Technische Universität München

Lehrstuhl für biomolekulare NMR-Spektroskopie

Department Chemie

Structural biology in peroxisome biogenesis

Emmanouilidis Leonidas

München, 2016

Technische Universität München

Lehrstuhl für biomolekulare NMR-Spektroskopie

Department Chemie

Structural biology in peroxisome biogenesis

Emmanouilidis Leonidas

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzender: Prof. Dr. B. Reif

Prüfer der Dissertation: Prof. Dr. M. Sattler

Prof. Dr. J. Buchner

Priv.-Doz. Dr. K. Hell

Die Dissertation wurde am 13.07.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 06.09.2016 angenommen.

Summary

Peroxisomes are essential organelles for all eukaryotic cells, since they accommodate reactions that require oxidative conditions. Malfunction of the organelle in the early development stage leads to severe disorders with few months life expectancy. Two main pathways, depending on the final compartment destination, target enzymes to the peroxisome membrane or lumen posttranslationally. Peroxisome-specific proteins, the peroxins or Pex#, orchestrate this routing. The membrane protein Pex14 interacts with the cytosolic chaperone Pex5 to ultimately transport enzymes into the organelle. On the other hand, proteins destined to the peroxisome membrane are targeted by Pex19, which acts as a chaperone and transport factor for the peroxisomal membrane protein (PMP) insertion pathway. How these components are synchronized and spatially arranged to form a pore and perform translocation remains elusive. In this work key peroxins involved in both pathways were investigated using biophysical and structural biology methods. The soluble receptors of the two pathways, Pex5 and Pex19, are in the focus of this work. For Pex5, the interactions with Pex14 at the membrane were characterized, while for Pex19 the role of farnesylation for binding to peroxisomal membrane proteins (PMPs) was studied.

Based on the Pex5/Pex14 interaction in trypanosoma structured-based drug development was pursued to develop novel therapy against the life-threatening parasite *Trypanosoma brucei*. The causative agent of sleeping sickness and huge loses in livestock worldwide, harbors – in the same organelle – peroxisomal and glycolytic reactions, thus the name glycosome. This unique feature sets glycosome biogenesis as an excellent target for drug development. Based on the

solution structure of the Pex14 protein and available data for its interaction with Pex5, we developed a molecule that selectively inhibits this interaction and kills parasites in cultures.

Chapter 1 introduces to biological background of peroxisome biogenesis and highlights previously known data about proteins of interest. Moreover, this chapter includes basic theory of NMR and some applications used in this study. Chapter 2 presents all experimental details and protocols from various techniques used.

Chapter 3.1 reports the biochemical and structural characterization of a novel interaction between the C-terminal domains of Pex14 and Pex5 in humans. Initial NMR titrations experiments revealed an additional interface, distinct from the previously well studied N-terminal interaction. A short unstructured peptide derived from Pex14 was identified to bind the structured C-terminal domain of Pex5 (TPR). Based on NMR chemical shift perturbations and intermolecular distance restraints obtained from the complex we determined the NMR structure. The structural model was further validated by mutagenesis on the interfaces using in vivo and in vitro assays. As the Pex5 C-terminal TPR domain is the cytosolic receptor of C-terminal peroxisomal targeting signal (PTS1) bearing proteins, we also investigated the potential influence of this newly described interaction in PTS1 recognition. Affinity measurement by fluorescence polarization revealed that the presence of PTS1-bound receptor shows worst affinity for the C-terminal region of Pex14. The potential function and future perspectives for this interaction are discussed.

Chapter 3.2 presents the structure-based drug development for targeting the Pex14 protein of the parasite *Trypanosoma brucei*. The solution NMR structure of tbPex14 N-terminal domain was solved and used as a model to design small molecules competing with the tbPex5 interaction. Initial hits, resembling the diaromatic nature of the native binder Pex5 WxxxF motifs (WF), were co-crystallized with tbPex14. The crystal structures confirmed the anticipated binding mode. It

was possible to prove that the compounds compete WF for Pex14 in vitro. Further compound optimization led to inhibitors considerable more toxic towards parasite than to human cells. Chapter 3.3 shows the structural characterization of the role of Pex19 farnesylation for PMP binding. NMR was used to map the interaction between farnesylated Pex19 and PMPs. Intermolecular NOEs and chemical shift perturbations were used to derive distance restraints to calculate a model of the complex. Helix α1 and a so-called ”lid” region were identified to form a hydrophobic cavity, enclosing the hydrophobic side chains of the PMPs. Mutations on the “lid” region, yield very poor peroxisomal localization in vivo, confirming the complex model.

In summary, this thesis presents novel insights for a key interaction in peroxisomal biogenesis Pex14-Pex5. As the steps involved in translocation remain unknown, the reported complex structure and biophysical characterization of the novel interaction is a first reported step that could potentially be regulated. The role of Pex19 farnesylation suggests a regulatory role for the PMP membrane insertion pathway. Since the nature of the peroxisomal pore is highly dynamic, the new findings regarding its regulation, are crucial to better understand its function.

’Δεν ελπίζω τίποτα,
δεν φοβάμαι τίποτα,
είμαι ελεύθερος”
Νίκος Καζαντζάκης

”I hope for nothing,
I fear nothing,
I am free”
Nikos Kazantzakis

Contents

Summary **iii**

1. Introduction **1**

1.1. Biological background 1

1.2. The ubiquitous and essential peroxisome organelle 2

1.3. Diseases and disorders related to peroxisome biogenesis 4

1.3.1. The protozoa *Trypanosoma brucei* 5

1.4. Matrix import pathways 6

1.4.1. Pex5 domain organization 9

1.4.2. Pex14 domain organization 11

1.5. Membrane insertion pathway 13

1.5.1. Pex19 chaperone 14

1.6. Nuclear magnetic resonance spectroscopy (NMR) 17

1.6.1. Theory of NMR 17

1.6.2. Relaxation 19

1.6.3. Nuclear Overhauser Effect (NOE) 20

1.6.4. Residual Dipolar Couplings 24

1.6.5. Protein NMR 24

1.6.6. Chemical exchange in NMR 26

1.7. Scope of Thesis 28

2. Materials and Methods **31**

2.1. Bacterial strains . . . 31
2.2. Vectors . . . 32
2.3. Buffers . . . 33
2.4. Molecular biology . . . 34
2.4.1. Quickchange mutagenesis . . . 35
2.4.2. Gene cloning . . . 35
2.4.3. Agarose gel electrophoresis . . . 36
2.4.4. Competent cells transformation . . . 36
2.4.5. Protein expression in pET systems . . . 37
2.4.6. Protein expression and isotope labeling . . . 38
2.4.7. Affinity chromatography purification . . . 38
2.4.8. Size exclusion chromatography purification . . . 39
2.5. Protein detection and evaluations . . . 40
2.5.1. SDS-PAGE gel electrophoresis under denaturing conditions . 40
2.5.2. Coomasie blue stain . . . 40
2.5.3. Protein concentration determination . . . 41
2.6. In vitro assays . . . 41
2.6.1. Fluorescence polarization (FP) . . . 41
2.6.2. AlphaScreen . . . 42
2.6.3. In vitro farnesylation . . . 43
2.6.4. Crystallization . . . 44
2.7. NMR methods . . . 45
2.7.1. Assignments . . . 45
2.7.2. NOESY experiments . . . 46
2.7.3. Chemical shift perturbations . . . 46
2.7.4. Secondary structure prediction . . . 47
2.7.5. Titrations . . . 47
2.7.6. Pex14-Pex5 inhibitors . . . 48
2.7.7. Residual dipolar couplings . . . 49

2.7.8. Structure calculation 49

3. Results **53**

3.1. Novel interactions between Pex14 and Pex5 in human 53

3.1.1. Characterization of Pex14 C-terminal 53

3.1.2. Pex14 CTDΔcc is mainly unstructured 54

3.1.3. Pex14 CTDΔcc binds Pex5 TPR in vitro 55

3.1.4. Backbone assignment of Pex5 TPR 56

3.1.5. Mapping of the Pex5/Pex14 interaction 58

3.1.6. Conformation of the WQI peptide bound to Pex5 TPR 62

3.1.7. Intermolecular NOEs between Pex5 TPR and WQI peptide . . 64

3.1.8. Structure calculation of the Pex5 TPR - WQI complex 66

3.1.9. Affinity measurement between WQI peptide and Pex5 TPR . 66

3.1.10. Refinement of TPR structure 68

3.1.11. Discussion 70

3.2. Structure-based drug design of a trypanosomal Pex14/Pex5 inhibitor 75

3.2.1. NMR structure of free tbPex14 NTD 75

3.2.2. Pharmacophore model and initial compounds 76

3.2.3. Structure-based inhibitor optimization 77

3.2.4. Mislocalization of PTS1 enzymes in parasite cultures (in collaboration with R.Erdmann lab) 82

3.2.5. Pex14/Pex5 inhibitor discussion 82

3.3. Role of Pex19 farnesylation 85

3.3.1. Mapping of PMP interaction 85

3.3.2. Intermolecular NOEs between PMP and Pex19 CTD$_{farn}$ . . . 86

3.3.3. Conformation of ALDP peptide 89

3.3.4. Docking of the ALDP peptide to Pex19 CTD$_{farn}$ 91

3.3.5. Mislocalization of PMPs in vivo (in collaboration with R.Erdmann lab) 93

3.3.6. Discussion: Role of farnesylation in PMP recognition 95

4. Conclusions and outlook **97**

Bibliography **99**

Appendix **117**

A. Appendix **117**

B. Abbreviations **125**

C. Acknowlegdements **127**

1. Introduction

1.1. Biological background

The science of biology can be parallelized as the top train of a three level train stack. Advances in mathematics, physics and chemistry are always preceded and required for biology to proceed. Therefore, while both macro- and micro world had started being described long ago, it was only last century when we first manage to conceive and prove the idea that all cells facilitate specialized molecules enabling them to exist and reproduce. The celebrity-like molecule of DNA was structurally characterized only in the second half of 20th century (Watson & Crick, 1953). The "Human Genome Project" was thought to revolutionize biology when completed in 2003, as much CERN does now to Physics (International Human Genome Sequencing, 2004). Unfortunately, biologists quickly recognized that knowing not always correlates with understanding, since the knowledge of the complete human genome did not yield any useful information at that point. The perplexing world of gene organization and the function/structure relationship of cell's molecular machineries in cells – proteins and RNA – are still the missing pieces of the puzzle.

1.2. The ubiquitous and essential peroxisome organelle

Cells are extremely complex entities, which show very high level of organization in their constituents. The inside of these entities includes an astonishing number of elements, each granted with a very specific role. This is how these magnificent systems manage to handle a plethora of different conditions and stimuli, and yet maintain their homeostasis.

As homeostasis is essential in all kingdoms of life, cells have tackled many problems by separating various compartments within them. Organelles like mitochondria, the ER or the nucleus serve as regions for specialized functions. Mitochondria, for example, are double membrane organelles present in all eukaryotic cells and operate as energy production facility. The existence of these compartments immediately creates an immense network of interactions that must be sorted out by the cell in order to live. Strikingly, in all living cells almost one third of the proteins have to be transported to the correct compartment to function properly (Dujon et al., 1994; Fraser et al., 1995; Oliver et al., 1992). Throughout millions years of evolution, cells have devised abundant methods to tackle this problem. In the vast majority of cases, a signal peptide motif is co-translated with a protein for targeting to a specific site. These motifs are subsequently recognized by large molecular machines to achieve the actual translocation. In the past, the molecular pathways that lead proteins to mitochondria, the nucleus or even chloroplasts have been well characterized. In contrast, the cascade of reactions that target proteins to peroxisomes or peroxisome biogenesis only recently started to be studied (Baker & Paudyal, 2014; Erdmann & Schliebs, 2005; Girzalsky et al., 2010; Hettema et al., 2014). Peroxisomes are organelles present in all eukaryotes and enclose reactions, which require oxidizing conditions (Hruban et al., 1972).

The first observation of peroxisomes was recorded in 1954 by J. Rhodin,

with initial name being microbodies (De Duve & Baudhuin, 1966). It was only until 1969 that the function of this organelle was linked to hydrogen peroxide metabolism (de Duve, 1969). The Nobel Prize award in 1974 to C. de Duve highlights the importance of peroxisome discovery. Further research revealed that peroxisomes can be classified in subclasses of glycosomes (Opperdoes & Borst, 1977), Woronin bodies (Jedd & Chua, 2000) and glyoxysomes (Tolbert & Essner, 1981). All subcategories share the same granular inner material, a single membrane and the absence of genetic material (Fig. 1). The differences between categories are the exact composition and enzyme function within the organelle.

