sequence is present in many purine-nucleotide triphosphate (ATP/GTP) binding proteins (nucleoside and nucleoside monophosphate kinases, ribosomal elongation factors, the β subunit of ATP synthases, myosin heavy chains, phosphoglycerate kinases and, last but not at least, the members of Ras family of small GTPases). However, it should be noted that this motif is not found in *all* ATP/GTP binding proteins, and, if used alone in a database mining, might sometimes "fish" false hits (Saraste *et al.*, 1990). The P-loop is also known as "the G-rich" loop, since, in H-Ras, for example, it contains four glycines in eight residues. What are the features of this loop? First, the loop wraps around the β and γ phosphates with the main chain amide nitrogen atoms pointing "inside"; thus it helps to balance the high negative charge of the phosphate groups. The same role is played by the conserved lysine, which is in contact with the β and the γ phosphates; in the GDP state, the lysine coordinates only the β phosphate. Second, the serine or threonine binds the Mg^{2+} ion which, in turn, is coordinated by the phosphates β and γ.

Motifs PM2 (**T**) and **PM3** (**D**xx**G**Q/TE, W**D**TA**G**QE in all Ypt family members) involve a conserved threonine, aspartate and glycine (underlined) that bind the γ phosphate of the GTP. Threonine 35 —if we refer to the well-known H-Ras p21 structure (Pai *et al.*, 1989; Schlichting *et al.*, 1990; Pai *et al.*, 1990; Milburn *et al.*, 1990)— or T40 —in Ypt7p numbering— and glycine 60 (G67 in Ypt7p) bind the γ-phosphate *via* their backbone nitrogen, while aspartate 57 (D64 in Ypt7p) is part of the second coordination sphere of the Mg^{2+} ion. These residues are responsible for the conformational changes which occur after the hydrolysis of GTP to GDP: since they cannot bind the β phosphate, the whole loop which contains the threonine (L2) as well as the one which contains the W**D**AT**G**QE motif (L4 *plus* α2) undergoes large rearrangements. In Ypt7p, the threonine and the glycine move after the GTP hydrolysis about 10Å from their original positions (see also **Figure I-3**). These rearrangements make the GTPase competent for interaction with other molecules. Since loops L2 and L4 plus helix α2 transmit to the outer world the state of the GTPase (**switched** *on* or *off*), they are known also as the "Switch I" and "Switch II" regions. Moreover, for

reasons discussed below (see I.2.1.b), "switch I" is known also as the "effector loop", while "switch II" is also known as the "catalytic loop".

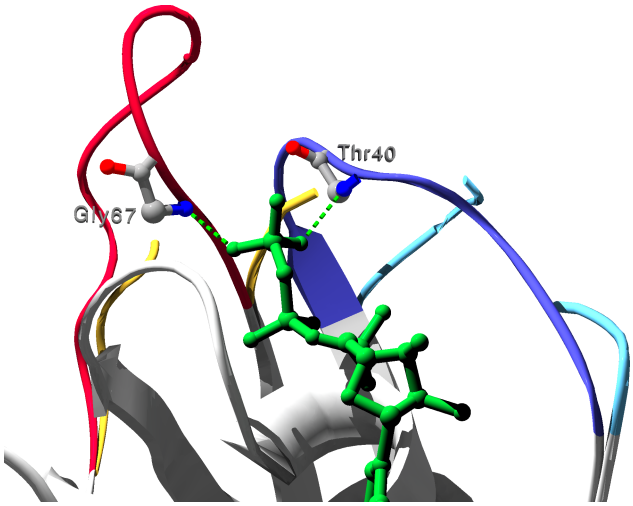


Figure I-3: Comparison of the hydrogen bonding pattern for the GTP and GDP states of Ypt7p. The parts that do not move significantly upon GTP hydrolysis are colored light gray. The nucleotide (here only the three phosphates and the sugar base visible) is green. "Switch I" region is depicted in dark blue for the GTP state, while it is light-blue in the GDP state. "Switch II" is red for GTP and yellow for the GDP state. Residues T40 and G67, responsible for anchoring the loops to the γ-phosphate, are represented in "ball and stick", with the H-bonds to the γ-phosphate indicated. Some parts of the loops (including residues T40 and G67) could not be traced in the electron density for the GDP case.

Motif G1 (F/Y) contains a phenylalanine or a tyrosine (in the Rab/Ypt family only the Rab7 and Ypt7p have a tyrosine) positioned perpendicular to the plane of the base. Strong interactions between the π electrons of the base and the aromatic aminoacid stabilize the contact with the purine base (see **Figure I-4**).

Motif G2 (NKxD) provides both hydrophobic and ionic interactions with the base. Two nitrogens from the base, N2 and, respectively, N1, are bound by the aspartate 119 (D129 in Ypt7p), with both its OD1 and OD2. This residue is very important, since it is one of the key elements involved in the high specificity of these proteins towards the guanosine base. Mutating this aspartate to asparagine (D→N mutation) switches the specificity of the GTPase from guanosine to xanthosine (Hwang & Miller, 1987; Hoffenberg *et al.*, 1995; Schmidt *et al.*, 1996). This altered specificity has proved to be a very useful way of controlling the activity of

a small G protein in a complex context, especially for *in vivo* experiments, where there are so many GTP-binding proteins involved and the activity of only one should be perturbed (Rybin *et al.*, 1996).

The same motif also provides a lysine which packs against the base by hydrophobic interactions and is H-bonding the O4' of the sugar ring with its terminal nitrogen.

Motif G3 (F/YxExSAK/L) does not contribute so much to the binding of the nucleotide. The serine H-bonds the O6 of the base (more weakly in the case of Ras than in the case of Ypt7p) while the lysine packs "above" the base, stabilizing it by hydrophobic interactions.

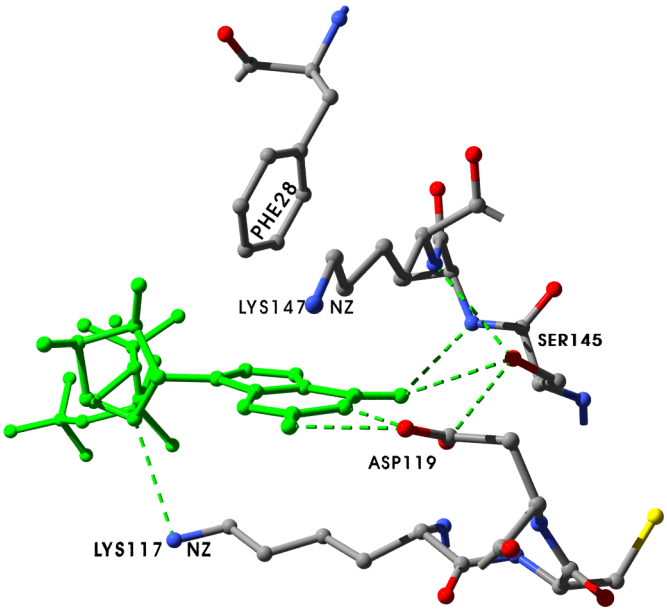


Figure I-4: The contacts between the base and the G1/G2/G3 motifs in H-Ras p21. Residue numbering according to H-Ras p21, coordinates derived from PDB entry 5P21.

The motifs PM1, PM3 and G2 were identified as early as 1987 (Dever *et al.*, 1987) as being the signature for all the GTP-binding proteins; they were named the "GTP binding domain" and had the consensus sequences: GxxxxGKT/S, DxxG and, respectively NKxD.

I.2.1.b. The *switch* regions

It is now generally accepted that all the G proteins shuttle between two states, a GDP-bound and a GTP-bound form. The balance between the time spent in the GDP and in the

GTP state is mainly dictated by the interaction of the protein with the GTPase activating proteins (GAPs) and the Guanine nucleotide Exchange Factors (GEFs), combined with the intrinsic GTPase activity. The GTP-bound is always the active state, while the GDP-bound is the inactive state. This is why the G proteins are considered to be "molecular switches" (Bourne *et al.*, 1991).

Following the determination of the X-ray structures of the elongation factor EF-Tu (also a GTPase domain closely related to Ras) in the GDP form (Jurnak, 1985) and the ones of p21 H-Ras in GDP form (de Vos *et al.*, 1988) and, later, in the GTP form (Pai *et al.*, 1989; Milburn *et al.*, 1990), it became clear that there are two main regions undergoing rearrangements as a consequence of GTP hydrolysis, namely the loops L2 and L4 and the helix α2. These regions have been named "switch I" (L2) and "switch II"(L4+α2).

In **switch I** the main element sensing the nucleotide state is the threonine residue from the PM2 motif, which binds the γ phosphate and the Mg^{2+} ion. Upon GTP conversion to GDP, this residue is not anymore part of the γ phosphate and Mg^{2+} coordination sphere and as a result the whole loop is displaced, in Ypt7p by more than 9Å (this work), which is a shift almost three times larger than that observed in Ras. In the triphosphate form, switch I is also stabilized in its conformation due to hydrophobic packing against Switch II. Since it was proven that, in the case of Ras, effectors interact with this loop, it was also named "effector loop". The structure of Rab3a bound to its effector molecule Rabphilin3 (Ostermeier & Brunger, 1999) shows, however, that the switch I region is only one of the many regions implicated in the complex formation, together with switch II and other regions (see the paragraph on "Rab3a-Rabphilin3 crystal structure" on page 17).

Switch II contains the "catalytic loop" L4 and the helix α2. The **D**TA**G**QE motif binds the γ phosphate *via* the glycine backbone nitrogen. Upon GTP hydrolysis, this loop also change its position dramatically and becomes highly unstructured; this is demonstrated by the very poor electron density and the high crystallographic thermal factors for its atoms. The glutamine residue (Q61 in p21 Ras) was shown to play a very important role in the catalysis, which led to the naming of loop L4 as "catalytic loop".

I.2.1.c. The C-terminus

Generally speaking, the C-terminus of almost all Ras-related GTPases (with the exception of Arf and Ran families) contains a prenylation signal to which an isoprenoid will be

attached; this posttranslational modification is essential for their reversible attachment to a biological membrane. Two features are special for the Rab/Ypt proteins: their C-terminus contains (with very few exceptions) **two** prenylated cysteine residues instead of the usual **one** (Seabra *et al.*, 1992). Moreover, a part of the C-terminus is responsible for the specific localization of the protein in the cell (Chavrier *et al.*, 1991).

The geranylgeranyl transferase type II adds, *via* a thioether linkage, a geranylgeranyl (C_{20} isoprenoid) moiety on each of these two cysteine residues (Rossi *et al.*, 1991; Ferro-Novick *et al.*, 1991). These two strongly hydrophobic groups are responsible for the reversible manner in which Ypt proteins are able to associate with the membrane.

A particularity of Rab proteins is the existence of a "*hypervariable C-terminal region*" which determines to which compartment the Rab protein will localize (Chavrier *et al.*, 1991; Stenmark *et al.*, 1994). It was shown that the last 34 residues (but not the last 13 or fewer) contained the targeting signal for Rab5. Similar results were observed with the yeast proteins Sec4p and Ypt1p: exchanging the last 40 residues of these two proteins showed that the localization was dictated by this part of the tail (Dunn *et al.*, 1993; Brennwald & Novick, 1993). However, the last 20-24 residues of Rab7 are flexible in solution (Neu *et al.*, 1997); the same holds true for crystals of Ypt51p, where at least the terminal 20 residues are not visible in the electron density (Esters *et al.*, 2000), as well as the last 22 residues in Ypt7p (this work), so this behavior cannot be explained by a certain structuring of the tail.

I.2.1.d. Variability among Rab/Ypt proteins

The striking feature of Rab/Ypt proteins is the fact that they have a high homology and, in spite of it, control a number of diverse processes. The experiments described in the paragraph I.2.1.c proved that the hyper-variable C-terminus is important for the localization of the proteins. However, in both cases, the tail was dictating only the localization of the protein, but not the function. The same studies, along with others, have attempted to define the regions of the proteins able to render the chimeric proteins functional. Three approaches were employed: construction of molecular chimeras and *in vivo* assessment of their effects, identification of consensus sequences by sequence analysis and interpretation of the data obtained from the structure determination of a Rab-Rab effector complex.

Construction of molecular chimeras:

In order to analyze the functional determinants of Ypt/Rab proteins, more complex chimeras were constructed, containing not only the C-terminus of Ypt1p grafted onto Sec4p, but also other different structural domains. The results showed that grafting loops L2 ("effector loop") and L7 together with helix α2 (part of the "switch II" region) of Ypt1p onto Sec4p rendered a protein able to behave like Ypt1p, both in terms of localization and function. Unexpectedly, the chimera between β3+α2+L5+L6+α3+L7 (the segment between the PM3 and G2 motifs) of Ypt1p and what was left from Sec4p was able to function as both Ypt1p and Sec4p, albeit with low efficiency. This evidence suggested, however, that the specificity determinants for each protein are not exactly the same.

In order to gain a better understanding of "what makes a Rab a Rab", Zerial and colleagues constructed several finer sampled chimeras between canine Rab5 and human Rab6 (Stenmark *et al.*, 1994). Based on sequence alignments of many Rabs (Valencia *et al.*, 1991), they first assessed the role of the N- and C-termini. Similar to the yeast system, the C-terminus alone was able to code for the correct localization, but not for functionality in the new context. Inserting the Rab5 α2 helix, part of the "switch II" region in the Rab6 led to an unexpected result: the chimera was not anymore attached to the membrane, but it was soluble. This suggests that Rab5 α2 might not be compatible with Rab6 α3; to test this hypothesis, a new chimera, containing the Rab5 α2/L5 and α3/L7 on the Rab6 frame (always with the Rab5 tail) resulted in a correctly localized protein, still missing the new functionality. The functionality could be partly (≈ 50%) restored only after adding also a part of the N terminus, namely the first 20 residues of Rab5 instead of the first 13 of Rab6.

The conclusion that could be drawn from these experiments was that there are many specificity determinants on the surface of any Rab/Ypt protein and, also, that they are not localized to a certain surface of the molecule, but on two "faces" of the protein: the N- and C-termini, together with the L7 on one side and mainly the α2 and α3 helices on the other.

Sequence analysis of the Rab family

Taking advantage of the increasing number of proteins deposited in the public databases (more than 700 Ras-related GTPases) Seabra and co-workers (Pereira-Leal & Seabra, 2000) performed a multiple sequence alignment on a representative of each Rab subfamily. Their main goal was to identify sequences which would be specific "tags" for all the

Rabs. This was not only because of the (still) lack of understanding of what defines a Rab/Ypt protein, but, as well, because different proteins of the family were submitted with "non-standard" names, so that there was a lack of nomenclature. Also, once identified a Rab consensus sequence, this might be used to search the database with the hope of finding new, as of yet not identified, members of the family.

The model set comprised 50 mammalian Rab GTPases. As expected, six regions aligned to almost 100% identity: the three phosphate-binding and the three guanine-binding motifs. However, five other regions could be identified, very conserved among Rabs and able to distinguish Rabs from Ras/Rho and Arf.

These five regions were denominated "Rab family motifs" and were labeled RabF1 to RabF5. RabF1 was localized in the *switch I* region (see **Figure I-5**). The other four family determinants were found to cluster in, or close to, the *switch II* region; RabF2 was localized in the sheet β3, RabF3 and RabF4 reside in the loop L4/α2 and RabF5 is in the β4. Based on these consensus sequences, a search in the databases identified 52 different Rabs, so two new.

This extensive alignment exercise delivered yet another conclusion: examining the phylogenetic tree, several clusters of "related" Rabs could be identified. The sequence analysis revealed four Rab **sub**family determinants, termed RabSF1 to RabSF4. They are characterized by very high aminoacid identities for the members of the same subfamily (between 80 and 90%), in high contrast with the sequence identity at the level of the whole Rab family (around 20-30%). Exception is RabSF4, for which the identity is 58% in the subfamily and only 14% in the whole family. These very low figures can be explained taking in account the fact that this subfamily determinant maps to the C-terminus of the proteins, in the so-called *hyper-variable* domain. RabSF1 was identified at the very N-terminus of the proteins. The other two subfamily motifs map to the α helices 1 and 3.

These results help to provide a more molecular explanation of the experiments performed by the groups of Zerial (Stenmark *et al.*, 1994) and Novick (Brennwald & Novick, 1993). In these experiments, the N- terminus, the α2/L5, α3/L7 and C-terminus were swapped between Rab5 and Rab6, in order to obtain a semi-functional chimeric protein. These regions are, by the definitions mentioned above, the RabSF1, RabF3 and RabF4, RabSF3 and, respectively, RabSF4.

It is, however, worth mentioning that the last two "consensus" sequences (for RabSF3 and RabSF4) cannot be easily mapped onto the members of the Ypt family, the yeast

counterparts. However, the Dunn (Dunn *et al.*, 1993) and Novick (Brennwald & Novick, 1993) experiments with Ypt1p/Sec4p chimeras showed that swapping the regions corresponding to RabSF2 and RabSF3 were necessary and sufficient to generate an active chimeric protein and that the C-terminal tail region of Ypt1p, termed the HV region in the Novick experiments, (and corresponding to RabSF4) was the solely necessary to drive Sec4p (a plasma membrane–associated protein) to the Golgi membranes, where Ypt1p resides normally.

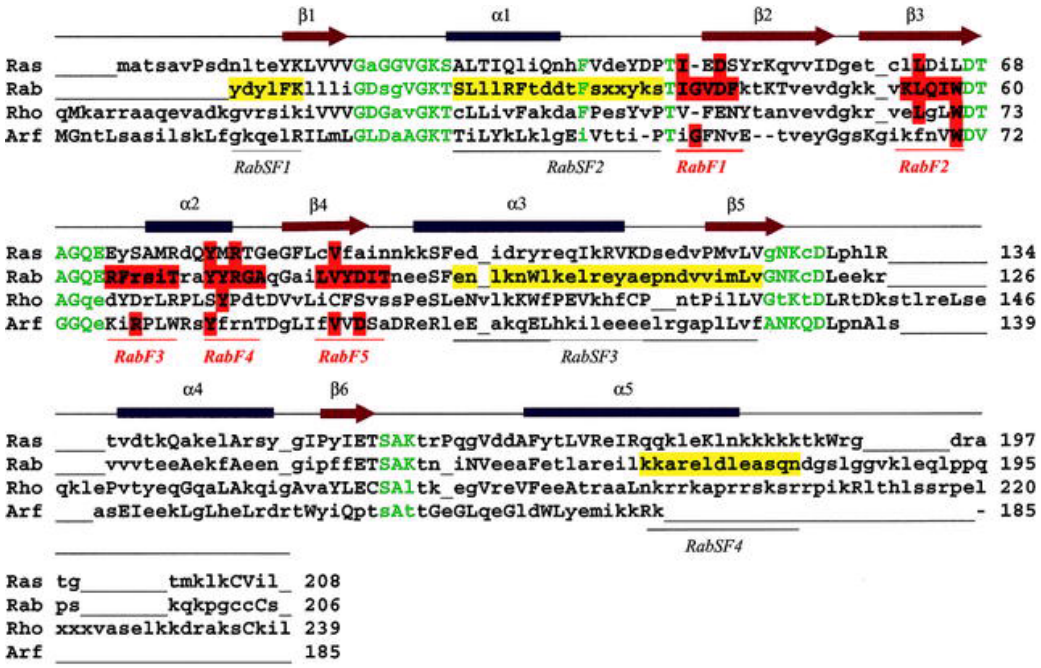


Figure I-5: Alignment of profile HMM model sequences (manually adjusted to accommodate structural considerations). The uppercase/lower case coding represents the results of the profile Hidden Markov Model method, in which uppercase characters were found at p>0.5. Residues found to be Rab specific are highlighted in red. When a position is conserved in other families, the corresponding position is also highlighted in red. Green characters denote the conserved nucleotide binding (PM/G) motifs. From (Pereira-Leal & Seabra, 2000).

Rab3a-Rabphilin3 crystal structure

The "final proof" concerning the structural specificity determinants was provided by the group of Axel Brunger which solved by X-ray diffraction the structure of the Rab3A:GTP-Rabphilin3 complex (Ostermeier & Brunger, 1999) (with a Rab3A mutant, Q81L, locked in the GTP state and only the effector domain of rabphilin-3A). Rabphilin3 is an "effector molecule" which binds Rab3A in the GTP state; not much is yet known about Rabphilin3 function downstream of Rab3A, but it is known that it is implicated in neurotransmiter release.

The structure of the complex explained many of the results described previously. Rabphilin3 binds Rab3A at two interfaces, roughly of the same size (1300Å^2 and, respectively, 1500Å^2). Most of the interactions are of hydrophobic nature, but some of them are also based on charge complementarity.

For interface A, Rab3A participates with the *switch* I and II regions, while Rabphilin3 participates with helix α1 and the C-terminus. The interface B is composed of Rabphilin3 C-terminal end of α2 helix and a motif next to it (SGAWFF), which binds the Rab3A N-terminus, C-terminus and the loops L5 and L7. These loops seem to be stabilized as a consequence of their interaction with Rabphilin, since in the structure of Rab3A alone (Dumas *et al.*, 1999) they exhibit a very weak electron density, indicative of flexible, non-structured regions. It should be mentioned that during the Rab3A-Rabphilin crystallization trials, an "engineered" version of Rab3A proved to be the best candidate; this version is lacking the first 18 aminoacid residues at the N-terminus, as well as the last three residues. However, the structure is still biologically relevant, since a ΔN18 version of Rab3A was able to bind Rabphilin3, while a ΔN21 version had no affinity at all for Rabphilin.

Corroborated with data on binding partners for other Rabs, the X-ray structure suggested that the N-terminus of Rab proteins, together with loops L5 and L7 and part of the C-terminal helix α5, form a "Rab Complementarity Determining Region" or, shortly, Rab CDR. It is noteworthy that the Rab CDRs for other Rab proteins have a sequence different from that of Rab3a and that these other Rabs cannot bind to Rabphilin3.

The data presented here strongly suggest that the Rab family determinants are present in some well defined regions and that the interactions with the effectors are based on two distinct mechanisms: the conserved *switch* regions are used mainly to transmit the GTP- or

GDP-bound state of the GTPase, while specific sequences in the hyper-variable N- and C-termini, as well as in loops L5 and L7, facilitate the interaction with a cognate partner, thus preventing other effectors from binding.

I.2.2. The prenylation machinery

Protein prenylation is a type of posttranslational modification consisting in the covalent attachment of an isoprenoid (be it farnesyl $–C_{15}–$ or geranylgeranyl $–C_{20}–$) via a thioether linkage to cysteine residues, usually near the C-terminus of the protein. These prenylated proteins play very important roles in signal transduction or intracellular traffic (Casey & Seabra, 1996), accounting for up to 2% of the cell proteins (Glomset & Farnsworth, 1994). The lipidic substrate is, in both cases, an ester with a pyrophosphate; hence, they are named farnesyl pyrophosphate (Fpp) or geranylgeranyl pyrophosphate (GGpp).

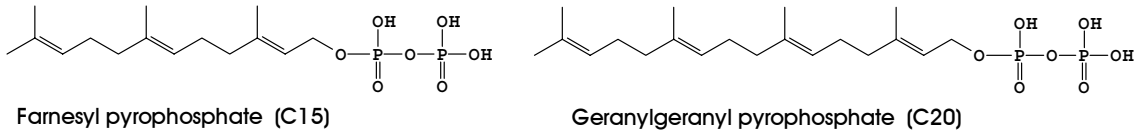


Figure I-6: Farnesyl- and geranylgeranyl pyrophosphates

There are three types of protein:prenyl transferases (here thereafter named, for simplicity, just prenyl transferases), grouped in two major classes: the CaaX prenyl transferases and the Rab geranylgeranyl transferase (Rab GGTase or, even shorter, GGTaseII).

I.2.2.a. CaaX prenyl transferases

This family includes farnesyl transferase (FTase) and geranylgeranyl transferase type I (GGTase I). Both enzymes recognize the C-terminal consensus sequence CaaX of the protein to be modified. The difference in specificity between FTase and GGTase I arises from the residue in the fourth position of the consensus sequence, namely the **X**. While FTase has as substrate peptides ending in Ser, Met, Ala or Gln, GGTase I recognizes peptides ending in Leu.

Farnesyl transferase was the first prenyl transferase to be identified and characterized (Reiss *et al.*, 1990), mainly in connection with the posttranslational modification of Ras-p21. The enzyme was shown to have specificity towards a tetrapeptide containing the prenylation signal Cys-aaX, where ***a*** signifies an aliphatic residue and ***X*** signifies a variety of residues. It was shown that the protein consists of two subunits, named initially α and β and weighting, respectively, 48kDa and 43kDa. The same study demonstrated that a peptide as short as four aminoacids was a good substrate for the enzyme. Different experiments showed that the β subunit was the one to which the proteic substrate was binding to, suggesting the α subunit to be the prenyl phosphate carrier (Reiss *et al.*, 1991). In time, different other studies and, more important, the X-ray structure (Park *et al.*, 1997) showed that the β subunit is the one binding both the lipid and the substrate. Among the proteins modified by this enzyme are the members of the Ras superfamily, nuclear lamins A and B, the γ subunit of transducin, rhodopsin kinase (Zhang & Casey, 1996; Glomset & Farnsworth, 1994).

Geranylgeranyl transferase type I was, later, shown to consist also from two subunits, again termed α and β. Remarkably, the α subunit is shared between FTase and GGTase I (Seabra *et al.*, 1991). This called for a small change in the names of the subunits, which are now known as $\alpha_{F/GGI}$ and β_F or β_{GGI}. Known substrates for the GGTase I are the members of the Rac/Rho family, most of the γ subunits of heterotrimeric G proteins (Casey & Seabra, 1996).

I.2.2.b. Rab geranylgeranyl transferase

Many substrates of GGTase I and FTase have enhanced mechanisms in order to help them correctly localize after the prenylation. Some of them, like Ras proteins (with the notable exception of K-(B)Ras), undergo another posttranslational modification, which is palmitoylation of cysteine residues upstream from the C-terminus. Others, like members of the Rho family, have stretches of polybasic aminoacid residues, able to electrostatically interact with the negatively charged phospholipids in the biological membranes. It appears that just having one C_{15} or C_{20} lipid attached to the C-terminus is not enough to keep the protein anchored to the membrane — although there are some exceptions (Rab8 and Rab13 have just one C-terminal cysteine (Zahraoui *et al.*, 1994)). Rab proteins seem to solve the problem in another way: they have *two* cysteines to be geranylgeranylated, instead of only one.

This special modification is conferred by another prenyl transferase, geranylgeranyl transferase II.

Unlike Ras/Rho/Rac, which exhibit the CaaX box as a signal for the prenylation, there is considerable heterogeneity in the C-terminal sequence of the Rab/Ypt proteins and there is no real consensus sequence for double geranylgeranylation. This might be one of the reasons for which GGTase II is not able to recognize/bind alone its proteic substrate and the prenylation of Rab/Ypt proteins needs an additional factor, a so-called "escort protein". This polypeptide is known as "Rab Escort Protein" (REP) in higher eukaryotes (Seabra *et al.*, 1992) and under the more cryptic name of Mrs6p in yeast (Jiang & Ferro-Novick, 1994). Mrs6p is a 66kDa soluble protein which appears to have multiple tasks. First, it binds/recognizes Ypt proteins and presents them to the transferase; secondly, once the Ypt proteins are modified by the addition of the hydrophobic moieties, it might shield the prenyl groups from the aqueous environment of the cytosol, hence preventing protein aggregation/ precipitation prior to its delivery to the donor-membrane; finally, Mrs6p is able to escort newly prenylated Ypt proteins to the target-membrane and deliver them there. However, new evidence suggests that Rep is not completely indispensable for delivery to the **correct** membrane. A recent study (Overmeyer *et al.*, 2001) employed a Rab1B mutant (Y78D) that was unable to bind to Rep or GDI. Replacing the C-terminal GGCC residues of Rab1B(Y78D) with a CLLL motif permitted the protein to be prenylated by GGTase I, but not by GGTase II, *in vitro* as well as *in vivo*. Surprisingly, this double mutant was proven to be correctly delivered to the ER and Golgi compartments, like its wildtype counterpart. Use of a dominant-negative mutant showed that the protein was fully functional and able to inhibit the ER to Golgi transport, demonstrating that the different prenylation did not have a major impact on its functionality.

Rab Geranylgeranyl transferase (or geranylgeranyl transferase II, or, even shorter, GGTaseII) is a heterodimeric protein; in yeast, the genes *BET2* and *BET4* (initially called *MAD2*) code for the two subunits, each about 35kDa in mass. The mammalian enzyme is composed of an α and a β subunit; these subunits (about 65 and, respectively, 42kDa) resemble the corresponding subunits from the farnesyl transferase and the GGTase I.

Indeed, the X-ray structure of rat GGTase II (Zhang *et al.*, 2000) shows that it has a fold similar with FTase. The α subunit is formed of 15 α-helices organized in a crescent-shaped, right-handed superhelix. The root mean square deviation (rmsd) between

corresponding Cα atoms in the α subunits of FTase and GGTase II is 2.1Å, at an aminoacid identity of only 22%. The β subunit has an α-α barrel fold composed of 12 helices; the center of the barrel forms a funnel-shaped pocket lined out with mostly aromatic residues, obviously the prenyl binding site. The bottom of the pocket is defined by a turn near the C-terminus. Compared with the β subunit of FTase, the rmsd is even smaller: 1.4Å for 280 superimposable Cα. However, there are, also differences between these two proteins. The GGTase II contains two extra domains, both located in the α subunit: an Ig-like and a leucine-rich repeat. It can be speculated that these extra domains are the ones implicated in the recognition of the REP-Rab complex.

Studies by Seabra (Desnoyers & Seabra, 1998) showed, before the publishing of the X-ray structure, that, surprisingly, there is a single prenyl binding site in the GGTaseII. This is somehow counterintuitive, since Rab proteins are known to be doubly prenylated (Seabra *et al.*, 1992). This finding suggested (together with the crystallographic structure of rat GGTase II) that the two prenyl groups are transferred in a successive manner onto the two C-terminal cysteines. The next question which arises is whether the two rounds are tightly coupled or not, namely whether the prenylated intermediate product is released from the complex or not. Recent evidence (Thomä NH, ms. in preparation) suggests that it is more likely that the prenylated intermediate does not leave the prenylation complex; hence, just some rearrangements take place at the active site, in order to shift the second cysteine in the active site and to give the possibility of reloading the enzyme with GGpp. This might also be a good reason for needing a third component for the prenylation machinery (apart from the transferase itself and its proteic substrate); this third component —the escort protein— would stabilize the interaction between the transferase and the small GTPase during the rearrangement. At this moment, it is not clear which is the cellular source of GGpp for the prenylation reaction: soluble, membrane-bound or the GGpp synthetase (Zhang *et al.*, 2000). In any case, the C-terminal tail of the Rab/Ypt protein has to move away from the active site, in order to leave access for the GGpp to the hydrophobic pocket. This means that the binding of the C-terminal tail of the substrate has to be more flexible and less tight than in the case of the CaaX prenyl transferases; in turn, this leads to the conclusion that the escort protein has the task of ensuring that, during this rearrangement process, the semi-prenylated Rab does not leave the prenylation machinery.

I.2.3. The cycles of a Ypt protein and its interacting partners

Rab/Ypt proteins are involved in different cycles. First of all, they cycle between a GTP-bound state and a GDP-bound state; in addition, they cycle between a membrane-bound and a "soluble" (cytosolic) form. Moreover, the membrane-bound form can be subdivided in vesicle- and target membrane-bound states.

I.2.3.a. GTP *vs.* GDP state

As the GTP switch hypothesis states (Bourne *et al.*, 1990), Ypt proteins cycle between a GTP state, in which they are "active" and able to fulfill their role, and a GDP state, in which they are "switched off" to prevent undesired membrane-fusion events. However, since Ypt proteins are slow GTPases, with rates of about $0.005min^{-1}$ (Wagner *et al.*, 1987), they need the help of a GTPase activating protein, GAP.