Along with lipid metabolism (oxidizing very-long-chain fatty acids and synthesis of bile acids) peroxisomes have a pivotal role in oxygen radicals detoxification (Gootjes et al., 2002). The deleterious side products of many reactions, hydrogen peroxide and molecular oxygen, are enzymatically reduced to non-toxic water molecules from catalase (Sies, 1974). On the other hand the very important cholesterol synthesis pathway is mediated by numerous acyl-CoA oxidases, thiolases, dehydrogenase and hydratase enzymes (Kemp & Wanders, 2007). Although most of the mentioned functions are conserved across species, certain cell types based on special needs, have acquired additional functions. Briefly, in plants synthesis of vitamin K1 and the photorespiration pathway occurs partially in the peroxisomes (Evert RF, 2006). In yeast methanol and amine oxidation as well as assimilation also are performed inside the organelle (van der Klei et al., 2006).

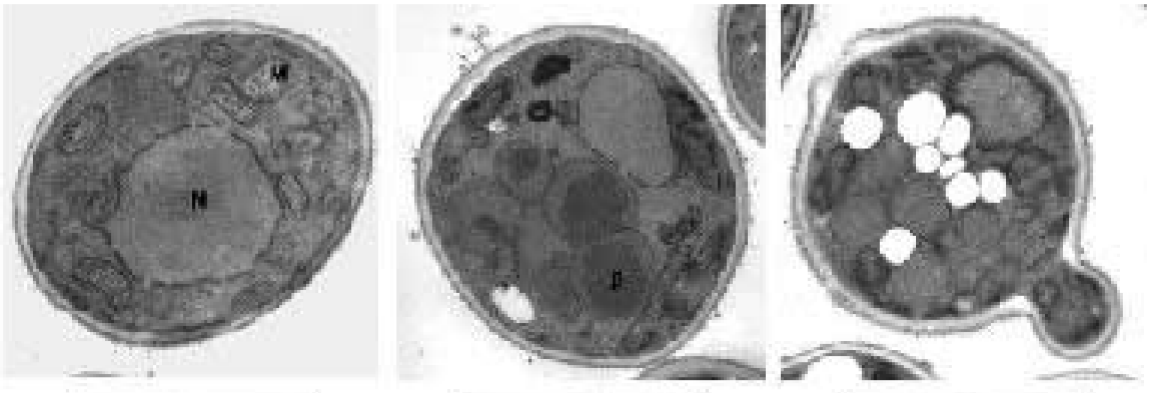


Figure 1. The peroxisome in yeast, Pichia pastoris.(left) Cells that grow in glucose containing medium enclose only 1-3 small peroxisomes. Upon change in composition of media (middle- methanol, right- oleate), cells adapt their metabolism of fatty acids, oxidation of methanol and reduction of deleterious hydrogen peroxide. Under these conditions the peroxisomes proliferate in number and size to facilitate all required additional metabolic enzymes to function. Peroxisome deficient cells are not able to respond in the same manner, ultimately affecting their survival. Adapted from Subramani 1998.

.

1.3. Diseases and disorders related to peroxisome biogenesis

The indispensable nature of the peroxisome is manifested by the numerous diseases and disorders, which are linked to malfunction or absent organelles. Depending on the origin and the nature of the malfunction, peroxisome biogenesis diseases (PBDs) or single enzyme deficiencies may lead to severe health issues. The range of the phenotypes is very wide, ranging from mild (hyperoxaluria type 1) to 6 months life expectancy after diagnosis (Zellweger Syndrome) (Salido et al., 2012). Zellweger Syndrome is a genetically unrelated group of PBDs with simi-

lar symptoms. Members of this group are the most sever Zellweger syndrome, the neonatal adrenoleukodystrophy (NALD) and the milder infantile Refsum disease. Moreover, single point mutations in crucial peroxisomal biogenesis factors, like Pex7, lead to rhizomelic chondrodysplasia punctate that is characterized by peroxisome-lacking cells. Imbalances in fatty acid metabolism result in abnormal levels of very long fatty acids and plasmalogens in the blood (Waterham & Ebberink, 2012).

1.3.1. The protozoa *Trypanosoma brucei*

The parasite of Trypanosoma genus infects humans and domestic animals causing mortality and severe economical losses (Perry & Sones, 2007). The most common disease to human is Human African trypanosomiasis (HAT) or Sleeping Sickness caused by *Trypanosoma brucei spp*, and Chagas disease induced by *Trypanosoma cruzi*. Although HAT is relatively well controlled, animal trypanosomiases thrive killing thousand livestock animals in sub-Saharan countries (Kristjanson et al., 1999; von Wissmann et al., 2011). On the other hand, Chagas disease is gradually becoming a global issue due to increased migration around the globe (Bayer et al., 2009; Requena-Mendez et al., 2015). It is characterized by widespread tissue damage, which can remain undetected for years after initial infection. The heart muscle is particularly susceptible, thus heart failure has fatal consequences.

Trypanosomatids couple the glycolytic and peroxisomal function in a single organelle, the glycosome. The latest is essential to bloodstream-form (Fig. 2) of the trypanosomes, since they rely almost exclusively on glycolysis of imported glucose from the host for ATP synthesis. Hindrance of glycosomal biogenesis has been assumed to lead to enzyme accumulation in the cytosol, which in turn causes runaway phosphorylation, ATP depletion and ultimately fatal metabolic imbalance for the parasite (Furuya et al., 2002; Haanstra et al., 2008; Kessler & Parsons, 2005). Therefore, this pathway is potentially interesting for drug discovery against

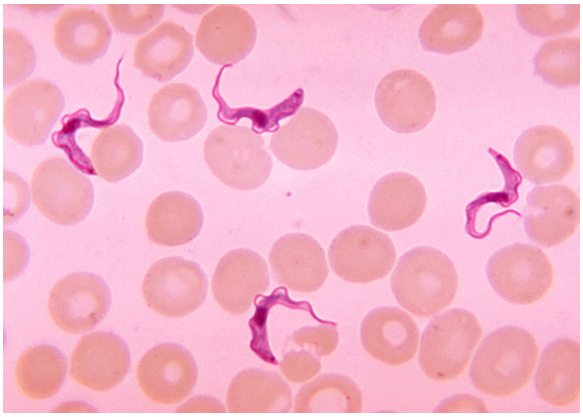


Figure 2. Blood sample from trypanosome-infected individual. Pink spheres depict red blood cells. The characteristic swirl shape of the parasite is shown in deep purple color. (Image credit: Dr. Myron G. Schultz. Public health Image Library).

sleeping sickness and many trypanosome-related diseases.

1.4. Matrix import pathways

As peroxisomes lack entirely genetic material in their interior, all matrix active enzymes are translocated post-translationally. Peroxins, abbreviated Pex, coordinate this routing. Although peroxins are not conserved among different species the general cascade of interactions are found in all eukaryotes (Baker & Paudyal, 2014; Erdmann & Schliebs, 2005; Hettema et al., 2014; Tabak et al., 2013). The import of enzymes into the lumen of the organelle exhibits some truly unique features. The very nature of the pore is well accepted to be highly dynamic, hence rationalizing how a broad collection of enzymes can be translocated, i.e. including oligomeric proteins and even 9 nm diameters nanoparticles (Brul et al., 1988; Glover et al., 1994; Keller et al., 1987; Titorenko et al., 2002; Walton et al., 1995; Zhang et al., 2003). It has been recently proven that the cargo proteins define the size of the translocating pore (Meinecke et al., 2010). Despite the lack of knowl-

edge on the actual translocation mechanism, the key peroxins participate in the pathway have been identified (Meinecke et al., 2010). Most of the peroxisomal matrix proteins are destined to the organelle via peroxisomal targeting signals (PTS1 or PTS2) located at either N-terminus (PTS2) or C-terminus (PTS1). Cargo proteins bearing either signal reach the matrix via distinct pathways, despite both pathways sharing some components. PTS1 pathway is far more frequently used than the PTS2, due to larger amount of PTS1 proteins. Pex5 mediates the transfer of all C-terminally tagged lumen enzymes from the cytosol to the peroxisome. It has been characterized as chaperone (Freitas et al., 2011) and receptor for PTS1 molecules (Rucktaschel et al., 2011; Terlecky et al., 1995). Pex5-Pex14 interaction plays a pivotal role in this pathway, since it links the cytosolic receptor Pex5 with the peroxisome membrane peroxins Pex14 and Pex13 (Albertini et al., 1997; Bottger et al., 2000; Niederhoff et al., 2005; Saidowsky et al., 2001; Schliebs et al., 1999). The complex is consisted of these three components (also Pex7 in yeast) form the docking complex (Fig. 3). Short helical motifs WxxxF (WF motifs) in the N-terminal region of Pex5 bind with nanomolar affinity to the globular structured N-terminal domain of Pex14. On the other hand, the interaction of the two membrane peroxins Pex14 and Pex13 seem to be involved at different stages of the translocation. While the first exhibit higher affinity for the cargo-loaded receptor, Pex13 prefers free Pex5 (Lanyon-Hogg et al., 2010; Mukai & Fujiki, 2006; Otera et al., 2002). Black lipid membranes used in functional assays depicted Pex14 and Pex5 as only factors for gating of the pore (Erdmann & Schliebs, 2005). Combined, these data indicate that the Pex14/Pex5 complex is responsible for the actual translocation, and Pex13 is associated with later stages, including receptor release back to the cytosol. The supplementary role of Pex17 in the yeast docking complex remains elusive (Huhse et al., 1998; Mast et al., 2010).

Much fewer peroxisomal proteins harbor the PTS2 signal, in their N-terminus. For those, the role of the shuttling receptor between the cytosol and the organelle is appointed to Pex7 (Rehling et al., 1996). Even though Pex7 highly

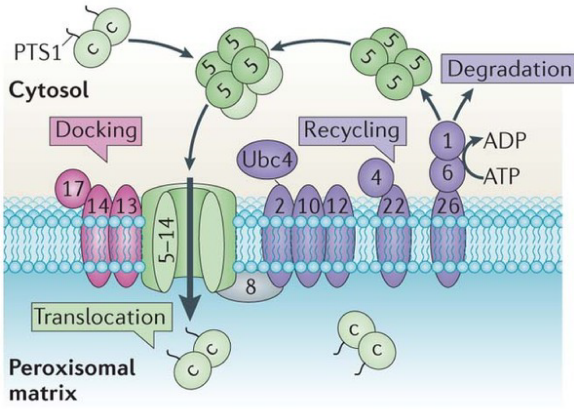


Figure 3. PTS1 import pathway components and interactions in yeast. Cargoes synthesized in the cytosol are recognized from Pex5 due to PTS1 signal. The receptor-cargo complex is docked on the peroxisomal membrane forming the docking complex with Pex14 and Pex13; Pex7 is yeast-specific component of this complex. Subsequent to pore formation and cargo release in the interior of the organelle, Pex5 is recycled back to the cytosol in a ubiquitin-related manner. System specific ubiquitination enzymes (Pex2,Pex10,Pex22,Pex12) and ATPases (Pex26,Pex1,Pex6) mediate this very last step of the pathway. Adapted from Smith & Aitchison 2013.

resembles Pex5, the absence of Pex14-binding WF motifs resulted the need of such feature from another protein. Depending on the organism, different molecules mediate Pex14 (Stein et al., 2002). For example, in humans a longer isoform Pex5(L) establishes the interaction, while Pex8 and Pex21 are used instead in yeast (Otera et al., 2000; Purdue et al., 1998). WF motifs and a Pex7 binding site are common elements in all these co-receptors (Matsumura et al., 2000). Following cytosolic initial interaction Pex7/co-receptors lead cargoes to the same docking complex as in PTS1.

1.4.1. Pex5 domain organization

Peroxin 5, together with 14, is the most essential components of import pathway. Pex5 is a soluble protein that recognizes PTS1-containing cargoes and drives them from the cytoplasm into the organelle (Gatto et al., 2000). The overall domain organization of the protein consists of a intrinsically disordered N-terminal domain (Pex5 NTD) and a globular C-terminal TPR domain (Fig. 4A) (Carvalho et al., 2006; Neuhaus et al., 2014). The previously mentioned WF motifs reside in the NTD. Recently, one additional WF-like motif serving the same Pex14 binding function has been identified. Although the new motif resembles the WxxxF/Y motif, its sequence lacks the di-aromatic feature (LVAEF) (Neuhaus et al., 2014). The binding affinity of both WF and the new motif to Pex14 N-terminal domain (Pex14 NTD) is in the high nanomolar range (Neuhaus et al., 2014; Saidowsky et al., 2001). The existence of multiple WF repeats in all species implicates that several Pex14 NTD molecules can bind to a single Pex5 NTD. In fact, it has been shown by Isothermal Titration Calorimetry (ITC) and Small Angle X-ray Scattering (SAXS) that the human Pex5:Pex14 complex has 1:8 stoichiometry in vitro (Neuhaus et al., 2014; Shiozawa et al., 2009). Whether this is the case also in the context of living cells remains unclear.