I.2.3.b. Membrane-bound *vs.* soluble

The GTP/GDP cycle is associated with another cycle: the shuffling of the protein from the membrane to the cytosol and back on the membrane. After being synthesized in the cytosol, Ypt proteins interact with the geranylgeranyl transferase and a Rab escort protein (Mrs6p in yeast, REP in higher eukaryotes) and are posttranslationally prenylated (Casey & Seabra, 1996). After this step, Mrs6p delivers the protein to the donor membrane. Following vesicle budding and fusion, the Ypt protein ends up on the target membrane, hydrolyzes the GTP to GDP, and can be extracted by a protein termed GDI (for GDP dissociation inhibitor) and delivered back to the donor-membrane.

I.2.3.c. Donor– *vs.* acceptor– membrane bound

As mentioned earlier, the Ypt GTPases are, after their prenylation, delivered to the donor membrane. From there, they travel, bound to the cargo vesicle, to the receptor membrane. When the cargo vesicle finds the proper target membrane, Ypt proteins help in the docking stage (Ungermann *et al.*, 2000; Ungermann *et al.*, 1998) (see below) and, following membrane fusion, end up on the target membrane. However, since they cannot supervise correct fusion while attached to the target membrane, they have to be extracted and recycled to the donor membrane.

I.2.3.d. Schematic overview

The three cycles presented above might be schematically represented as in Figure I-7:

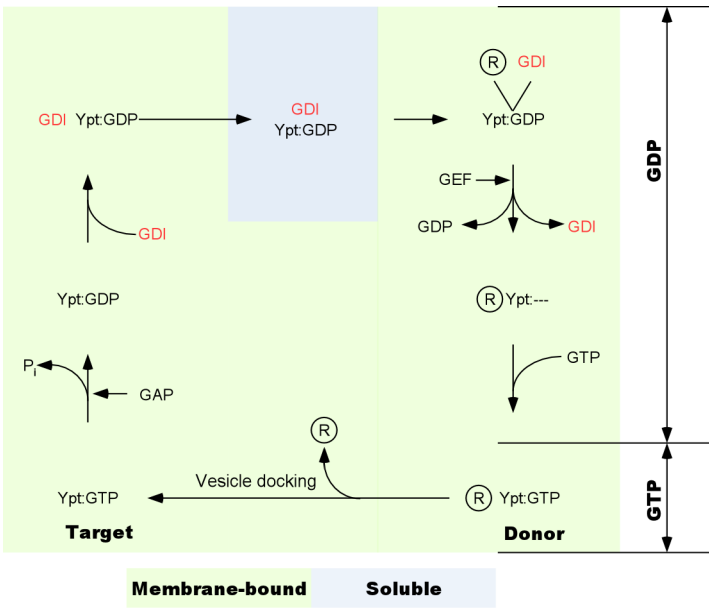


Figure I-7: Overlapping the GTP/GDP-bound, target/vesicle and membrane-bound and soluble cycles of a Ypt protein. ® = putative Ypt receptor (After (Gotte *et al.*, 2000))

It will be noted from the scheme above that there is overlap between the three cycles. Starting with the newly-prenylated Ypt protein (which is, only in the first round, escorted by Mrs6p, not Gdi1p) in a soluble form, the protein is then inserted in the appropriate membrane upon recognizing/being recognized by a putative membrane receptor. Once inserted into the membrane, the GDP molecule has to be exchanged for a GTP molecule in order to yield an active, "ready to fuse" Ypt protein. This is accomplished with the help of a guanine-nucleotide exchange factor (GEF, (Jones *et al.*, 1995)). A known GEF for Ypt proteins is the one for Sec4p, termed Sec2p. As its name suggests, this protein was found (like its substrate) while screening for mutants which affect the *sec*retory pathway; the *SEC2* gene is also essential, as is the case with *SEC4*.

Once in the GTP form, the Ypt protein is activated and can participate in the fusion process. However, after it plays its role, the Ypt has to be "switched off"; to this end, since the intrinsic GTPase rate is extremely low, the involvement of a GTPase activating protein

appears necessary. The first GTPase stimulating activity was discovered for Ypt6p in the group of Dieter Gallwitz (Strom *et al.*, 1993) and termed Gyp6p— **G**TPase activating protein (GAP) for **Y**pt **p**roteins. Later on, *GYP7* and *GYP1* genes were identified and cloned (Vollmer & Gallwitz, 1995). These Gyps accelerate the slow GTPase rates ($10^{-4}..10^{-3}s^{-1}$) by a factor of ca. 10^5 (Albert *et al.*, 1999)

Although not essential for the cell's viability (Rybin *et al.*, 1996; Eitzen *et al.*, 2000), the Ypt cycle between compartments is completed once the protein is returned to the starting membrane. Responsible for extracting the Ypt proteins from the target membrane and reinserting them in the donor membrane is a protein termed Gdi1p (first known as Sec19p). As the name suggests, the protein was found to block release of GDP from the inactive Ypt protein; however, its main feature is the capability of extracting **only** the GDP form of the protein from the membrane (Wu *et al.*, 1996).

I.3. The fusion process and main proteins implicated

Membrane fusion can be defined as the merger of two (phospho)lipidic bilayers present in an aqueous medium. In the case of artificial lipid bilayers, fusion can proceed by itself, without the help of any "catalyst"; on the other hand, this implies that there is neither specificity nor the chance of preventing fusion of vesicles which should not fuse.

Biological membrane fusion is a far more complex process, requiring a large number of proteins. However, the three *major* players known at the moment are the Rab/Ypt proteins, Sec1/Munc18 homologues (SM proteins) and proteins of the SNARE family. If the latter give the force necessary for the fusion to proceed, the Rabs are thought to be the ones to provide the specificity. For intracellular fusion events, there are two model systems studied: one is the yeast vacuole and the other is the mammalian synapse.

The use of vacuole system relies on the power of yeast genetics, enabling manipulation of the proteins that are expressed *in vivo* and monitoring the results; moreover, there is an *in vitro* assay for fusion that can be easily monitored (Ungermann *et al.*, 1998). The system is based on two yeast strains; one strain is deficient in vacuolar lumenal proteases, while the other one has the alkaline phosphatase gene deleted. Hence, one cell will have a vacuole containing pro-alkaline phosphatase, while the other one lacks the phosphatase, but has the proteases required to process the proenzyme. Upon fusion of vacuoles from both strains, the

proenzyme is converted very fast to the fully active enzyme; the extent of the reaction can be easily monitored.

The advantage of the mammalian synapse system is its ability to combine biochemistry, (mouse) genetics and electrophysiology (which allows the real-time measurement of the progress of the fusion reaction, thing not possible in the yeast system). No matter how different these two systems are, the conclusions one gets from investigating them are almost the same, suggesting that common mechanisms are used for fusion in organisms as different as fungi and superior mammals. However, since the order of the events is different in synapse as opposed to the vacuole system, I will refer here mainly to the yeast system.

The fusion process can be divided in three major steps: priming, docking and fusion. Priming consists of different reactions rendering the vesicles "ready for fusion". During docking, the primed vesicles attach to each other; this is followed by opening and extending a fusion pore — the fusion itself. The schematic drawing in **Figure I-8** should help in following these steps.

SNAREs are membrane proteins with high propensity for forming coiled-coil structures. Although weakly conserved on the sequence level, they share a so-called "SNARE motif" of about 60 aminoacid residues; these SNARE motifs mediate the assembly of SNAREs in the so-called "core complex" during fusion. This comprises an extremely stable coiled-coil structure built from four α-helices. If all the members of the core complex are anchored to the same membrane, the complex is called *cis*, while, if one member is on a different membrane, the complex is called *trans*. The term SNARE stands for **s**oluble **N**SF **a**ttachment protein **re**ceptor, where NSF is the abbreviation for *N*-ethylmaleimide-sensitive factor; this factor, actually a protein, was proven to be an ATPase that, in conjunction with SNAP proteins, breaks the coiled-coil structure of *cis* SNARE pair during the priming step. According to the *SNARE hypothesis* (Sollner *et al.*, 1993), the assembly of the "core complex", which consists of SNARE molecules placed on both vesicle and target membrane (hence *v-* and *t-* SNAREs), provides the necessary energy for bringing the lipid bilayers close enough so that the fusion can occur. However, SNAREs by themselves are not sufficient to provide specificity (Gotte & von Mollard, 1998) and SNARE pairing is far from being either the only or the last step in vesicle fusion (Price *et al.*, 2000).

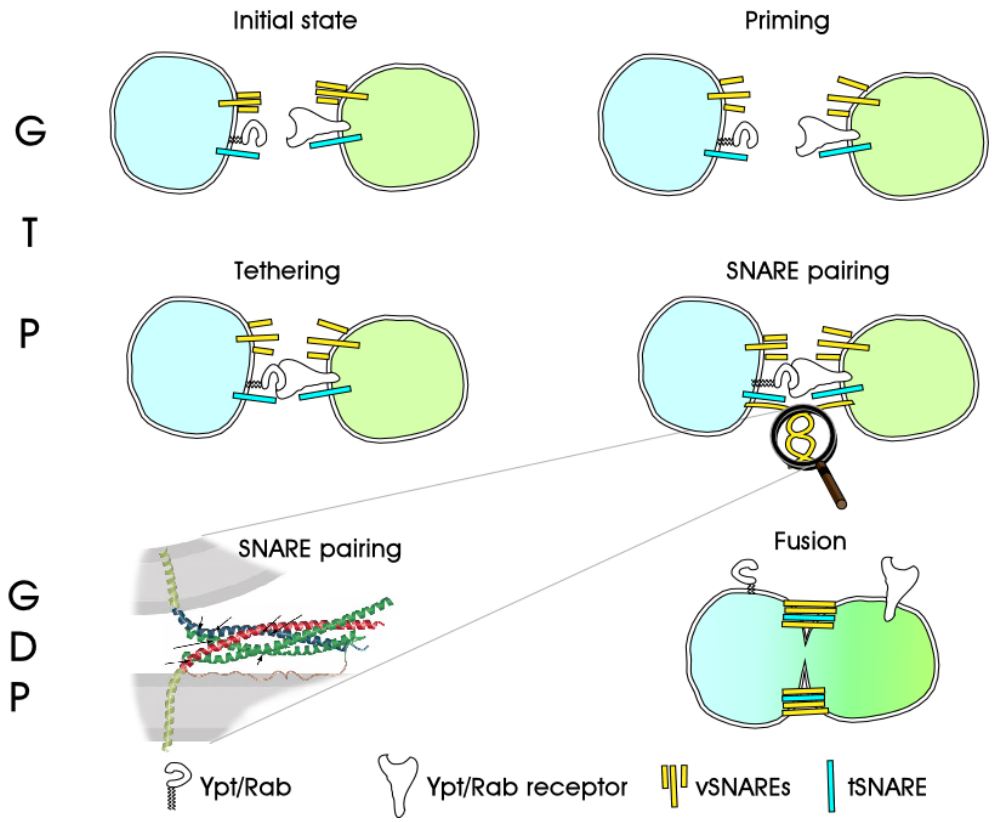


Figure I-8: Schematic representation of the fusion events. For the sake of simplicity, only Ypt and the SNARE proteins were represented. The coiled coil represented in the "SNARE pairing" step is based on actual X-ray structure (with the exception of the yellow, membrane-inserted fragments) and is reproduced from (Sutton *et al.*, 1998).

Sec1/Munc18 proteins are a class of proteins about which very little is known. It appears that they interact with the Q-type of SNAREs (SNAREs bearing, in the middle of the region responsible for the forming of the coiled-coil, a glutamine residue, as opposed to R-SNAREs, bearing an arginine), also known as syntaxins. Despite the fact that mutations in yeast homologues of Sec1/Munc18 led to a complete block in fusion, apparently providing an effect to be easily monitored, it is not yet clear how they act. The fact that syntaxins cannot simultaneously bind Sec1/Munc18 proteins and other SNAREs suggests that these proteins might prevent premature formation of the core complexes in *cis*, giving them a chance of forming *trans*, productive, complexes.

Ypt proteins are shown by different lines of evidence to be responsible for the specificity of the fusion process (Lazar *et al.*, 1997). According to the current model (Ungermann *et al.*,

1998), the docking event can be divided into the substeps of *tethering*, which is Ypt-dependent and reversible and *trans-SNARE pairing*, which is irreversible. Tethering can be reversed by simply diluting the vacuoles used in the fusion assay and continues even in excess of Sec18p and Sec17p which (by the use of ATP) continuously disassemble the newly-formed *trans* SNARE pairs, before these can accomplish their role. It seems also that the GTP form (i.e. "on") of Ypt is necessary till the moment of SNARE pairing; after this step, the Ypt protein can be switched "off" by either GTP → GDP hydrolysis or even physically removing it from the membrane (in the GDP form) with the help of Gdi. However, currently there are no clues concerning the identity of the "Rab receptor" with which Rab proteins are supposed to interact in the tethering step nor evidence on the exact molecular role played by the Ypt proteins during the fusion.

I.4. The members of the Rab/Ypt Family in yeast

Results of different studies performed *in vivo* on yeast (mainly generation of temperature-sensitive mutants) has led to the idea that the Ypt proteins can be divided in two major categories: those which are essential for cell survival and those which are not. The first class contains Ypt1p, Sec4p and *both* Ypt31p and Ypt32p, while in the second class fall Ypt51/52/53p, Ypt6p, Ypt7p, Ypt10p and Ypt11p. (See also **Figure I-9** for schematic localization of each Ypt protein in the cell).

I.4.1. Ypt1p

The first Ypt/Rab to be discovered and classified as a GTPase was exactly Ypt1p (initially termed Yp2p) (Gallwitz *et al.*, 1983). Discovered during the analysis of the surroundings of the actin gene, it was proven that its deletion is lethal for the yeast cell.

Further analysis showed Ypt1p to be involved in vesicular transport (hence the name **Y**east **P**rotein **T**ransport), more exactly in the step linking ER to Golgi. Blocking *YPT1* expression or use of *ypt1* temperature-sensitive mutants (Bacon *et al.*, 1989; Jedd *et al.*, 1995) leads to severe trafficking defects between the endoplasmic reticulum and the Golgi apparatus. This was proven both by EM studies, showing ER-derived transport vesicles of about 50nm in diameter, as well as by realizing that the vesicles are mainly populated with proteins which are

core-glycosylated, suggesting that they underwent only ER glycosylation but not maturation in the Golgi cisternae.

Moreover, it seems that Ypt1p is also involved in intra-Golgi traffic, mainly *cis-* to *medial-* compartments (Jedd *et al.*, 1995) defining it as being essential for the first two steps of the yeast secretory pathway. However, since Ypt31/32p are also involved in *intra* Golgi traffic (see below), it is understandable that overexpression of Ypt31/32p can partly rescue the dominant lethal *YPT1* mutations (Yoo *et al.*, 1999).

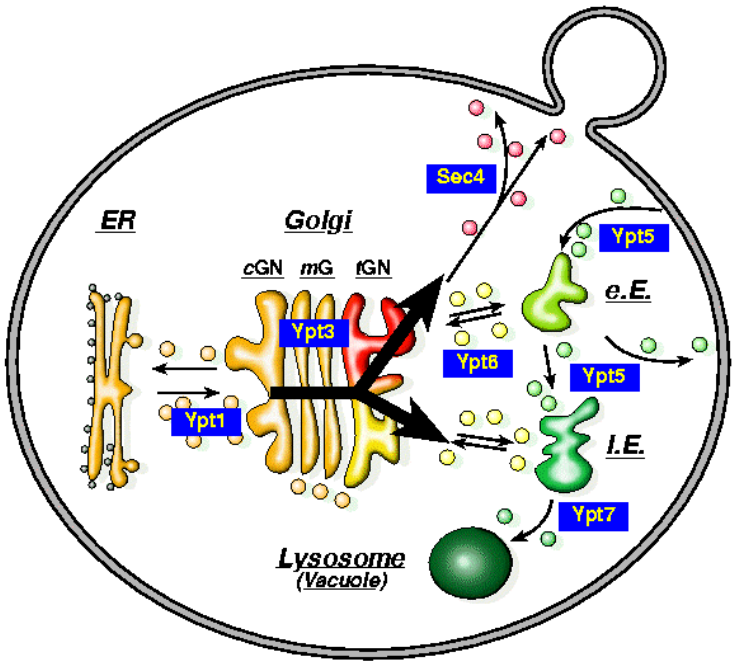


Figure I-9: Schematic representation of a yeast cell and the trafficking steps controlled by different Ypt proteins. ER–endoplasmic reticulum; cGN–*cis* Golgi Network; mG–medial Golgi; tGN–*trans* Golgi Network; eE–early endosome; lE–late endosome. Secretory pathway is marked in red, endocytic pathway is marked in green (from (Lazar *et al.*, 1997)).

I.4.2. Sec4p

Sec4p was discovered in a genetic screen for disorders in the *sec*retory pathway (Novick *et al.*, 1980). Five minutes after shifting the *sec4-8* temperature-sensitive mutant (the best characterized) from the permissive temperature to the non-permissive temperature, the

cells start displaying a blockade of the transport of the invertase, corroborated with the accumulation of vesicles as large as 100nm in diameter, filled with completely processed secretory proteins.

Localization studies showed Sec4p to be distributed in three compartments: a large pool was membrane-bound and associated with the plasma membrane, and smaller pools were found to be either "soluble" or associated with secretory vesicles; *sec4* temperature-sensitive mutants were found to have the main Sec4p pool located on the secretory vesicles at non-permissive temperature (Goud *et al.*, 1988). These two lines of evidence show that Sec4p is involved in the last stage of the secretory pathway, namely Golgi to plasma membrane.

I.4.3. Ypt31p and Ypt32p

Search for *Saccharomyces cerevisiae* homologues of *Schizosaccharomyces cerevisiae YPT3* gene yielded (with the help of a low-stringency hybridization protocol) two open reading frames; by virtue of their significant sequence identity (81% at protein level), and conforming to already-agreed rules (Kahn *et al.*, 1992), the newly-found proteins were termed Ypt31p and Ypt32p (Benli *et al.*, 1996).

Although disruption of either of them does not affect *Saccharomyces cerevisiae*, disruption of both is lethal. However, a more detailed study shows that depletion of Ypt31p (which is believed to be the more abundant isoform) leads to accumulation of highly glycosylated proteins in the cell, good proof that the intra-Golgi trafficking pathway is hindered. This is supported further by the finding that the majority of Ypt31p is found in subcellular fractions enriched in Golgi marker-enzymes. It is, however, not known in which part of the intra-Golgi trafficking pathway these proteins are implicated: retrograde or anterograde.

I.4.4. Ypt51p, Ypt52p and Ypt53p

These three isoforms were proven to be responsible for the endocytic delivery of the α mating factor and the lucifer yellow CH (two endocytic markers) from the external medium to the vacuole — yeast's equivalent of the mammalian lysosome (Singer-Kruger *et al.*, 1994). There is evidence linking the delivery of Golgi-derived vesicles to the vacuole. Mutations in the Ypt5 "family" also generate other defects, like problems in vacuolar protein maturation,

partial missorting of vacuolar hydrolases, alterations in the morphology and acidification of the vacuole.

Moreover, Ypt51p was previously discovered as Vps21p, where vps stands for "vacuolar protein sorting [defects]". However, even triple-deletion mutants (Ypt51p, Ypt52p, Ypt53p) are viable (under laboratory conditions), although the phenotype is markedly affected.

I.4.5. Ypt7p

The most striking phenotype achieved upon disrupting the *YPT7* gene is the fragmentation of the vacuole (Wichmann *et al.*, 1992). Cell-fractionation assays (Haas *et al.*, 1995a) localized the majority of Ypt7p on the membrane of the vacuole.

This vacuole fragmentation, in turn, leads, biochemically speaking, to a block of maturation of some vacuolar enzymes, for example carboxypeptidase Y and alkaline phosphatase, as well as problems in the degradation of the internalized α mating pheromone (Wichmann *et al.*, 1992). This and other data suggest that Ypt7p is involved in late endosome-to-vacuole transport as well as in the homotypic vacuole fusion (Haas *et al.*, 1995b). This process is important for a phenomenon know as "vacuole inheritance". During the division of *Saccharomyces cerevisiae*, many organelles —and the vacuole is among them— become highly fragmented and a part of them is passed over to the daughter-cell. At the end of the division process, the organelles reassemble in both daughter– and mother– cell. This is the moment when homotypic vacuole fusion takes place: two small vacuoles have to fuse, in order to generate a larger one and so on, till there is just one vacuole left in the cytoplasm, its size being around 1/4 of that of the cell.

Using a vacuole *in vitro* fusion assay{Haas, Conradt, et al. 1994 229 /id}, Ypt7p was shown to be required on **both** membranes (Haas *et al.*, 1995b) and the distinct steps involved in the Ypt7p-mediated vacuole fusion are partly resolved. Currently available evidence shows that fusion occurs in the ordered subreactions of priming, docking and fusion, docking being further subdivided into two steps: the tethering and the SNARE pairing (Ungermann *et al.*, 1998). Newer results (Price *et al.*, 2000) show that Ypt7p is the anchoring site for a 38S complex that is involved in vacuole docking. The main players in this complex are two proteins named Vam2p/Vps41p and Vam6p/Vps39p (in brief Vam2/6p) which are initially bound to the *cis* SNAREs, as a part of a 65S complex. After priming, catalyzed by Sec18p (NSF) and Sec17p (α-SNAP), the 65S complex partly dissociates, forming a smaller, 38S,

complex which has the ability to associate with Ypt7p, in a nucleotide-sensitive manner, i.e. it binds to Ypt7p•GTP but not to Ypt7p•GDP.

This would mean that the 38S complex meets all the requirements of a Ypt effector, adding it to the family of multiprotein effectors such as TRAPP (*cis*-Golgi transport protein particle) (Sacher *et al.*, 1998) and the exocyst (Bowser *et al.*, 1992; Kee *et al.*, 1997). Further experiments suggested that the Vam2/6p complex is present at the docking site and hence the role of Ypt7p would be to keep this complex associated with the vacuolar membrane during the tethering process, helping *trans*-SNARE pairing.

I.4.6. Ypt6p

Ypt6p *null* mutants manifest themselves also by a fragmented vacuole. Other ultrastructural modifications consist of the accumulation of vesicles of around 50nm diameter, as well as some membrane-bounded structures 150…300nm in diameter (Tsukada & Gallwitz, 1996). Other authors (Hengst *et al.*, 1995) report immuno-fluorescence staining around the Golgi apparatus. It might therefore be conceivable for Ypt6p to function in a step involving retrieval of the proteins in a late Golgi stage, perhaps the trans-Golgi network.

I.4.7. Ypt10p and Ypt11p

These two proteins were found *in silico*, after the sequencing of the yeast genome, based on sequence-homology with other Ras-like GTPases (Maurer *et al.*, 1995; Doignon *et al.*, 1993). None of them could be assigned a role in the intra-cellular traffic scheme, since there is no obvious phenotypic effect upon deleting/inactivating their corresponding genes.

Ypt11p is the most peculiar, being about 50% bigger than the other proteins (it contains 355 aminoacids, in contrast to 200..220 for all the other Ypt proteins; see also the alignment depicted in **Figure I-1** on page 7); it is possible that Ypt10p and Ypt11p are products of gene duplication and redundancy, but more studies must be done before coming to this conclusion.

Chapter II.

Materials and Methods

II.1. Chemicals

Reagents for general use in the laboratory were purchased from Riedel-de-Haën (Seelze, Germany). Enzymes for molecular biology experiments were purchased from Pharmacia and New England Biolabs. Kits for molecular biology were purchased from Qiagen (Hilden, Germany) or Macherei Nagel (Düren, Germany).

Acrylamide for SDS-PAGE was purchased from Applichem (Darmstadt, Germany). Acetonitrile for HPLC was from JT Baker and was of HPLC grade.

Reagents for crystallography were purchased either from Fluka (Seelze, Germany) and the pre-made screens (factorial or grid-screens) from Hampton Research (Laguna Niguel, CA, USA). PEG 3350 was, however, from Sigma (Taufkirchen, Germany).

*Mant*GDP and *mant*GppNHp were sinthetized in MPI Dortmund by Andrea Beste. Compounds for fluorescently labelling the proteins (N-(Iodoacetaminoethyl)-1-naphthyl-amine-5-sulfonic acid (1,5 I-A-EDANS) and tetramethylrhodamine-5-iodoacetamide dihydro-iodide (5-TMRIA)) were obtained from Molecular Probes Europe (Leiden, The Netherlands).

MALDI matrices (sinapinic acid (3,5-Dimethoxy-4-hydroxy cinnamic acid) for proteins and CHCA (α-cyano-4-hydroxycinnamic acid) for peptides) were from Sigma. CHCA was recristallized from ethanol in house.

II.2. Materials

PCR reactions were performed using a PE 9700 thermocycler from Applied Biosystems (Weiterstadt, Germany), while sequencing was performed using the BigDye™ kit from the same company; the sequencing products were analyzed by a ABI Prism 373XL machine, also from Applied Biosystems, run by a core facility.

Protein purification was usually performed on Waters 650 system (Millipore, Eschborn, Germany), using Pharmacia columns for Ni^{2+} affinity purification, as well as for gel filtration. Ni^{2+} resin was also obtained from Qiagen. The ion-exchange protocols were optimized using a BioCad Perfusion Chromatography workstation (PerSeptive Biosystems, Weiterstadt,

Germany) and Poros resins, from the same supplier. The same Poros resin was used in the scaled-up protocol.

Proteins were concentrated in devices (Microcon, Centricon and Centriprep) from Millipore (Eschborn, Germany). Other system for concentrating or dializing the proteins was the "ultrathimble" from Schleicher&Schuell (Essen, Germany).

The HPLC columns were supplied either by Bischoff (Leonberg, Germany) or by Phenomenex (Aschaffenburg, Germany). The HPLC system used for isocratic conditions was from Latek (Eppelheim, Germany) with a Shimadzu data recorder and integrator, while for gradient methods Beckman System Gold (Krefeld, Germany) or Waters 600s HPLC (Millipore, Eschborn, Germany) were used.

Linbro 24-well crystallization trays, as well as all the other equipment for crystal handling were from Hampton Research (Laguna Niguel, CA, USA). The in-house X-ray data collection was composed from an Enraf-Nonius 591 rotating anode (Delft, The Nederlands) with mirror optics and a MAR345 (X-ray Research GmbH, Norderstedt, Germany) image plate detector. Crystals were cryocolled using Oxford Cyrosystems 600 Series (Oxford, UK).

Long time base fluorescence spectra were aquired on a AMINCO SLM 8100 spectro-photometer (Silver Spring, MD, USA), while stopped flow experiments were performed in a High Tech Scientific SF61 apparatus (Salisbury, UK). Primary data analisys was performed with the help of the sowftare provided with the machine. Secondary data analysis was performed with the program Grafit 3.0 (Erithacus Software, Surrey, UK). Competitive titration data interpretation and global fitting were performed with Scientist 2.0 (Micromath, Salt Lake City, UT, USA).

The MALDI-TOF apparatus used for mass spectrometry analysis was Voyager DE Pro (PerSeptive Biosystems, Weiterstadt, Germany). Data was first evaluated in the Data Explorer v3.5 software provided with the machine. The centroid calculations were done in a custom Excel (Microsoft, Redmond, VA, USA) worksheet.

II.3. Bacterial strains used

Strains used for cloning:

XL1-Blue	*Rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17(r_K^-, m_K^-), *sup*E44, *rel*A1, *lac*, [F', *pro*AB$^+$, *lac*I^{q}ZΔM15, ::Tn10(Tet^r)
DH5α	Φ80d*lac*ZΔM15, *rec*A1, endA1, *gyr*A96, *thi*-1, *hsd*R17(r_K^-, m_K^-), *sup*E44, *rel*A1, *deo*R, Δ(*lac*ZYA-*arg*F)U169

Strains used for expression:

BL21(DE3)	*E. coli* F$^-$ *ompT*, *hsdS*$_B$, (r_B^- m_B^-) *dcm*, Tet^r *gal, dcm,* λ(DE3)
BL21-CodonPlus(DE3)-RIL:	*E. coli* B F$^-$ *ompT*, *hsdS*$_B$, (r_B^- m_B^-) *dcm*, Tet^r *gal, dcm,* λ(DE3), Hte, [*argU ileY leuW* Cam^r]

II.4. Protocols

II.4.1. Growth and storage of bacterial stocks

Bacteria were tipically grown at 37°C (some expression protocols required lower temperatures), with antibiotics, where appropriate. Typical antibiotics concentrations used were 100μg/ml ampicillin, 35μg/ml chloramphenicol, 30μg/ml kanamycin. Media were sterilized by autoclaving at 121°C for 20 minutes. Thermolabile compounds (such as antibiotics) were added as sterile-filtered solutions.

For stock preparation, a log-phase culture (OD_{600}≈0.7) was 1:1 diluted with a solution of 40% (v/v) sterile glycerol in an Eppendorf tube and snap-frozen in liquid nitrogen. The Eppendorf tube was kept at –80°C. When needed, the tube was removed from the freezer, scratched at the surface with a Pt loop and placed again in the freezer, care being taken not to melt the tube contents.

II.4.2. DNA methods

II.4.2.a. Cloning

Cloning was performed as described in (Sambrook *et al.*, 1989). Restriction digestion of DNA was performed in concordance with the instructions supplied with the restriction enzyme. Usually, about 0.5..1μg of DNA was digested with about 10U (supplier's definition) of restriction enzyme(s). Resulting restriction fragments were separated in agarose gels.

Gel electrophoresis of DNA fragments was performed either in agarose or in low melting point agarose (for gel extraction) 0.75..1% with ethidium bromide (0.1μg/ml); the gels were run submerged in TAE buffer at 10V/cm. The desired bands were extracted from gel using kits from Qiagen.

For ligation, the relative amounts of vector and insert were estimated from the intensity of the bands as seen under UV light after staining with ethidium bromide and a 5 fold excess of insert over plasmid was used.

II.4.2.b. Preparation of electrocompetent bacteria

- Make an O/N culture
- Dilute 1/100 (v/v) the preculture in 500ml of fresh LB- <antibiotic>
- Incubate till $OD_{595} \approx 0,7$
- Pellet the cells (≈1500-2000x g, 15')
- Wash in 500 ml MilliQ water, sterile, cold
- Pellet
- Wash in 500 ml MilliQ water, sterile, cold
- Pellet
- Wash with 20 ml of 10% (v/v) glycerol
- Pellet
- Resuspend in 2 ml 10% glycerol
- Aliquot as 50μl; snap-freeze and store at -80ºC

II.4.2.c. Electroporation transformation

For electroporation transformation, ≈50μl of electrocompetent bacteria were added in an ice-cooled cuvette. DNA was added (in a volume not exceeding 10μl and about 10ng DNA) and a 1.8kV pulse was applied, using a BioRad *E.coli* pulser. Subsequently, the cells were resuspended in 1ml of warm LB media and incubated with shaking (200rpm) at 30°C for 1hour; after this time, the cells were plated on Petri plates with LB-Agar supplemented with the appropriate antibiotic(s).