Many biophysical and biochemical experiments have proven that the Pex5 NTD is disordered (Carvalho et al., 2006; Neuhaus et al., 2014; Shiozawa et al., 2009; Su et al., 2009). Nevertheless, a Nuclear Magnetic Resonance (NMR) study detected the helical propensity for some of the WF motifs (Neufeld et al., 2009). Furthermore, SAXS measurements revealed that the NTD, even in the presence of cargo protein, remains flexible (Shiozawa et al., 2009). Therefore, it is well accepted that the NTD does not contribute to cargo binding (Shiozawa et al., 2009). Although SAXS and static light scattering (SLS) defined Pex5 as monomeric in solution (Shiozawa et al., 2009), a recent study from Pichia pastoris placed an intermolecular disulfide bond forming cysteine at the NTD. The authors postulated

that the dimerization of NTD is crucial for redox-dependent regulation of the import (Ma et al., 2013).

One additional function of the Pex5 NTD is the recycling of the receptor. As soon as the translocation finishes, the peroxin must be exported from the membrane and become available to more cycles. Interestingly, the export of Pex5 from the organelle's membrane is the only ATP driven step of the whole translocation pathway. The cell regulates this crucial step by ubiquitination of Pex5 NTD. System specific ubiquitin-conjugating and ligase enzymes, Pex12 and Pex4, attach ubiquitin at a conserved cysteine at position 4 (Carvalho et al., 2007; Smith & Aitchison, 2013). Monoubiquitinated Pex5 protein is extracted from the membrane in a ATP-dependent manner (Platta et al., 2013). If for any reason this step is inhibited, pulyubiquitination and proteosomal degradation of Pex5 is pursued (Kiel et al., 2005).

Seven tetratricopeptide repeats (TPRs) compose the C-terminal cargo-binding domain of Pex5 (Pex5 TPR). Two TPR triplets (1-3 and 5-7) are organized in a ring forming the PTS1 binding site (Fig. 4B). Both triplets are linked by a non-canonical TPR (4) that is only partially structured (Gatto et al., 2000; Stanley et al., 2006). The PTS1 peptide contributes most in the cargo-binding event, as it is the determinant factor of affinity. The initially discovered PTS1 sequence serine-lysine-leucine (SKL) was later expanded to variants containing [S/A/C]-[K/H/R]-[L/M] (Gould et al., 1987; Lametschwandtner et al., 1998; Reumann et al., 2009). Depending on the nature of alternative residues the PTS1 peptide affinities range from low nanomolar to low milimolar for TPR. Conserved asparagines in the TPR groove enable the grasping of PTS1 backbone by formation of hydrogen bonds (Gatto et al., 2003; Stanley et al., 2006). Interestingly, binding of PTS1 to TPR domain causes moderate alteration in the general arrangement of the two TPR triplets via relative movement between TPRs (1-3) and (5-7) (Stanley et al., 2007; Williams et al., 2011). Although this observation is valid for all complex structures available, the potential regulatory role it possesses is unknown.

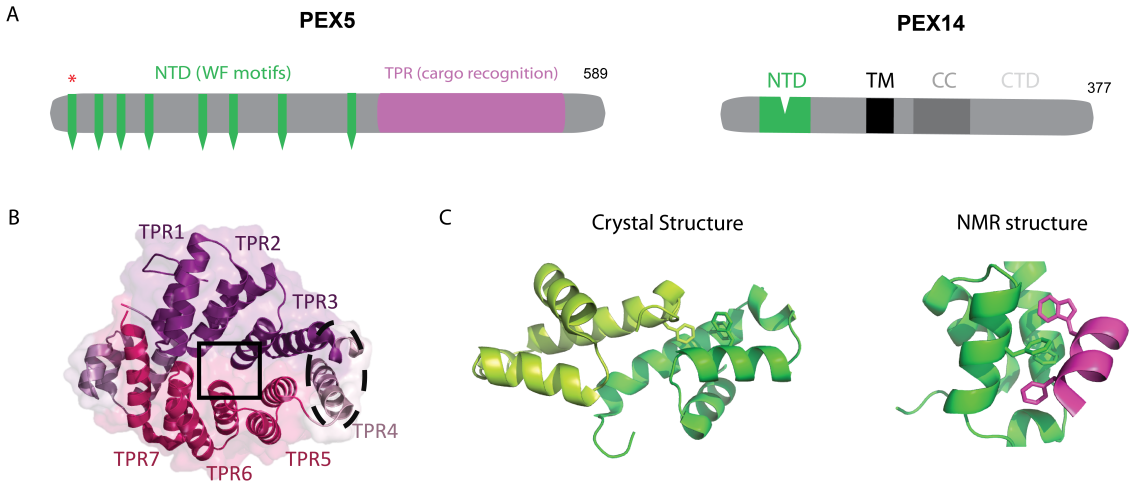


Figure 4. Pex14 and Pex5 protein structures. A) Domain organization of human Pex5 and Pex14 proteins. The highly unstructured NTD of Pex5 contains the WF motifs which upon binding to Pex14 NTD fold to -helices. The C-terminal TPR domain of Pex5 is responsible for recognition of PTS1-containing proteins. Contrary to Pex5 NTD, the NTD of Pex14 is folded. TM = transmembrane region, CC=coiled coil. B) Cargo recognition domain of Pex5 contains 7 TPR repeats folded in a circular manner. The black rectangular highlights the PTS1 binding site. The half-TPR (TPR4) separates equally the two TPR triplets (1-3 and 5-7) and enables relative movements with respect to each other. C) Crystal and NMR structure of Pex14 NTD. The NMR structure the complex with a WF motif (magenta) of Pex5. Adapted from Emmanouilidis et al. 2016.

1.4.2. Pex14 domain organization

Docking on peroxisome membrane follows the initial cargo recognition in the cytoplasm. The membrane pre-anchored docking complex of Pex14 and Pex13 mediates this step (Albertini et al., 1997; Bottger et al., 2000). Experimental data and sequence analysis of Pex14 homologues expose four features of the protein: 1) the N-terminal helical domain (Pex14 NTD), 2) the hydrophobic region, 3) the coiled-coil domain and 4) a C-terminal region (Fig. 4A).

The best studied region of the protein is the structured NTD. This small

soluble domain has been shown to interact not only with the PTS-mediated pathways but also with peroxisomal membrane proteins (PMPs) translocation pathways (Neufeld et al., 2009; Neuhaus et al., 2014). Lately, the rather unexpected binding partner tubulin was added to that list (Bharti et al., 2011). The crystal structure of rat NTD unveiled the 3-helical bundle fold of the domain (Fig. 4C) (Neuhaus et al., 2014; Su et al., 2009). Two stacked phenylalanines at the core of the structure fix the fold and create two hydrophobic pockets exposed on the surface of the protein. Protein packing in the crystal resulted in artificial dimerization of the molecule, while solution (NMR) indicated that it is predominantly monomeric (Fig. 4C) (Neufeld et al., 2009).

Complex NMR structures of the prototypic WF motif, non-canonical new motif found in Pex5 NTD and FFxxxF motif found in membrane insertion chaperone Pex19, revealed the mode of interaction between Pex14 NTD and its binding partners. Two hydrophobic pockets on the Pex14 NTD surface are filled by aromatic (or aliphatic) side chains of the ligands. Interestingly, the peptide derived from Pex5 and Pex19 have opposite orientation when bound on Pex14 NTD (Neufeld et al., 2009). This overlapping binding site of Pex14 for Pex5 and Pex19 depict a competitive activity, the physiological role of which remains unclear.

Any information regarding the structure or function beyond the NTD is scarce. In the primary sequence the hydrophobic patch, responsible for membrane binding, follows NTD. The general protein arrangement with respect to the membrane remains a disputed topic in field. Although it is well accepted that the coiled-coil region and the C-terminal are facing the cytoplasm, for the relative position of the NTD exist contradictory results (Brocard et al., 1997; Johnson et al., 2001; Shimizu et al., 1999; Will et al., 1999). Strikingly, different assays have placed the NTD in all possible positions with respect to the membrane (i.e. outer-,inner-side of the membrane associated, and integral membrane protein). Furthermore, a lumen loop of yeast Pex13 showed an interaction with Pex14 (Schell-Steven et al., 2005). This finding implies that at least at some stage of the translocation a Pex14

region is indeed inside the organelle.

1.5. Membrane insertion pathway

Although many scenarios exist regarding the biogenesis of the peroxisome (de novo formation, fission and fusion, ER-derived vesicles), all require at some stage the enrichment of the membrane with peroxins (Motley & Hettema, 2007; Purdue & Lazarow, 2001; Schrader et al., 2012; Titorenko et al., 1997). The correct spatial topology of peroxisomal membrane proteins (PMPs) in the organelle's membrane is one of the most demanding processes in the organelle's biogenesis and maintenance. Three peroxins – Pex3, Pex16, and Pex19 in mammals – are capable of restoring by complementation the absence of peroxisomes in peroxisome deficient cells (Fig. 5) (Fujiki et al., 2006; Kim et al., 2006). Thus, these three factors are essential in PMP insertion in peroxisome membranes (Kim & Hettema, 2015). Although essential, Pex16 has not been linked yet to any particular function. Interestingly, the topology and role of Pex16 has been reported to vary upon the tested organism. In yeast for example; data indicate that it is involved in fission suppression while in mammals in early biogenesis steps (Honsho et al., 2002).

Pex3 is a integral membrane protein that associates with Pex19. Thus, it functions as Pex19-dependent receptor on the membrane (Fang et al., 2004). Although as integral membrane protein Pex3 is a PMP, yet the membrane insertion does not rely on Pex19 (Jones et al., 2004). Pex3 spans once through the membrane but more than 90% of the protein is facing the cytoplasm. The very short luminal N-terminal segment and the transmembrane region resemble the typical ER signal anchor-like motifs (Thoms et al., 2012). The unexpected binding partner Sec61 from the endomembrane channel system recognizes these Pex3 elements, providing insights to the origin of the organelle in yeast (Hoepfner et al., 2005; Kragt et al., 2005; Tam et al., 2005; Thoms et al., 2012). The crystal structure of the cytoplasmic C-terminal domain of the Pex3 protein revealed a helical bundle fold (Fig. 5) (Sato

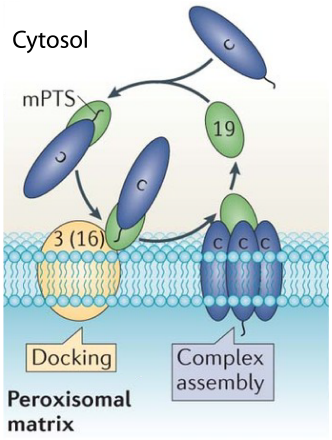


Figure 5. Peroxisomal membrane proteins insertion in the membrane. Membrane anchored proteins Pex16 and Pex3 recruit Pex19 from the cytoplasm. Pex3 spans the membrane once, with the C-terminal region facing the cytosol. Pex19 bears farnesylation site, CAAX box, and binds PMPs in the cytosol. The relative position of Pex3, Pex19 and PMP on the membrane is not known, therefore the picture is apparent schematic representation. Adapted from Smith & Aitchison 2013.

et al., 2010; Schmidt et al., 2010). The structure was determined in the presence of a Pex19-derived peptide, revealing the mode of interaction. Notably, the Pex19 peptide was found on the farthest location from the anticipated transmembrane anchor (Fig. 5).

1.5.1. Pex19 chaperone

Arguably, Pex19 is the central component of PMP targeting to peroxisome membranes. It is an essential, multifunctional peroxin ubiquitously found in all eukaryotes. In the different scenarios of peroxisome biogenesis it accomplishes district roles. For example, in the proliferation and division model, Pex19 recognizes nascent PMPs via the membrane Peroxisome Targeting Signal (mPTS), acting as quality control node for proper folding, prior to organelle delivery in a

Pex3-dependent manner (Halbach et al., 2006; Jones et al., 2004). These findings assigned Pex19 as both a PMP shuttling receptor and chaperone, with recycling ability after PMP release, like Pex5. In fact, Pex19 and Pex5 share a lot of common features, other than recycling, including the domain organization and localization.

Similar to PTS1 receptor Pex5, Pex19 can also be divided into two parts, the unstructured N-terminal region, and a structured C-terminal (Pex19 CTD) (Fig. 6). The N-terminal part is generally less conserved than the Pex19 CTD (Altschul et al., 1990). Although the length of Pex19 deviates across species, function is expected to be similar. The sole element that it is absolutely conserved is the so-called CAAX box at the C-terminal end of the protein. This motif, consisting of a cysteine followed by two aliphatic amino acids and a variable, is subject to post-translational modification. Indeed, Pex19 has been shown experimentally to be farnesylated in vivo, at least for human and yeast cells (Fujiki et al., 2006). The farnesyl group, a 15-carbon isoprenoid, is covalently attached to the cysteine by the enzyme farnesyltransferase (FTase).

Albeit the structure of the full-length protein remains unknown, insights about the structural composition of fragments are available. A high-resolution crystal structure of the human Pex19 CTD, without the CAAX box, has been solved and revealed a largely helical domain (Schueller et al., 2010). For the disordered N-terminal counterpart only stretches of few residues in the human protein exhibit helix formation upon Pex3 binding (Schmidt et al., 2010).

The mechanism by which Pex19 recognizes PMPs is only poorly understood. The first helix $\alpha 1$ of the Pex19 CTD showed significant contribution in PMP binding, however in this study the CAAX box was absent (Schueller et al., 2010). Since PMPs are hydrophobic proteins, the interaction was expected to been governed by hydrophobic interactions (Fransen et al., 2001; Jones et al., 2001). Indeed, mutations of exposed hydrophobic amino acids in the $\alpha 1$ helix resulted severe reduction of PMP binding. The binding affinity of PMPs to Pex19 CTD was found to be in the micromolar range. Interestingly, the presence of the CAAX motif en-

hances the binding tenfold (Rucktaschel et al., 2011). This result led to the conclusion that farnesylation has an important role in PMP binding. Despite these results, the physiological relevance of Pex19 farnesylation is still controversial due to contradicting results. For example, Pex19 variant with mutated farnesylation site was still able to exhibit peroxisomal localication of PMP catalase like wildtype Pex19 (Sacksteder et al., 2000).

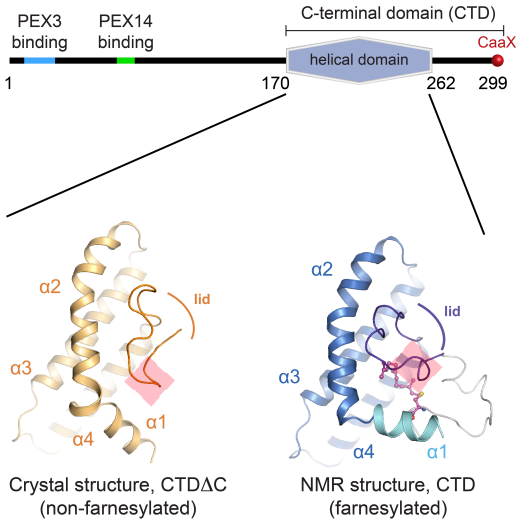


Figure 6. Domain organization and experimental structures of Pex19. Unstructured N-terminal region of the protein comprises Pex13 and Pex14 binding sites. C-terminal domain of the protein, which is responsible for PMP binding, was found to form a helical structure both by NMR and crystallography. At the end of the sequence a cysteine of CaaX box is subject to farnesylation. This feature was missing in the construct used for the crystal structure, while NMR structure is of farnesylated protein. Differences between two structures are mainly found in $\alpha 1$ helix orientation and the lid region.

1.6. Nuclear magnetic resonance spectroscopy (NMR)

1.6.1. Theory of NMR

During the past fifty years nuclear magnetic resonance (NMR) spectroscopy has evolved greatly and today, together with X-ray crystallography and cryo-EM, is considered a major tool for biomolecule structure determination. When compared to other spectroscopy techniques, like mass spectroscopy, larger sample amounts are needed but the method is non destructive, enabling a plethora of experiments done with just one sample.

Most isotopes have a characteristic spin quantum number (I). This number can be zero (i.e. ^{12}C,^{16}O), non-zero integer (i.e. ^{2}H) or a fraction (i.e. ½ for ^{13}C ^{15}N). The nuclei have the unique characteristic of possessing an intrinsic nuclear angular momentum. In the presence of an external magnetic field (B_o), two spin states exist, +1/2 and -1/2. These two states differ in the alignment with the external field. The lower energy +1/2 is aligned to the field, while the higher energy -1/2 is opposed. The energy difference between the two states depends on the strength of the external magnetic field, yet it is always very small. When expressed in frequency terms, the difference corresponds to a distinct frequency, the Larmor frequency. The Larmor frequency depends on the type of the nucleus via the gyromagnetic ratio, and the external field (Eq. 1.1).

$$\Delta E = \gamma \hbar B_o = \hbar \omega_o \qquad \text{(Eq. 1.1)}$$

with $\hbar$ = reduced Planck constant, γ = gyromagnetic ratio, B_o magnetic field strength, ω_o = Larmor frequency

The two spin system is an oversimplification for easier understanding, in reality many different ones are present in the sample. The vector model was developed to adequately describe and make it conceptually easier to conceive how a

large number of spins behave in a magnetic field. According to this model a bulk magnetic moment vector is the sum of the individual magnetic moment vectors. This vector is called bulk magnetization of the sample or magnetization vector. In equilibrium the difference between spins aligned to the external field and opposed to, create a net magnetization that is aligned to the external magnetic field (z axis). Upon application of radio frequency (*rf*) pulse, the net magnetization vector tilts away from the z-axis. The vector precesses about the field with the motion defined as Larmor precession.

Population follows the Boltzmann equation:

$$\frac{N\alpha}{N\beta} \approx 1 - \frac{\gamma\hbar B_o}{kT} \qquad \text{(Eq. 1.2)}$$

with $N\alpha$ and $N\beta$ number of spins in α or β state, $\hbar$ = reduced Planck constant, γ = gyromagnetic ratio, B_o magnetic field strength, k = Boltzmann constant, T = temperature

In the NMR spectrometer the bulk magnetization is manipulated with a electromagnetic field in the transverse plane, x-y. Since the spectrometer consists of a large superconductive magnet, it is not applicable to alternate so drastically the field. Therefore, an additional B_1 field perpendicular to B_o is used to create frequency pulses at near-Larmor frequency. This B_1 field generates radio frequency in form of a *rf* pulse. As soon this pulse is applied, the vector deviates accordingly but as soon as the external influence (*rf*) is stopped, it will return to the z-axis, at equilibrium. The "artificially" induced deviation from equilibrium gradually decays giving rise to free induction decay (FID), which is recorded. This decay signal is in time-domain and it is composed by the summary of cosine waves of all frequencies present in the sample. Ultimately, the time-domain signal is converted to frequency-domain by Fourier Transformation.

The Larmor frequency depends on the gyromagnetic ratio of each nucleus

and the magnetic field. The NMR spectra should appear identical for each nucleus type. This is not the case since, every nucleus experiences a slightly different chemical environment due to its position in the sample. This shift from the default Larmor frequency is called chemical shift. Usually this deviation is very small and it its measured in parts per million (ppm). In order to be able to compare chemical shift values from different instruments, the reference molecule tetramethylsilane (TMS) is used to calibrate the chemical shifts for ^{11}H and ^{13}C NMR (Eq. 1.3).

$$\delta_{ppm} = 10^6 * \frac{\nu - \nu_{\text{TMS}}}{\nu_{\text{TMS}}} \qquad \text{(Eq. 1.3)}$$

with ν and ν_{TMS} the absolute resonance frequency of measured sample and reference sample TMS

1.6.2. Relaxation

As previously stated, subsequently to a short *rf* pulse application, net magnetization will resettle to equilibrium along the z-axis. This process of returning to equilibrium is called relaxation and plays crucial role in all aspects of NMR spectroscopy. Mainly two aspects contribute to relaxation in solution: dipole-dipole interactions and anisotropy. Two distinct processes govern relaxation. The first spin-lattice or T1 relaxation is the mechanism by which the magnetization vector return to equilibrium along the axis of the static applied magnetic field, the z-axis. The second spin-spin or T2, describes the decay, or fan out, of magnetization vector coherence in the x-y plane. In biological systems, since x-y coherence decay is followed by the return of the magnetization in the z-axis, usually T1>T2.

Recovery to the z-axis is exponential with the time constant T1, while transverse magnetization (M_{xy}) decays exponentially with a time constant T2 (Eq. 1.4 and 1.5)

$$\frac{dM_z(t)}{dt} = -\frac{1}{T_1}[M_z(t) - M_z^0] \quad \text{(Eq. 1.4)}$$

$$\frac{dM_{xy}(t)}{dt} = -\frac{1}{T_2}M_{xy}(t) \quad \text{(Eq. 1.5)}$$

with M_z, M_{xy} magnetic field component on z-axis or x-y plane, M_z^0 equilibrium value of magnetic field component.

The two relaxation processes depend on the motion of the molecule in solution. Any induced coherence in the x-y plane decays faster for slow-tumbling large molecules, while the opposite is true for fast-tumbling small molecules.

1.6.3. Nuclear Overhauser Effect (NOE)

In 1953 Albert Overhauser described a phenomenon, which later on revolutionized NMR spectroscopy, especially for large biomolecules like proteins. The Nuclear Overhauser Effect (NOE) is the phenomenon where the relaxation of one spin effect relaxation of nearby spins too, due to dipole-dipole interactions. This cross-relaxation involve double quantum and zero quantum transitions giving rise to positive or negative NOEs respectively. The type of transition is largely defined by the size of the molecule, with small molecules preferring double quantum while large zero quantum (Fig. 7). The significance of the NOE, lies in the fact that the magnitude of how much one spin will effect a neighbor, depends on distance r. In fact, the NOE between two spins declines rapidly further away from the radiated spin, with $1/r^6$ dependence. By measuring the NOE between spins we can estimate the distance between then, this is used a distance restraint to build a structure model of the biomolecule. The NMR experiment to record is named NOESY.

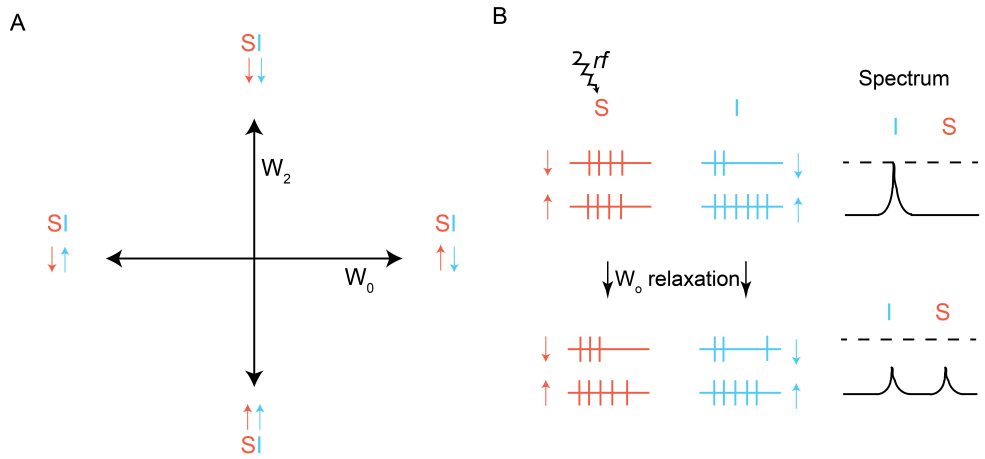


Figure 7. Cross relaxation and the NOE effect. A) Different energy levels due to spin orientations. Two spins S and I with only dipolar coupling may undergo double quantum (W_2) or zero quantum (W_0) transition during relaxation. B) Irradiation of spin S causes the two spin states to be equally populated and therefore no NMR signal is observed in the spectrum. In proteins, after the rf pulse is stopped, through W_0 transition S gradually re-establishes equilibrium and the NMR signal reappears. While the I spin the higher energy state becomes more populated, as result the differences between states is lesser, therefore there is decrease in the intensity of the signal of I spin in the spectrum.

Filtered NOESY

As previously described, only certain isotopes are NMR active, in the case of proteins ^{1}H, ^{13}C and ^{15}N are of interest. Formation of a protein complex AB with uniformly ^{15}N labeled A and unlabeled B will lead to several types of ^{1}H NOEs, namely intermolecular and intramolecular. Although all NOEs are important, usually the intermolecular ones define the interaction interface and contribute to complex structure calculation. In the described experimental setup the intermolecular NOEs are between protons, which are attached to ^{15}N on A, and ^{14}N on B. Such intermolecular NOEs can be recorded by using multidimensional NMR experiments

containing X-filter element. Due to the presence of scalar coupling only in ^{1}H-^{15}N, it is possible to differentiate NOEs originating from the two correlations.