II.4.2.d. Preparation of DMSO-competent bacteria

- Grow a 5ml inoculus o/n in LB at 37°C
- Next day innoculate 100ml SOB with the o/n inoculus and grow it to OD600=0.6 at 30°C
- Chill cells 10-20 minutes on ice
- Spin cells down at 4°C (10-15min at 4000x g)
- Gently resuspend in 25ml TB and incubate 10 minutes on ice
- Spin gently at 4°C (15min at 3500 x g)
- Resuspend in 4.65ml TB and add 0.35ml DMSO
- Aliquote as 50-100μl and shock-freeze

Solutions:

SOB: LB + 10mM MgSO4 + 10mM MgCl2

10x TB: 2.38g HEPES + 18.6g KCl + 15ml 1M CaCl2; add to pH 6.7 with KOH and add 10.88g MnCl2·4H2O. Add water to 100ml. Filter sterile

II.4.2.e. Transformation in DMSO-competent bacteria

- Thaw DMSO competent cells on ice.
- Add 100ng of circular plasmid to 50μl cells.
- Incubate 5-10minutes on ice.
- Heat shock 60seconds at 42°C
- Resuspend in 1ml LB and grow in the shaker 20-60minutes
- Plate on plates with the appropriate antibiotic

II.4.2.f. Sequencing of DNA using the BigDye terminator kit

- Prepare a sequencing mix as follows:

  BigDye Premix 8,0μl

  DNA Template 0,5μl (500ng)

  Primer 0,5μl (3-10pmol)

  Dist. water ad 20,0μl

- Run 25 cycles as follows: 15" at 96 °C, 10" at 50 °C (or appropiate Tm), 3' at 60 °C
- Purification:

add 1μl Dextran blue (colour and pp of DNA)+ 2μl AcONa 3M, pH 4,6 + 50μl 100% EtOH (ice cold both)

Incubate 20' on ice

20' at max speed in a cool centrifuge (12-14.000 x g)

Wash 2-3 times w/ 75% EtOH (get rid of ddNTPs, salt)

Dry (SpeedVac)

- Send it to the in house sequencing facility.

II.4.2.g. PCR-screening of colonies

- Pick colonies using a sterile toothpick. Thoroughly rub the toothpick against the bottom of a PCR Eppendorf vial, after that against a small area of an LB-Antib. agar plate.
- Set up the following PCR mix:

For ONE colne	
Taq 10x Buffer	1.25μl
dNTPs	0.25μl
Primer 1 (32pmol/μl)	0.16μl
Primer 2 (32pmol/μl)	0.16μl
DW *ad* 12,5ul	10.56μl
Taq Polimerase	0.12μl
Total volume	12.50μl

(multiply by no. of clones+1, make the mix and distribute 12,5μl in each Eppendorf)

- Run the PCR protocol:3' at 96°C, 25× (20" at 96°C, 30" at 50°C (or appropiate melting temperature), 40" at 72°C),∞ at 4°C
- Load the PCR products on the gel

II.4.3. Protein methods

II.4.3.a. Expression and purification of wildtype and mutants/truncants of Ypt7p

- Transform pET11aYPT7 in E.coli BL21(DE3)cells. Select on LB-Amp 125mg/l plates.
- Grow an O/N preculture (25ml), spin the cells down, wash with fresh media and inoculate it into 5l of LB-Amp. Grow to OD_{600} ≈0.8 at 37°C.
- Induce with 0,4mM IPTG (final concentration) and continue growth at 37°C for 5 hours.
- Spin down the cells. Resuspend in **lysis buffer**. Run twice through the fluidizer. Clear the lysate by ultracentrifugation (min. 30min @100.000xg). Filter through ZapCaps or similar.
- Load on a Q-Sepharose column (column volume ≤400ml) preequilibrated with buffer **IExA**; wash the column with the same buffer. Elute with linear gradient of 50..450mM NaCl in 6.5CV. YPT7 will elute around 150-200 mM NaCl.
- Do an SDS-PAGE to identify the fractions which contain YPT7. Pool them and add ammonium sulphate till 70% saturation (0,472g/ml). Pellet 20-30min @ 10.000xg in Falcon tubes. Resuspend in **gel-filtration buffer**.
- Run gel-filtration on a Superdex G75 column preequilibrated with **gel-filtration buffer**. Flow rate ≤0,75ml/min. Identify fractions containing YPT7 by SDS-PAGE. Concentrate and snap-freeze.

Solutions used:

Lysis buffer: 20mM TRIS pH 8.0, 10mM $MgCl_2$, 1mM NaN_3, 1mM PMSF, 1% Triton X100, 5mM DTE

IExA: 20mM TRIS pH 8.0, 50mM NaCl, 5mM DTE

IExB: 20mM TRIS pH 8.0, 500mM NaCl, 5mM DTE

Gel-filtration buffer: 20mM TRIS pH 8.0, 150mM NaCl, 10mM $MgCl_2$, 5mM DTE, 25μM GDP

II.4.3.b. Expression and purification of Ypt31p

- Transform pET3aYPT31 in E.coli BL21(DE3) cells. Select on LB-Amp 125mg/l plates.
- Grow an O/N preculture (25ml), spin the cells down, wash with fresh media and inoculate it into 2l of LB-Amp. Grow to OD_{600}≈0.8 at 37°C.

- Induce with 0,4mM IPTG (final concentration) and continue growth at 37°C for 5 hours.
- Spin down the cells. Resuspend in **lysis buffer**. Run twice through the fluidizer.
- Centrifuge at 100.000 x g, filter the supernatant through a ZapCap.
- Load on a Poros HQ column (10ml or more) preequilibrated with buffer 95%**IexA** + 5%**IexB** (100mM NaCl); wash the column with the same buffer till stable baseline. Elute with a gradient of 100 to 500mM NaCl in 20 column volumes (5%**IExB** to 25%**IExB**). YPT31 will elute around 250-300mM NaCl.
- Do a SDS-PAGE to identify the fractions that contain YPT31. Pool them and add ammonium sulfate till >80% saturation (0,6 g/ml). Pellet 20-30min @ 10.000xg in Falcon tubes. Resuspend in minimal volume of **gel-filtration buffer**.
- Run gel-filtration on a Superdex G75 column preequilibrated with **gel-filtration buffer** (elution volume is approx. 150-160ml for a 26/60 column). Flow rate ≤0,75ml/min, collect 2ml fractions. Identify fractions containing YPT31 by SDS-PAGE. Concentrate and snap-freeze.

Solutions used:

Lysis buffer: 20mM TRIS pH 8.5, 2mM $MgCl_2$, 20mM NaCl, 1mM NaN_3, 1mM PMSF, 2mM βME

IExA: 20mM TRIS pH 9.0, 2mM βME, at least 1l

IExB: 20mM TRIS pH 9.0, 2M NaCl, 2mM βME

Gel-filtration buffer: 20mM TRIS pH 8.0, 40mM NaCl, 5mM $MgCl_2$, 5mM DTE, 25μM GDP

II.4.3.c. Expression and purification of yeast GGTaseII

- Transform BL21 (DE3) with plasmids pGATEVmod/Bet2 (#187) and pBCUC-MAD2 (#59) (either in two steps or in one step).
- Pick one colony, let it grow in 50ml LB-Amp125+Kn35 O/N at 37°C. In the morning pellet the cells, resuspend in fresh media, repellet and resuspend in 5ml and inoculate in 5-10l (100-200μl/1l).
- Grow @37°C till OD_{600}≈0.7-0.8
- Reduce temperature to 20°C and induce with 100μM IPTG. Grow them over night.

- Pellet and lyse the cells (with the fluidizer) in "Bug Buffer" with 1mM PMSF, 2mM β–mercapto ethanol.
- Spin them down 45' at 30.000 rpm in a Beckman Ti45 rotor.
- Filter the SN through a ZapCap and load the flow-through on a 20ml Hi-Trap Ni column preequilibrated with Bug Buffer with 9mM Imidazole.
- Wash the column with "Bug Buffer" + 9mM imidazole until you get back to baseline (approx 150ml) at a flow rate of 3ml/min.
- Elute in a gradient of 9mM to 300mM imidazole in "BugBuffer" in 120ml at 3ml/min. Collect 1.5ml fractions.
- Run SDS-PAGE and identify the fractions of interest (approx. 200mM imidazole).
- Cleave the HIS-tag with TEV protease: to the pooled fractions of protein, add EDTA to 0.2mM (TEV is inhibited by M^{2+} ions) and add TEV protease to 1:20 mol/mol (approx 0.9mg TEV for 60mg protein). Incubate O/N at 4°C.
- In the morning, add $MgCl_2$ to 2mM in the mixture (to block EDTA that would chelate Ni^{2+} from the column). Load the cleavage mixture on Ni^{2+} column equilibrated with "Bug Buffer". Collect the flow-through.
- Concentrate the flow-through: use a CentriPrep MWCO 30kDa (clear base) and spin it to 1500 x g.
- Gelfiltration: equilibrate a G200 26/60 preparative column with **gel filtration buffer**. Load the concentrated protein (≤ 4ml) and run it at a flow rate of 0.75-1.0ml/min.
- Identify the peaks by SDS-PAGE (the second one is Bet2 alone). Concentrate the protein to around 8mg/ml or above, aliquote and snap-freeze it.

Solutions used:

"Bug Buffer": 50mM NaPi pH8, 300mM NaCl

Gel filtration buffer: 25mM HEPES Na, 25-50mM NaCl, 5mM DTT

II.4.3.d. Expression and Purification of untagged Mrs6p

- Transform pET30b(+)—MRS6 into BL21(DE3) cells and select on LB-Kan (50mg/l) plates.
- Grow a preculture (50ml), spin the cells down, wash with fresh media and inoculate it into 5l of LB-Kan. Grow to $OD_{600} \approx 0.7$.

- Place on ice and induce with 0.1mM IPTG (final conc.). Grow 6h at 18°C.
- Harvest cells by centrifugation and wash once in **PBS**. Resuspend the cells in "BugBuffer". Brake cells with fluidizer twice. Add a new portion of PMSF. Clear lysate at 40.000 x g 1h at 4°C in a Ti45 rotor.
- Add solid ammonium sulphate to 25% saturation. Centrifuge at 12.000rpm 20-30min 4°C in a Beckman JA12 rotor. Discard the pellet. To the supernatant add solid ammonium sulphate to 40% saturation. Centrifuge again. Resuspend the pellet in 100ml of 50mM TRIS HCl pH 7.2, 10mM β mercapto-ethanol. Centrifuge again as above (to eliminate insoluble material).
- Subject supernatant to anion exchange chromatography on 20ml PorosHQ column, equilibrated with buffer **IExA**. Flow rate is 5ml/min. Wash the column with buffer **IExA** containing 50mM NaCl.
- Elute protein with following gradient: (buffer **IExA** plus) 50 to 400mM NaCl in 10CV and 5CV of 500mM NaCl. Collect 5ml fractions. Run SDS-PAGE to identify fractions.
- Dyalize O/N against buffer **IExB**.
- Load the protein on a PorosHS 10ml column equilibrated with buffer **IExB** at the flow rate of 5ml/min. Collect flow-through. Run SDS-PAGE to identify fractions.
- Precipitate with ammonium sulphate to 40% saturation. Centrifuge at 12.000rpm 20-30min 4°C in a Beckman JA12 rotor. Resuspend in 5ml of **gel-filtration buffer**. Centrifuge as above to eliminate the insoluble material.
- Load on a Superdex G200 26/60 column preequilibrated with **gel-filtration buffer**. Flow-rate: 1ml/min. Run SDS-PAGE.
- Concentrate and snap-freeze.

Solutions used:

"BugBuffer": 50mM NaH_2PO_4/Na_2HPO_4 pH 8.0, 0.3M NaCl, 1mM PMSF, 10mM β mercapto-ethanol

Buffer IExA: 50mM TRIS HCl pH 7.2, 10mM β mercapto-ethanol

Buffer IExB: 25mM TRIS HCl pH 8.6, 10mM NaCl, 10mM β mercapto-ethanol

Gel-filtration buffer: 25mM HEPES pH 7.2, 25mM NaCl, 10mM DTE

II.4.3.e. Nucleotide exchange using alkaline phosphatase

- This method is appropriate just for non-hydrolyzable nucleotide analogues. The success of the protocol depends on using alkaline phosphatase, calf intestinal and not *E.coli*, since the latter is able to break diphosphate or triphosphate linkages.
- Equilibrate a buffer exchange column (NAP5/ NAP10/ PD10) with **NE buffer** by filling the column three times.
- Load 300μl GTPase (approx 80-100nmoles) onto the column in no more than V_{max} sample.
- Add **NE buffer** up to the difference between V_{max} sample and the sample volume; let it sit at RT for 5 minutes.
- Elute the column with V_{elute} buffer.
- Collect the protein and add 2…4 times molar excess of non-hydrolyzable nucleotide analogue.
- Run a sample on HPLC to check the molar ratio.
- Add 5U of alkaline phosphatase for each mg of protein.
- Incubate at RT for about 40 minutes.
- Check the hydrolysis extent by HPLC. If everything is ok:
- Load another buffer exchange column, preequilibrated with **NE buffer** with the protein solution. Repeat steps 3 and 4 but this time collect fractions (3-4 droplets). Check each fraction by Bradford.
- Repeat steps 5 to 9 if this is the case (try to have more 95% nucleotide exchanged).
- Add $MgCl_2$ to 10mM, in order to stabilize the protein. Run the mixture through a column equilibrated with **final buffer**. Keep sample volume smaller than V_{max} sample.
- If necessary, add **final buffer** up to the difference between V_{max} sample and the sample volume; let it sit at RT for 5 minutes.
- Elute the column with V_{elute} buffer.
- Concentrate, measure the protein concentration (both Bradford and nucleotide concentration assayed by HPLC) and snap-freeze it.

HPLC ANALYSIS

CH_3CN gradient

• Gradient of 0 to 65% **HPLC-B** in 13 minutes on a C18 column 250×4.6mm equilibrated with **HPLC-A**, flow rate is 1.5ml/min. Go back to 0%**HPLC-B** in 2 min and hold it for 10min. Detection at 254nm.

Isocratic mode

• Flow rate 1…1.5ml/min on an isocratic machine with **IsoA** buffer. Also on C18 column. Detection at 254nm.

Solutions used:

NE buffer: 50mM HEPES 7.2, 3mM DTE, 5mM EDTA

Final buffer: 50mM HEPES 7.2, 3mM DTE, 5mM MgCl2

HPLC-A: 50mM KH_2PO_4/K_2HPO_4, pH 6.5

HPLC-B: 50% HPLC-A 50% CH3CN

IsoA: 10mM tetra-n-butyl ammonium bromide, 100mM KPi pH 6.5, 8%CH3CN.

Degas all buffers prior to use!!

II.4.3.f. Nucleotide exchange using EDTA

- This method is usable for any nucleotide substitution; it relies on the ability of EDTA and SO42- ion to decrease the affinity of the G protein for the nucleotide.
- Equilibrate three columns (NAP 5) with **buffer A**.
- Load 300μl GTPase (approx 80-100nmoles) onto the column in no more than 500μl.
- Add **buffer A** to make up the difference between 500μl and the sample volume; let it sit at RT for 5 minutes.
- Eluate the column with 1ml **buffer A**.
- Collect the protein and add 10× molar excess of desired nucleotide and EDTA to 10mM final concentration.
- Incubate at RT for 40 minutes.
- Reload the column and re-add nucleotide and EDTA.
- Repeat steps 6 & 7.
- Run through a buffer-exchange column equilibrated with **buffer B**.
- Remove the excess of nucleotide by passing through a NAP5 column pre-equilibrated with **buffer B**.

- Concentrate, measure the protein concentration (both Bradford and nucleotide concentration) and snap-freeze it.

HPLC ANALYSIS

CH3CN gradient

- Gradient of 0 to 65% **HPLC-B** in 13 minutes on a C18 column 250×4.6mm equilibrated with **HPLC-A**, flow rate is 1.5ml/min. Go back to 0%**HPLC-B** in 2 min and hold it for 10min. Detection at 254nm.

Isocratic mode

- Flow rate 1.5ml/min on an isocratic machine with IsoA buffer. Detection at 254nm.

Solutions used:

Buffer A: 50mM HEPES 7.2, 3mM DTE, 5mM EDTA

Buffer B: 50mM HEPES 7.2, 3mM DTE, 2mM $MgCl_2$

HPLC-A: 50mM KPi pH 6.5

HPLC-B: 50% HPLC-A 50% CH_3CN

IsoA: 10mM tetra-n-butyl ammonium bromide, 100mM KPi pH 6.5, 8%CH_3CN.

Degas all buffers prior to use!!

Chapter III.

Crystallographic analysis of Ypt7p

III.1. Introduction

Many biochemical experiments reveal different properties of the proteins being studied: specificity, ability to interact in a certain manner with some partners but not with other etc. These observations determine the start of a wealth of other biochemical experiments trying to pinpoint areas or even residues important for the observed functions. But the best and most convincing evidence can, usually, come from actually seeing *how* the structure of the protein looks. Eventually, in combination with structural analysis of different mutants, this can provide ultimate answers to questions regarding the specificity and the reaction mechanism.

Currently there are two main ways of determining the atomic structure of a protein: X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy. The main advantage of NMR is that it uses molecules in solution, so it is not limited to those proteins that crystallize well. This is a great advantage over the X-ray method, in which the sample has to be in the form of a crystal, sometimes extremely difficult to obtain, sized at least $50\times50\times50\mu m^3$. On the other hand, the usual methods limit the NMR to solving structures of proteins that are smaller than ca. 40kDa. Protein crystallography, on the contrary, does not have any size limitations: very large proteic complexes such as virus capsids (Tsao *et al.*, 1991), RNA polymerase (Cramer *et al.*, 2001), the photosystem I (Jordan *et al.*, 2001) and even ribosomes (Ban *et al.*, 2000; Wimberly *et al.*, 2000), to name just the biggest and most impressive, were crystallized and their structures solved at atomic resolution.

Two of the "small G protein" family members were crystallized fairly early, the ribosomal elongation factor EF-Tu (Jurnak, 1985; Berchtold *et al.*, 1993) elongation factor and the oncogenic product H-Ras p21 (de Vos *et al.*, 1988; Pai *et al.*, 1989; Schlichting *et al.*, 1990). This led to a better understanding of the processes underlying their capacity to act as "molecular switches" and to speculate about the way they hydrolyze GTP to GDP, although the fine details of the mechanism are even now not perfectly understood. Some of the problems can be addressed using molecular modeling might be a good tool, but this approach is still very much dependent on the correctness of the sequence alignment and it cannot accurately model neither insertions in loops nor N- and C-terminal extensions. All the currently known 700 members of the "small G proteins" family share a fairly high sequence

homology, but there are small differences in their structures. Exactly these differences enable them to perform different functions through interactions with different partners, in a word, confer them specificity.

The Ypt family consists, in the yeast *Saccharomyces cerevisiae,* of eleven members (Gotte *et al.*, 2000), controlling all the vesicular transport pathways in the cell. What is so striking in the family, and is even more pronounced compared to other proteins, is the specificity of the members. Every single member of the family interacts with certain common proteins (the constituents of the prenylation machinery and Gdi1p), whereas they remain extremely specific in the interaction with other proteins. For example, TRAPP interacts specifically with Ypt1p (Wang *et al.*, 2000), and not with other partners. Other proteins (such as the GTPase activating proteins, Gyp) have a broader specificity spectrum, still distinguishing between different members of the family. For example, Gyp7p is able to stimulate the GTPase activity of Ypt7p, as well as the activity of Ypt31/32p and Ypt1p (albeit to a lower extent), but it does not interact with Sec4p or Ypt6p. These variations arise in both the GTP- as well as the GDP- bond states of the proteins.

Despite different sequence analyses, some of them very sophisticated (Pereira-Leal & Seabra, 2000), the significant differences outlined above cannot be explained simply based on sequence differences between the members of the subfamily. This stresses the need for high-resolution structural information on a number of the members of the family.

Ypt7p is a member of the Ypt family (see also Section I.4.5) involved in the fusion machinery implicated in late endosome-to-vacuole transport and homotypic vacuole fusion. Together with Ypt51p and Ypt52p and Ypt53p, it is involved in the endocytic pathway, but *YPT7* mutants behave differently from the *YPT5X* mutants, showing that they have different roles and that interacting partners can distinguish them reasonably well. Also, these interacting partners are able to recognize and bind Ypt7p in one of the two states: GTP- or GDP-bound; this is the case for Gdi1p, which can bind significantly better the GDP form, or the case for Gyp7p, Gyp1p and effectors which bind more strongly the GTP form. It is clear that the structural modifications arising from GTP→GDP hydrolysis have deep consequences. At the time this work was started, there was no Rab/Ypt structure available and there were only speculations on what changes the GTP→GDP transitions would involve, suppositions based mainly on the high similarity with H-Ras p21. One of the purposes of this work was to crystallize Ypt7p, if possible in both forms (complexed with GDP and a non-hydrolyzable GTP

analogue) and to characterize it, together with the structural changes which accompany the GTP→GDP transition.

III.2. Crystallization and model building for Ypt7pΔC10

III.2.1. Design of Ypt7p for crystallization studies

For crystallization of Ypt7p, it was decided that a C-terminal truncation would be advantageous. This decision was based on previous results with H-Ras and structural studies performed on canine Rab7, the homologue of Ypt7p.

Crystallization attempts on H-Ras were successful only after a C-terminal truncation was performed (Ras $p21_{1-166}$); it was shown that this truncation by 23 aminoacids does not influence neither the G domain structure nor its biochemical properties (John *et al.*, 1989). A second indication came from investigations performed on Rab7. It was shown by NMR studies that the last 25 residues of Rab7 are not structured in solution (Neu *et al.*, 1997). The study by Neu *et al.* showed that: *i)* approximately the last 25 residues at the C-terminus were flexible and *ii)* nucleotide binding and hydrolysis kinetics were not affected by the truncations. Point *i)* suggests a potential hindrance to crystallization, while point *ii)* guarantees that the crystal structure will not reflect artifacts induced in the active site or important loops by the truncation at the C-terminus. On the other hand, crystallization studies (Brachvogel *et al.*, 1997) showed that a 10 residue C-terminal truncation of Rab7 yielded better-quality crystals than a 25 residue C-terminal truncation. Based mainly on the Rab7 results and the high homology between it and Ypt7p (62% identity, 75% strongly similar), the decision was taken to shorten the C-terminus of Ypt7p by 10 residues.

In order to obtain this Ypt7p truncation, PCR with degenerated primers was performed on the original plasmid pET11a *YPT7* (kindly provided by Dr. Dieter Gallwitz, Göttingen). The 5' primer was the standard T7 promoter primer (Novagen catalogue no. 69348-3) and the 3' primer had the sequence 5'-ATTTTC**GGATCC**TTAGCGAATA-3', containing a *Bam*H I restriction site (bold) and a STOP codon (underlined). Both the amplicon and the pET11a*YPT7* vector were cut with *Nde* I (recognition site of which contains the START codon for *YPT7*) and *Bam*H I and ligated together. The vector obtained in this manner was called pET11a *YPT7*ΔC10. The protein encoded by this vector will be referred to as

Ypt7pΔC10 and was purified by a two step protocol, involving ion exchange and gel filtration (see the protocol described in section II.4.3.a, at page 39 for details).

III.2.2. Crystallization of Ypt7pΔC10 and improvement of the crystals

As structural results with other GTPases have already shown, in the GTP form, the loops L2 and L4 are kept in place by interactions between their amide hydrogen and the γ phosphate of the GTP. In the GDP case, these loops become extremely flexible, thus hindering the crystallization. Because of this, and since one of the purposes of this work was to find out differences between the GTP- and GDP- bound states, a pool of Ypt7pΔC10 was loaded with the non-hydrolyzable GTP analogue GppNHp (5P-guanosyl-β-γ-imido-triphosphate). This analogue, utilized also for the crystallization of p21 Ras, has a stereochemistry very similar to that of GTP and can still participate in H-bonding contacts with its β-γ NH, as opposed to $GppCH_2p$. The protocol used for nucleotide exchange (see Section II.4.3.e) is based on the phosphatase activity of calf intestinal phosphatase, which can convert GTP and GDP to GMP and then guanosine, while not being able to hydrolyze the pNHp linkage efficiently.

Using multifactorial crystallization screens (Jancarik & Kim, 1991) crystals of Ypt7pΔC10•GppNHp appeared under the following conditions: 0.1M TRIS pH 8.0, 2M $(NH_4)_2HPO_4$, 20°C. Crystals obtained under these conditions were still too small for diffraction experiments, so optimization was necessary even for the first tests. A remark has to be made: in incomplete factorial screens (like the Jancarik and Kim), there is no "optimization" of the conditions based on the chemical knowledge. In this particular case, one has to note that the pH of the reservoir (and of the crystallization drop) is dictated not by the 0.1M TRIS, but by the 2M $(NH_4)_2HPO_4$ and that small pH modifications have to be done by altering the pH of the ammonium phosphate solution. Two approaches were used for improving the quality of the crystals in a first phase: variation of the pH and salt concentration on the one hand and seeding on the other. Seeding (in this case microseeding) is performed with the purpose of lowering the activation energy for forming crystals.

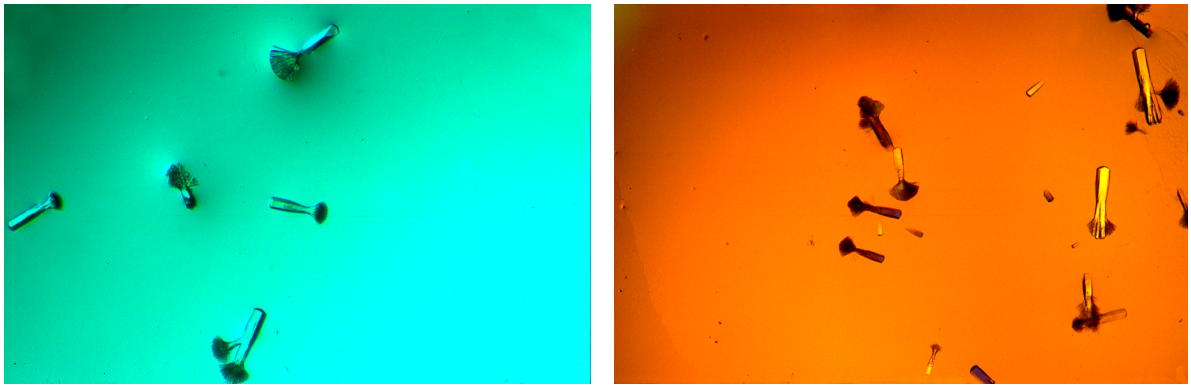


Figure III-1: One of the first crystals of Ypt7pΔC10•GppNHp. Colors are different due to polarization filters. Crystals are about 100-120μm on the longest dimension.

In a vapor diffusion crystallographic setup, a droplet is prepared by mixing (usually equal volumes of) protein sample and precipitant. The droplet is suspended over a reservoir containing the same solution of the precipitant. Hence, at the start of the experiment, the precipitant concentration in the droplet is just ½ of the concentration in the reservoir (protein concentration is rarely above 1mM, which means roughly 20mg/ml for Ypt7p, while precipitant concentration is in the molar range for any salt used). Because the system will tend to equilibrate, water vapors from the droplet will diffuse towards the reservoir. However, since the reservoir volume is about 300-500 times bigger than that of the droplet, the effect will be slow shrinkage of the droplet and thus slow concentration of protein and precipitant.

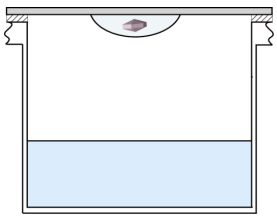


Figure III-2: Schematic representation of a vapor diffusion setup in "hanging drop" geometry.

The crystals obtained (see **Figure III-1**) showed dendritic growth (a tree-like disorder at the growing end), usually a sign of disorder caused by a too fast growth. In the case when the precipitant used is salt, equilibration is completed to more than 90% in about 25 hours (Mikol

et al., 1990). On the other hand, the rate-limiting step in crystallization is the appearance of a critical nucleus; this nucleus can appear only in suprasaturated states, relatively far from equilibrium. By seeding, the critical nucleus is provided "from the exterior" and crystallization can start earlier in time, at a state of lower supersaturation. In turn, this means that the crystallization process will be slower (the driving force for the crystallization will be smaller), resulting in a higher chance of forming well-ordered crystals. A second advantage of seeding is that by doing serial seedings one can try to control the number of crystals which will appear in a droplet. The fewer crystals, the bigger they will have the chance to grow, since all the proteic material available for crystallization will be used to grow fewer crystals. This mixed approach worked (although not extremely well), yielding crystals of type shown in **Figure III-3**. They were bigger, but still manifested the same dendritic growth defect. On a very intense beamline, like the microfocus beamline ID13 at EMBL/ESRF Grenoble, their regular "stem" was able to diffract up to a Bragg *d* spacing of 2.1Å.

Figure III-3: Crystal of Ypt7pΔC10•GppNHp. Conditions, in this case, were: 2.2M $(NH_4)_2HPO_4$, 0.1M TRIS, pH 8.5, 18°C. The stem is about $80\times80\times150\mu m^3$.

Considering that it would be helpful to obtain Ypt7pΔC10•GppNHp crystals of better quality and bigger size, I tried a different approach. The crystals seem to be extremely sensitive to temperature. Although it is not uncommon for proteic crystals to be affected by changes in the temperature (which drives changes in the solubility of the protein, not always in a predictable direction), the sensitivity of Ypt7pΔC10•GppNHp crystals was extreme: variations of 5°C were able to trigger their complete dissolution. This unusual sensitivity was exploited in the following way: setups were prepared as usual at 20°C with a reservoir solution

consisting of 1.9M $(NH_4)_2HPO_4$, 0.1M TRIS, pH 9.0, 2mM L-Cys, left to equilibrate and crystals to form. Once crystals appeared (with the already-known defects), the setup tray was moved to a 15°C incubator, with a temperature buffer (dummy plates) to prevent condensation and rapid temperature changes. After 48 hours, all the crystals were completely dissolved. Then, the plate was shifted back to 20°C, with the same temperature buffer, so that the change in temperature will take place over days, and left completely still for several days. This approach yielded crystals measuring about $125\times125\times1250\mu m^3$ with no visible defects (**Figure III-4**). Unfortunately, bigger is not always better. The diffraction pattern of these crystals showed increased mosaicity, due most probably to handling problems. Crystals of this size present two major difficulties: they tend to bend relatively easily when collected in a cryo-loop and freezing such a large crystal fast enough that only amorphous ice is formed becomes a problem.

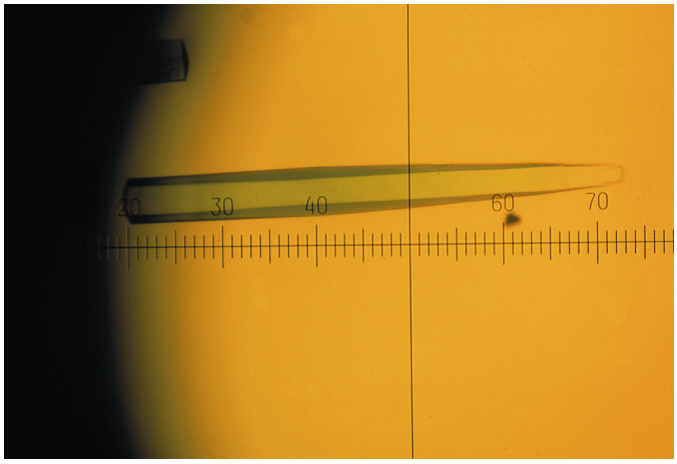


Figure III-4: Crystal of Ypt7pΔC10•GppNHp obtained by temperature-induced "recrystallization". One mark on the scale corresponds to 25μm.

The data set collected at the beamline ID13 (see data collection statistics in **Table III-1**) could be indexed in the space group P6. However, systematic absent reflections suggested that the real space group also contains a screw axis.

Table III-1: Data collection statistics for Ypt7pΔC10•GppNHp

A. Data statistics

Beamline	ID13/EMBL-ESRF
Temperature	100K
Area detector	MAR CCD
Wavelength	0.782Å
Space group; cell dimensions	$P6_1$; $a = b = 76.689$Å, $c = 117.962$Å
Number of recorded reflections	223,356
Average redundancy	9.84
R_{merge} [a]	6.9%
Intensities (I/σ) [b]	13.12 (3.58)

B. Refinement statistics

Resolution limit	2.2Å
Number of unique reflections	19769
Completeness of data (%) [c]	99.0 (98.6)
R_{cryst} [d] / R_{free} [e] (after the last refinement round in $P6_1$)	0.2721 / 0.2976
Rms bond lengths (Å)	0.006451
Rms bond angle (deg.)	1.20121
Mean/rms on B factors (Å^2)	
Backbone	50.50/24.25
Side-chain	52.65/23.25
GppNHp	38.37/2.84
Solvent (30 water molecules)	42.68/4.45

[a] $R_{merge} = 100 \times \sum I - \langle I \rangle / \sum I$, I is the observed intensity and $\langle I \rangle$ is the average intensity calculated from multiple observations of symmetry related reflections; [Reported for all reflections with signal/noise >= -3.0, all res. range]

[b] The value in parentheses is calculated for the highest resolution shell collected (2.2-2.1Å); [all values are reported for reflections with signal/noise >= –3.0]

[c] The value in parentheses is calculated for the highest resolution shell used in refinement (2.3–2.2Å);

[d] $R_{cryst} = \sum |F_o - F_c| / \sum_{hkl} |F_o|$, where F_o and F_c are, respectively, observed and calculated amplitudes;

[e] R_{free} is an R_{cryst} calculated using 10% of the processed data (1977 reflections), chosen randomly, kept constant and omitted from all subsequent structure refinement steps.

III.2.3. Phasing, model building and the problems

The Matthews coefficient (Matthews, 1968) suggested a solvent content consistent with the existence of two molecules per asymmetric unit. In concordance with this, after extensive molecular replacement trials, a solution was found by searching for two molecules. Phasing was achieved using molecular replacement as implemented in the program AMoRe (Navaza J, 1994), using Ras p21 and the recently solved Ypt51p structure (Esters *et al.*, 2000), PDB entry code 1EK0, as models. The successful solution could be obtained using the Ypt51p structure with all the atoms (not just a polyalanine model). The space group proved to be $P6_1$.

Building of the Ypt7pΔC10 molecule was achieved by alternate cycles of manual rebuilding against σ-weighted electron density maps in the program O (Jones T.A. *et al.*, 1991) and positional and B factor refinement in the CNS package (Brunger *et al.*, 1998). However, a problem announced itself from the beginning and after six rounds of rebuilding/refinement proved to be a crucial one. Due to unfortunate crystal contacts, two adjacent molecules were facing each other, with their "switch" loops strongly interacting. As a consequence, the residues 34 to 40 and 65 to 77 could not be traced in the electron density map. This meant that the most important features of the protein could not be seen and, even if seen, interpretation might be flawed because of the crystal packing artifact.

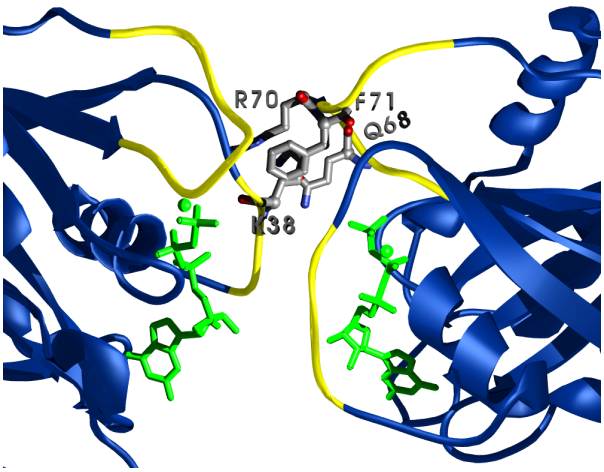


Figure III-5: Clashes between two molecules of Ypt7pΔC10•GppNHp. Two molecules of Ypt7pΔC26•GppNHp (see below) were positioned similar to the crystal packing for the Ypt7pΔC10•GppNHp form. Regions marked yellow could not be seen in the electron density for the ΔC10 case. It can be easily noted that Lys38 of one molecule clashes with residues 68-71 of the other.

Furthermore, for residues 183-187 only the main-chain trace could be interpreted and residues 188-198 were not visible at all. Electrospray ionization mass spectrometry (ESI-MS) analysis revealed that solely the first amino acid (Met1) was lost, probably due to usual posttranslational modification, but no further degradation of the protein occurred.

After the sixth round of rebuilding, the crystallographic R factor (showing the degree of concordance between the crystallographic model and the experimental data) and the R_{free} factor, an unbiased R_{cryst} , were still very high (27% and, respectively, 30%) for a fairly high resolution data set and there was no prospect of major improvement. The only solution seemed to take a few steps back before making some others forward.

III.3. Crystallization and model building for Ypt7pΔC26

III.3.1. Designing the new construct

Based on the interpretation of the observed electron density and the comparison with the recent crystal structures of Rab3a (visible residues 18-186 (Dumas et al., 1999)) and Ypt51p (visible residues 5-175 (Esters et al., 2000)), a new construct was designed including only residues 1-182, thus lacking the last 26 residues of Ypt7p. This was also in accordance with the H-Ras $p21_{1-166}$ construct and, more importantly, with the Metcalf experiments (Neu *et al.*, 1997) showing that the last 25 residues of Rab7 are flexible in solution.

In order to obtain this construct, PCR with degenerated primers was performed on pET11a *YPT7*ΔC10. The 5' primer was the T7 promoter primer, while the 3' primer was 5'-C**GGATCC**TCGAGTTAAGCTTGATTCTG-3'; it contained a *Bam*HI site (bold) and a premature STOP codon (underlined). The PCR amplicon was inserted into pET11a *YPT7*ΔC10, yielding pET11a *YPT7*ΔC26. The protein encoded by this vector will be referred to as Ypt7pΔC26. The integrity of the open reading frame for this construct was, like for pET11a *YPT7*ΔC10, checked for PCR errors by DNA sequencing (BigDye terminator kit, PE Applied Biosystems). After purification (which proceeded by the same protocol as the one for Ypt7pΔC10), a sample of Ypt7ΔC26 was loaded with GppNHp.

III.3.2. Crystallization of Ypt7pΔC26•GDP and Ypt7pΔC26•GppNHp

With this second construct of Ypt7p, we were able to obtain crystals in complex with GppNHp as well as with GDP (the latter was not possible in the ΔC10 case), both of them in the space group $P2_12_12_1$. Despite the fact that the protein and the space group were the same, the crystals obtained for the two forms had different cell dimensions and crystal packing.

First crystals of Ypt7pΔC26•GDP were found in 20% PEG 3350, 200mM NaCl. Further trials to improve crystal quality led to crystals able to diffract (at beamline ID14-1, ESRF Grenoble) up to 1.3Å. For these crystals, conditions were: 0.1M HEPES pH 7.0, 1M LiCl, 21% PEG 6K. For data collection, the crystals were flash-frozen in liquid nitrogen. As cryo-protectants 10% glycerol in the reservoir solution, or perfluorinated mineral oil or microscopy immersion oil were used.

Ypt7pΔC26•GppNHp crystals were found under conditions similar to those used for Ypt7pΔC26•GDP, namely 1M LiCl, 20-30%PEG 6K over a wide pH range: 0.1M of citrate pH 5.0, MES pH 6.0, HEPES pH 7.0, TRIS pH 8.0 or Bicine pH 9.0. Attempts to fine-tune the conditions yielded rod-shaped crystals of about $30\times30\times500\mu m^3$ (**Figure III-6**) able to diffract to a Bragg spacing of up to 1.7Å. Conditions for these crystals were: 0.1mM Bis-Tris Propane pH 7.0, 1M LiCl, 26% PEG 3350.

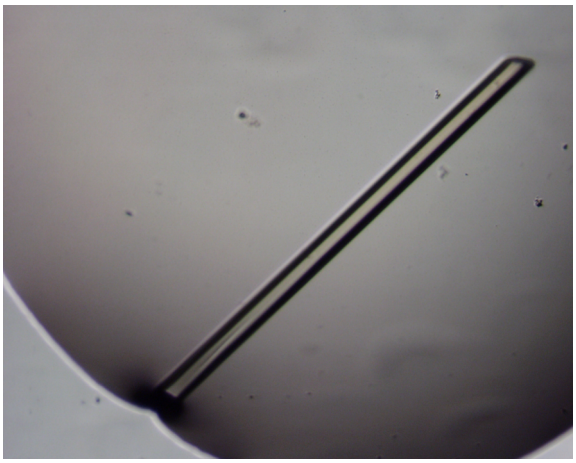


Figure III-6: Crystal of Ypt7pΔC26•GppNHp grown from 0.1mM Bis-Tris Propane pH 7.0, 1M LiCl, 26% PEG 3350. Size is about $30\times30\times500\mu m^3$.

Further optimization of the crystallization conditions led to crystals sized $80\times80\times100\mu m^3$ (**Figure III-7**) which were grown using 0.1M Bicine pH 9.0, 1M LiCl and 25% PEG 3350; these crystals were able to diffract (at beamline ID14-1 at ESRF) to 1.57Å. For data collection, the crystals were flash-frozen in liquid nitrogen. No cryo-protectant was needed.

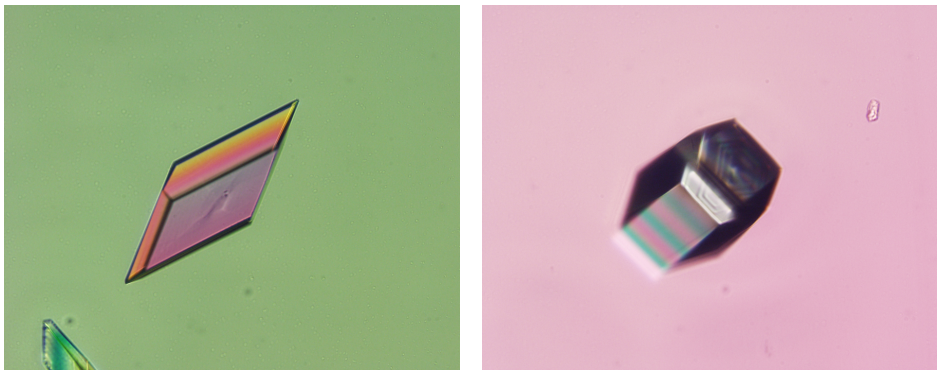


Figure III-7: Crystals of Ypt7pΔC26•GppNHp grown from 0.1M Bicine pH 9.0, 1M LiCl and 25% PEG 3350. The one on the right is about $80\times80\times100\mu m^3$ and diffracted up to 1.57Å.

Table III-2: Data statistics for Ypt7pΔC26•GppNHp

A. Data statistics

Beamline	ID14-1/ ESRF
Temperature	100K
Area detector	MAR CCD
Wavelength	0.934Å
Space group; cell dimensions	$P2_12_12_1$; a = 40.137Å b = 60.667 Å, c = 73.537Å
Number of recorded reflections	124,357
Average redundancy	5.15
R_{merge}	10.3%
Intensities (I/σ)	8.08 (2.75)

(continued on the next page).

Table III-2 (contd.): Data statistics for Ypt7pΔC26•GppNHp

B. Refinement statistics

Resolution limit	1.57Å
Number of unique reflections	23450
Completeness of data (%)	96.2 (95.7)
R_{cryst} / R_{free}	18.32/23.87
Rms bond lengths (Å)	0.008
Rms bond angle (deg.)	1.824
Mean/rms on B factors (Å^2)	
Backbone	17.44/9.71
Side-chain	23.71/16.58
GppNHp + Mg^{2+}	12.78/2.87
Solvent (284water molecules)	43.08/19.81

Table III-3: Data statistics for Ypt7pΔC26•GDP

A. Data statistics

Beamline	ID14-1/ESRF
Temperature	100K
Area detector	MAR CCD
Wavelength	0.934Å
Space group; cell dimensions	$P2_12_12_1$; a = 49.508Å b = 55.181Å, c = 60.200Å
Number of recorded reflections	263110
Average redundancy	6.47
R_{merge}	9.2 %
Intensities (I/σ)	10.81 (1.90)

B. Refinement statistics

Resolution limit	1.3Å
Number of unique reflections	40524
Completeness of data (%)	98.2(94.2)
R_{cryst} / R_{free}	20.91 / 25.88
Rms bond lengths (Å)	0.014
Rms bond angle (deg.)	2.130
Mean/rms on B factors (Å^2)	
Backbone	24.08 / 14.33
Side-chain	28.98 / 16.20
GppNHp + Mg^{2+}	22.04 / 3.73
Solvent (225water molecules)	41.94 / 16.11

III.3.3. Phasing and building the model

Obtaining phases for the Ypt7pΔC26 crystals (in both GDP and GppNHp forms) was an easy task, the partly refined model of Ypt7ΔC10 (without residues 182-187) being an ideal molecular replacement model.

From the first type of Ypt7pΔC26•GppNHp crystals (the long needle morphology), data were collected *in house*, using a rotating anode and CuK_α radiation, to a resolution of 1.7Å; however, for phasing and initial refinement, it was used only up to 2.0Å, in order to obtain very good data statistics. The model obtained in this intermediate refinement step was used with the data up to 1.57Å for the final refinement.

The same strategy was applied to both crystal forms (GDP and GppNHp). At the beginning, rounds of manual rebuilding were alternated with positional and individual B-factor refinement in the program CNS. For the more advanced stages of the refinement, I decided to switch to the program SHELX (Sheldrick & Schneider, 1997). The main reason for this was its better capacity to deal with alternate conformations. At high resolution it might be noted that some residues are not positioned in a single place; their side chains may have two or more alternate positions, in each of them interacting with other partners able to satisfy the need for H-bonding, ionic interactions or hydrophobic interaction. Although sometimes disordered residues were found even in the active site of an enzyme (Scheidig *et al.*, 1999), usually they do not have very important structural consequences. Nevertheless, modeling this disorder enables the crystallographer to obtain a better model for the structure to be solved. Unlike CNS, SHELX has a more natural way of dealing with this problem. Not only is defining these residues easier, but SHELX keeps the sum of occupancies of each individual conformer equal to 1, while giving them crystallographic thermal factors as close as possible, which is a more correct approach.

Another reason for which SHELX was chosen is its ability to define anisotropic thermal factors for each atom in part. At very high resolutions (usually below 1.5Å) it can be shown that each atom has preferential directions of movement, such that the probability of finding the atom at a certain position can no longer be described by a sphere, but by an

ellipsoid. In this case, there are six thermal parameters (anisotropic displacement components) describing the positioning of the atoms.

III.3.4. Anisotropic displacement factors: to use them or not?

The 1.57Å resolution of the GppNHp structure would, in general, be considered too low to justify modeling of thermal anisotropy for each atom in part and even the 1.3Å resolution of the GDP form is considered to be a borderline case.

On the other hand, using anisotropic displacement parameters might help better interpreting the experimental data. A first, rather pragmatic approach, would be to consider that any protocol which leads to the lowering of the crystallographic R factor ***and*** the free R factor (Brunger, 1992) is a valid protocol.

The crystallographic R factor is defined as

$$R_{Cryst} = \sum |F_o - F_c| / \sum_{hkl} |F_o|$$

where F_o and F_c represent, respectively, the recorded (observed) structure factor amplitudes and the ones calculated on the basis of the model. Basically, it is an indicator which shows how well does the model explains the diffraction data. Unfortunately, the R value can be made (willingly or not) arbitrarily low, simply by increasing the number of parameters which are used to describe the model. The direct consequence is that a low R value is not necessarily an indicator of a good fit and that using more parameters might artificially lower it. For this very reason, trying to assess the opportunity of modeling anisotropic displacement factors just by checking the R value would not have been appropriate.

To escape this bias, Brünger ported from the field of statistics the concept of the cross-validation method, introducing the *free* R value (Brunger, 1992). The R_{free} is calculated exactly in the same manner the crystallographic R is calculated. The only difference is that it uses about 5..10% of the data (usually 2000 reflections) which were *never* used, neither in the refinement process, nor in the modeling stage. Hence, one can, at any time, refer to the R_{free} and see how well the model predicts (as opposed to explains) the experimental data. This is possible because each recorded reflection is, in fact, the result of the diffractive contribution of every atom in the crystallographic unit cell. It can safely be said that each reflection contains information about the content of the *whole* unit cell, not just a part of the molecule. This makes the R_{free} an ideal candidate for evaluating different refinement protocols (Adams *et al.*,

1997). It has to be also noted that, in the absence of systematic errors, the R_{free} should always be higher than the conventional crystallographic R, but that a difference which is too big is a sign of over-refinment, when the R is artificially lowered, but the R_{free} cannot be.

The R and R_{free} values for Ypt7ΔC26•GppNHp and Ypt7ΔC26•GDP were checked before and after a SHELX run in which anisotropic displacement was modeled. The problem which has to be kept in mind is that during an anisotropic refinement the number of parameters to be fitted increases from four to nine. This means that the ratio between the observations (reflexes) and the parameters to be fitted decreases drastically. The results of the two types of refinement are presented in the table below:

Aniso	Ypt7ΔC26•GppNHp				Ypt7ΔC26•GDP			
	R	R_{free}	observed	parameters	R	R_{free}	observed	parameters
no	0.1832	0.2387	21465	6831	0.2091	0.2588	37935	5237
yes	0.1476	0.2299	21465	15351	0.1637	0.2301	37935	15975

The conclusion which might be drawn here is the following: in the case of the GppNHp structure, the lowering of R_{free} as a consequence of modeling anisotropic thermal factors is very modest, especially when compared with the large decrease in the crystallographic R value, showing that there are not enough observations to enable one to use such a refinement protocol. In the GDP case, however, the results suggests that use of atomic thermal anisotropy may lead to a more accurate modeling of the recorded data.

III.4. Results

Ypt7p has a fold very similar to that of other small GTP-binding proteins consisting of a six-stranded β-sheet surrounded by α-helices. However, there are only four (and not five, as usual) helices (see also **Figure I-2** and **Figure III-8**). It has essentially the same main-chain trace as other Ypt/Rab proteins, including the glycine-induced bulge at the end of the effector loop, L2, first described by Dumas (Dumas *et al.*, 1999). Glycine 42 (Ypt7p numbering) represents an insertion, compared to Ras, found in all members of the Rab/Ypt family (see the discussion about Rab Family 1 determinant, page 14 and figure **Figure I-5**). This glycine forces the main chain to "bulge" towards helix α2 (or the extended loop L4 in the case of Ypt7p, see below).

The conformational differences between the structure of Ypt7pΔC26•GppNHp and the structures of Rab3A•GppNHp (Dumas *et al.*, 1999), Ypt51•GppNHp (Esters *et al.*, 2000) and Sec4•GppNHp (Stroupe & Brunger, 2000) are small but significant and concern the main-chain and side-chain conformations primarily in the two switch regions, including the effector loop L2 —between strands β2 and β3— and the catalytic loop L4, as well as loop L5 —between helix α3 and strand β5. In addition, the comparison of the Ypt7pΔC26•GppNHp structure with the Ypt7pΔC26•GDP structure underlines the L2 and L4 loops as the major conformational switches. These differences and the associated changes will be primarily discussed. In order to present and discuss these differences, the high-resolution structure of Ypt51p•GppNH will be primarily referred to, since both proteins come from *Saccharomyces cerevisiae*, are involved in the endocytic pathways of vesicular transport, and they might be localized on the same organelle at some points in their life-cycle. Therefore, despite the fact that they have some interacting proteins in common (the components of the prenylation machinery, Gdi1p etc.), they have to be able to differentiate between specific binding partners. As an example, GDP release from Ypt7p cannot be stimulated by the guanine nucleotide exchange factor Vps9p, which is an exchange factor for Ypt51p (Hama *et al.*, 1999; Esters *et al.*, 2001).

III.4.1. Ypt7pΔC26•GppNHp

III.4.1.a. Main chain trace

The major differences in the main-chain trace occur at the N- and C-termini (residues 3-9 and 172-182), the β-turn L3 between β-sheets β2 and β3 (residues 49-59), loop L7 between α-helix α3 and β-sheet β5 (residues 110-119) and the tight turn between helix α4 and β-sheet β6 (residues 148-151, L9) (see also **Figure III-8**). Also, loop L8 (between β5 and α4) and the region between residues 69 and 82 (normally loops β3/α2 —L4— and α2/β4 —L5, together with helix α2) show major differences when compared with other Ypt structures. The Ypt7p loops L3 and L9 have one, and the loop L7 four residues inserted, compared to Ypt51p (see the alignment on page 7). These differences in the size of the loop, taken together with the differences in the amino-acid sequence, explain very well the observed conformational difference. These loops are in close neighborhood to the N- and C-termini, together forming a characteristic epitope on the side opposite the nucleotide binding pocket of

the protein. The differences are Ypt7p-specific, since the comparison between Ypt51p, Sec4p, Rab3a and Rab5c reveal significantly fewer differences in the main-chain trace for the above-discussed regions. However, these proteins display significant variations in the amino-acid sequence for this surface epitope.

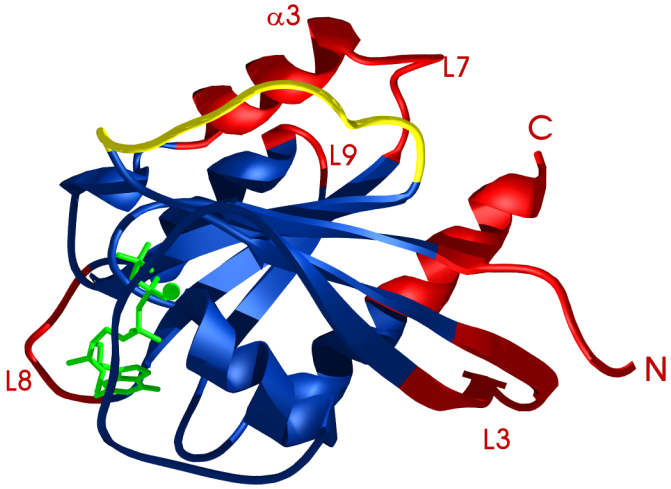


Figure III-8: Ypt7pΔC26•GppNHp. Regions showing main chain differences (RMSD > 1.0Å from Ypt51p coordinates) compared to the structures of other Rab/Ypt family members are depicted in red. The extended loop L4 (usually L4/α2/L5) is yellow.

It has to be borne in mind that loops L5 and L7, together with the N terminus and the terminal half of helix α5 were shown, in the case of Rab3A complexed with Rabphilin3 (Ostermeier & Brunger, 1999), to be major players in the complex interface, forming what was called the "Rab Complementarity Domain" (see also the discussion about the Rab3a-Rabphilin3 crystal structure on page17).

The electrostatic surface potential of this area is more or less neutral and does not display significant positively or negatively charged patches. In addition, as already discussed, in the crystallized Ypt7p construct, the last 26 residues were removed, since it is known (Neu *et al.*, 1997) that about the last 25 residues of the C-terminus are flexible in solution. Based on this and the observed charge distribution, it seems unlikely that this surface is oriented towards the negatively charged membrane to which all Ypt/Rab-proteins are anchored *via* their geranylgeranylated C-terminus; more probably, this surface is rather accessible for specific interacting proteins.

The interface between the switch I and switch II region displays significant differences in the main-chain trace as well as in the side-chain orientation compared to the described structures of rat Rab3a, mouse Rab5c, yeast Ypt51p and yeast Sec4p. The structural plasticity in this region was first highlighted by Merithew (Merithew *et al.*, 2001) in mouse Rab5c. This interface is mainly determined by the interaction of three aromatic amino acid side-chains. Phenylalanine at position 45 and tryptophan at position 63 are absolutely conserved in all members of the Ypt/Rab-family. The last member of this hydrophobic triad (at position 78) is phenylalanine in Ypt7p and Rab7, whereas it is a tyrosine in all other Ypt/Rab-family members. The side-chain of Phe78 in Ypt7p•GppNHp is oriented towards the interior of the protein, whereas the tyrosines at the corresponding position are oriented towards the surface of the respective Ypt/Rab-protein. As a consequence, the main-chain trace between residues 42 and 45 displays a different trace compared to all other Ypt/Rab-protein structures. The main-chain trace is shifted around 2 Å towards the switch II region, displaying the maximum shift at the position Ala43. Together with slightly different side-chain orientations of the aminoacids Phe45 and Trp63, this feature is further evidence for the structural plasticity in the RabF2/RabF3/RabF4 epitope.

A peculiar feature of Ypt7p is the extended loop L4 that starts at position Ala66 and ends at position Gly81. In all other structures of Rab/Ypt proteins and other members of the Ras-superfamily, this region contains a well defined α helix. However, the region between residues 73 to 78 of Ypt7p•GppNHp is in an extended loop conformation, and not in the helical conformation adopted by the corresponding residues of all other small G proteins.

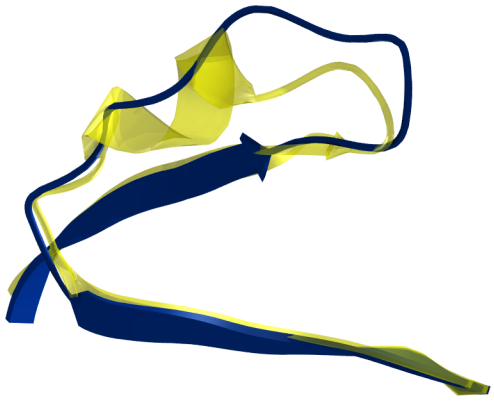


Figure III-9: Comparison between Ypt7p•GppNHp (blue) and Rab3A•GppNHp (transparent yellow), showing the unusual disorganization of helix α2 in the case of Ypt7p

There are two main reasons which suggest that this feature is not due to crystal packing artifacts: first, there are no obvious crystal packing interactions in the $P2_12_12_1$ crystal form which could induce the loss of the helical structure. The region is involved in some minor crystal contacts, but the electron density map is very well defined. Consistent with this, the crystallographic thermal factors for the residues in that area are around 15Å^2, even lower than the B factors for the main chain, 17.5Å^2.

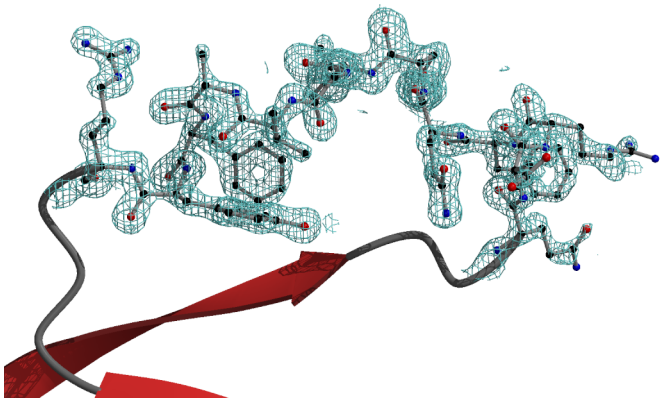


Figure III-10: Electron density around residues 68 to 80 of Ypt7p•GppNHp is very well defined. Map contoured at 2σ.
Orientation is similar to that in **Figure III-9**.

Second, the respective loop could not be traced at all in the $P6_1$ crystal form (Section III.2.3, page 55). However, in the $P6_1$ crystal form the two *switch* regions of neighboring molecules were facing and possibly disturbing each other. This suggests two hypotheses: either the helix was partly unstructured before crystal contacts were created, and the extended loop could fit in various positions, or the helix was actively unwound during the crystallization. Nevertheless, it seems to be less probable that the crystallization-promoting forces would be strong enough to actively drive the "melting" of a helix, unless it is intrinsically unstable. Therefore, the lack of a helical structure from beginning seems to be most likely.

In order to test this hypothesis, molecular dynamics simulations were run using the program GROMACS (Berendsen *et al.*, 1995). GROMACS is a molecular dynamics program that solves Newton's equations of motion for particles (atoms) in small time steps, in cartesian space. The system is coupled to an external temperature bath and is pressure-controlled. As

with any other molecular dynamics (MD) program, the main problem is the fact that the simulation uses classical mechanics in order to describe the motion of atoms; in fact, it would be impossible today to solve the quantum mechanical ones for systems as large as proteins. Nevertheless, the use of constrains for describing bond lengths and angles proved to be a reasonable workaround in many different MD programs.

Three test runs were with GROMACS. For the first one, a molecular chimera was prepared. The L4/α2/L5 region of Rab3A (residues 81-97) was used to replace the region between residues 68 and 84 in Ypt7p. Rab3A was chosen because this region does not have any insertions compared to Ypt7p. The residues were mutated, such that the sequence was the one of Ypt7p. For each mutation, one of the most frequent rotamers was used, care being taken not to generate clashes. This chimera was placed in a periodic box, keeping 5Å distance between the molecule and the boundaries of the periodic box, solvated and equilibrated during a short energy minimization protocol using a steepest descent algorithm and then submitted to a MD simulation. The time span was 4ps and snapshots were analyzed for time steps spaced 0.5ps apart. After the first picosecond, the α helix spanned only residues 73-77 (compared with 73-78 at the beginning of the simulation) and after two picoseconds no α helical structure could be detected. At some time points, residues 75-77 could be seen again in an α helical conformation, but this did not have a long lifetime. During the whole 4ps simulation, helix α2 was the only one to loose its conformation, whereas all other secondary structure elements remained well defined, only shrinking or extending by 1-2 residues during some time periods.

In the control simulation, Rab3a was submitted to the same MD protocol. The conformation of loop L4 varied slightly during the MD run, but, overall helix α2 did not disorganize, although it extended or shrank by one or maximally two residues. These MD-simulations are in harmony with the conclusion that the observed non-helical conformation is a genuine feature of Ypt7p and is not due to crystal packing artifacts.

From all the reasons exposed above, we do not think that the observed non-helical conformation is due to crystal packing artifacts, but that it represents a genuine feature of Ypt7p. This might be a consequence of the different packing between the above-described aromatic aminoacids Phe45, Trp63 and Phe78. In particular, the orientation of Phe78 prevents this loop from adopting a helical backbone conformation. However, it is improbable that this is the only reason, because a third MD simulation, with the Ypt7p-Rab3A chimera,

having Phe78 mutated to Val did not behave significantly different: the helix was still disorganized.

Therefore, in addition to amino acids of the RabF2/RabF4 epitope, the amino acids of the RabF3 epitope display significantly different orientations. These described structural differences contribute to a substantial difference in charge distribution and surface topology of Ypt7p, and most probably of Rab7, in view of the sequence identity between Ypt7p and Rab7.

In the structure of Ypt7p bound to GDP, the electron density in this area is poor and does not allow confident tracing of some of the amino acids of the switch I (residues 38 to 41) and switch II regions (residues 67 to 77). The switch II region becomes ordered at the well defined residue Phe78 which, in concert with Tyr79, is involved in hydrophobic interactions that include the conserved Trp63 (see also **Figure III-12**).

III.4.1.b. Side chain conformations

Tyr 33 and positioning of the purine base

Tyr33 is localized on top of the guanine base, forming an edge-to-face interaction with it. In most members of the ras superfamily, this residue is a phenylalanine (residue Phe28 in H-Ras p21). Comparison with Ypt51p, Sec4p, Rab3a and Rab5c reveals that the aromatic rings of phenylalanine and the corresponding tyrosine overlap. However, the guanine base and the ribose sugar moiety of the bound nucleotide are slightly shifted and turned in the case of Ypt7p (see Figure 5), whilst keeping the β- and γ-phosphate at the same positions. The side chain hydroxyl group of Tyr33 forms a hydrogen bond with the nucleotide sugar oxygen O2' (d=2.8 Å) and with the side chain OE1/NE2 (d=2.9 Å) of Gln35. Both side chain atoms of Gln35, OE1/NE2, are also involved in hydrogen bonding with Lys160-NZ (d=3.1 Å). The same interactions are seen in the GDP-form of Ypt7p. In both structures, the GXP base plane is tilted by about 11.3 ° compared to the position of the guanine-nucleotide base in other well-defined structures.