The transfer NOE

Many more methods have been developed to study interactions between proteins by NMR, since single structures usually do not answer biological questions. For that purpose in addition to filtered NOESY experiment, a method was established to obtain distance restrains under usually occurring conditions. As large biomolecules usually tumble slowly in solution they exhibit strong negative NOEs, while a smaller molecule tumbles faster showing near zero to small positive NOEs (Fig. 8). This difference is exploited by the transfer-NOESY (tr-NOESY) experiment to reveal otherwise unseen NOEs (Balaram et al., 1972; Post, 2003). Transferred NOEs is very helpful tool not only for ligand binding determination but also conformation information of the bound ligand.

Usually in a protein-peptide complex, the peptide's resonances are severely broadened due to rapid relaxation when bound to a much larger protein. In contrast, the population fraction of the free peptide exhibits much slower relaxation, hence stronger signals. By recording a NOESY spectrum with relative long mixing time, we allow NOEs from the bound state to transfer to the free state. The transfer of the NOEs from the bound state to the free state depends on the on/off rates and the affinity of the binding. In fact the following inequality (Eq. 1.6) must be fulfilled in order transferred NOE to occur :

$$|N_b\sigma_b| >> |N_f\sigma_f| \qquad \text{(Eq. 1.6)}$$

with σ the cross-relaxation rate and N number of molecules in free (f) and bound (b) state.

In practice, the mentioned prerequisites are fulfilled by a small ligand

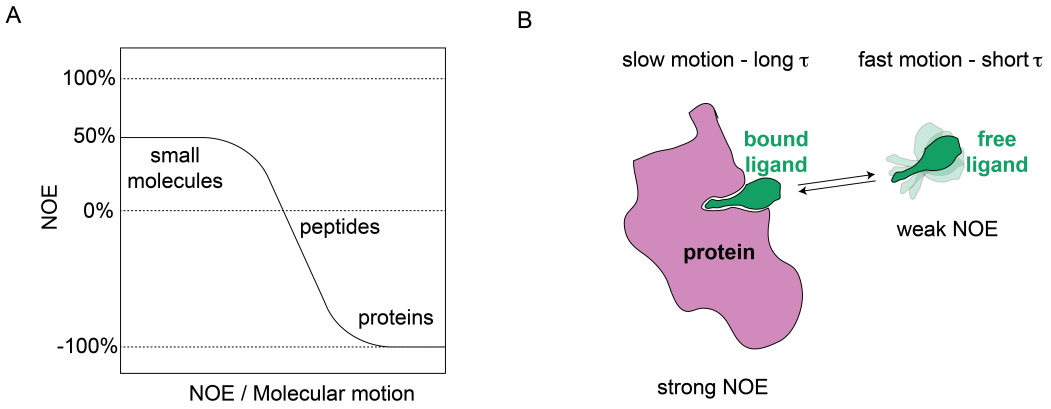


Figure 8. NOE intensity in relation with molecular motion and transfer NOE scheme. A) Graph of NOE dependence on molecular motion. B) Free ligands show due to fast motion in solution show weak or even zero NOE, while when bound to a much larger protein it acquires the characteristics of a larger molecule. As result NOE built while bound still exists when the ligand dissociates.

binding to a large protein with micromolar to milimolar range of affinity. Moreover, the ligand must be in high molar excess over the protein, usually 10 and 30 to 1. Comparison NOESY peptide spectra in the presence and in absence of binding protein, additional NOEs should originate from the bound conformation of the peptide. The discrimination between free and bound NOEs can be achieved by comparing signal intensities, since usually free peptides would have near zero NOE values. A major drawback of the tr-NOESY experiment is spin diffusion. The relative long mixing times cause not only enhancement of NOE between nearby protons, but also indirect via protein protons. This may lead to ”relay” NOEs and wrong interpretation. Additionally, NOEs obtained from such experiments cannot be used straightforward as distance restraints, as their intensities are greatly affected by the long mixing times used. In all cases, various mixing times should be tested to determine the optimal value for every protein/ligand system.

1.6.4. Residual Dipolar Couplings

Since NOEs report distances of maximum 5-6 Å, they lack any structural information regarding much longer distances. A 50 kDa protein which consists of two domains would give the same set of NOEs regardless of the relative position of each domain with respect to the other. Due to the very short range of NOEs, only few ,if any, interdomain NOEs can be obtained depending on the distance between the two domains. Consequently, it is usually not possible to detect larger changes only with NOEs. In contrast, residual dipolar couplings (RDCs) give information on relative orientation of bond vectors. A protein in isotropic solution would tumble freely, therefore all vectors formed by dipole-dipole interactions are averaged out to zero. However, for the same protein when weakly (3-5%) aligned, i.e. in filamentous phages, incomplete averaging occurs. In such anisotropic conditions, dipole-dipole interactions are no longer zero. By measuring the residual dipolar couplings, for example for ^{1}H-^{15}N bond vectors of known bond lengths, we can extract information about the relative angle between a nucleus-nucleus vector and the external magnetic field. Ultimately, residual dipolar couplings reveal the relative orientation of regions of the molecule, which are not necessarily close in space.

1.6.5. Protein NMR

The theoretical aspects described above have been successfully applied to biological systems to address questions regarding structure-function relations. Although this study focuses on the aspects proteins and peptides are the only biomolecules studied by NMR, the structure and dynamics of nucleic acids can also be investigated. Currently, conventional NMR techniques can yield protein structures for sizes up to 20-30 kDa. However, high field NMR spectrometers and more sophisticated methods can push this limit further. Many advances occurred since the beginning of the technique in order to overcome some major burdens. Probably the biggest issue spectroscopists dealt was the immense signal overlap

that occurs from the large numbers of same atom types in a protein. In contrast to small organic molecules, one-dimensional spectra of the naturally abundant isotope ^{1}H lacks resolution, thus little information can be extracted. Moreover, as previously described, large biomolecules relax fast, resulting quick loss of signal.

The first milestone in protein NMR was the incorporation of ^{15}N and ^{13}C isotopes in the sample. Recombinantly expressed proteins from bacteria growing in media containing these isotopes led to uniform enrichment with these NMR-active isotopes. This expanded the type of one-dimensional spectra that could be acquired, but more importantly enabled multidimensional experiments. The later significantly increased the required resolution. Overtime more expensive and sophisticated methods were developed to heighten even more the upper limit of the protein size studied. Partial or complete deuteration greatly improved relaxation properties of the molecules, by a significant reduction of dipole-dipole interaction. In combination with residue-, or even chemical group-specific labeling constitutes the latest and most powerful tool to study large biomolecules. In parallel to sample preparation improvements, many advances in electronics, physics and materials boost the construction of more powerful instrumentations. In this context, gigahertz magnets are available nowadays which enhance even more both sensitivity and resolution.

As previously mentioned, multidimensional experiments enhances the amount and quality of information extracted by NMR, yet already simple 1D proton spectra may reveal important aspects about the tested biomolecule. Lineshape and dispersion in 1D spectra are indicative of overall structure integrity, and presence of any fold. One dimensional ^{1}H spectra and 2D ^{1}H,^{15}N heteronuclear single quantum correlation spectra (HSQC) comprise the two most commonly used experiments in biomolecular NMR. The HSQC spectrum correlates all amides, creating a unique spectrum for each protein, thus also referred to as the “fingerprint” spectrum. Usually all resonances of this spectrum are assigned to specific amides in the protein by facilitating triple resonance experiments. In such experiments,

while the two dimensions are still the amide proton and the amide nitrogen, yet the additional third dimension enables to identify specific residue within the polypeptide chain. Depending on the type of the experiment, this third dimension includes resonances of scalar coupled carbons to the amide. In the HNCA experiment for example one plane contains ^{1}H-^{15}N correlations , while the third dimension includes information about C_α chemical shift values. Combination of spectra that include CO, C_α or C_β it is possible to unambiguously assign all backbone atoms of the protein, since carbon resonance frequencies are characteristic for each amino acid. The side-chains are also assigned similarly by TOCSY spectrum, which correlate all ^{1}H and ^{13}C frequencies of all atoms within the same residue.

All experiments described so far in this section, are based on net magnetization transfer through bond, thus are useful only for resonances assignments. In the NOESY experiment though, magnetization is transferred through space via dipolar interactions between nearby atoms. Although ^{1}H,^{1}H NOESY spectra can be sufficient to resolve NOEs in a peptide or small molecule, this is not the case for larger than few kilo Dalton proteins. Isotope-edited NOESY-HSQC spectra, in which the HSQC sequence selects protons bound to ^{15}N or ^{13}C, were developed to tackle this problem. Such experiments not only can further assign resonances previously unknown, but also provide distance information between protons – the basis for structure calculation by solution NMR.

1.6.6. Chemical exchange in NMR

NMR is a very versatile tool, being able to determine high-resolution protein structures, study dynamics but also interactions. All kind of interactions have been extensively studied by NMR, including protein-protein, protein-nucleic acid and protein-small molecule. Although it is possible to validate if a small molecule interacts with an unlabeled protein already by 1D proton spectra, the most frequently used technique is the 2D NMR titration. It consists of a series of ^{1}H,^{15}N

or ^{1}H,^{13}C HSQC spectra of a protein in the presence of increasing ligand concentration. If interaction occurs, chemical microenvironment of atoms involved in the interaction changes, causing chemical shift perturbations. In such experiments results vary greatly due to kinetics of the complex formation. The frequencies difference in free and bound state, as well as association and dissociation rates defines the outcome of the titration. Three exchange regimes are possible:

$$\begin{array}{cc} \text{slow} & \Delta\nu >> k_{ex} \\ \text{intermediate} & \Delta\nu \approx k_{ex} \\ \text{fast} & \Delta\nu << k_{ex} \end{array}$$

with $\Delta\nu$ = frequency difference of free and bound state, k_{ex} = exchange rate (Eq. 1.7)

$$k_{ex} = k_{on} + k_{off} \qquad \text{(Eq. 1.7)}$$

with k_{on} and k_{off} the rates of conversion between bound and free states.

In the slow exchange regime, only two sets of resonances are visible, corresponding to free and bound states respectively. The intensity of the each signal correlates with to the relative population of the two states (Fig. 9A). This type of exchange, although heavily depend on the exchange rates, typically indicates a tight interaction. In contrast, more loose interactions in µM-mM affinity range, exhibit fast exchange. When in fast exchange, signals correspond to the average chemical shift value between population weighted free and bound state. Hence, it is possible to track the perturbation of chemical shifts (Fig. 9B). Accordingly, when the exchange lies in-between slow and fast exchange regime, a combination of both behaviors is observed, line broadening and chemical shift perturbation (Fig. 9C).

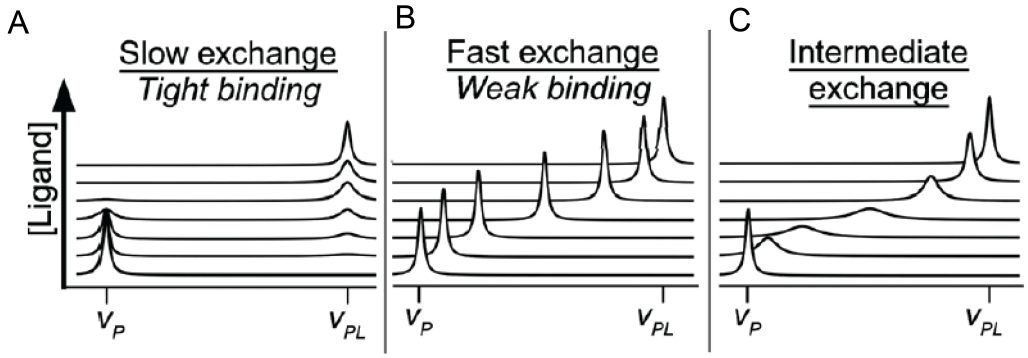


Figure 9. Schematic overview of exchange regimes. ν correspond to free ($_P$) and bound ($_{PL}$) resonances. A) In slow exchange regime upon increasing concentration of ligand the signal which corresponds to free state gradually decreases while bound frequency signal increases. B) In contrast, weak binding is usually in fast exchange regime resulting a shift of frequency observed as more ligand is present. C) Interactions following intermediate exchange exhibit frequency shift and change in intensity of the signal upon addition of ligand. Adapted from Kleckner & Foster 2011.

1.7. Scope of Thesis

Almost one third of all proteins in a cell have to be targeted to the correct compartment for proper function. The mechanisms by which proteins are translocated on or inside the lesser-known peroxisome remains elusive, despite the numerous disorders linked to the malfunction mechanism. The highly dynamic nature of the pore involved in the process so far has limited the applicable techniques to study such systems. Therefore, in this thesis I applied solution NMR spectroscopy to address questions regarding peroxisome biogenesis.