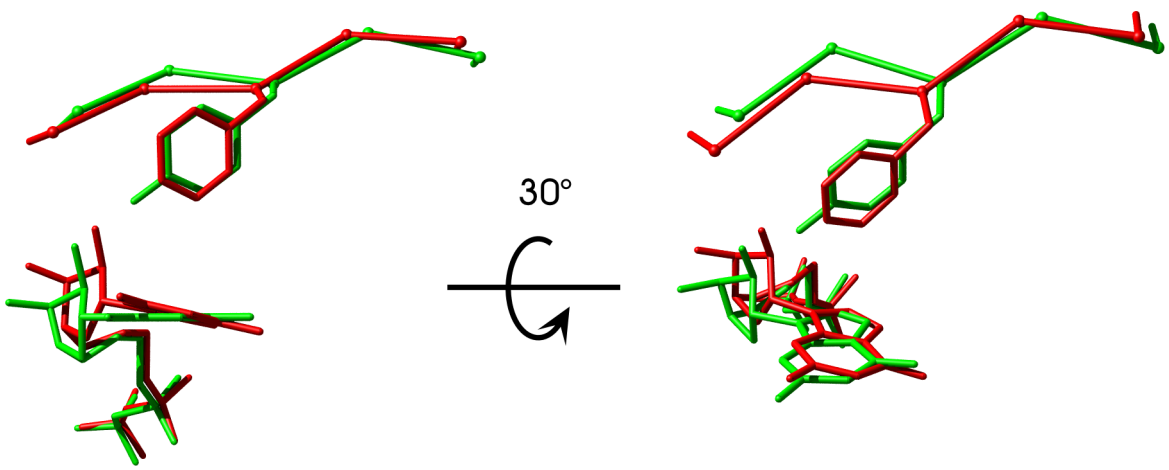


Figure III-11: Tilting of the purine base due to the Phe32→Tyr33 substitution in Ypt7p•GppNHp (green), shown *versus* Ypt51p•GppNHp (red). The oxygens O1A and O2A were removed for the sake of clarity. In the left image, the base is seen "edge on". The Cα atoms are represented like blobs.

Tyr 37

Tyr37 of Ypt7p (equivalent to Tyr32 of H-Ras p21) is oriented away from the bound nucleotide, similar to Lys36 of Ypt51p (Esters *et al.*, 2000) and Gln50 in Rab5c (Merithew *et al.*, 2001). Thus, it does not interact with GTP, but forms a hydrogen bond between its side chain hydroxyl group and the side chain of His26, creating part of a hydrophobic surface patch. However, the corresponding pair in Ypt51p and Rab5c (Leu25/Lys36 and Leu39/Gln50, respectively) contains a hydrophobic and a charged residue. Tyr37 does not adopt a conformation facing the solvent and covering the nucleotide binding pocket, as observed for Phe51 in Rab3a and Phe49 in Sec4p (see Figure 6). This might indicate a conformational difference between Ypt/Rab proteins of the endocytic pathway and those of the exocytic pathway. Following this rationale, we expect the above described conformational difference in the switch I region to be shared with Ypt1p and Rab1a, since both these proteins have tyrosines at the same position.

The above-described aromatic residues at positions 33 and 37 determine a significant variation in the surface topology when compared with Ypt51p. In addition, the three isoforms of Ypt5p (Ypt51p, Ypt52p and Ypt53p) have a positively charged residue at position 37 and therefore the charge distribution of this surface patch is significantly changed. These two

features might form an important determinant for specificity of binding partners for Ypt5Xp *versus* Ypt7p, enabling specific effectors to discriminate between Ypt7p and Ypt5Xp.

III.4.2. Differences between the GDP and GppNHp forms

As was already mentioned at the beginning of this work, one of the reasons for trying to crystallize both the GTP and GDP forms of Ypt7p was to gain a better understanding of exactly what changes take place after GTP hydrolysis. Although the generic differences between the GTP- and GDP-bound states are known for Ras p21 and now even for Sec4p (Stroupe & Brunger, 2000), small variations among several Rab/Ypt proteins in the same nucleotide-bound state are crucial for their function. Therefore, the determination of their individual structures with high resolution is needed for a detailed understanding.

GTP to GDP hydrolysis leads to the loss of the γ-phosphate group and, as a direct consequence, two hydrogen bonds, between O1G and the amide hydrogen atom of Thr40 (part of loop L2) and between O2G and the amide hydrogen atom of Gly67, part of loop L4, are lost (see **Figure I-3**). This changes lead to large structural changes localized in the loops L2 and L4 (which, as discussed above, is unusually long and is followed directly by sheet β4).

III.4.2.a. Switch I region

As a direct consequence of the loss of the hydrogen bond between the γ-phosphate and residue Thr40, the whole loop L2 is not fixed anymore, and residues Glu36 to Phe45 undergo displacements larger than 1Å. Furthermore, residues 38, 39 and 40 could not be traced in the electron density at all. We can confidently define the "switch I" region in Ypt7p as the region between residues 36 and 45. This is also in accordance with the Ras definition of "switch I" region, which spans residues 30-38 and with Sec4p (residues 48 to 56, equivalent to 48-57 in Ypt7p).

The biggest displacement, of 8.7Å, is seen for Ile41; this is almost three time as large as the one observed in the case of Ras p21, where Ile36 is moved away by 3.1Å only.

III.4.2.b. Switch II region

The second region which undergoes serious rearrangements upon GTP hydrolysis is the extended loop L4. Residues 67-77 could not be traced at all in the electron density in the GDP state, whereas residues Thr65, Ala66 and Phe78 to Gly81 have RMSDs above 1.0Å

when comparing the GDP structure with the GTP one. The switch II region appears to be slightly smaller than the corresponding region in Sec4p (equivalent residues would have been Thr65 to Ala82). Further more, the loop regions of the two forms come "in register" earlier in the Ypt7p sequence than in the Sec4p case. This is mostly due to a hydrophobic interaction between Trp63 and Phe78, which pack in an "edge-to-face" manner in the GTP state and in a "face-to-face" mode in the GDP case. The corresponding residues in Sec4p (Trp74 and Tyr89) are too far apart to interact. This difference can be easily explained if one notes the fact that Tyr89 of Sec4p is part of an α helix with a defined structure, whereas Phe78 of Ypt7p is part of a loop able to adopt virtually any conformation admitted by the Ramachandran plot.

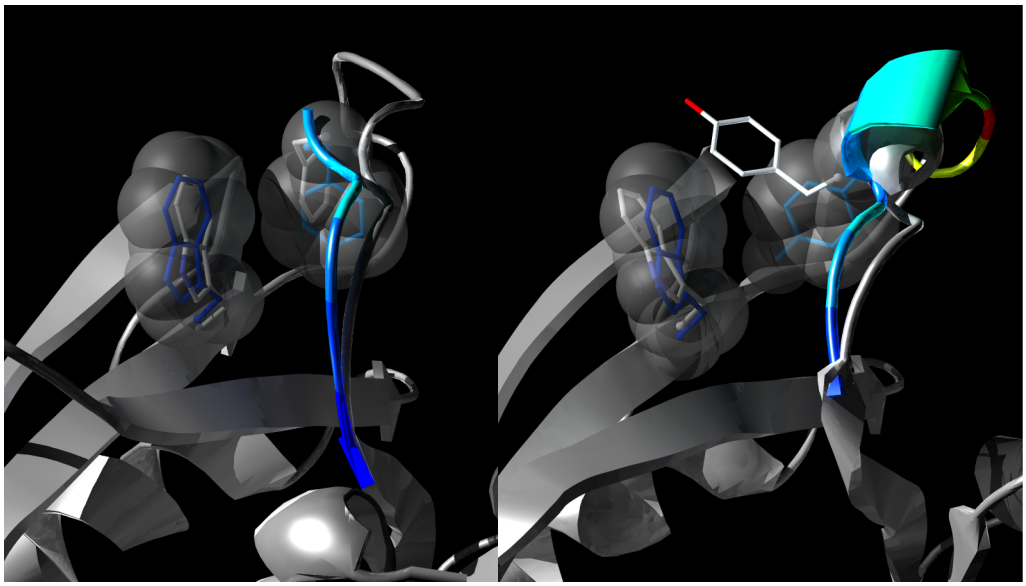


Figure III-12: Comparison between the changes in the switch II region of Ypt7p (left) and Sec4p (right) upon GTP hydrolysis. In the GppNHp form, the conserved Trp and the Phe/Tyr (respectively) are colored according to the CPK coloring scheme. In the GDP form, the loop L4 (Ypt7p) and L5 (Sec4p) are gradient-colored, in accordance with the displacement from corresponding atoms of the GppNHp structure. Colors range from dark blue to light blue, cyan, green and yellow for displacements less than 1Å, 2Å, 3Å, 5Å, 6Å to red for more than 6Å. The displayed residues are colored according to the same scheme. The semitransparent surfaces represent spacefilling models of the residues in the GDP form. Note the edge-to-face interaction in the Ypt7p case and the far bigger distance between equivalent residues in Sec4p.

III.5. Conclusions

The high-resolution crystal structures of Ypt7p in the GTP- and the GDP-bound conformations have been determined and described. The first engineering of Ypt7p yielded well-diffracting crystals of Ypt7p in the GTP form, but later it was shown that the *switch* regions, responsible for the conformational differences between the GDP and GTP conformations, were not visible in the electron density due to unfortunate crystal packing. In addition, the end of the C-terminal helix was not visible in the electron density. A second construct was designed with the hope that these problems would be circumvented. The strategy proved successful: the protein could be crystallized and the structures of both the GTP and the GDP conformations solved to very high resolution (1.6Å in the GppNHp-bound case and 1.3Å in the case of the GDP structure). Although it was postulated that the differences between the GDP and GTP state are similar to the ones observed in Ras, only recently the structure of Sec4p (Stroupe & Brunger, 2000) in the GTP and GDP states showed that qualitatively the same differences can be seen in the Ypt family as well. Different Ypt proteins have to be, nevertheless, distinguished from each other in either the GDP- or GTP-bound states. Small variations among several Ypt/Rab-proteins in the different nucleotide-bound states are crucial for their specific function. Therefore, the determination of their individual structures with high resolution are needed for a comprehensive understanding of effector binding specificity.

The presented differences in main-chain and side-chain conformations of Ypt7p compared to Ypt51p, Sec4p, Rab3a and Rab5c, in addition to the sequence differences, contribute to structural plasticity and variations in both the surface charge distribution and surface topology within the Ypt/Rab-family of proteins. The observed structural differences are clustered in areas which were recently highlighted as Rab-specific or Rab-subfamily specific regions (Pereira-Leal & Seabra, 2000). For the RabSF2 region (part of the switch I region) Ypt7p displays, in the GTP-bound form, a shifted conformation when compared with other Ypt/Rab structures and a different surface topology, mainly as a consequence of the two tyrosine residues Tyr33 and Tyr37. In the structure of Ypt7p•GppNHp, helix α2 is unwound, which results in an extended loop spanning from β-strand β3 to β4, including the RabF3 region. Molecular dynamics simulations revealed that the intrinsic sequence of Ypt7p within this region determines this feature. These two conformational changes associated with the

switch I and the switch II region are likely to contribute to Ypt7p specific binding of effector proteins and presumably help to discriminate between Ypt7p and Ypt51p when they are located on the same membrane. Furthermore, due to insertions in loops L3 and L7, the neighboring RabSF1 and RabSF4 regions are different in their conformation compared with other Ypt/Rab proteins. Therefore, it is probable that this surface patch can serve as an additional binding site for Ypt7p-specific effector proteins. These variations, each taken in part, are small, but they add up and most probably become a specificity tag allowing differentiation between the otherwise sequence- and fold-similar Ypt/Rab proteins.

It is a common feature that the GDP-bound form of Ras-related small GTP-binding proteins has a less well-defined structure than the GTP-bound form. In particular, the important switch I and switch II regions are very flexible when the protein is in complex with GDP, lacking the γ-phosphate group which fixes these regions. This higher conformational mobility in the GDP-bound form might help proteins which interact promiscuously with Ypt/Rab proteins in the GDP-bound form to overcome variations in the amino acid sequence and topology at the interaction surface (e.g. GDP dissociation inhibitor, GDI, (Garrett *et al.*, 1994)).

With the structure in hand, the results of kinetic characterization of different Ypt7p mutants can be more easily understood. The GTPase activity of Ypt7p is most effectively enhanced (8×10^5 fold) *in vitro* by the GAP Gyp7p (Vollmer *et al.*, 1999) and to a lesser extent by the GAPs Gyp6p and Gyp1p (Tickle *et al.*, 1998). This acceleration of the GTPase activity by Gyp7p is reduced about 5 times when Ypt7p-I41M, Ypt7p-D44N and Ypt7p-Q68L mutant proteins are used as substrates (Vollmer *et al.*, 1999). The reduced activity of the Ypt7p-Q68L mutant can be rationalized by the fact that this glutamine is supposed to be an essential part of the active site for the GTP-hydrolysis reaction in analogy to the Gln61 of H-Ras p21 (Frech *et al.*, 1994). It is worth mentioning that activity is not completely abolished in the Ypt7p-Q68L mutant, as it is observed for the similar H-Ras p21-Q61L mutant. It can be hypothesized that in the binary complex between Ypt7p and its cognate GAP, the glutamine 68 might not take the same crucial role as Gln61 does in H-Ras p21. The other two positions (41 and 44) of Ypt7p are in the postulated interface between the Ypt protein and GAP (Rak *et al.*, 2000) and perturbation of the surface topology might be responsible for the weakened interaction between the two proteins.

It would have been more rewarding and instructive to gain a structural insight on the complex between Ypt7p and Mrs6p or even the ternary complex between the GGTase II, Ypt7p and Mrs6p. Unfortunately, neither of the complexes was amenable to crystallization. No crystals could be obtained for the ternary complex (A. Rak, personal communication), while the Ypt7p-Mrs6p complex yielded only two types of crystallization results: precipitates or phase separations. However, to gain a better understanding of this complex system, I set out to investigate the interactions between different Ypt proteins and Mrs6p, using steady-state and pre steady-state methods. The results of these experiments are described in the next chapter.

Chapter IV.

Kinetic analysis of interaction between Mrs6p and different Ypt proteins

IV.1.Introduction

Although much is known, at the molecular level, about the prenylation machinery and how it works, the quantitative description of the system is still missing. More exactly, the control steps involved in this process are far from being understood. Characterizing the system at the level of the interaction between its components might shed more light on the way it is regulated.

To this end, I tried to use various techniques in order to understand and characterize the system. Equilibrium titrations were used in the first place, since they can provide information about the affinity between the components of the system. The different combinations used showed that, concerning the interaction with Mrs6p, *Saccharomyces cerevisiae* Ypt proteins can be divided into two groups: a set of Ypt proteins binding tightly (in the low nM range) and a second set with affinities lower by ca. one order of magnitude. It was also shown that this difference applies only to the GDP form and not to the GTP form. Trying to understand better where this difference arises from, I performed pre steady state analysis to learn about the differences in the association and dissociation rates. Last, but not least, a recently developed method was used in a trial to structurally characterize the interaction interface between Mrs6p and one Ypt protein.

IV.2.Equilibrium relationships

IV.2.1. Equilibrium titrations: the principle[1]

Taken a system consisting of a ligand, L, and its receptor, R. The following reaction will take place:

$$L + R \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} LR$$

[1] Most of the theory presented in this chapter is based on (Gutfreund, 1995)

During an equilibrium titration experiment, the researcher tries to find out how much receptor is bound to ligand at certain ligand and receptor concentrations. This will give information about how tight the interaction between the two binding partners is, namely the dissociation constant k_D of the equilibrium. Two mass action rate constants will describe the system: k_{12} will be named the association rate constant, while k_{21} will be referred to as the dissociation rate constant. The mass action law states that *the rate of reaction is directly proportional to the product of the concentrations of the reactants* and the proportionality constant is the rate constant. If we denote the concentration of free ligand by c_L and the one of free receptor by c_R, then we can write the following equations:

$$\textit{rate of association} = \left[\frac{dc_{LR}}{dt}\right]_{association} = k_{12}c_L c_R \qquad \text{(IV.1)}$$

$$\text{and } \textit{rate of dissociation} = -\left[\frac{dc_{LR}}{dt}\right]_{dissociation} = k_{21}c_{LR}. \qquad \text{(IV.2)}$$

Summing up the dissociation and association rates, we obtain the rate by which the concentration of the LR complex varies over time:

$$\left[\frac{dc_{LR}}{dt}\right] = k_{12}c_L c_R - k_{21}c_{LR}. \qquad \text{(IV.3)}$$

It is important to note that, at equilibrium, this rate of change equals zero, that is: the amount of LR complexes formed in the time unit equals the amount of LR complexes that dissociate in the same time period. It is easier to use the fraction (or proportion) of LR which is formed at a certain moment (and the fraction of free R at that moment), such that

$$p_R + p_{LR} = 1, \text{ where } p_R = \frac{c_R}{c_{R0}} \text{ and } p_{LR} = \frac{c_{LR}}{c_{R0}} \qquad \text{(IV.4)}$$

Applying this substitution in equation (IV.3), and taking into the account that the equilibration proceeds over time and that concentrations of ligand and receptor will vary during the equilibration, one can rewrite equation (IV.3) as

$$\left[\frac{dp_{LR}(t)}{dt}\right] = k_{12}c_L(t)p_R(t) - k_{21}p_{LR}(t) \qquad \text{(IV.5)}$$

or, using (IV.4),

$$\left[\frac{dp_{LR}(t)}{dt}\right] = k_{12}c_L(t)(1\text{-}p_{LR}(t)) - k_{21}p_{LR}(t) = -(k_{12}c_L(t) + k_{21})p_{LR}(t) + k_{12}c_L(t) \quad \text{(IV.6)}$$

so that the dependency of the concentrations and bound fractions with time is explicitly written. At equilibrium ($t=\infty$), the rate of change of the concentration of LR (and of the proportion of LR) equals zero, so that:

$$k_{12}c_L(t)(1-p_{LR}(t)) - k_{21}p_{LR}(t) = 0 \quad \text{for } t=\infty \quad \text{(IV.7)}.$$

It is usual to perform measurements using a vast excess of ligand over receptor. In this case, the variation of the free ligand concentration over time is insignificant (since just a very small amount of ligand can be bound by the receptor, even at full saturation of the receptor) and the reaction, albeit a second order one, seems to behave like a first order reaction, with an exponential decay for the free receptor concentration and, consequently, with an exponential rise for the RL complex concentration. Because of this, it is said that the experiment is performed under ***pseudo first order conditions***. If this is the case, then $c_L(t)$ at any time can be approximated to $c_L(0)$ and, for simplification, it will be noted as c_L. Equation (IV.7) can be written as

$$k_{12}c_L(1-p_{LR}(t)) - k_{21}p_{LR}(t) = 0 \quad \text{for } t=\infty \quad \text{(IV.8)}$$

which leads to

$$p_{LR}(\infty) = \frac{k_{12}c_L}{k_{12}c_L + k_{21}} = \frac{c_L}{c_L + k_D} \quad \text{(IV.9)}$$

where k_D is called the ***dissociation constant*** of the equilibrium (or the *equilibrium dissociation constant*) and is defined as the ratio between the dissociation and the association rate constants

$$k_D = \frac{k_{21}}{k_{12}} \quad \text{(IV.10)}$$

Inspection of equation (IV.9) reveals that k_D, under these pseudo first order conditions, can be defined as the concentration of ligand at which half of the receptors are in complex with the ligand. The graphical representation of the equation (IV.9) yields a part of a limb of a rectangular hyperbola.

It is important to note that the definitions above consider that $c_L(\infty) = c_L(0)$ (noted by c_L), which is the case only if the experiment is performed under pseudo first order conditions. Unfortunately, in practice this is not easy to acquire, hence the need for deriving

a more robust calculation for the fraction of receptor involved in receptor-ligand complexes. Using equations (IV.1) and (IV.2) and equating them, one can obtain

$$k_D = \frac{k_{21}}{k_{12}} = \frac{c_L c_R}{c_{LR}} \quad \text{(IV.11)}.$$

The problem is that, usually, one does not know the amounts of *free* ligand and receptor, but only the total amounts of L and R and the concentration of the complex LR (for which a proportional signal is read out). Using the p_{LR} notation and the mass conservation rules ($c_L + c_{LR} = c_{L\,total} = c_{L0}$ etc.), the following substitutions can be made: $c_L = c_{L0}(1-p_{LR})$ and $c_R = c_{R0}(1-p_{LR})$. But the concentration of bound L equals the one of bound R and this is the concentration of the LR complex, so we can write:

$$c_{LR} = p_{LR} \cdot c_{R0}\,; \quad c_L = c_{L0}(1 - c_{R0} \cdot p_{LR})\,; \quad c_R = c_{R0}(1 - p_R) \quad \text{(IV.12)}$$

Substituting these terms in the k_D equation (IV.11), we can write

$$k_D = \frac{(c_{L0} - p_R c_{R0}) c_{R0}(1 - p_R)}{p_R c_{R0}} \quad \text{(IV.13)}$$

which leads to a quadratic equation in p_R with the solutions:

$$p_R = \frac{(c_{L0} + c_{R0} + k_D) \pm \sqrt{(c_{L0} + c_{R0} + k_D)^2 - 4 c_{L0} c_{R0}}}{2 c_{R0}} \quad \text{(IV.14)}$$

It is easy to realize that only the quadratic equation solution in which the radical is subtracted has physical meaning ($0 \le p_R \le 1$). The other one, with the radical preceded by the plus sign, will always yield values bigger than 1.

The advantage of calculating the proportion of receptor participating in ligand-receptor complexes according to formula in (IV.14) is that the calculation is valid for any concentrations of L and R and for any k_D.

In a practical situation, one starts with a certain quantity of receptor, adds ligand stepwise and determines the concentration of the LR complex at that particular ligand concentration. The obtained values are fitted to equation (IV.14), resulting in the determination of the value of k_D.

In the cases presented in the next sections, the concentration of the LR complex was determined using fluorescence as a readout. The use of fluorescence as a signal has several advantages: usually the signal over noise ratio is high and the fluorescence undergoes

significant changes upon binding of the ligand to the receptor. In a fluorescence experiment, samples are illuminated (excited) by light at a certain wavelength and the emitted fluorescent light is detected at right angle to the exciting light. This means that the (usually faint) emitted light will be read against a practically dark background; in contrast, when using transmittance readings, a small variation has to be read against the high background of the sample, resulting in a poorer signal-to-noise ratio.

A second advantage of fluorescence is that it can easily be used as an interaction reporter. This arises from two different reasons: first, fluorescence of some dyes might change dramatically upon changing the environment of that dye; secondly, there can be a fluorescence resonance energy transfer between a donor on one molecule and an acceptor on a second molecule.

Fluorescence resonance energy transfer (FRET) is a distance-dependent interaction between the electronic excited states of two dye molecules in which excitation is transferred from a donor molecule to an acceptor molecule **without** emission of a photon. For this reason is also known as ***non-radiative energy transfer***.

The efficiency of FRET depends on the inverse sixth power of the distance between the donor/acceptor pair:

$$E = \frac{R_0^6}{R_0^6 + R^6}$$

where R_0 is the Förster distance (or radius), which is specific for each donor/acceptor pair and depends on the properties of the donor and the acceptor, like the dipole orientation of the two, the fluorescence quantum yield of the donor, the refractive index of the media and the spectral overlap of the two species. Usually, the Förster radius is around 30-60Å. This means that with a well-chosen pair of donor (on one interaction partner) and acceptor (on the second interaction partner), a situation can be achieved in which, upon binding, either the appearance of sensitized fluorescence of the acceptor or the quenching of donor fluorescence can be monitored. Proteins have natural fluorophores, namely tryptophan and phenylalanine residues, which, both, absorb in the range 275-295nm and emit in the range of 310-340nm, but the tryptophan residues absorb about five times more strongly than tyrosine residues, making them the most important ones.

The choice of using a fluorescence signal for the kinetic experiments was made also because this kind of signal is able to report in real time on the interactions between the proteins, making it the most useful tool for fast kinetics experiments.

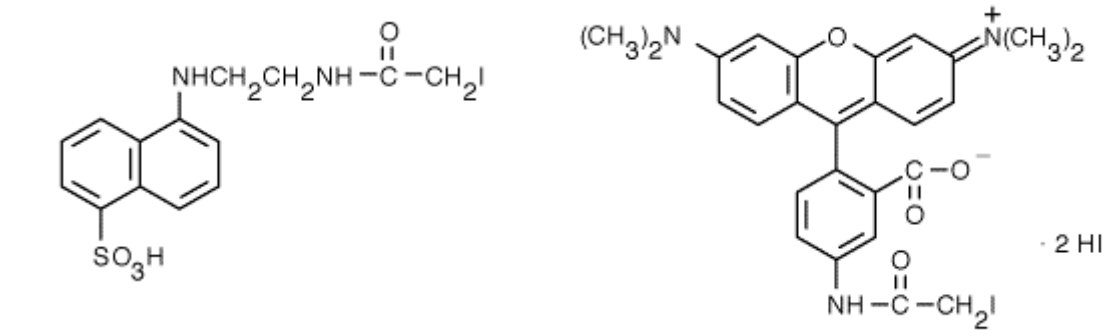
In this work, the following strategy was employed for equilibrium titrations: a small amount of Ypt protein, fluorescently labeled, was diluted in 1ml of buffer solution (25mM HEPES Na pH 7.2, 50mM NaCl, 5mM DTT, 1mM $MgCl_2$ and 10μM GDP) and set up in a stirred fluorescence cuvette, thermostated at 25°C. The final concentration was in the range of 30-500nM. Mrs6p was subsequently added in small aliquots and fluorescence was read out during the whole experiment. The recorded values were plotted against the total concentration of Mrs6p and a least squares fitting routine as implemented in the program Grafit was used to derive the k_D for the reactions. In the case of a competitive titration, a different approach was employed (see SectionIV.2.3.a).

IV.2.2. Interaction between Mrs6p and Ypt7p and Rab7, GDP form

As a continuation of the structural work done, the first Ypt protein to be characterized for its ability to bind Mrs6p was Ypt7p. Two different fluorophores were used in equilibrium titrations: rhodamine and dansyl, both as haloalkyl reagents, which react readily, in a selective manner, with thiol groups of proteins, yielding thioether-coupled products. The two fluorophores used were naphthylamine-5-sulfonic acid (DANS) and tetramethylrhodamine (TMR or rhodamine) (**Figure IV-1**). The fluorescently labeled proteins obtained will be referred to as YptX dans and YptX rhodamine.

It is known that the C-terminal tail of Ypt/Rab proteins is not structured in solution ((Neu *et al.*, 1997) and discussion in Section III.2.1). However, since the tail contains two cysteine residues which subject to geranylgeranylation, large structural rearrangements are required before and during the prenylation reaction (see also the discussion about the GGTase II in Section I.2.2.b). Hence, a label placed in this part of the protein should be able to report on the environment changes and thus on the interactions with other proteins, in particular if these proteins are able to bind the C terminus of Ypt, for example Mrs6p or GGTase II. Ypt7p contains four cysteine residues. Two of them were seen in the electron density map and are buried in the hydrophobic core. The other two are placed at the C terminus, so are the only ones likely to be labeled with a water-soluble reagent capable of reacting with free cysteines.

This is also in accordance with other results, showing that in Rab7 only two cysteines (C205 and C207) were likely to be labeled when challenged with a water-soluble reagent (P. Metcalf, unpublished data). For this aim, Ypt7 was incubated with 1,5-IAEDANS or 5-TMRIA for 2 hours at 4°C. After this, the buffer was exchanged and the excess label removed by passing the reaction mixture over a PD-10 column preequilibrated with 20mM HEPES pH 7.2, 10mM NaCl and 2mM DTE, 10μM GDP; the resulted protein was concentrated and stored in aliquots at –80°C. Incorporation of the label was assessed by ESI-MS and fluorescence yield measurements, revealing that both cysteines were quantitatively labeled.



a b

Figure IV-1: Two reagents used for fluorescently labeling different Ypt proteins
a– N-(Iodoacetaminoethyl)-1-naphthylamine-5-sulfonic acid (1,5 I-A-EDANS)
b– tetramethylrhodamine-5-iodoacetamide dihydroiodide (5-TMRIA)

For the Ypt7p-dans•GDP interaction with Mrs6p, two different types of signals were used: the direct signal (exciting the dansyl fluorophore at 346nm and reading out the fluorescence signal at 490nm) and the FRET signal, using tryptophan as donor and dansyl as acceptor and monitoring the increase in dansyl fluorescence (the monochromator for excitation was set at 280nm, while the one for emission was set at 490nm). The result of this titration was somehow unexpected: the obtained k_D, of about 480nM (**Figure IV-2**) was far above the k_D of the mammalian homologous system (Rab7 against Rep1), which displayed affinities in the order of 1-2nM (Alexandrov *et al.*, 1998). In order to test whether this result was correct, a second experiment was performed, in which the FRET signal was used.

Using FRET with tryptophan as donor involves a problem: at every step in the titration, the excitation background is raised, since the added protein also contains tryptophans. Due to this, the titration curve will not reach an obvious plateau. This behavior can be corrected in the following manner: if enough points have been recorded after the real saturation of the receptor (Ypt7p dansyl in this case), they will reflect only the contribution of the tryptophans added. A linear fit can be performed on this region of the plot, deducing the amount by which the recorded fluorescence increases during each titration step, and subtracting it from the recorded values. Such an approach proved efficient when titrating Ypt7p dansyl against Mrs6p using FRET from Trp to dansyl, as can be seen in **Figure IV-3**, where the corrected data reaches a plateau. The result obtained from this experiment was in agreement with the one obtained from the first experiment, suggesting a k_D value in the range of 300-500nM.

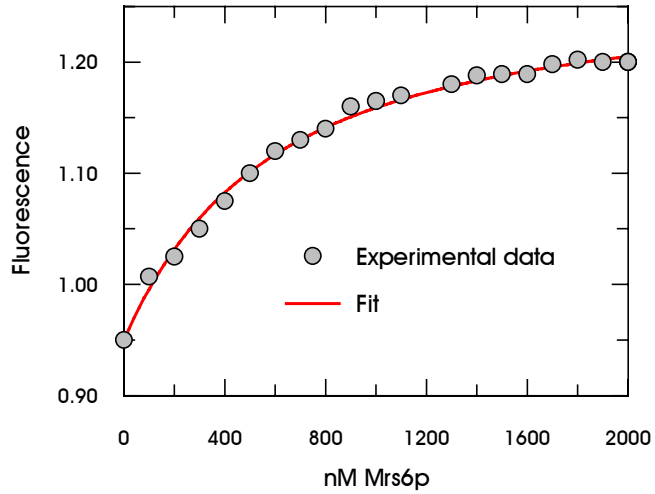


Figure IV-2: Titration of Ypt7 dansyl with Mrs6p, using direct fluorescence signal. The fitted curve corresponds to a k_D of 480±40nM. Concentration of Ypt7p•GDP dansyl was 136nM.

To be more confident about the results obtained, a third titration was performed, this time with Ypt7p labeled with tetramethylrhodamine. The fit to this titration curve indicated a k_D value of 350±33nM, also in agreement with the previously obtained values.

These first results suggested that the affinity between the components of the yeast prenylation system were about two orders of magnitude lower than the ones of the mammalian

system. In order to clarify this, an experiment was performed which, at a first glance, might seem strange: assessing the interaction between Rab7 and Mrs6p, a "molecular chimera".

To this end, a strategy similar to that used in the case of Ypt7p was used: Rab7 was labeled with dansyl and the interaction was characterized by reading the increase in the fluorescence of dansyl following FRET from tryptophan. Much to our surprise, the data obtained showed that the hybrid system has a k_D of about 13nM (**Figure IV-4 a**).

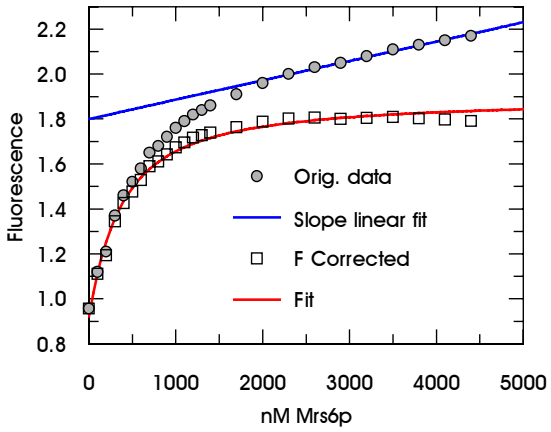


Figure IV-3: Titration of Ypt7p•GDP with Mrs6p. The squares represent the data corrected for the contribution of the intrinsic tryptophan fluorescence. The red curve corresponds to a fit for k_D=300±25nM. Concentration of Ypt7p•GDP dansyl was 136nM.

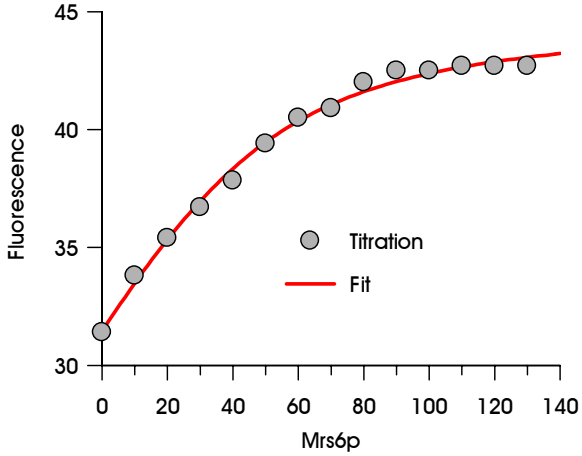


Figure IV-4: Titration of Rab7dans•GDP with Mrs6p. In the fitted curve corresponds to a k_D of 13.5±3.2nM. Concentration of Rab7dansyl was 50nM.

From these experiments, the conclusion which could be drawn was that Mrs6p *per se* is able to bind Ypt/Rab proteins strongly, but Ypt7p is not able to bind tightly to Mrs6p. If Mrs6p was able to bind substrates with affinities of about 10-15nM. It was therefore interesting to see if there are, in yeast, binding partners able to take advantage of this binding capacity.

IV.2.3. The yeast tight binders: Ypt1p and Sec4p

One such substrate proved to be Ypt1. Titration of Ypt1 dansyl with Mrs6p yielded a titration curve which suggested a k_D of about 11nM. Preliminary data suggested that Sec4p was also able to bind Mrs6p very tightly, but the exact affinity could not be measured using this approach. The problem when doing equilibrium titrations is that the concentration of the labeled partner (which is kept constant during the whole experiment) should not be too much higher than the k_D of the system. If this condition is not satisfied, in the sense that the concentration is much higher than k_D, the titration curve will rise linearly to the endpoint, exhibit a sharp kink at a ligand concentration equal with that of the receptor and then plateau; the only information which can be derived from such a curve is the stoichiometry of the reaction. The k_D obtained from such a curve is prone to large errors.

This condition could not be satisfied in the case of Ypt1p and not even reached in the Sec4p case. In principle, since the titration signal is fluorescence, read against a black background, the concentration of the label can be lowered as much as desired, in order to reach the condition $c_R < k_D$. In practice, however, a problem arises: the solvent scatters light by two mechanisms: one is the Rayleigh scattering, which occurs at a wavelength equal to that of the excitation and the other one is the Raman scattering at wavelengths longer that the one of the exciting light. If the concentration of the labeled compound is very low, the fluorescent light emitted is very faint and can become swamped by the background noise. The intensity of the excitation light cannot be increased without limits, since too intense light will result in "photobleaching", the destruction of the fluorophore. A possible solution for such a situation is to use a competitive titration strategy.

IV.2.3.a. Competitive titrations: the principle

In cases when one cannot lower the concentration of the label in such a manner that will be able to have a readout and still be at a concentration low enough compared with k_D,

the competitive titration might prove to be a very good solution. In principle, for a competitive titration, one mixes a labeled ligand, an unlabeled one and their (common) binding partner. One of the ways of performing such an experiment (and which was employed in this study) consists in mixing together the labeled and the unlabeled ligands and titrating them with the binding partner.

Generally speaking, if a competition titration experiment is to be performed, the following equilibria have to be considered (U is the unlabeled ligand, while L is the labeled one):

$$E + U \underset{}{\overset{k_{DU}}{\rightleftharpoons}} EU$$

$$E + L \underset{}{\overset{k_{DL}}{\rightleftharpoons}} EL$$

with the k_Ds described as:

$$k_{DU} = \frac{[E][U]}{[EU]} \qquad \text{(IV.15)}$$

$$k_{DL} = \frac{[E][L]}{[EL]} \qquad \text{(IV.16)}$$

The mass conservation equations

$$[Eo] = [E] + [EU] + [EL] \qquad \text{(IV.17)}$$

$$[Uo] = [EU] + [U] \qquad \text{(IV.18)}$$

$$[Lo] = [EL] + [L] \qquad \text{(IV.19)}$$

can be introduced in the above definitions,

$$k_{DU} = \frac{([E_o]-([EU]+[EL]))([U_o]-[EU])}{[EU]} \qquad \text{(IV.20)}$$

$$k_{DL} = \frac{([E_o]-([EL]+[EU]))([L_o]-[EL])}{[EL]} \qquad \text{(IV.21)}$$

Writing [EL] as a function of k_{DU}, $k_{DL,}$ $[E_o]$, $[U_o]$ and $[L_o]$ one obtains the three solutions of a cubic equation, out of which only one has a physical meaning. The simulations in **Figure IV-5** show what happens in different situations, when the ratio between the affinities of the receptor for the unlabeled ligand and labeled ligand varies by four orders of magnitude. When the k_D of the labeled ligand is lower than the one of the unlabeled, the curve appears hyperbolic. In the other cases (and especially when k_{DU} is lower than k_{DL}) the curve has a

sigmoidal aspect. This is because at the start of the titration the receptor will bind preferentially the ligand with a higher affinity. If the ligand with higher affinity is not labeled, at the beginning of the titration there will be a "lag" phase with no signal. The curves depicted in **A** and **B** possess more information content than the one in **C** and, especially, the one in **D**. Fitting data to a sigmoidal titration curve will result in better defined and more trustworthy k_Ds than the results based on a hyperbolic curve, since there is in principle more information content.

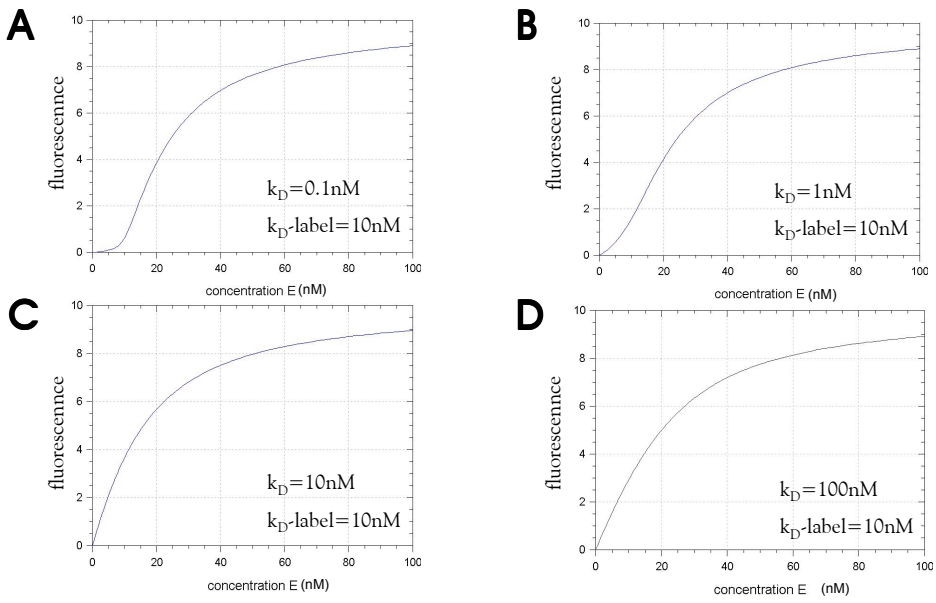


Figure IV-5: Simulated competitive titration curves. In all cases 10nM of receptor in the presence of 100nM fluorescently labeled ligand were titrated against increasing concentrations of unlabeled ligand. The k_D of labeled ligand for binding to the protein was fixed at 10nM. The k_D for the non labeled ligand was 100nM (**A**), 10nM (**B**), 1nM (**C**) and 0.1nM (**D**).

Although there are analytical solutions for equations (**IV.20**) and (**IV.21**), they are difficult to introduce in many data fitting programs, using their equation editors. A more simple approach is to use for this purpose a program in which equations (**IV.15**) to (**IV.19**) can be introduced in their explicit form, such as Scientist. In such a program, the equations are introduced exactly in the form presented above. The fluorescence is defined as being proportional with the concentration of the EL species. An input script for the program Scientist dealing with such a competitive titration is presented below

```
/ Competitive Titration
// L (fluorescent substrate)
// U (non fluorescent)
// Reactions  E+L <>EL : KdL -dans fluorescence
//            E+U <>EU : KdU
// Signal on EL

IndVars: E0
DepVars:  L, U, F, EL, EU, E
Params: KdL, KdU, gain, L0, U0, offset

// Equations
// Equilibrium
EL=(L*E)/KdL
EU=(U*E)/KdU

// Mass conservation
L0=L+EL
U0=U+EU
E0=E+EL+EU

// constraints
0<L<L0
0<U<U0
0<EL<L0
0<EU<U0
0<E<E0

// Signal
F=offset+gain*(EL)

// Intial Conditions
E0=0
***
```

IV.2.3.b. Results of the competitive titrations for Ypt1p and Sec4p

Based on this system, the following experiment was performed: Sec4p was mixed with Ypt7p dansyl labeled, yielding 250nM of each. This mixture was titrated with Mrs6p. The result of the fitting procedure, performed as described above, suggested that the k_D for the Sec4p – Mrs6p interaction is 26nM and the one between Ypt7p dansyl and Mrs6p is 500nM. A similar approach yielded a k_D of 10nM for the Ypt1p – Mrs6p interaction in the presence of Ypt7p dansyl labeled, value in very good agreement with the one determined by titrating Ypt1p dansyl labeled with Mrs6p (Section IV.2.3 at page 85). This result proved also that the dansyl labels attached to the C terminus of different proteins did not affect the affinities between the Ypt protein and Mrs6p, since Ypt1p used in this case was not labeled.

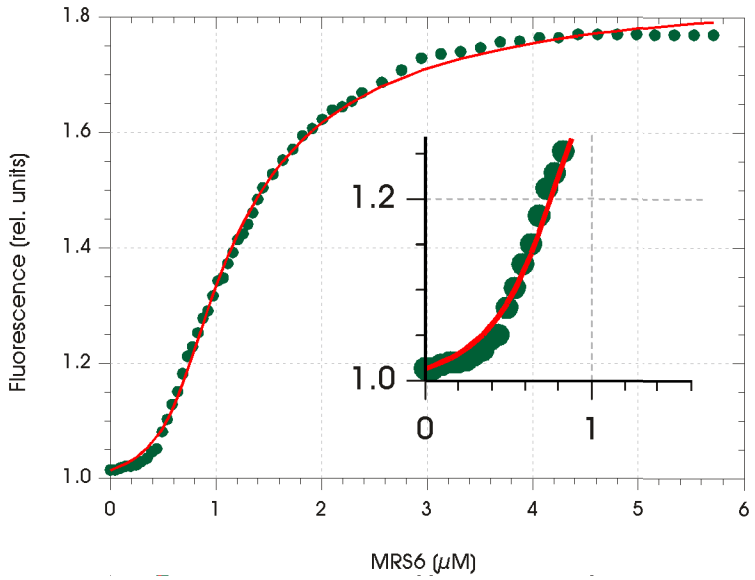


Figure IV-6: Competitive titration between Sec4p, Ypt7p dansyl and Mrs6p. The inset is a blow-up of the "lag" region from which the k_D between Sec4p and Mrs6p was determined. 270nM of Ypt7p•GDP-dansyl and 270nM Sec4p•GDP were used.

The results obtained so far (and the ones obtained from the pre steady state kinetic determinations, see next section) suggested the fact that the members of the Ypt family can be divided in two categories, based on their interaction with Mrs6p: loose binders and tight binders. All the results presented so far were obtained with the proteins in the GDP form. Although it is generally believed that prenylation takes place with the Ypt proteins in the GDP form, it was interesting to see what happens in the case of GTP-loaded Ypt proteins. Since GTP is, even at the low GTPase rates of the Ypt proteins, hydrolyzed in time, a non-hydrolyzable analogue was used, the same one used in the crystallography part of this work: GppNHp.

IV.2.4. Interaction between Ypt1p•GppNHp and Mrs6p

Different titration strategies were tried, with different Ypt proteins in the GTP state, labeled with different fluorophores. Unfortunately, no usable signal could be detected during these equilibrium titration experiments. Two techniques were used to obtain, however, the information. One was competitive titration. The other one consisted in employing pre steady state kinetics in order to derive the binding affinities.

For the case of Ypt1•GppNHp, two competitive titration experiments were performed, both having Ypt7 dansyl labeled as competitor for the Mrs6p binding site. Each of them, separately, could not yield the information desired, because the shape of the titration curves was approaching a hyperbola, looking much like the curve in **Figure IV-5 D**. The power of the method consisted in using the information from *both* titration curves *simultaneously*, in a procedure known as **global fitting**. Owing to the easy way in which explicit equations can be implemented in the program Scientist, the following model was used:

```
// Competitive Titration
// L (fluorescent substrate)
// U (non fluorescent)
// Reaction  E+L <>EL : KdL-dans
//           E+U <>EU : KdU
// Signal on EL

IndVars: E0
DepVars:  EL1, EU1, L1, U1, F1, E1, EL2, EU2, L2, U2, F2, E2
Params: L01, KdL, KdU, gain1, gain2, L02, U01, U02, offset1, offset2

// First competition                    // Second competition
// Equilibrium                          // Equilibrium
EL1=(L1*E1)/KdL                         EL2=(L2*E2)/KdL
EU1=(U1*E1)/KdU                         EU2=(U2*E2)/KdU

// Mass conservation                    // Mass conservation
L01=L1+EL1                              L02=L2+EL2
U01=U1+EU1                              U02=U2+EU2
E0=E1+EL1+EU1                           E0=E2+EL2+EU2

// Signal                               // Signal
F1=offset1+gain1*(EL1)                  F2=offset2+gain1*(EL2)

// Restraints                           // Restraints
0<L1<L01                                0<L2<L02
0<U1<U01                                0<U2<U02
0<EL1<L01                               0<EL2<L02
0<EU1<U01                               0<EU2<U02
0<E1<E0                                 0<E2<E0


// Initial Conditions
E0=0
***
```

As can be seen, the two equilibrium titrations are treated separately, but there is a single set of k_{DA} and k_{DB} which ties the systems together. The k_D that was deduced from this experiment for the interaction between Mrs6p and Ypt1•GppNHp is 1100nM (compared with 10nM for Ypt1•GDP), which is also surprising, compared to the mammalian system, where the affinity difference between the GDP and the GppNHp forms of Rab7 and Rep1 is only one order of magnitude (1nM *versus* 20nM for the GTP form (Alexandrov *et al.*, 1998)).

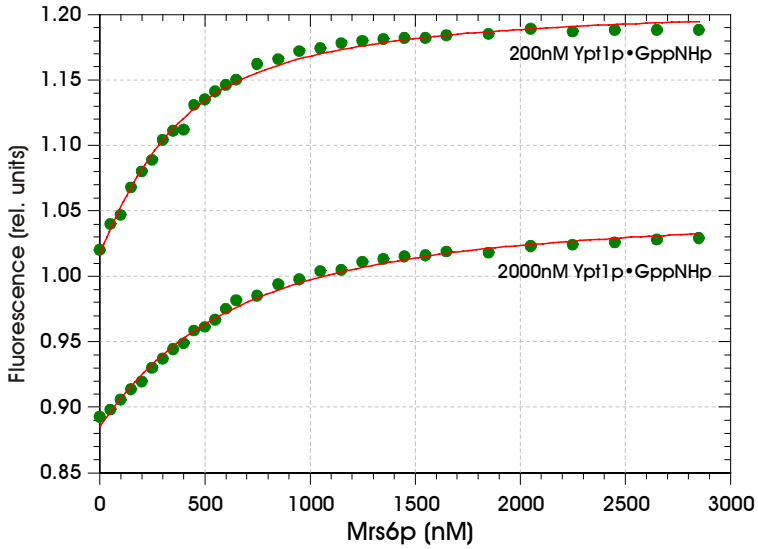


Figure IV-7: Global fitting of the data obtained from two competition experiments. 200nM Ypt7p•GDP-dansyl and the indicated amount of Ypt1p•GppNHp were titrated with Mrs6p, yielding a k_D of 1100nM for the interaction between Ypt1p•GppNHp and Mrs6p.

Having gained some data about the interaction between Mrs6p and different Ypt proteins from the equilibrium titrations, next step was characterisation of the interaction in the pre-equilibrium step, by employing pre steady-state kinetics methods, such as stopped flow.

IV.3. Pre steady-state kinetics of interaction between Mrs6p and different Ypt proteins

IV.3.1. Stopped flow measurements: the principle

Performing steady state kinetic measurements, it is possible to obtain information about the affinities between two (or more) interacting partners and about the k_{cat}, which can correspond to a single rate constant, but might be, as well, the result of the combination of multiple rate constants, too fast to be perceived individually. In order to examine the intermediate steps and the rate constants linking them, as well as to detect the transient intermediates, the kineticist has to perform measurements which give information about the

rate of approach to the steady state. Since, usually, the steady state is attained in time periods of less than a second, the use of specialized apparatus is required.

One way of observing pre steady state events involves usage of a stopped flow machine. Such a machine (which is schematically represented in **Figure IV-8**) is usually able to measure pre steady state events for reactions as fast as $700s^{-1}$. The two driving syringes (A and B) are pushed by the drive and express a defined volume (during this work, the volume was set to $75\mu l$ for each syringe). The reactants flow from each syringe through the mixing chamber to the observation chamber. At the beginning, they start "aging" and still travel to the stop syringe, such that in the optical cell there is always a "new" mixture of a constant age, dictated by the linear velocity of the flow. Once the stopping syringe is pushed far enough, it reaches a position where it stops. In the same moment a microswitch is activated and data recording starts. Thus, the data will be recorded for a mixture which was mixed a certain time period before and is now aging, providing in this way information about the reactions taking place at short time intervals after the mixing moment. In this work, the readout is a fluorescence signal, coming either from the intrinsic tryptophan fluorescence or from the fluorescence of an artificial reporter group.

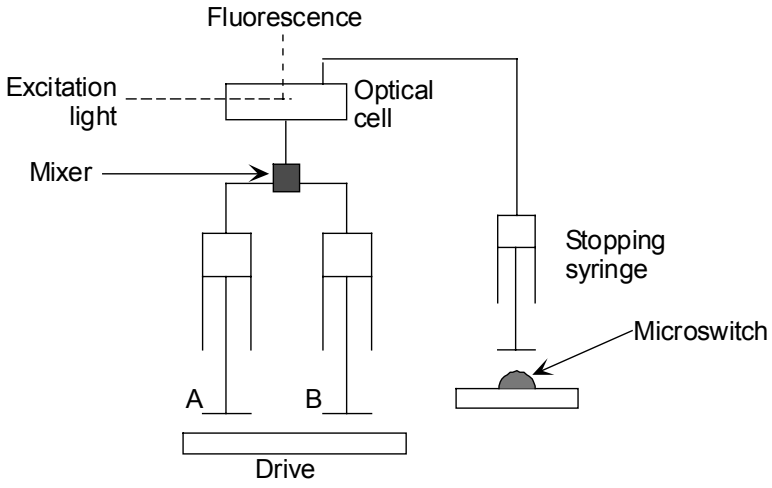


Figure IV-8: Schematic representation of a stopped flow machine.

Two types of experiments were performed and will be described in this work. The first one consisted in studying the association kinetics between Mrs6p and different Ypt/Rab substrates, whereas the other one studied the dissociation of the Rab proteins from Mrs6p.

IV.3.1.a. Association kinetics

In an association kinetics determination, the two interacting partners are in the two different syringes and mixed together in the mixing chamber. At this moment they start to associate and (hopefully) an optical signal can be read out. The experimental conditions are often chosen such that one reagent is in large excess over the other one. Imposing such a condition results in linear rate equations describing the kinetic behavior of the system. If we consider a reaction of the form:

$$L + R \xrightarrow{k} LR$$

then the rate of conversion of production of LR is given by

$$\frac{dc_R}{dt} = -kc_L(t)c_R(t) \qquad \textbf{(IV.15)}$$

(see also **(IV.1)** and the mass action law). However, if L is in large excess over R (usually ten fold or more), $c_L(t)$ will be, to a first approximation, constant and will equal $c_L(0)$. Under these ***pseudo first order*** conditions, the apparent rate constant is $kc_L(0)=k_{obs}$ which has the dimensions of a first order rate constant (s^{-1}). Equation **(IV.15)** can be integrated to yield

$$c_R(t) = c_R(0)\exp[-kc_L(0)\cdot t] \qquad \textbf{(IV.16)}$$

which is an exponential time-dependence. A set of measurements can be carried out at different c_L concentrations. A least square fit to the experimental data will yield, for each case, the value of the observed pseudo first order rate constant, namely $kc_L(0)$. Plotting the obtained values against $c_L(0)$, a straight line should be fitted through these values. The slope of this line will be exactly the k value, the second order rate constant.

For a reversible reaction:

$$L + R \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} LR$$

equation **(IV.3)** can be used to calculate the amount of LR at any time:

$$\frac{dc_{LR}}{dt} = k_{12}c_L(t)c_R(t) - k_{21}c_{LR}(t) \qquad \textbf{(IV.17)}$$

which, keeping in mind that for pseudo first order conditions, $c_R(t)=c_R(0)-c_{LR}(t)$ and $dc_L/dt=0$, can be rewritten as:

$$\frac{dc_{LR}}{dt} = k_{12}c_L(c_R(0) - c_{LR}(t)) - c_{LR}(t)k_{21} = k_{12}c_Lc_R(0) - c_{LR}(t)(k_{12}c_L + k_{21}) \quad \textbf{(IV.18)}.$$

This can be regarded as a first order differential equation with constant coefficients, which was proven to have as solution the equation:

$$c_{LR}(t) = c_{LR}(\infty) + [c_{LR}(0) - c_{LR}(\infty)] \cdot e^{-t/\tau} \quad \textbf{(IV.19)}$$

where $\tau = 1/(k_{12}c_A + k_{21})$ is called the ***time constant*** of the reaction. To find out the value of $c_{LR}(\infty)$, it is enough to note that at equilibrium ($t = \infty$) $dc_{LR}/dt = 0$. Hence:

$k_{12}c_Lc_R(0) - c_{LR}(\infty)(k_{12}c_L + k_{21}) = 0$ and thus

$$c_{LR}(\infty) = c_R(0)\frac{c_L}{c_L + \frac{k_{21}}{k_{12}}} = c_R(0)\frac{c_L}{c_L + k_D} \quad \textbf{(IV.20)}$$

Combining (**IV.19**) and (**IV.20**), and taking in account that $c_{LR}(0) = 0$, will result in:

$$c_{LR}(t) = c_R(0)\frac{c_L(0)}{c_L(0) + \frac{k_{21}}{k_{12}}}\{1 - \exp[-k_{12}c_L(0) + k_{21}] \cdot t\} \quad \textbf{(IV.21)}.$$

Hence, it can be said that recording the time course of a complex formation we will obtain an exponential whose observed rate constant will be

$$k_{obs} = k_{12}c_L + k_{21} \quad \textbf{(IV.22)}.$$

If the mechanism involves a single step, then a plot of k_{obs} against c_L for different $c_L(0)$ will result in a straight line. Its slope will be the second order rate constant k_{12}(or the association rate constant, k_{on}) and the intercept with the y axis (k_{obs}) will represent the value of k_{21}, namely the dissociation rate constant, k_{off}.

Since k_D is defined by the ratio between k_{21} and k_{12} (**IV.11**), the ratio between these two values should be in accordance with the value of k_D obtained from equilibrium titrations, providing a self-consistency check and showing whether the presumed model is or not a correct one.

IV.3.1.b. Dissociation kinetics

A second approach was used in a trial to obtain a better defined k_{off}, since deducing the value of k_{off} from the intercept of the linear fit of variation of the observed rate constants with ligand concentration (equation (**IV.22**)) is not always extremely accurate, due to experimental

errors and small values of k_{off}. In a displacement reaction, two equilibria have to be considered:

$$R + X \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} RX$$

$$R + Y \underset{k_{31}}{\overset{k_{13}}{\rightleftharpoons}} RY$$

The idea of the experiment is to find out k_{21}, the dissociation constant of X from the RX complex, by measuring the decay of RX or the rise of RY. To this end, two rate equations can be written:

$$\frac{dc_{RX}}{dt} = k_{12}c_X(t)c_R(t) - k_{21}c_{XR}(t) \qquad \text{(IV.23)}$$

$$\frac{dc_{RY}}{dt} = k_{13}c_Y(t)c_R(t) - k_{31}c_{YR}(t) \qquad \text{(IV.24)}.$$

In order to solve these equations analytically, several conditions are required:

a) c_X and c_Y should he high enough such that it can be assumed that $c_X(t) = \text{constant}$ and $c_Y(t) = \text{constant}$

b) c_R will be low enough (the free ligand binding sites), and it can be assumed that $c_R(0) = c_{RX}(t) + c_{RY}(t)$, which means that at any time there is no free receptor, but only complexed with any of the two ligands. This leads to

$$\frac{dc_{RX}}{dt} = -\frac{dc_{RY}}{dt}.$$

c) The displacing ligand (Y) is in large excess over the other one (more correctly, $k_{13}c_Y(0) \gg k_{12}c_X(0)$) .

If these conditions are fulfilled, then it can be shown that the first order rate constant for the dissociation of X from the XR complex will be given by a single exponential showing either of the fall of the RX complex with time or the rise of the YR complex with time:

$$c_{RX}(t) = c_{RX}(0)\exp(-k_{21}t) \qquad \text{(IV.25)}$$

$$c_{RY}(t) = c_{RX}(0)\,[1 - \exp(-k_{21}t)] \qquad \text{(IV.26)}$$

Analyzing equations (IV.25) and (IV.26), an important idea arises: one can monitor either the signal of the dissociation between X and R or the signal resulted from the association between Y and R. Both approaches were used in the determinations of the dissociation constants between different Ypt proteins and Mrs6p. When possible, it was tried to monitor the rise in the signal from the YR association. This was done with the belief that the results obtained in this way will be free from any artifacts induced by the existence of a fluorescent reporter group. Two of the three fluorescent reporter groups used in these assays (dansyl and rhodamine) are highly hydrophobic and placed on the C-terminal part of the Ypt proteins, on the same cysteine residues to which the geranylgeranyl is normally attached. The possibility that these hydrophobic compounds would preferentially bind to the presumed hydrophobic pocket of Mrs6p should have been taken into account. By using an unlabeled Ypt protein as the ligand whose k_{off} was assessed, and using a labeled Ypt protein as the displacement agent, one could be sure to avoid any influence from the fluorescence reporter group.

The experiment consisted in mixing together, in one syringe, Mrs6p with the Ypt protein whose k_{off} was to be determined. In the second syringe was the displacing agent (also a Ypt protein, preferably with a k_D lower than that of the unlabeled one) bearing a fluorophore. The concentrations were chosen such that, after mixing, the labeled compound was in large excess over the unlabeled one.

IV.3.2. Interaction of Ypt51p and Ypt7p with Mrs6p

Since no information could be gained from equilibrium titrations between Ypt51p and Mrs6p, I hoped that using pre steady state kinetics measurements would yield the information needed. To this end, a series of stopped flow experiments were performed, using Ypt51 labeled with *mant*GDP. The fluorescent group 3'-*O*-(*N*-methylanthraniloyl) is bound to the 2' and the 3' OH of the guanosine diphosphate and can be excited either directly (λ_{ex}=356nm, λ_{em}=440nm) or by using tryptophan quenching (λ_{ex}=280nm, λ_{em}=440nm).

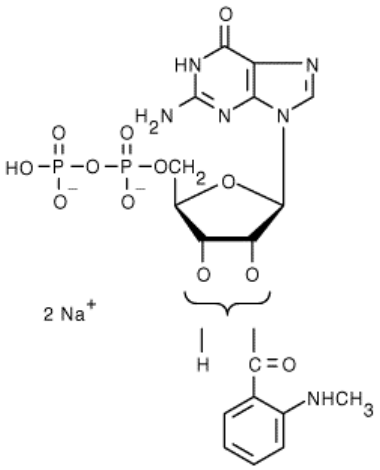


Figure IV-9: 3'-*O*-(*N*-methylanthraniloyl)-2'(3')-guanosine diphosphate (*mant*GDP)

For the association kinetics, constant amounts of Ypt51p•*mant*GDP (300nM) were mixed with increasing quantities of Mrs6p (1000, 1500, 2000, 2500, 3000 and 5000nM) and the fluorescent trace due to complex formation was recorded. Multiple traces were averaged and exponential fits gave the values for the observed rate constant for each concentration. The linear fit (**Figure IV-10 A**) yielded values of $0.0017 s^{-1} nM^{-1}$ for k_{on} and $0.7 s^{-1}$ for k_{off}. Thus, the k_D for the Mrs6p – Ypt51p•GDP was calculated as 386nM. Although, in this particular case, the linear fit was excellent, a displacement reaction was monitored to measure the dissociation rate directly. In this reaction, Ypt51•*mant*GDP was mixed with Mrs6 in one syringe of the stopped flow machine. In the other one, Ypt51•GDP was used, in excess over Mrs6p and Ypt51p. The exponential fit to the observed transient (**Figure IV-10 B**) indicated that $k_{off} = 0.565 s^{-1}$, in very good agreement with the one previously determined. Increasing the concentration of the unlabeled Ypt51p did not lead to an increase in the observed rate, proving that the observed rate was the real dissociation rate. Based on this, a value of 315nM could be calculated for the k_D of the Ypt51p-Mrs6p complex.

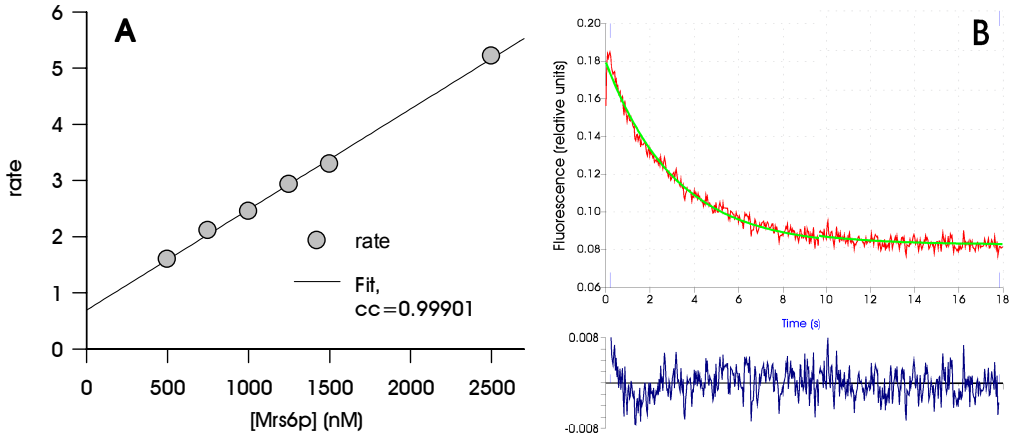


Figure IV-10: Pre steady state kinetics results from the interaction of Ypt51p with Mrs6p. In the results from the association kinetics are presented, whereas in B the stopped flow trace from a displacement reaction and the exponential fit are depicted. The lower panel in **B** shows the residuals of the fit. In **A,** 300nM Ypt51p•GDP-mant were mixed with increasing concentrations of Mrs6p. In **A,** 300nM Ypt51p•GDP-mant and 300nM Mrs6p were mixed in one syringe, while the other contained 1800nM Ypt51•GDP

IV.3.3. Interaction between Ypt1p and Rab7 and Mrs6p

Based on similar approaches, the interaction between Ypt1p and Mrs6p and Rab7 and Mrs6p were characterized.