One of the key players in this process is membrane peroxin, Pex14. Together with the cytosolic receptor Pex5, it has been shown recently to be required

for pore activity. Our knowledge regarding the function and interactions of Pex5 and Pex14 is rather limited. In fact, the only structurally characterized region of Pex14 is the N-terminal domain, which forms a three-helical bundle. This domain represents less than one quarter of the full-length protein. A hydrophobic patch of amino acids and a putative coiled-coil domain are predicted from sequence analysis, but their molecular structure and function is so far unknown. The remainder of the protein has not yet been associated with any function. Therefore, the C-terminal region of Pex14 is studied in this thesis, by characterizing its conformation, dynamics and molecular interactions with Pex5 using biophysical methods, NMR spectroscopy and functional studies.

The interaction between the N-terminal domain of Pex14 and WxxxF motifs in Pex5 is important for the cargo import into peroxisomes and, similarly, glycosomes in trypanosomes. Structure-based drug design is therefore employed to develop antiparasitic drugs that target *Trypanosoma brucei*. This work follows a previous suggestion that targeting the import of enzymes into glycosomes is a viable target for drug development in these organisms. The aim is to demonstrate that potent and specific inhibitors of the Pex14/Pex5 interaction can be developed towards a novel therapy to treat trypasome-diseases.

Finally, an open question regarding the peroxisomal membrane protein insertion pathway is addressed. Enrichment of peroxisome membrane with proteins depends on the cytosolic transport receptor Pex19. This chaperone recognizes peroxisomal membrane proteins (PMPs) in the cytoplasm and targets them to the membrane. The domain for this interaction has been identified in the C-terminal region. However, the interface of this interaction is largely unknown. Moreover, given the hydrophobic nature of these interactions it is intriguing how Pex19 can differentiate PMPs from other hydrophobic proteins. Most importantly, Pex19 is farnesylated in vivo, but the role of the modification remains controversial and has not been characterized at the structural level. The aim of this project is to investigate the mode of interaction between Pex19 and PMPs. Therefore, both forms of

the protein, farnesylated and non-farnesylated, are used to study the role of the modification for PMP recognition and peroxisome function.

2. Materials and Methods

2.1. Bacterial strains

Table 1. Bacterial strains

E.coli Strain	Genotype
BL21 (DE3)	F– ompT gal dcm lon hsdSB(rB- mB-) (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])
DH5a	F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 80dlacZM15 (lacZYA-argF)U169, hsdR17(rK–mK+), λ–
XL1	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F′[::Tn10 proAB+ lacIq (lacZ)M15] hsdR17(rK- mK+)

2.2. Vectors

Table 2. Plasmids

Plasmid	Vector	Insert	Reference
FTase	pETM-11	His6-RAM1 RAM2	provided by Dr. S. Holton, EMBL Hamburg
Pex19 CTD	pETM-11	His6 -Pex19 (161-299)	Schueller et al. 2010
pTUML30	pETM-11	His6 -Pex19 (161-299) P273R/P274R	this study
pTUML31	pETM-11	His6 -Pex19 (161-299) P273F/P274F	this study
pTUML07	pETM-11	His6 - Pex14	this study
pTUML22	pETM-11	His6 -Trx- Pex14 (225-377)	this study
Pex5 TPR	pETM-11	His6 -GST- Pex5 (315-639)	this study
pTUML32	pETM-11	His6 -GST- Pex5 (315-639) Q491A/L494A	this study
tbPex14	pETM-11	His6 - Pex14 (1-84) T.brucei	provided by Prof. R. Erdmann, RUB
pTUML10	pETM-11	His6 - TrpPex14 (19-84) T.brucei	this study

2.3. Buffers

Lysogeny Broth LB (1L)	**Mineral Medium M9 (1L)**
Tryptone 10 g	Na_2HPO_4 6 g
Yeast Extract 5 g	KH_2PO_4 3 g
NaCl 10 g	NaCl 0.5 g
NaOH 1N to pH 7.2	$^{15}NH_4Cl$ 0.5 g
	glucose 4 g
	$MgSO_4$ (1M) 1 ml
	$CaCl_2$ (1M) 0.3 ml
	Biotin 1 mg
	Thiamin 1 mg
	100x trace elements stock solution 10ml
100x trace elements stock solution (1L)	**Chromatography buffer**
EDTA, pH: 7.5 5 ml	Tris.HCl 50 mM
$FeCl_3$ x 6 H20 8.3 mg	NaCl 50 mM
$ZnCl_2$ 84 mg	pH 8
$CuCl_2$ x 2 H20 13 mg	filtered / degassed
$CoCl_2$ x 6 H20 10 mg	
H_3BO_3 10 mg	
$MnCl_2$ x 6 H20 0.6 mg	
Lysis buffer	**High salt wash buffer**
Tris.HCl 50 mM	Tris.HCl 50 mM
NaCl 500 mM	NaCl 700 mM
Imidazole 1 mM	Imidazole 1 mM
pH 8	pH 8
TAE 10X (1L)	**Low salt wash buffer**
Trizma base 48.4 g	Tris.HCl 50 mM
Glacial acetic acid 11.44 mL	NaCl 50 mM

EDTA 0.5M pH 8.5 20 mL	pH 8
Laemmli buffer	**Elution buffer**
SDS 4%	Tris.HCl 50 mM
Glycerol 20 %	NaCl 50 mM
Tris.HCl 120 mM	Imidazole 250 mM
pH 6.8	pH 8
Dialysis buffer (TEV)	**NMR buffer**
Tris.HCl 50 mM	Sodium phosphate 20 mM
NaCl 150 mM	NaCl 50 mM
β-mercaptethanol 5 mM	pH 6.5
pH 8	

All culture media were supplemented with sterile filtered 50 mg/L kanamycin. All chemicals were dissolved in D_2O and sterile filtered for perdeuterated protein production.

2.4. Molecular biology

Molecular biology techniques were performed according to the following established manuals: A) "MOLECULAR CLONING" laboratory manual, J. Sambrook, E.F. Fritsch,T. Maniatis and B) "Current protocols in Molecualr biology", F.Ausubel, R. Brent, R.Kingston, D.Moore, J.Seidman, J.Smith, K.Struhl.

Commercially available kits and protocols for PCR and cloning were purchased from New England Biolabs and Promega. All constructs were sequenced by GATC-Biotech laboratories.

2.4.1. Quickchange mutagenesis

Quickchange mutagenesis kit was used throughout the study to introduce desired mutations. Ready-made reagents were purchased from Agilent Technologies. PCR reaction was set up with 50 ng of template plasmid, mutation-bearing primers in final concentration of 10 pmol/mL and Pfu high fidelity polymerase in 2.5 u/µL final units concentration. Following the PCR program as described in Quickchange kit, the mixture is incubated with DpnI enzyme. The enzyme digested the parental methylated DNA. As result, the non-methylated PCR product with the mutation was the only intact DNA molecule in the mixture. Subsequent to 1 h DpnI digestion at 37 ^{o}C, competent DH5a or XL1 E.coli cells were transformed with the mixture. The cells were plated in LB agar plates containing kanamycin. The colonies grown were sequenced for the mutation validation.

2.4.2. Gene cloning

In this study most constructs are heavily optimized for gene expression and protein stability. To achieve that genes have to be cloned in special expressing vector. DNA oligonucleotides (primers) are designed to complement the region of interest in the target gene. In addition, the oligonucleotides have flanking sequences consisting of restriction enzymes recognition sites, compatible with the vector enzymes. The result of a PCR reaction with these primers and the template DNA is a double stranded DNA molecule with the gene of interest coupled with flanking nucleotides. This product and the expressing vector are digested with the restriction enzymes to produce compatible single stranded overhangs in both molecules. Ligation reaction of both DNA molecules result a circular expression vector bearing the gene of interest and all necessary elements for plasmid replication and protein expression.

For DNA digestion, 10 µg of DNA is incubated at 37 ^{o}C with one or two restriction enzymes at final concentration of 1-2 u/µL, in final 50 µL reaction vol-

ume. The digestion is confirmed by agarose gel electrophoresis, in comparison with the untreated DNA. Mixing purified vector and target gene at various molar ratios ranging from 1:1 to 1:6 and ligase enzyme at 1 u/µL sets up ligation reaction, with vector being 50 ng in total 10 µL reaction volume. Ligation reaction is incubated overnight at 16 ^{o}C. Following day competent cells are transformed with whole ligation mixture and plated in LB agar plates containing antibiotic.

2.4.3. Agarose gel electrophoresis

Agarose gel electrophoresis is a DNA separation method based on the molecular weight. DNA samples are mobilized through the gel due to external electric field. Smaller molecules are migrating faster than big molecules, since their movement is hindered less by the porous nature of the gel matrix. The distinct bands consisted of different size DNA molecules are stained by ethidium bromide (EtBr) agent, which is ingredient of the gel. The result molecules are visible under UV light.

For the agarose gel preparation, 1% agar solution is made with TAE 1X buffer and heated to near-boil temperature in microwave oven. Following initial cooling of the mixture to below 56 ^{o}C, EtBr is added to final concentration of 0.5 mg/ml. Dye and glycerol is added to DNA samples prior loaded for handling convenience. After the desired migration of the dye in the gel is accomplished, the gel is subjected to UV radiation to inspect the DNA sample.

2.4.4. Competent cells transformation

Transformation in competent cells is an convenient way to import any gene/vector to the desired cell strain. Provided the genetic material bears all elements for replication and self gene expression, the cell replication and expression machineries can use this new material. Hence we can insert any desired vector in the cell and produce recombinantly target genes.

Approximately 100 µL of competent cells are mixed with 2-10 ng of the vector, which carries not only the gene of interest but also an antibiotic resistance gene. The cells are gradually thawed on ice for 30 min and subjected to 42 ^{o}C heat pulse for 45 sec, and returned back to ice for 2 more minutes. 700 µL of LB media is added, and they are incubated for 30 min at 37 ^{o}C. After the recovery, the cells are plated on a LB agar plate, which contains the suitable antibiotic for colony screening.

2.4.5. Protein expression in pET systems

Plasmids from the pET expression system are vectors which enable the recombinant production of any target protein in E.coli cells. It consists of the important elements of: lacI gene encoding the lac repressor, the T7 promoter – which is compatible only with T7 RNA polymerase and not with bacterial RNA polymerases, the f1 origin of replication and the desired antibiotic resistance gene. In the presence of T7 RNA polymerase, when lac suppressor is not suppressed, the transcription of the inserted gene proceeds in very quick manner. The short length of the T7 promoter enables this fast reaction, as long as T7 RNA polymerase exists. The production of the target protein is proportional related to the size of the transcribed mRNA. Within few hours the protein of interest will be the dominant molecule inside the cell. Since the existence of T7 RNA polymerase is of paramount importance for the pathway, it is crucial the cell strains to contain the specific gene, usually referred as (λDE3). The polymerase is also under the control of lac transcription elements. For purification purposes, all pET vectors used in this study express target proteins fused with solubility enhancement tags, like Thioredoxin (Trx), and poly-histidine tail in the N-terminal end of the polypeptide chain.

2.4.6. Protein expression and isotope labeling

All proteins studied were expressed and purified from pET vectors cells with the proceeding protocols. Expressing strain of E.coli BL21(DE3) was transformed with the plasmid of interest. A single transformed colony was suspended in culture media, containing the antibiotic of choice, for overnight growth. The dense preculture was diluted twenty times in larger volume medium and incubated at 37 ^{o}C till OD_{600} reaches 0.6-0.8. With the desired density and amount of cells, cells were induced with 200 µM IPTG at 20 ^{o}C for 16 hours.

For stereospecific assignments media was supplemented with 10% U-[^{13}C]-D-glucose (Neri et al., 1989). For Leu, Val and Ile methyl labeled Pex5 TPR or Pex19 CTD samples, cells were grown in perdeuterated medium containing U-[^{2}H]-D-glucose at 37 ^{o}C. At OD_{600} =0.5, 100 mg of respective precursors were added in the culture. The precursors used were α-ketobutyric acid $^{13}C_4$,3,3-d_2 for (^{1}H-$^{1}\delta$ methyl)-isoleucine and 2-keto 3-methyl ^{13}C-butyric 4-^{13}C, 3d for (^{1}H-$^{1}\delta$ methyl) -leucine and (^{1}H-$^{1}\gamma$ methyl)-valine labeling. One hour after the addition of the precursors, expression was induced with 200 µM IPTG at 20 ^{o}C for 16 hours. Finally, cells were harvested by centrifugation (4000 g, 4 ^{o}C, 15 min).