For the interaction between Rab7 and Mrs6p, two labels were used: dansyl labels attached to the C-terminal cysteines and the *mant* label attached to GDP. The k_{on} values obtained with these two approaches were in very good agreement (0.003015nM^{-1}s^{-1} and, respectively, 4.196μM^{-1}s^{-1}). The experimentally determined k_{off}s were also in good agreement (0.0993s^{-1} and 0.1127s^{-1}). The calculated k_D value for this interaction was in the range of 35±2nM, close to the value of 15nM determined by equilibrium titrations (Section IV.2.2, page 81).

The pre steady state interaction between Ypt1p and Mrs6p was characterized also by using a dansyl label attached to the C terminus of Ypt1p. The on and off rates for this interaction were found to be, respectively, 0.00365nM^{-1}s^{-1} and 0.331s^{-1}, resulting in a k_D value of 90nM, somewhat higher than the one obtained by equilibrium titrations either directly or by using a competitive titration approach (Section IV.2.3).

IV.3.4. Mrs6p *versus* Ypt7p•GppNHp

It is known that in *Saccharomyces cerevisiae* the ratio between GTP and GDP varies from 3.9 to 1.7 (mid-exponential growth phase and, respectively, late exponential growth) (Rudoni *et al.*, 2001), but GTP is always in excess over GDP. In this context, it is possible that, after synthesis, Ypt proteins will be loaded with GTP and not GDP. Thinking that it would be interesting to find out whether Mrs6p makes any differences based on the status of nucleotide loading, I tried to characterize the interaction between Mrs6p and Ypt proteins loaded with the non-hydrolyzable GTP analogue 5'-[β,γ-imido]-guanosintriphosphate (GppNHp). A global fit analysis for the case of Ypt1p•GppNHp revealed that its affinity towards Mrs6p was very low (k_D=1100nM).

I wanted also to characterise the interaction between a Ypt family member which, in the GDP form, binds less tightly to Mrs6p. To this end, Ypt7p was again chosen. Using dansyl as fluorescent reporter group, the rates of association and dissociation between Ypt7p•GppNHp and Mrs6p were found to be $0.273\mu M^{-1}s^{-1}$ and $0.085s^{-1}$, leading to a calculated k_D value of 310nM.

IV.4. Proton/Deuteron exchange experiments and mapping the Mrs6p–Sec4 interaction interface

The results obtained thus far are somewhat puzzling: it appears that, with respect to their interaction with Mrs6p, there are two classes of Ypt proteins: some are able to bind Mrs6p with affinities in the range of 10 to 20nM, whereas the others bind by an order of magnitude less tightly. These results are especially difficult to explain, since the sequence homology and identity between the members of the family is high and analysis of the crystal structures of three members (Ypt51p (Esters *et al.*, 2000), Sec4p (Stroupe & Brunger, 2000) and Ypt7p (this work)), two of which had the structure solved both for the GDP and the GTP conformations, did not reveal any obvious reason for this behavior.

Unfortunately, the efforts to crystallize Ypt7p in complex with Mrs6p were unsuccessful, the crystallization trial drops exhibiting only precipitates or phase separations. Attempts to crystallize the ternary prenylation complex between the yeast geranylgeranyl transferase II, Mrs6p and various Ypt proteins were also unsuccessful (Alexey Rak, personal

communication). Due to the size of Mrs6p (66kDa) the problem would also not be amenable to magnetic nuclear resonance spectroscopy (NMR) methods.

Under these circumstances, a different approach was tried. It was based on a method developed at the University of California, San Diego (Mandell *et al.*, 1998b; Mandell *et al.*, 1998a). The method showed that protein-protein interfaces could be identified by decreased amide proton solvent accessibility and the use of a matrix-assisted laser desorption/ionization time-of-flight(MALDI-TOF) mass spectrometer. The principle is that amides exhibiting slow proton/deuteron exchange rates when the proteins are in complex and fast exchange rates in the uncomplexed state are positioned at the protein-ligand interface. The rate of amide proton/deuteron (H/D) exchange is temperature- and pH- dependent and a change of pH from 7 to 2.5 lowers the exchange rate by three to four orders of magnitude (depending on the peptide studied); this leads to a half-life of 30..120 minutes at 0°C.

In order to map the protein-protein interface of the Mrs6p/Ypt complex, two types of determinations were performed: in one series of experiments, each member of the complex (lyophilized protein) was incubated in buffered D_2O at pD 7.25. Subsequently, the protein-protein complex was diluted ten fold in H_2O and allowed to back-exchange for various time periods and subsequently quenched by lowering the pH to 2.5 and the temperature to 0°C. The samples were subjected to digestion with pepsin and the resulting peptide mixture was analyzed by MALDI-TOF. The second series of experiments started also by incubating each member of the complex in D_2O. After a defined time, the proteins were mixed, enabling them to build the complex, and the complex was diluted in H_2O and subjected to the same treatment as the individual components. Mass spectra were acquired and the mass shift for each peak was assessed (see below).

IV.4.1. The method

IV.4.1.a. Sample preparation

Information about the affinity between members of the Ypt family and Mrs6p proved to be very important, since it enabled me to choose the right interacting partner (Sec4p in this case, since its three dimensional structure is also known) and calculate (using the equation (**IV.14**) and the k_Ds previously determined) protein concentrations at which more than 95% of the protein is involved in the complex. Also, since the proteins had to be, at the beginning,

lyophilized, it was useful to know whether this process had a negative influence on one (or both) members of the complex. In order to examine this, competitive titrations were performed with the protein before and after lyophilization. The affinity constants determined in both cases were equal, showing that lyophilization did not have any adverse effect.

The buffer in which Mrs6p and Sec4p were dissolved was exchanged for one consisting of 30mM potassium phosphate, pH 7.5; after that, the proteins were lyophilized in aliquots of 750pmoles (Sec4p) and 1875pmoles (Mrs6p). For the experiment, each aliquot was redissolved in 10μl of D_2O and allowed to exchange the amide proton for deuterons for 10 minutes. After this time period, the sample was diluted 10 fold with H_2O and allowed to exchange back the deuterons for protons for 0, 1, 2, 3, 5 and 10 minutes. The H/D exchange was quenched by addition of 7.5μl TFA 0.05% (pH 2.5), transferred to ice and incubated for 10 minutes with an equimolar amount of pepsin. At the end of the digestion period, samples were snap-frozen in liquid nitrogen and kept at –80°C until further analysis. For the complex, each protein was dissolved separately in 10μl D_2O, allowed to exchange the protons for deuterons for eight minutes only and then the two aliquots mixed together. After an extra two minutes, the complex was diluted with H_2O and all the subsequent steps were carried out as indicated above.

IV.4.1.b. Data collection and analysis

In order to minimize the loss of deuterons during data recording, each sample was analyzed separately. Each sample was thawed as fast as possible, mixed with cold MALDI matrix (5mg/ml α-cyano-4-hydroxycinnamic acid in 1:1:1 acetonitrile:ethanol: 0.05%TFA) and spotted onto a chilled MALDI target. To accelerate the drying process, the MALDI target was placed in an exicator and a moderate vacuum was applied. The time span for acquiring each sample was very carefully kept constant.

Spectra were acquired with a Voyager DE-PRO workstation using delayed extraction with a 100ns pulse delay, the guide wire was set at 0.02% and the grid voltage at 70%.

The acquired spectra were calibrated once with an external calibration and then a second, internal, calibration was performed. A typical spectra is shown in **Figure IV-11**. The mass centroid of each isotopic peak cluster was calculated using a Microsoft Excel worksheet.

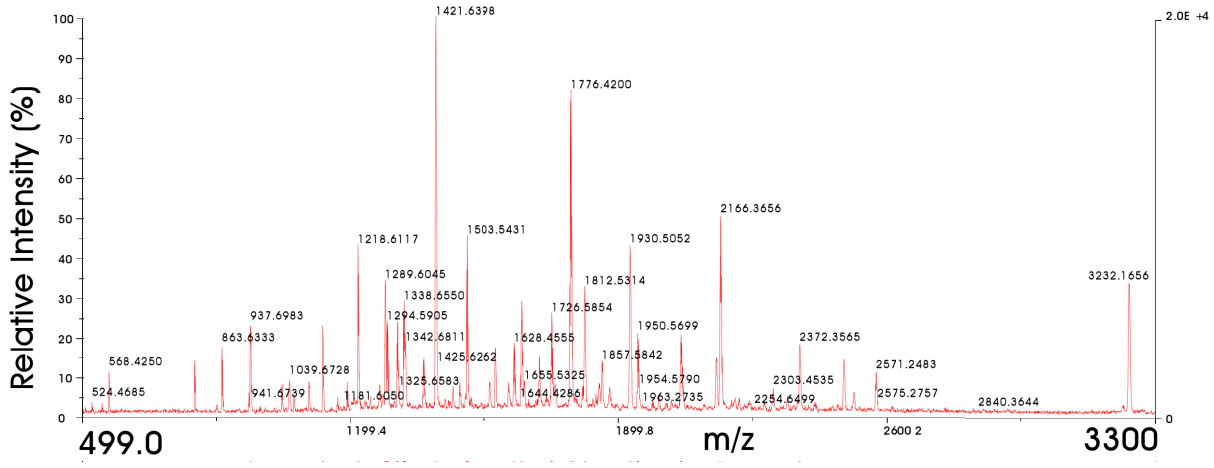


Figure IV-11: A typical mass spectra obtained after digesting a mixture of Mrs6p and Sec4p with pepsin

For each isotopic peak cluster investigated, the variation of the centroid position with the back-exchange time was plotted and fitted to the exponential

$$D = D_{max} + (D_{max} - D_0) \cdot \exp(-k \cdot t).$$

where D is the number of deuterons retained in a particular sample, D_{max} is the number of deuterons retained in the sample corresponding to the timepoint zero and *t* is the time elapsed before quenching the analyzed sample. The number of deuterons was calculated as the difference between the centroid of the isotopic peak cluster for the deuterated sample and the centroid of the undeuterated control. Two spectra displaying such a shift can be seen in **Figure IV-12** and two exponential fits in the **Figure IV-13**.

Based on the exchange rates obtained from the exponential fits, a *protection factor* was defined, as the ratio between the decay rate in the uncomplexed case and the rate in the complexed case. Thus, the higher the protection factor, more efficient was the observed "protection" of the corresponding peptide.

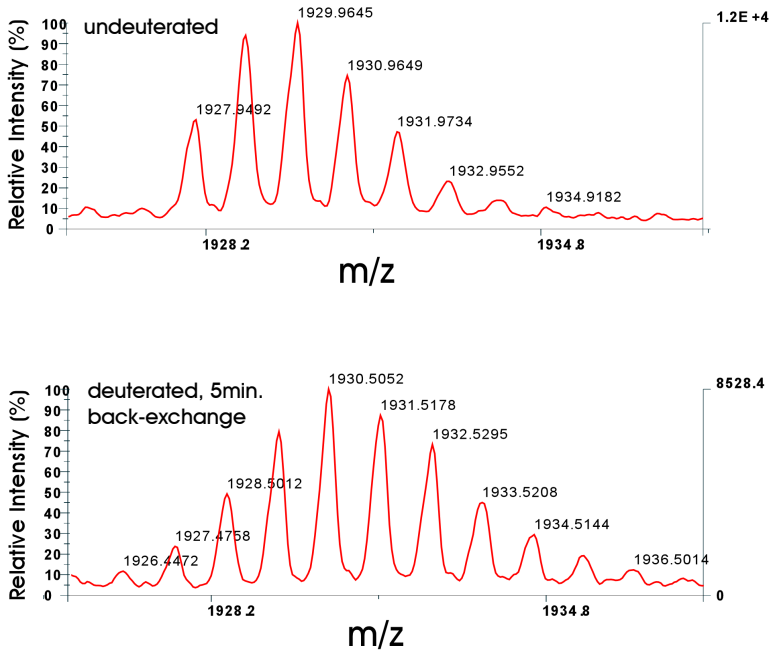


Figure IV-12: Shift of the isotopic peak cluster for the peptide of monoisotopic mass 1927.91 occurring after 5 minutes of back-exchange, compared with the isotopic peak cluster obtained for the undeuterated protein.

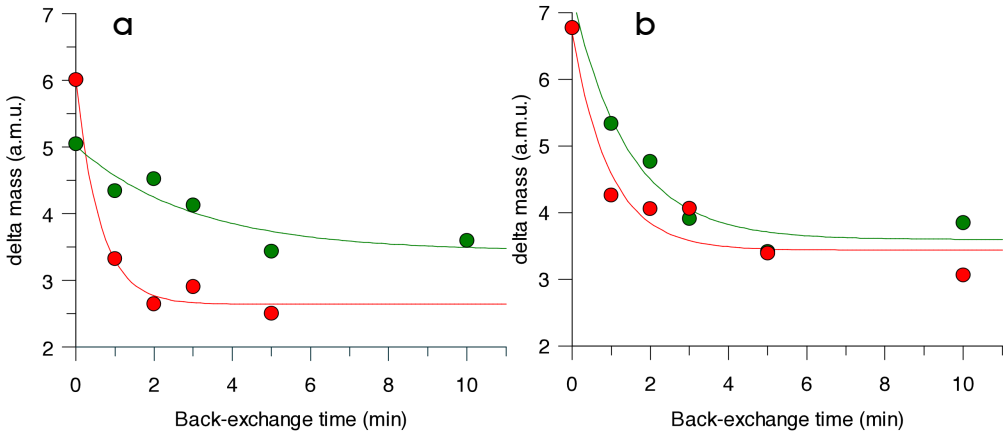


Figure IV-13: Exponential fits for data acquired for two different peptides. Red is for the protein digested uncomplexed and green is for the protein digested in complex. The spectra in **a** corresponds to the peptide of monoisotopic mass 1927.9159 and the spectra in **b** corresponds to the peptide of mass 2190.2511, both originating from Mrs6p.

IV.4.1.c. Assigning the peaks

An important problem was assigning the peaks. The method used for assigning the peaks was the following: the mass spectrum of each protein digested separately was subjected to the deisotoping process. During this process, the Voyager software calculates the monoisotopic mass of each isotopic peak cluster found and outputs a list containing only the deisotoped masses. The Protein Prospector web server at http://prospector.ucsf.edu/ performs an *in silico* enzymatic digestion of a protein sequence and calculates the mass of each peptide, reporting also the monoisotopic masses. The two lists were compared and, for each match, the corresponding peptide was mapped on the sequence of the protein.

The accuracy of the MALDI machine used is about 0.2 atomic mass units for a sample with the mass of 2000 atomic mass units. In other words, a sample with the mass 2000.0000 could be reported as having the mass anywhere between 1999.9000 and 2000.1000 atomic mass units. Thus, if a peak matched more than one putative peptide in the list output from Protein Prospector, there was, in a first run, no way of deciding which peptide it really is, leading to ambiguities in the assignments. However, a significant region of each of the two interaction partners could be covered.

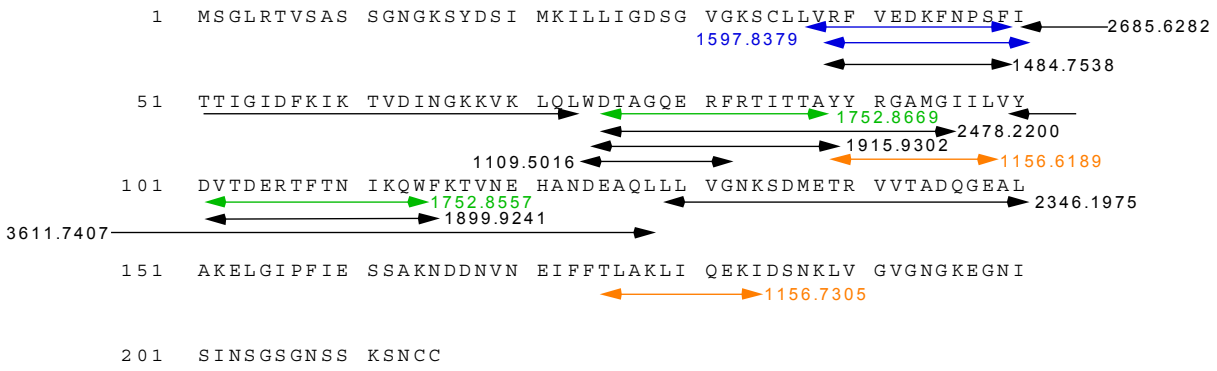


Figure IV-14: Amino acid sequence of Sec4p indicating the position of each peptic fragment that was identified in the mass spectrum of undeurated Sec4p digest. The colored arrows indicate peptides which could not be unambiguously identified.

IV.4.2. Results

Each protein was digested undeuterated, in order to assign the peaks and then the time course of deuterium loss for each peptide was recorded. The same determinations were done for the digest originating from the complex between Sec4p and Mrs6p. An unfortunate situation appeared: many of the peaks found in the digest of separate proteins could not be found in the spectra acquired using the digest of the Sec4p–Mrs6p complex. It is known that some peptide peaks are suppressed in digest mixtures (Mandell *et al.*, 1998b), but the basis of this effect is not clear, especially because the physics of the matrix desorption is not totally understood. In a trial to circumvent this problem, different dilutions of the sample were tested. I found that diluting the sample as much as 60 fold with matrix could yield spectra displaying some of the peaks which were previously not observed.

In an attempt to resolve some of the ambiguously assigned peaks, two methods of sequencing were tried. In the first one, named MS/MS, the digested sample was analyzed in a nanospray ionization mass spectrometer. During an MS/MS experiment, ions of a certain mass (more correctly m/z ratio) are trapped in an ion trap filled with helium. A radiofrequency signal is used to drive the ions in a circular motion. Under certain conditions, the energy can be set in such a manner that, after colliding with the He ions in the ion trap, the peptides will break exactly at a peptide bond, yielding two subfragments. Because these subfragments have a different m/z ratio, they will escape from the ion trap and will be detected. Since, in general, *any* peptide bond has almost the same chance of be broken, the method allows one, in principle, to determine the sequence of the peptide.

The second method used was post source decay (PSD). The method allows obtaining structural information about a sample by analyzing the fragment ions generated from the original ions in the flight tube.

Regrettably, none of the methods could yield sequence information about the peaks which could not be assigned previously due to ambiguities. However, the number of peptides that could be assigned and characterized for both the digest of the single protein and the digest of the complex allowed some information to be extracted.

In order to obtain structural information, a model of the two interacting partners was necessary. Coordinates of Sec4p were already available (Stroupe & Brunger, 2000), but for Mrs6p there was no crystal structure, nor a model. The closest relative of Mrs6p for which structural information was available was bovine Gdiα (Schalk *et al.*, 1996). Unfortunately, the

degree of identity between these two proteins is less than 20%, which is a serious problem for today's homology modeling programs. It is today generally accepted that models based on less than 30% sequence identity have significant alignment errors, resulting in large errors in main chain positions (Vitkup *et al.*, 2001). Modeling was attempted with two homology modeling programs: Modeller (Sali & Blundell, 1993) and Swiss-Model (Peitsch, 1996). Both programs use as input coordinates of the template (bovine Gdiα in this case) and a sequence alignment. The sequence alignment was generated with T-Coffee (Notredame *et al.*, 2000), an algorithm which uses a combination of local and global pair-wise alignments to generate a library of alignment information which will be then used for generating the alignment. The algorithm is especially recommended for difficult cases, i.e. cases in which the sequence homology/identity is very low. This obtained alignment was further improved manually, by attempting to align the insertions/deletions with flexible structural elements (loops) in Gdiα. However, neither of the modeling programs used succeeded in modeling the N terminal extension of Mrs6p and the long insertions in the Mrs6p sequence in a satisfactory manner. These regions will be referred to as "untrusted regions".

The peptides arising from Mrs6p which could be assigned with certainty were mapped on the surface of the models obtained. The following peptides could be mapped on the surface of Mrs6p:

Region	Sequence	Protection factor
149–159	IKSRVHQYLEF	5.33
184–209	FTDQNLPLMTKRNLMKFIKFVLNWEA	5.55
251–268	NVKVPEALQRIRRYLTSF	1.5
352–367	QVHRLTCIVENPCTEW	4.7

For the case of Sec4p, only one peptide could be unambiguously assigned; that was the peptide of mass 1597.8379, which maps to the switch I region of Sec4p. The "protection factor" for this peptide was 3.35.

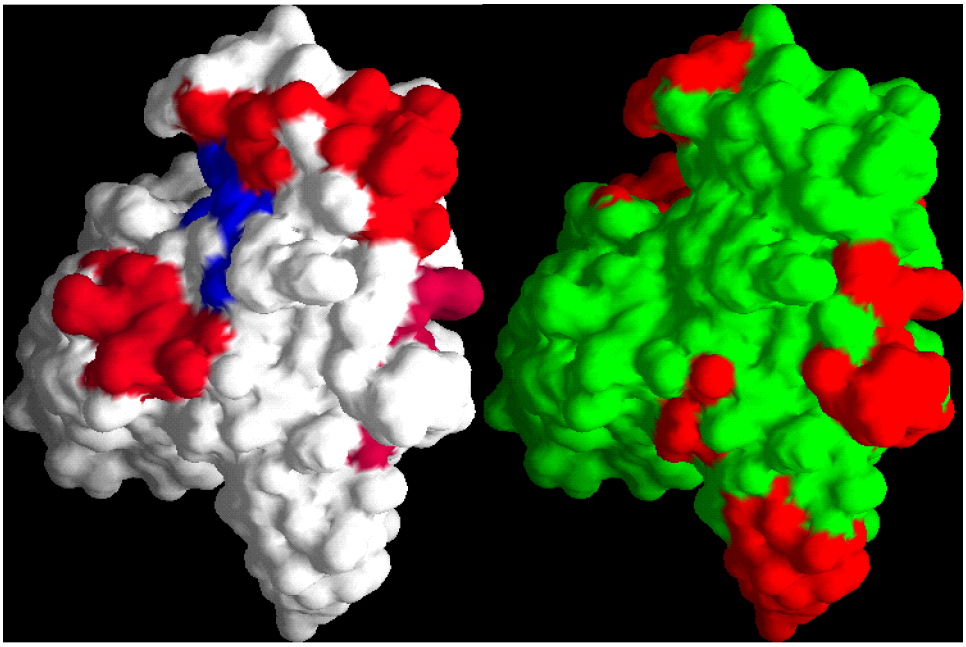


Figure IV-15: Model of Mrs6p. In the left panel, the known peptides are mapped on the surface. The color scale indicates the "protection factor", with red being high (good protection) and blue being low (no protection). In the right panel the untrusted areas are marked in red. Note that the mapped peptides do not fall in untrusted areas of the model.

IV.5. Conclusions

The kinetic experiments performed with different members of the Ypt family showed that they can be divided in two categories, with respect to their interaction with Mrs6p. Members of the first category are able to bind Mrs6p tightly, with affinities in the range of 10 to 30nM. Although not as strong as that in the mammalian system, this interaction can still be considered tight, especially when compared with the affinity of the members of the second category, which is lower by an order of magnitude. The pre steady state experiments showed that this effect arises from differences in both the association and dissociation rates, but the variation in the association rates is more significant than the one in dissociation rates. The data obtained from the pre steady state and steady state kinetic determinations is summarized in **Table IV-1** and **Table IV-2**.

Protein	k_{on}	k_{off}	k_{off} intercept	k_D calc	k_D exp
Ypt7p•GDP	0.000314	0.164	0.154	521.0	354.0
Rab7•GDP	0.004196	0.099	0.309	23.7	13.0
Ypt51p•GDP	0.001790	0.565	0.691	315.7	N.D.
Sec4p•GDP	N.D.	N.D.	N.D.	N.D.	4.2
Ypt1p•GDP	0.00365	0.331	0.488	90.7	11.2

Table IV-1: Overview of the data obtained from kinetic determinations for the interaction between different Ypt/Rab•GDP proteins and Mrs6p. N.D. – the values could not be determined; k_{off} intercept – dissociation rate obtained from the intercept between the liniar fit of the observed rate-*versus*-concentration plot and the Y (observed rate) axis of the graph; this value is usually susceptible to large errors; k_{off} – experimentally determined dissociation rate; k_D calc – affinity between the interaction partners calculated as the ratio between k_{on} and k_{off}; k_D exp – affinity experimentally determined from equilibrium titration experiments. Units: $[k_{on}] = s^{-1}nM^{-1}$, $[k_{off}] = s^{-1}$, $[k_D] = nM$

Protein	k_{on}	k_{off}	k_{off} intercept	k_D calc	k_D exp
Ypt1p •GppNHp	N.D.	N.D.	N.D.	N.D.	1100.0
Ypt7p •GppNHp	0.000273	0.085	0.010	311.5	N.D.

Table IV-2: Overview of the data obtained from kinetic determinations for the interaction between different Ypt/Rab•GppNHp proteins and Mrs6p. Notations as in **Table IV-1**

It is interesting to note that the only tight binders are those encoded by vital genes. Different studies showed that Sec4p and Ypt1p are vital (Schmitt *et al.*, 1986; Walworth *et al.*, 1989), whereas Ypt7p, Ypt51p and Ypt31p are not (Gallwitz *et al.*, 1983; Wichmann *et al.*, 1992; Singer-Kruger *et al.*, 1994; Benli *et al.*, 1996). It is known that Ypt proteins are normally prenylated in the cell (Alory & Balch, 2000; Miaczynska *et al.*, 2001) in ideal growth conditions. However, under less favorable conditions, it is possible that the cell has to decide which are the most important proteins to be prenylated. Recently (Alory & Balch, 2000), it was shown that $Mrs6p^{ts}$ mutants were able to sustain growth if the cells were shifted at hourly

intervals for less than two minutes from the non-permisive temperature (37°C) to the permisive (30°C) temperature. It was also shown that, in cells overexpressing the MRS6 gene, Ypt proteins accumulate and remain partially unprenylated (Alory & Balch, 2000).

Under stress conditions, like the ones described above, it is conceivable that the prenylation machinery has to make critical decisions in terms of which Ypt family members will be prenylated first. It is possible that, in this context, the ability to discern between the vital members of the family and the less important ones starts playing an important role. The results of the kinetic determinations would provide a good explanation for such a behavior.

The above-mentioned discrimination does not hold for the GTP state of the proteins. On the contrary, in the GTP state the vital Ypt proteins are bound at least three times less tightly than the non-essential ones. This might be also part of a regulatory system, by which the cell avoids the delivery of GTP-loaded, active, Ypt proteins with essential roles in the membrane. In this way, the most important Ypt proteins have less chances of evading the tight regulatory system functioning in the cell.

In an attempt to understand the molecular and structural basis of this behavior, the interaction interface between Mrs6p and Sec4p was partially mapped. Although the coverage is not extensive, the results are in accordance with previous results and also bring new information.

Out of the four peptides that could be identified and mapped with certainty on the protein surface of Mrs6p, three were shown to be implicated in the Mrs6p–Sec4p interface. Peptide 149-159, one of the two most-protected peptides, maps on Gdiα to residues 96-106. These residues are part of a loop in contact with helix α9; it was shown (Schalk *et al.*, 1996) that residues 248,249,250 of Gdiα, part of this helix, when mutated, induce a lowering of the affinity between Gdiα and Rab3A by about 25 times.

Residues 352-367 are less protected in Mrs6p and map to region 296-310 in Gdiα. The study mentioned above showed that mutation of residue 309 had a minor effect in altering the affinity between Rab3A and Gdiα, suggesting that this area participates in the interaction, but is less important.

Mutation of R195 and K199 in Mrs6p to alanine reduces the growth rate of a *MRS6Δ* strain (Alory & Balch, 2000). Consistent with this, residues 184-209 were shown to have the

highest protection factor, suggesting that the mutations described affect the interaction interface with the Ypt proteins.

Unfortunately only one peptide could be mapped on the structure of Sec4p. However, this peptide (region 38-50, the blue arrow in **Figure IV-14**) is part of the *switch* I region and exhibits a relatively high protection factor (3.35), an indication that the switch regions are important for the interaction with Mrs6p and offering an explanation for the discrimination between the GDP and the GTP states.

Taken together, all these data provide extra information on the system controlling the prenylation of the members of the Ypt family and, thus, controlling the whole vesicular traffic machinery of the cell.

Chapter V.

Summary

Ypt/Rab proteins are small Ras-related GTPases involved in intracellular traffic. The current data suggest a model in which Ypt proteins (in yeast) and Rab proteins (in higher eukaryotes) play a major role in the catalysis/specificity of the intracellular traffic, from vesicle budding to vesicle transport, docking and fusion (Segev, 2001). However, the exact subreactions in which Ypt/Rab proteins are required and the exact role they play remain unknown. Also, we do not know much about *how* the specificity is achieved and which are the molecular mechanisms that control the very complex endomembraneous system within the cell. Neither do we know what makes these Ypt proteins, structurally so much alike, the key players in a specificity process which requires the differentiation between different compartments. Another question which was unanswered at the beginning of this work was how big are the structural differences between the GTP- and the GDP- bound forms of the members of the family and if defining the "switch" regions based solely on the homology with Ras was a correct decision or not.

This work was done with the points outlined above in mind and with the hope of gaining some insight into the molecular basis of this specificity. To answer these questions, crystallographic analysis was combined with extensive kinetic characterization of different Ypt family members.

For the crystallographic analysis, Ypt7p was C-terminally truncated by 10 residues. This Ypt7p truncation mutant yielded well-diffracting crystals in the GTP form, but later it was shown that the *switch* regions, responsible for the conformational differences between the GDP and GTP conformations, were not visible in the electron density due to unfortunate crystal packing. In addition, the end of the C-terminal helix was not visible in the electron density and crystal-packing analysis suggested that it interacts with the *switch* regions of two other molecules. A second construct, lacking the last 26 aminoacid residues of Ypt7p was designed with the hope that these problems would be circumvented. The strategy proved successful: the protein could be crystallized and the structures of both the GTP and the GDP conformations solved to very high resolution (1.6Å in the case of the GTP-bound form and 1.3Å in the case of the GDP structure). Although it was postulated that the differences between the GDP and GTP state are similar to the ones observed in Ras, only recently the structure of Sec4p (Stroupe & Brunger, 2000) in the GTP and GDP states showed that qualitatively the same differences can be seen in the Ypt family as well. Different Ypt proteins have to be, nevertheless, distinguished from each other in either the GDP- or GTP-bound

states. Small variations among several Ypt/Rab-proteins in the different nucleotide-bound states are crucial for their specific function. Therefore, the determination of their individual structures with high resolution is needed for a comprehensive understanding of effector binding specificity.