2.4.7. Affinity chromatography purification

As all proteins were fused N-terminally to six-histidine tag, it was possible to exploit the robust affinity chromatography purification protocol using nickel-bearing beads. The special resin is widely used for recombinant protein purification and consists of agarose beads coated with nitrilotriacetic acid (NTA). In neutral pH, the acid has three deprotonated carbonyls, which can chelate bivalent ions like nickel. Since immobilized nickel has still two unoccupied ionic bonds, it possesses the ability to chelate amino acid aromatic side chains, histidines in particular. The stationary phase – nickel beads – binds selectively to the six-histidine epitope, while the rest of cell extract is not reactive, thus resulting separation. Pro-

tein then can be recovered from the beads, by washing the column with a higher affinity binding molecule for nickel, imidazole.

The cell paste was resuspended in lysis buffer supplemented with protease inhibitor PMSF at final concentration of 2 mM. Cell lysis was achieved via sonication at 4 ^{o}C. Following step included separation of cell debris and intact cells membranes by centrifugation (20 min, 4 ^{o}C, 7000 g). The supernatant contains all soluble proteins together with the protein of interest. This is loaded to a Ni-NTA chromatography column, previously equilibrated with 10 column volumes of lysis buffer. With constant flow rate 1 ml/min the column was washed with 5 volumes of each high and low salt buffers. Protein was eluted with 5 column volumes of elution buffer. After purity and yield checked in SDS gel, the sample was dialyzed overnight, in the presence of 5 mg/mL TEV protease, against dialysis buffer to remove the imidazole. Separation of cleaved protein and tag was achieved with additional nickel column purification. Purified proteins were tested for freeze-thaw stability, and showed that they can withstand freezing. Hence all proteins were stored at -20 ^{o}C in dialysis buffer or were furthermore purified using size exclusion chromatography. For crystallography and NMR sample preparation tag removal was achieved with TEV protease cleavage while dialysis in dialysis buffer. In these cases, the dialysis bag contained also 1 mg/mL TEV protease and 1 mM DTT.

2.4.8. Size exclusion chromatography purification

To reach >95% purity level of protein preparations in addition to affinity chromatography, size exclusion chromatography (SEC) was applied. As another chromatographic technique, molecules are separated based on their hydrodynamic radius in a column containing the stationary phase. The latest consists of porous material, which traps transiently smaller molecules, while bigger particles pass through unaffected. At the exit of the column a UV detector calibrated

at 280 nm detects and records the proteins traces. An automated machine collects the output in many fractions. The resulting chromatogram comprises larger proteins in the beginning, since they are not interacting with the matrix, and gradually smaller molecules in the end. For large scale protein preparations long and wide columns were facilitated to achieve maximum resolution and separation.

Preparation of sample for size exclusion chromatography included concentration up to 0.5-1 mL. After the sample was loaded in the loop, it was injected in a previously equilibrated column with SEC buffer and the run developed with 1 mL/min flow rate. The eluted fractions were collected and pooled together upon chromatogram validation.

2.5. Protein detection and evaluations

2.5.1. SDS-PAGE gel electrophoresis under denaturing conditions

Proteins can be sorted by their molecular weight in a polyacrylamide gel with the application of current. A multi-protein sample is mixed with laemmli buffer, which contains SDS ionic detergent in such a concentration to destroy all non-covalent interactions that keep proteins folded. To ensure complete denaturation the buffer consists also of β-mercaptethanol to reduce any disulfide bonds present. SDS-bound proteins are denatured and largely negatively charged, which enables them to migrate through the gel when electric field is applied. The complex SDS-proteins have roughly the molecular weight of the protein; therefore the separation by protein molecular weight is possible.

2.5.2. Coomasie blue stain

Detection of the proteins, which are embedded inside a SDS-PAGE gel, was accomplished via staining of the gel with Coomasie Briliant Blue (R-250). The

dye stains with blue color protein bands with protein content higher than 0.1 mg.

2.5.3. Protein concentration determination

Protein concentrations throughout the study were determined using Nanodrop 2000 (ThermoScientific). The principle of this method is the UV light absorption from aromatic amino acids. Specifically, aromatic rings of mainly Tryptophan, secondarily Phenylalanine and Histidine exhibit absorbance peak at 280 nm. Since the structure elements of the polypeptide chain affect the absorbance, factors like pH and salt concentration may affect the absorbance spectrum. Concentration measurement with Nanodrop 2000 requires only 2-3 μL of buffer, as blank, and same amount of protein sample. Division of the absolute 280 nm absorbance by the extinction coefficient factor calculates the final concentration.

2.6. In vitro assays

2.6.1. Fluorescence polarization (FP)

Fluorescence polarization is an affinity measurement method, which can be used in high throughput schemes. Fluorescein isothiocyanate (FITC) is a fluorescent molecule, which absorbs at 495 nm and emits at 519 nm. When FITC is fused covalently to a ligand it tumbles fast in solution, thus the polarization of emitted light is reduced. In contrast, when the small molecule is bound to a much larger receptor, it adopts the kinetics of a large molecule. The mixture of FITC-ligand and receptor is illuminated with 495 nm wavelength polarized light and detector records the polarity of the emitted signal. Since this technique requires ligand chemically fused with tag, the possibility of the tag itself binding to the protein should be excluded by additional controls. Titration with many different receptor concentrations reveals the saturation point, where no more polarization

is gained upon complex formation.

Peptides, fused N-terminally with FITC, were purchased from Peptide Specialty Laboratories GmbH as lyophilized powder. The powder was initially dissolved in 1 mL water producing high concentrated peptide, at concentrations 5-10 mM depending on the delivered amount. Since the peptides were purified with trisfluoroacetic acid (TFA), which might alter severely the pH of any experimental conditions, overnight dialysis against water removed all residual TFA. For the determination of each binding curve a 12 point titration was performed with constant concentration of 20 nM labeled peptide and by increasing concentrations of receptor. Starting concentration of Pex5 TPR variants was 3 mM with 12 serial dilutions reaching 1.5 nM. FITC-WQI peptide concentration was 20 nM. Reaction mixtures was then transferred into 96-well Optiplate and measured as triplicates in an EnVision plate reader. Following equations were used for the affinity calculation.

$$K_d = \frac{[P][L]}{[PL]} \qquad \text{(Eq. 2.1)}$$

$$[PL] = \frac{K_d + [L_T] + [P_T] - \sqrt{(K_d + [L_T] + [P_T])^2 + 4[P_T][L_T]}}{2} \qquad \text{(Eq. 2.2)}$$

with [P], [L] and [PL] the free protein, free ligand and complex concentrations, [L_T], [P_T] total ligand and protein concentrations.

2.6.2. AlphaScreen

Amplified Luminescence Proximity Homogeneous Assay (AlphaScreen) is a luminescence method to measure biological interactions and estimate affinities. AlphaKit is the commercially available kit from PerkinElmer. The technique consists of two type of beads, donors and acceptors. Donor beads are pthalocya-

nine molecules, which upon excitation with 680 nm light release singlet oxygen. If acceptor beads are closer than 200 nm in space, through a cascade of reactions, 615 nm light is emitted from them. In biological applications, and in the case of this study, donor and acceptor beads are conjugated with streptavidin and anti-histidine epitope respectively. For the assay to function, tested ligands were biotinylated and receptors His-tagged. The high affinities of streptavidin-biotin and epitope-His-tag ensure that if ligand-receptor complex is formed, the beads will be close enough to emit and detect signal output. Combination of a labeled ligand with a non-labeled one, results a competition assay. Inhibition constant (Ki) can be calculated by plotting relative fluorescence versus concentration of non-labeled ligand.

The concentration of the beads used in all experiments was 2 µg/ml. Initially the Pex14-labelled WF peptide complex was formed in 20 mL total low salt buffer by mixing in 1:1 ratio the two components, at 100 µM concentration each. In a 96-well plate 40 µL of DMSO was added to the first column, while the rest of the wells were filled with 25 µL. In the first well of each row 10 µL of 50 mM ligand stock solution were supplemented to complete 50 µL total mixture volume. Serial dilutions were carried out by transferring 25 µL from each well to the next in every row. To every well, which contained various ligand concentrations, 197.5 µL of preformed complex were added and incubated for 1 hour at room temperature in dark conditions. Each well reaction was divided into 3 to fill a 384-well white Optiplate and measured as triplicates in an EnVision plate reader. The resulting curves were analyzed with software Prism7.

2.6.3. In vitro farnesylation

In vitro farnesylation of Pex19 CTD variants was achieved enzymatically by farnesyltransferase. 50 µM of purified protein was mixed with the substrate farnesyl pyrophosphate to a ratio 1:1.2. The enzymatic reaction was set up with 1

µM enzyme, 5 mM $MgCl_2$, 10 µM $ZnCl_2$ and 3.5 mM β-mercaptethanol in 10 mL reaction volume of buffer containing 50 mM Tris.HCl at pH 8. The samples were incubated for 1 hour at 37 ^{o}C. To separate the His-tagged enzyme, samples were loaded on affinity chromatography column. The flow-through was enriched with farnesylated Pex19 CTD and overnight dialysis against NMR buffer finalized the preparation.

2.6.4. Crystallization

At high concentrations and under specific conditions molecules adopt higher ordered structures, the crystals. X-ray diffraction pattern of such crystals reveals the distribution of electrons inside the crystal unit, thus enabling the 3D reconstruction of the electron density map and the structure of the molecule. Protein samples are subjected to precipitation under controlled environment, testing many different buffer conditions.

For determination of complex structures, Pex14 was mixed with 10-fold excess of ligand and co-concentrated up to 4 mM. Since the solubility of most ligands lies well below this number, occurred precipitation was removed by centrifugation. Pex14 complex with compound 2 was achieved in 0.1 M SPG buffer pH 9, 25% (w/v) PEG 1500 in two weeks, while with compound 3 in 0.1 M SPG buffer pH 6, 25% (w/v) PEG 1500 in three weeks. The crystals were measured at ESRF, Grenoble at ID23-2 beamline. Collected data were processed by XDS and XSCALE software (Kabsch, 2010). Molecular replacement was used to solve the structures, with rat Pex14 (PDB: 3FF5) structure as molecular replacement probe (Su et al., 2009). Models and ligands were created using COOT program (Emsley et al., 2010) followed by Refmac5 refinement (Vagin et al., 2004).

2.7. NMR methods

All NMR experiments in this study were performed at 298 K in NMR buffer, unless otherwise stated.

2.7.1. Assignments

Backbone assignments for hsPex14 CTD, hsPex5 TPR and tbPex14 NTD protein was achieved with standard triple resonance experiments HNCA, HNCACB, HNCO, HNCACO, CBCA(CO)NH (Sattler et al., 1999). Due to signal overlap in the case of 35 kDa TPR domain of hsPex5, all experiments were TROSY version and additional HCC(CO)NH experiment was recorded. Spectra of tbPex14 were obtained from Bruker Avance 500 MHz spectrometer, while for TPR and hsPex14 CTD from Bruker Avance 900 MHz. ^{1}H, ^{13}C HSQC and (H)CC(CO)NH-TOCSY recorded spectra assisted for the stereospecific assignments of tbPex14 NTD. Fully assignments of unlabeled peptides were extracted from natural abundance ^{1}H,^{1}H TOCSY, ^{1}H,^{1}H NOESY, ^{1}H-^{13}C HSQC 2D spectra recorded on a Bruker Avance III 500 MHz spectrometer. Sample concentrations in all cases were ranging 0.8-1 mM.

Methyl Ile, Leu and Val assignments for hsPex5 TPR domain were obtained from ^{13}C- and ^{15}N- edited simultaneous NOESY-HSQC spectra in combination with ^{1}H-^{13}C ctHMQC, ^{1}H-^{15}N TROSY-HSQC and HCC-FLOPSY (Kadkhodaie et al., 1991). The NOESY mixing time was 100 ms, while for the FLOPSY 22 ms.

NMRpipe software (Delaglio et al., 1995) was used for the processing of all spectra, while analysis and assignments were done on Sparky (Goddard & Kneller, 1996). Backbone assignment procedure was assisted by semiautomatic assignment program MARS (Jung & Zweckstetter, 2004).

2.7.2. NOESY experiments

^{13}C- and ^{15}N- edited NOESY-HSQC spectra were recorded with 80 ms mixing time on uniformly labeled tbPex14 NTD on Bruker Avance 900 MHz spectrometer equipped with TXI cryo-probe head.

Intermolecular NOEs between Pex19 CTD$_{farn}$ and ALDP peptide were detected from a ω1-filtered/ω2-edited NOESY-HSQC experiment with 100 ms. The spectrum was recorded on Bruker Avance III 500 MHz spectrometer with TXl cryo-probe head. The sample composition was 300 µM protein and 1mM ALDP peptide.

Intermolecular NOEs between hsPex5 TPR and WQI peptide were detected from ω1-filtered/ω2-edited NOESY-HSQC experiment with 110 ms mixing time. The spectrum was recorded on Bruker Avance 900 MHz spectrometer with TXl cryo-probe head. Sample consisted of 400 µM HsPex5 TPR and 4 mM of WQI peptide.