The presented differences in main-chain and side-chain conformations of Ypt7p compared to Ypt51p, Sec4p, Rab3a and Rab5c, in addition to the sequence differences, contribute to structural plasticity and variations in both the surface charge distribution and surface topology within the Ypt/Rab-family of proteins. The observed structural differences are clustered in areas which were recently highlighted as Rab-specific (RabF) or Rab-subfamily specific (RabSF) regions (Pereira-Leal & Seabra, 2000). For the RabSF2 region (part of the switch I region) Ypt7p displays, in the GTP-bound form, a shifted conformation when compared with other Ypt/Rab structures and a different surface topology. In the structure of Ypt7p•GppNHp, helix α2 is unwound, which results in an extended loop spanning from β-strand β3 to β4, including the RabF3 region. Molecular dynamics simulations were performed and revealed that the intrinsic sequence of Ypt7p determines this feature, proving that it is not induced by crystal-packing interactions. These two conformational changes associated with the switch I and the switch II region are likely to contribute to Ypt7p specific binding of effector proteins and presumably help to discriminate between Ypt7p and Ypt51p when they are located on the same membrane. Furthermore, due to insertions in loops L3 and L7, the neighboring RabSF1 and RabSF4 regions are different in their conformation compared with other Ypt/Rab proteins. Therefore, it is probable that this surface patch can serve as an additional binding site for Ypt7p-specific effector proteins. These variations, each taken in part, are small, but they add up and most probably become a specificity tag allowing differentiation between the otherwise sequence- and fold- similar Ypt/Rab proteins.

It is a common feature that the GDP-bound form of Ras-related small GTP-binding proteins has a less well-defined structure than the GTP-bound form. In particular, the important switch I and switch II regions are very flexible when the protein is in complex with GDP, lacking the γ-phosphate group which fixes these regions. This higher conformational mobility in the GDP-bound form might help proteins which interact promiscuously with Ypt/Rab proteins in the GDP-bound form to overcome variations in the amino acid sequence and topology at the interaction surface (e.g. GDP dissociation inhibitor, GDI, (Garrett *et al.*, 1994)).

The purpose of the kinetic analysis was to gain a better understanding of the way prenylation is controlled. To this end, equilibrium titrations as well as pre steady state kinetics methods were employed. By making use of a large array of techniques, the characterization of the interaction between the escort protein Mrs6p and a variety of members of the *Saccharomyces cerevisiae* Ypt family was achieved. The results obtained helped us to suggest a mechanism by which cells differentiate between the more and the less important members of the family.

The kinetic experiments performed with different members of the Ypt family showed that they can be divided in two categories, with respect to their interaction with Mrs6p. Members of the first category are able to bind Mrs6p tightly, with affinities in the range of 10 to 30nM. Although not as strong as that in the mammalian system, this interaction can still be considered tight, especially when compared with the affinity of the members of the second category, which is lower by an order of magnitude. It is interesting to note that the only tight binders are those encoded by vital genes.

It is known that Ypt proteins are normally prenylated in the cell (Alory & Balch, 2000; Miaczynska *et al.*, 2001) in ideal growth conditions. However, under less favorable conditions, it is possible that the cell has to decide which are the most important proteins to be prenylated. Recently (Alory & Balch, 2000), it was shown that Mrs6p^{ts} mutants were able to sustain growth if the cells were shifted at hourly intervals for less than two minutes from the non-permisive temperature (37°C) to the permisive (30°C) temperature. It was also shown that, in cells overexpressing the MRS6 gene, Ypt proteins accumulate and remain partially unprenylated (Alory & Balch, 2000).

Under stress conditions, like the ones described above, it is conceivable that the prenylation machinery has to make critical decisions in terms of which Ypt family members will be prenylated first. It is possible that, in this context, the ability to discern between the vital members of the family and the less important ones starts playing an important role. The results of the kinetic determinations would provide a good explanation for such a behavior.

The above-mentioned discrimination does not hold for the GTP state of the proteins. On the contrary, in the GTP state the vital Ypt proteins are bound at least three times less tightly than the non-essential ones. This might be also part of a regulatory system, by which the cell avoids the delivery of GTP-loaded, active, Ypt proteins with essential roles in the

membrane. In this way, the most important Ypt proteins have less chances of evading the tight regulatory system functioning in the cell.

In an attempt to understand the molecular and structural basis of this behavior, the interaction interface between Mrs6p and Sec4p was partially mapped using variations in amide proton solvent accessibility as monitored by proton-deuteron exchange and MALDI-TOF analysis. Although the coverage is not extensive, the results are in accordance with previous results and also bring new information.

Only one peptide could be mapped on the structure of Sec4p. However, this peptide is part of the *switch* I region and is relatively highly protected, an indication that the switch regions are important for the interaction with Mrs6p. This results offers an explanation for the discrimination between the GDP and the GTP states.

Out of the four peptides that could be identified and mapped with certainty on the protein surface of Mrs6p, three were shown to be implicated in the Mrs6p–Sec4p interface. Peptide spanning residues 149-159, one of the two most-protected peptides, maps on Gdiα to residues 96-106. These residues are part of a loop in contact with helix α9; it was shown (Schalk *et al.*, 1996) that residues 248,249,250 of Gdiα, part of this helix, when mutated, induce a lowering of the affinity between Gdiα and Rab3A by about 25 times.

Residues 352-367 are less protected in Mrs6p and map to region 296-310 in Gdiα. The study mentioned above showed that mutation of residue 309 had a minor effect in altering the affinity between Rab3A and Gdiα, suggesting that this area participates in the interaction, but is less important.

Mutation of R195 and K199 in Mrs6p to alanine reduces the growth rate of a *MRS6Δ* strain (Alory & Balch, 2000). Consistent with this, residues 184-209 were shown to have the highest protection factor, suggesting that the mutations described affect the interaction interface with the Ypt proteins.

Taken together, the data presented in this work provide extra structural and kinetic information on the system controlling the prenylation of the members of the Ypt family and, thus, controlling the whole vesicular traffic machinery of the cell.

Publications resulted from this work

H. Esters, K. Alexandrov, **A. T. Constantinescu**, R. S. Goody and A. J. Scheidig. "High-resolution Crystal Structure of S. cerevisiae Ypt51ΔC15-GppNHp, a Small GTP-binding Protein Involved in Regulation of Endocytosis". *J.Mol.Biol.* 298 (1):111-121, 2000.

A. Kalinin, N. H. Thoma, A. Iakovenko, I. Heinemann, E. Rostkova, **A. T. Constantinescu** and K. Alexandrov. "Expression of Mammalian Geranylgeranyltransferase Type-II in Escherichia coli and Its Application for *in vitro* Prenylation of Rab Proteins". *Protein Expr.Purif.* 22 (1):84-91, 2001.

A. T. Constantinescu, A. Rak, K. Alexandrov, R. S. Goody and A. J. Scheidig."Rab-subfamily specific regions of Ypt7p are structurally different from other RabGTPases". Submitted to *Structure with Folding and Design*.

Chapter VI.

References

Adams PD, Pannu NS, Read RJ, Brunger AT: **Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement**. *Proc.Natl.Acad.Sci.U.S.A* 1997, **94:**5018-5023.

Albert S, Will E, Gallwitz D: **Identification of the catalytic domains and their functionally critical arginine residues of two yeast GTPase-activating proteins specific for Ypt/Rab transport GTPases**. *EMBO J.* 1999, **18:**5216-5225.

Alexandrov K, Simon I, Iakovenko A, Holz B, Goody RS, Scheidig AJ: **Moderate discrimination of REP-1 between Rab7 x GDP and Rab7 x GTP arises from a difference of an order of magnitude in dissociation rates**. *FEBS Lett.* 1998, **425:** 460-464.

Alory C, Balch WE: **Molecular Basis for Rab Prenylation**. *J.Cell Biol.* 2000, **150:**89-104.

Bacon RA, Salminen A, Ruohola H, Novick P, Ferro-Novick S: **The GTP-binding protein Ypt1 is required for transport in vitro: the Golgi apparatus is defective in ypt1 mutants**. *J.Cell Biol.* 1989, **109:**1015-1022.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: **The complete atomic structure of the large ribosomal subunit at 2.4 A resolution**. *Science* 2000, **289:**905-920.

Benli M, Doring F, Robinson DG, Yang X, Gallwitz D: **Two GTPase isoforms, Ypt31p and Ypt32p, are essential for Golgi function in yeast**. *EMBO J.* 1996, **15:** 6460-6475.

Berchtold H, Reshetnikova L, Reiser CO, Schirmer NK, Sprinzl M, Hilgenfeld R: **Crystal structure of active elongation factor Tu reveals major domain rearrangements**. *Nature* 1993, **365:**126-132.

Berendsen HJC, van der Spoel D, van Drunen R: **GROMACS: A message-passing parallel molecular dynamics implementation**. *Comp.Phys.Comm* 1995, **91:**43-56.

Bourne HR, Sanders DA, McCormick F: **The GTPase superfamily: a conserved switch for diverse cell functions**. *Nature* 1990, **348:**125-132.

Bourne HR, Sanders DA, McCormick F: **The GTPase superfamily: conserved structure and molecular mechanism**. *Nature* 1991, **349:**117-127.

Bowser R, Muller H, Govindan B, Novick P: **Sec8p and Sec15p are components of a plasma membrane-associated 19.5S particle that may function downstream of Sec4p to control exocytosis**. *J.Cell Biol.* 1992, **118:**1041-1056.

Brachvogel V, Neu M, Metcalf P: **Rab7: crystallization of intact and C-terminal truncated constructs complexed with GDP and GppNHp**. *Proteins* 1997, **27:**210-212.

Brennwald P, Novick P: **Interactions of three domains distinguishing the Ras-related GTP-binding proteins Ypt1 and Sec4**. *Nature* 1993, **362:**560-563.

Brunger AT: **Free R value: a novel statistical quantity for assessing the accuracy of crystal structures**. *Nature* 1992, **355:**472-475.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL: **Crystallography & NMR system: A new software suite for macromolecular structure determination**. *Acta Crystallogr.D.* 1998, **54 (Pt 5):**905-921.

Casey PJ, Seabra MC: **Protein prenyltransferases**. *J.Biol.Chem.* 1996, **271:**5289-5292.

Chavrier P, Gorvel JP, Stelzer E, Simons K, Gruenberg J, Zerial M: **Hypervariable C-terminal domain of rab proteins acts as a targeting signal**. *Nature* 1991, **353:**769-772.

Cramer P, Bushnell DA, Kornberg RD: **Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution**. *Science* 2001, **292:**1863-1876.

de Vos AM, Tong L, Milburn MV, Matias PM, Jancarik J, Noguchi S, Nishimura S, Miura K, Ohtsuka E, Kim SH: **Three-dimensional structure of an oncogene protein: catalytic domain of human c-H-ras p21**. *Science* 1988, **239:**888-893.

Desnoyers L, Seabra MC: **Single prenyl-binding site on protein prenyl transferases**. *Proc.Natl.Acad.Sci.U.S.A.* 1998, **95:**12266-12270.

Dever TE, Glynias MJ, Merrick WC: **GTP-binding domain: three consensus sequence elements with distinct spacing**. *Proc.Natl.Acad.Sci.U.S.A* 1987, **84:**1814-1818.

Doignon F, Biteau N, Crouzet M, Aigle M: **The complete sequence of a 19,482 bp segment located on the right arm of chromosome II from Saccharomyces cerevisiae**. *Yeast* 1993, **9:**189-199.

Dumas JJ, Zhu Z, Connolly JL, Lambright DG: **Structural basis of activation and GTP hydrolysis in Rab proteins**. *Structure.Fold.Des* 1999, **7:**413-423.

Dunn B, Stearns T, Botstein D: **Specificity domains distinguish the Ras-related GTPases Ypt1 and Sec4**. *Nature* 1993, **362:**563-565.

Eitzen G, Will E, Gallwitz D, Haas A, Wickner W: **Sequential action of two GTPases to promote vacuole docking and fusion**. *EMBO J.* 2000, **19:**6713-6720.

Esters H, Alexandrov K, Constantinescu AT, Goody RS, Scheidig AJ: **High-resolution Crystal Structure of S. cerevisiae Ypt51(DeltaC15)-GppNHp, a Small GTP-binding Protein Involved in Regulation of Endocytosis**. *J.Mol.Biol.* 2000, **298:**111-121.

Esters H, Alexandrov K, Iakovenko A, Ivanova T, Thoma N, Rybin V, Zerial M, Scheidig AJ, Goody RS: **Vps9, rabex-5 and dss4: proteins with weak but distinct nucleotide-exchange activities for rab proteins**. *J.Mol.Biol.* 2001, **310:**141-156.

Ferro-Novick S, Newman AP, Groesch M, Ruohola H, Rossi G, Graf J, Shim J: **An analysis of BET1, BET2, and BOS1. Three factors mediating ER to Golgi transport in yeast**. *Cell Biophys.* 1991, **19:**25-33.

Frech M, Darden TA, Pedersen LG, Foley CK, Charifson PS, Anderson MW, Wittinghofer A: **Role of glutamine-61 in the hydrolysis of GTP by p21H-ras: an experimental and theoretical study**. *Biochemistry (Mosc).* 1994, **33:**3237-3244.

Gallwitz D, Donath C, Sander C: **A yeast gene encoding a protein homologous to the human c-has/bas proto- oncogene product**. *Nature* 1983, **306:**704-707.

Garcia-Ranea JA, Valencia A: **Distribution and functional diversification of the ras superfamily in Saccharomyces cerevisiae**. *FEBS Lett.* 1998, **434:**219-225.

Garrett MD, Zahner JE, Cheney CM, Novick PJ: **GDI1 encodes a GDP dissociation inhibitor that plays an essential role in the yeast secretory pathway**. *EMBO J.* 1994, **13:**1718-1728.

Glomset JA, Farnsworth CC: **Role of protein modification reactions in programming interactions between ras-related GTPases and cell membranes**. *Annu.Rev.Cell Biol.* 1994, **10:**181-205.

Gotte M, Lazar T, Yoo JS, Scheglmann D, Gallwitz D: **The full complement of yeast Ypt/Rab-GTPases and their involvement in exo- and endocytic trafficking**. *Subcell.Biochem.* 2000, **34:**133-173.

Gotte M, von Mollard GF: **A new beat for the SNARE drum**. *Trends Cell Biol.* 1998, **8:**215-218.

Goud B, Salminen A, Walworth NC, Novick PJ: **A GTP-binding protein required for secretion rapidly associates with secretory vesicles and the plasma membrane in yeast**. *Cell* 1988, **53:**753-768.

Gutfreund H: *Kinetics for the life sciences: receptors, transmitters and catalysts*. Cambridge University Press; 1995.

Haas A, Scheglmann D, Lazar T, Gallwitz D, Wickner W: **The GTPase Ypt7p of Saccharomyces cerevisiae is required on both partner vacuoles for the homotypic fusion step of vacuole inheritance**. *EMBO J.* 1995b, **14:**5258-5270.

Haas A, Scheglmann D, Lazar T, Gallwitz D, Wickner W: **The GTPase Ypt7p of Saccharomyces cerevisiae is required on both partner vacuoles for the homotypic fusion step of vacuole inheritance**. *EMBO J.* 1995a, **14:**5258-5270.

Hama H, Tall GG, Horazdovsky BF: **Vps9p is a guanine nucleotide exchange factor involved in vesicle- mediated vacuolar protein transport**. *J.Biol.Chem.* 1999, **274:**15284-15291.

Hengst L, Grabowski R, Gallwitz D: **Ypt6p**. In *Guidebook to the small GTPases*. Edited by M Zerial and LA Huber. Oxford University Press; 1995:403-404.

Hoffenberg S, Nikolova L, Pan JY, Daniel DS, Wessling-Resnick M, Knoll BJ, Dickey BF: **Functional and structural interactions of the Rab5 D136N mutant with xanthine nucleotides**. *Biochem.Biophys.Res.Commun.* 1995, **215:**241-249.

Hwang YW, Miller DL: **A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein**. *J.Biol.Chem.* 1987, **262:**13081-13085.

Jahn R, Sudhof TC: **Membrane fusion and exocytosis**. *Annu.Rev.Biochem.* 1999, **68:**863-911.

Jancarik J, Kim SH: **Sparse matrix sampling: A screening method for crystallization of proteins**. *J.Appl.Crystallog.* 1991, **24:**409.

Jedd G, Richardson C, Litt R, Segev N: **The Ypt1 GTPase is essential for the first two steps of the yeast secretory pathway**. *J.Cell Biol.* 1995, **131:**583-590.

Jiang Y, Ferro-Novick S: **Identification of yeast component A: reconstitution of the geranylgeranyltransferase that modifies Ypt1p and Sec4p**. *Proc.Natl.Acad.Sci.U.S.A.* 1994, **91:**4377-4381.

John J, Schlichting I, Schiltz E, Rosch P, Wittinghofer A: **C-terminal truncation of p21H preserves crucial kinetic and structural properties**. *J.Biol.Chem.* 1989, **264:**13086-13092.

Jones T.A., Zou JY, Cowan SW, Kjeldgaard M: **Improved methods for building protein models in electron density maps and the location of errors in these models**. *Acta Crystallogr.D.* 1991, **47:**110-119.

Jones S, Litt RJ, Richardson CJ, Segev N: **Requirement of nucleotide exchange factor for Ypt1 GTPase mediated protein transport**. *J.Cell Biol.* 1995, **130:**1051-1061.

Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N: **Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution**. *Nature* 2001, **411:**909-917.

Jurnak F: **Structure of the GDP domain of EF-Tu and location of the amino acids homologous to ras oncogene proteins**. *Science* 1985, **230:**32-36.

Kahn RA, Der CJ, Bokoch GM: **The ras superfamily of GTP-binding proteins: guidelines on nomenclature**. *FASEB J.* 1992, **6:**2512-2513.

Kee Y, Yoo JS, Hazuka CD, Peterson KE, Hsu SC, Scheller RH: **Subunit structure of the mammalian exocyst complex**. *Proc.Natl.Acad.Sci.U.S.A.* 1997, **94:** 14438-14443.

Kirchhausen T, Bonifacino JS, Riezman H: **Linking cargo to vesicle formation: receptor tail interactions with coat proteins**. *Curr.Opin.Cell Biol.* 1997, **9:**488-495.

Lazar T, Gotte M, Gallwitz D: **Vesicular transport: how many Ypt/Rab-GTPases make a eukaryotic cell?** *Trends Biochem.Sci.* 1997, **22:**468-472.

Mandell JG, Falick AM, Komives EA: **Identification of protein-protein interfaces by decreased amide proton solvent accessibility**. *Proc.Natl.Acad.Sci.U.S.A.* 1998a, **95:**14705-14710.

Mandell JG, Falick AM, Komives EA: **Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry**. *Anal.Chem.* 1998b, **70:**3987-3995.

Matthews BW: **Solvent content of protein crystals**. *J.Mol.Biol.* 1968, **33:** 491.

Maurer KC, Urbanus JH, Planta RJ: **Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV carrying a ribosomal protein gene cluster, the genes encoding a plasma membrane protein and a subunit of replication factor C, and a novel putative serine/threonine protein kinase gene**. *Yeast* 1995, **11:**1303-1310.

Merithew E, Hatherly S, Dumas JJ, Lawe DC, Heller-Harrison R, Lambright DG: **Structural Plasticity of an Invariant Hydrophobic Triad in the Switch Regions of Rab GTPases Is a Determinant of Effector Recognition**. *J.Biol.Chem.* 2001, **276:**13982-13988.

Miaczynska M, Wagner W, Bauer BE, Schweyen RJ, Ragnini A: **Ypt protein prenylation depends on the interplay among levels of Rab escort protein and geranylgeranyl diphosphate in yeast cells**. *Yeast* 2001, **18:**697-709.

Mikol V, Rodeau JL, Giege R: **Experimental determination of water equilibration rates in the hanging drop method of protein crystallization**. *Anal.Biochem.* 1990, **186:**332-339.

Milburn MV, Tong L, deVos AM, Brunger A, Yamaizumi Z, Nishimura S, Kim SH: **Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins** . *Science* 1990, **247:**939-945.

Navaza J: **AMoRe: an automaed package for molecular replacement**. *Acta Crystallogr.A.* 1994, **50:**157-163.

Neu M, Brachvogel V, Oschkinat H, Zerial M, Metcalf P: **Rab7: NMR and kinetics analysis of intact and C-terminal truncated constructs**. *Proteins* 1997, **27:**204-209.

Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment**. *J.Mol.Biol.* 2000, **302:**205-217.

Novick P, Field C, Schekman R: **Identification of 23 complementation groups required for post- translational events in the yeast secretory pathway**. *Cell* 1980, **21:**205-215.

Ostermeier C, Brunger AT: **Structural basis of Rab effector specificity: crystal structure of the small G protein Rab3A complexed with the effector domain of rabphilin- 3A**. *Cell* 1999, **96:**363-374.

Overmeyer JH, Wilson AL, Maltese WA: **Membrane Targeting of a Rab GTPase That Fails to Associate with Rab Escort Protein (REP) or Guanine Nucleotide Dissociation Inhibitor (GDI)***. *J.Biol.Chem.* 2001, **276:**20379-20386.

Pai EF, Kabsch W, Krengel U, Holmes KC, John J, Wittinghofer A: **Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation**. *Nature* 1989, **341:**209-214.

Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A: **Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 A resolution: implications for the mechanism of GTP hydrolysis**. *EMBO J.* 1990, **9:**2351-2359.

Palade G: **Intracellular aspects of the process of protein synthesis**. *Science* 1975, **189:**347-358.

Park HW, Boduluri SR, Moomaw JF, Casey PJ, Beese LS: **Crystal structure of protein farnesyltransferase at 2.25 angstrom resolution**. *Science* 1997, **275:**1800-1804.

Peitsch MC: **ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling**. *Biochem.Soc.Trans.* 1996, **24:**274-279.

Pereira-Leal JB, Seabra MC: **The Mammalian Rab Family of Small GTPases: Definition of Family and Subfamily Sequence Motifs Suggests a Mechanism for Functional Specificity in the Ras Superfamily**. *J.Mol.Biol.* 2000, **301:**1077-1087.

Price A, Seals D, Wickner W, Ungermann C: **The docking stage of yeast vacuole fusion requires the transfer of proteins from a cis-SNARE complex to a Rab/Ypt protein**. *J.Cell Biol.* 2000, **148:**1231-1238.

Rak A, Fedorov R, Alexandrov K, Albert S, Goody RS, Gallwitz D, Scheidig AJ: **Crystal structure of the GAP domain of Gyp1p: first insights into interaction with Ypt/Rab proteins**. *EMBO J.* 2000, **19:**5105-5113.

Reiss Y, Goldstein JL, Seabra MC, Casey PJ, Brown MS: **Inhibition of purified p21ras farnesyl:protein transferase by Cys-AAX tetrapeptides**. *Cell* 1990, **62:**81-88.

Reiss Y, Seabra MC, Armstrong SA, Slaughter CA, Goldstein JL, Brown MS: **Nonidentical subunits of p21H-ras farnesyltransferase. Peptide binding and farnesyl pyrophosphate carrier functions**. *J.Biol.Chem.* 1991, **266:**10672-10677.

Rossi G, Yu JA, Newman AP, Ferro-Novick S: **Dependence of Ypt1 and Sec4 membrane attachment on Bet2**. *Nature* 1991, **351:**158-161.

Rudoni S, Colombo S, Coccetti P, Martegani E: **Role of guanine nucleotides in the regulation of the Ras/cAMP pathway in Saccharomyces cerevisiae**. *Biochim.Biophys.Acta* 2001, **1538:**181-189.

Rybin V, Ullrich O, Rubino M, Alexandrov K, Simon I, Seabra MC, Goody R, Zerial M: **GTPase activity of Rab5 acts as a timer for endocytic membrane fusion**. *Nature* 1996, **383:**266-269.

Sacher M, Jiang Y, Barrowman J, Scarpa A, Burston J, Zhang L, Schieltz D, Yates JR, III, Abeliovich H, Ferro-Novick S: **TRAPP, a highly conserved novel complex on the cis-Golgi that mediates vesicle docking and fusion**. *EMBO J.* 1998, **17:**2494-2503.

Sali A, Blundell TL: **Comparative protein modelling by satisfaction of spatial restraints**. *J.Mol.Biol.* 1993, **234:**779-815.

Sambrook J, Fritsch EF, Maniatis T: *Molecular cloning, a laboratory manual*, edn 2nd. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press; 1989.

Saraste M, Sibbald PR, Wittinghofer A: **The P-loop--a common motif in ATP- and GTP-binding proteins**. *Trends Biochem.Sci.* 1990, **15:**430-434.

Schalk I, Zeng K, Wu SK, Stura EA, Matteson J, Huang M, Tandon A, Wilson IA, Balch WE: **Structure and mutational analysis of Rab GDP-dissociation inhibitor**. *Nature* 1996, **381:**42-48.

Scheidig AJ, Burmester C, Goody RS: **The pre-hydrolysis state of p21(ras) in complex with GTP: new insights into the role of water molecules in the GTP hydrolysis reaction of ras- like proteins**. *Structure.Fold.Des* 1999, **7:**1311-1324.

Schlichting I, Almo SC, Rapp G, Wilson K, Petratos K, Lentfer A, Wittinghofer A, Kabsch W, Pai EF, Petsko GA, .: **Time-resolved X-ray crystallographic study of the conformational change in Ha-Ras p21 protein on GTP hydrolysis**. *Nature* 1990, **345:**309-315.

Schmidt G, Lenzen C, Simon I, Deuter R, Cool RH, Goody RS, Wittinghofer A: **Biochemical and biological consequences of changing the specificity of p21ras from guanosine to xanthosine nucleotides**. *Oncogene* 1996, **12:**87-96.

Schmitt HD, Wagner P, Pfaff E, Gallwitz D: **The ras-related YPT1 gene product in yeast: a GTP-binding protein that might be involved in microtubule organization**. *Cell* 1986, **47:**401-412.

Seabra MC, Brown MS, Slaughter CA, Sudhof TC, Goldstein JL: **Purification of component A of Rab geranylgeranyl transferase: possible identity with the choroideremia gene product**. *Cell* 1992, **70:**1049-1057.

Seabra MC, Reiss Y, Casey PJ, Brown MS, Goldstein JL: **Protein farnesyltransferase and geranylgeranyltransferase share a common alpha subunit**. *Cell* 1991, **65:**429-434.

Segev N: **Ypt and Rab GTPases: insight into functions through novel interactions.** *Curr.Opin.Cell Biol.* 2001, **13:**500-511.

Sheetz MP: **Motor and cargo interactions.** *Eur.J.Biochem.* 1999, **262:**19-25.

Sheldrick GM, Schneider TR: **SHELXL: high resolution refinement.** *Methods Enzymol.* 1997, **277:**319-343.

Singer-Kruger B, Stenmark H, Dusterhoft A, Philippsen P, Yoo JS, Gallwitz D, Zerial M: **Role of three rab5-like GTPases, Ypt51p, Ypt52p, and Ypt53p, in the endocytic and vacuolar protein sorting pathways of yeast.** *J.Cell Biol.* 1994, **125:**283-298.

Sollner T, Bennett MK, Whiteheart SW, Scheller RH, Rothman JE: **A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion.** *Cell* 1993, **75:**409-418.

Stenmark H, Valencia A, Martinez O, Ullrich O, Goud B, Zerial M: **Distinct structural elements of rab5 define its functional specificity.** *EMBO J.* 1994, **13:**575-583.

Strom M, Vollmer P, Tan TJ, Gallwitz D: **A yeast GTPase-activating protein that interacts specifically with a member of the Ypt/Rab family.** *Nature* 1993, **361:**736-739.

Stroupe C, Brunger AT: **Crystal Structures of a Rab Protein in its Inactive and Active Conformations.** *J.Mol.Biol.* 2000, **304:**585-598.

Sutton RB, Fasshauer D, Jahn R, Brunger AT: **Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution.** *Nature* 1998, **395:** 347-353.

Tickle IJ, Laskowski RA, Moss DS: **Rfree and the rfree ratio. I. Derivation of expected values of cross- validation residuals used in macromolecular least-squares refinement.** *Acta Crystallogr.D.* 1998, **54 (Pt 4):**547-557.

Tsao J, Chapman MS, Agbandje M, Keller W, Smith K, Wu H, Luo M, Smith TJ, Rossmann MG, Compans RW, .: **The three-dimensional structure of canine parvovirus and its functional implications**. *Science* 1991, **251:**1456-1464.

Tsukada M, Gallwitz D: **Isolation and characterization of SYS genes from yeast, multicopy suppressors of the functional loss of the transport GTPase Ypt6p**. *J.Cell Sci.* 1996, **109 (Pt 10):**2471-2481.

Ungermann C, Price A, Wickner W: **A new role for a SNARE protein as a regulator of the Ypt7/Rab-dependent stage of docking**. *Proc.Natl.Acad.Sci.U.S.A.* 2000, **97:**8889-8891.

Ungermann C, Sato K, Wickner W: **Defining the functions of trans-SNARE pairs**. *Nature* 1998, **396:**543-548.

Valencia A, Chardin P, Wittinghofer A, Sander C: **The ras protein family: evolutionary tree and role of conserved amino acids** . *Biochemistry (Mosc)*. 1991, **30:**4637-4648.

Vitkup D, Melamud E, Moult J, Sander C: **Completeness in structural genomics**. *Nat.Struct.Biol.* 2001, **8:**559-566.

Vollmer P, Gallwitz D: **High expression cloning, purification, and assay of Ypt-GTPase-activating proteins**. *Methods Enzymol.* 1995, **257:**118-128.

Vollmer P, Will E, Scheglmann D, Strom M, Gallwitz D: **Primary structure and biochemical characterization of yeast GTPase- activating proteins with substrate preference for the transport GTPase Ypt7p**. *Eur.J.Biochem.* 1999, **260:**284-290.

Wagner P, Molenaar CM, Rauh AJ, Brokel R, Schmitt HD, Gallwitz D: **Biochemical properties of the ras-related YPT protein in yeast: a mutational analysis**. *EMBO J.* 1987, **6:**2373-2379.

Walworth NC, Goud B, Kabcenell AK, Novick PJ: **Mutational analysis of SEC4 suggests a cyclical mechanism for the regulation of vesicular traffic**. *EMBO J.* 1989, **8:**1685-1693.

Wang W, Sacher M, Ferro-Novick S: **TRAPP Stimulates Guanine Nucleotide Exchange on Ypt1p**. *J.Cell Biol.* 2000, **151:**289-296.

Wendland B, Emr DS, Riezman H: **Protein traffic in the yeast endocytic and vacuolar protein sorting pathways**. *Curr.Opin.Cell Biol.* 1998, **10:**513-522.

Wichmann H, Hengst L, Gallwitz D: **Endocytosis in yeast: evidence for the involvement of a small GTP- binding protein (Ypt7p)** . *Cell* 1992, **71:**1131-1142.

Wickner W, Haas A: **Yeast homotypic vacuole fusion: A Window on Organelle Trafficking Mechanisms**. *Annu.Rev.Biochem.* 2000, **69:**247-275.

Wieland F, Harter C: **Mechanisms of vesicle formation: insights from the COP system**. *Curr.Opin.Cell Biol.* 1999, **11:**440-446.

Wimberly BT, Brodersen DE, Clemons WM, Jr., Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V: **Structure of the 30S ribosomal subunit**. *Nature* 2000, **407:**327-339.

Wu SK, Zeng K, Wilson IA, Balch WE: **Structural insights into the function of the Rab GDI superfamily**. *Trends Biochem.Sci.* 1996, **21:**472-476.

Yoo JS, Grabowski R, Xing L, Trepte HH, Schmitt HD, Gallwitz D: **Functional implications of genetic interactions between genes encoding small GTPases involved in vesicular transport in yeast**. *Mol.Gen.Genet.* 1999, **261:**80-91.

Zahraoui A, Joberty G, Arpin M, Fontaine JJ, Hellio R, Tavitian A, Louvard D: **A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cells but colocalizes with the tight junction marker ZO-1 in polarized epithelial cells**. *J.Cell Biol.* 1994, **124:**101-115.

Zhang FL, Casey PJ: **Protein prenylation: molecular mechanisms and functional consequences**. *Annu.Rev.Biochem.* 1996, **65:**241-269.

Zhang H, Seabra MC, Deisenhofer J: **Crystal structure of Rab geranylgeranyltransferase at 2.0 A resolution**. *Structure.Fold.Des* 2000, **8:**241-251.