Resonances of the bound conformation of WQI peptide were obtained from 2D ^{1}H,^{1}H trNOESY (Post, 2003) spectra with 250 ms mixing time. The experiment was recorded on Bruker Avance III 500 MHz spectrometer on sample containing 50 µM unlabeled hsPex5 TPR and 2 mM unlabeled WQI peptide.

2.7.3. Chemical shift perturbations

Chemical shift perturbations obtained from ^{1}H, ^{15}N HSQC spectra were calculated from the following formula:

$$\Delta\delta(NH) = \sqrt{(\Delta\delta^{1}H)^{2} - (\frac{\Delta\delta^{1}H)}{6})^{2}} \qquad \text{(Eq. 2.3)}$$

with $\Delta\delta$ = difference of chemical shift in different spectra.

2.7.4. Secondary structure prediction

Secondary structure prediction for ALDP peptide was based on ^{13}Ca and ^{13}Cb assigned chemical shifts. The difference between the observed values and the average chemical shifts of random coil was calculated according to:

$$\Delta\delta(^{13}C_\alpha - ^{13}C_\beta) = (\delta^{13}C_\alpha(obs)) - (\delta^{13}C_\alpha(rc)) - [(\delta^{13}C_\beta(obs)) - (\delta^{13}C_\beta(rc))] \quad \text{(Eq. 2.4)}$$

with obs and rc being the experimental and random coil chemical shift values respectively.

2.7.5. Titrations

All peptides where purchased from Peptide Specialty Laboratories GmbH as lyophilized powder. They were all solubilized to 1 mL ddH2O, resulting in milimolar solution, depending on the molecular weight (MW) and the amount delivered. This solution was dialyzed overnight against NMR buffer using 300 Da cut off dialysis bag at 4 ^{o}C. The final concentration of this stock solution for each peptide was calculated based on the peptide miligrams, MW and the final volume of each solution.

ALDP and Pex13 peptides final stock concentration was 3.1 mM and 2.6 mM respectively. For ^{1}H,^{15}N and ^{1}H,^{13}C HSQC titration with those peptides to Pex19 CTD and Pex19 CTD$_{farn}$ initial protein concentration was 75 µM with increasing concentration of peptides reaching 1:10 ratio.

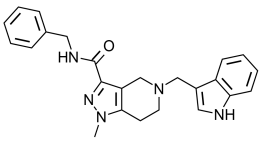
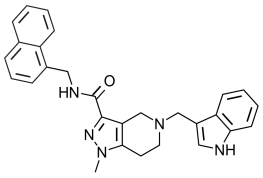
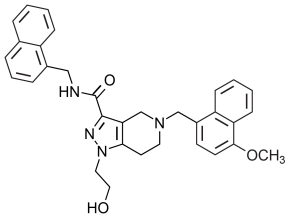
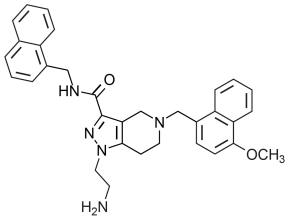
For both hsPex14 WQI peptides the final stock concentration was 12 mM. For ^{1}H,^{15}N TROSY-HSQC and ^{1}H,^{13}C ctHSQC titration with those peptides to U-[^{2}H], U-[^{15}N] (^{1}H-$^1\delta$ methyl)-ile, (^{1}H-$^1\delta$ methyl)-leu (^{1}H-$^1\gamma$ methyl)- val labeled HsPex5 TPR initial protein concentration was 300 µM with increasing concentration of peptides reaching 1:10 ratio.

Table 4. Peptide sequences

Peptide from	Sequence	Length (aa)
ALDP	GMNRVFLQRLL	11
Pex13	ALVRTIRYLYRRLQR	15
hsPex14 WQI (long)	SAPKIPSWQIPVKSPSPS	18
hsPex14 WQI	KIPSWQIP	8

2.7.6. Pex14-Pex5 inhibitors

Table 5. Pex14-Pex5 inhibitors

Inhibitor	Formula	MW (Da)
1		399
2		450
3		504
4		519

The tbPex14-tbPex5 inhibitors were either purchased from Chemibridge or in-house chemically synthesized. Stock solutions were made in all cases by dissolving the needed amount in DMSO to 50 mM final concentration. ^{1}H,^{15}N HSQC titrations were performed with 100 µM ^{15}N labeled protein and increasing concentration of compounds up to 1:10 ratio.

2.7.7. Residual dipolar couplings

Residual dipolar couplings were measured on a Bruker Avance 800 MHz spectrometer. $^{1}H^{N}$-^{15}N J-couplings were extracted from difference in ^{15}N frequency between the $^{1}H^{N}$-^{15}N HSQC and the ^{1}H-^{15}N TROSY-HSQC recorded at the same temperature 298 K (Bax et al., 2001). For the ^{15}N-^{13}CO and ^{13}Ca-^{13}CO J-couplings two 3D TROSY-HNCO experiments were measured (Bax et al., 2001). For the isotropic reference measurements 450 µM perdeuterated ^{15}N labeled hsPex5 TPR was mixed with 5 mM unlabeled WQI peptide, while for the RDCs 15mg/ml Pf1 phages (ASLAbiotech) were supplemented. Analysis of all spectra was achieved with Sparky and MODULE (Dosset et al., 2001).

2.7.8. Structure calculation

CYANA 3.0 was used to calculate structure of tbPex14 NTD and automatic NOE assignment (Guntert, 2009). In addition, torsion angles restrains by TALOS+ (Shen et al., 2009) were implemented. The lowest energy structure was water refined by ARIA2 (Nilges et al., 1997).

The bound conformation of WQI peptide onto hsPex5 TPR was calculated based on the NOEs (2.10.2) that were assigned by CYANA 3.0. Both bound conformations of the protein and the peptide were docked using the intermolecular NOEs obtained (2.10.2) and CSPs. HsPex5 TPR residues with CSP $>$ 0.05 ppm (Lys472, Glu473, Leu476, Val479, Glu480, Asp482, Thr483, Ser485, Ile486, Gln491, Leu494, G495, Phe498, Lys506, Asp509, Ala513, Ser516,) and the Ile238, Pro239,

Ser240, Trp241, Gln242, Ile243 in the WQI peptide were used as active residues and those within 6.5 Å were treated as passive residues. The unambiguous intermolecular NOEs included are shown in Table 6. The solvated docking was performed on HADDOCK web server (Dominguez et al., 2003). In iterations 0, 1 and during water refinement 2000, 400 and 400 structures were calculated, respectively. Minimum cluster size was set to 10.

Table 6. Unambiguous NOE restrains for Pex5 TPR- WQI docking

Pex5 TPR	Pex14 (WQI pept)
Ile486 ^{1}H-δ1 methyl	Ile238 or Ile243 HG2#
Ile486 ^{1}H-δ1 methyl	Ile238 or Ile243 HG1#
Ile486 ^{1}H-δ1 methyl	Ile238 or Pro244 HA
Ile486 ^{1}H-δ1 methyl	Pro239 HG2
Ile486 ^{1}H-δ1 methyl	Pro239 HG3
Val479 ^{1}H-γ1 methyl	Ile238 or Ile243 HG2# or Ile238 or Ile243 HD1#
Val479 ^{1}H-γ1 methyl	Pro239 HG# or Pro244 HB# or Ile238 HB or Ile243 HB
Leu476 ^{1}H-δ1 methyl	Ile238 HB or Ile243 HB
Leu476 ^{1}H-δ1 methyl	Trp241 all aromatic protons

Based on the NMR titrations and NOESY experiments both ambiguous and unambiguous restrains were obtained for the complex Pex19 CTD$_{farn}$ – ALDP peptide. The NMR structure of farnesylated Pex19 CTD, and the helical ALDP peptide (sequence: Asn68-Arg-Val-Phe71-Leu-Gln-Arg-Leu-Leu76) were used as first and second molecule respectively. Pex19 residues with CSP $>$ 0.05 ppm (Leu172, Met175, Ser177, Asn181, Leu182, Leu183, Ser184, Lys185, Y189, L192) and the Phe71 in the ALDP peptide were used as active residues and those within 6.5 Å were treated as passive residues. In iterations 0, 1 and during water refinement 2000, 400 and 400 structures were calculated, respectively. The minimum cluster size was set to 10. Unambiguous restraints were derived from ω1-filtered/ω2-

edited NOESY experiment and involved methyl protons of Pex19 Leu182 or Leu183 and Pro273 or Pro274 to Phe71 aromatic protons.

3. Results

3.1. Novel interactions between Pex14 and Pex5 in human

3.1.1. Characterization of Pex14 C-terminal

Human Pex14 is a 41 kDa membrane protein and as such it was never successfully isolated. Therefore I designed various constructs lacking the lipophilic putative transmembrane region of the protein. Pex14 CTD consists of the complete C-terminal part beyond the transmembrane region 138-377, while Pex14cc and Pex14 CTDΔcc comprise the coiled-coil and C-terminal domain without the coiled-coil region respectively (Fig. 10A). All three constructs were successfully cloned and purified as soluble proteins, yet the structural studies were not possible for both constructs containing the coiled-coil domain. As expected, the coiled-coil region seems to promote polymerization of the protein resulting in either large soluble oligomers (Pex14cc) or less soluble aggregates (Pex14 CTD). This behavior was observed in gel filtration chromatograms, where both the 9 kDa construct of Pex14cc and the Pex14 CTD were eluted in volumes corresponding to much larger molecules (Fig. 10). In fact, the C-terminal region was found primarily in the void volume of the column, indicating aggregation. In contrast, Pex14 CTDΔcc elutes as much smaller species from the column and proved to be less oligomerization prone. Solubility, size and yield of the CTDΔcc made this the construct of choice.

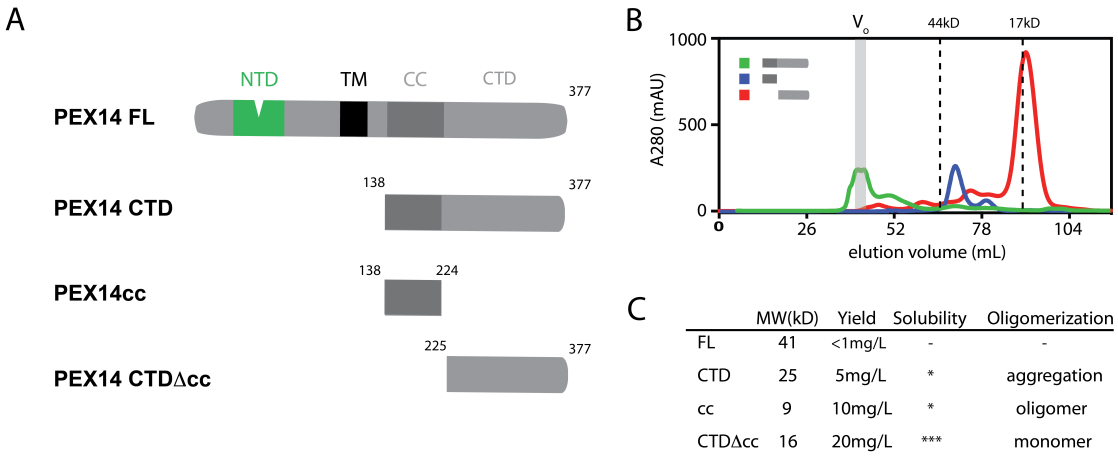


Figure 10. Protein profiles of various C-terminal Pex14 variants. A) Different Pex14 construct with their boundaries. NTD N-terminal domain, TM Transmembrane, CC coiled-coil, CTD C-terminal domain. B) Overlay of UV absorbance from size exclusion chromatographs of different constructs. The void volume of the column (Vo), as well as the elution volume of standard proteins with their respective molecular weights are marked on top of the graph. C) Table of molecular weight, yield of expression (mg/L LB culture), solubility and oligomeric state of different Pex14 constructs. * 200-300 µM, *** >1 mM solubility limit respectively.

3.1.2. Pex14 CTDΔcc is mainly unstructured

Pex14 CTDΔcc was recombinantly expressed and purified as described in 2.4.6, 2.4.7 and 2.4.8. A uniformly ^{15}N and ^{13}C labeled sample was produced and standard backbone assignment NMR experiments were recorded and analyzed. The peak dispersion in ^{1}H,^{15}N HSQC spectra was found to be between 8 and 8.5 ppm, indicates that the protein is mostly disordered. The poor dispersion leads to severe signal overlap in multidimensional experiments (Fig. 11A). Additionally, the primary structure of the protein comprises many repeat sequences and this results in signal overlap. Backbone assignment was achieved up to 72%.

Backbone assignments is a prerequisite for structural analysis of the do-

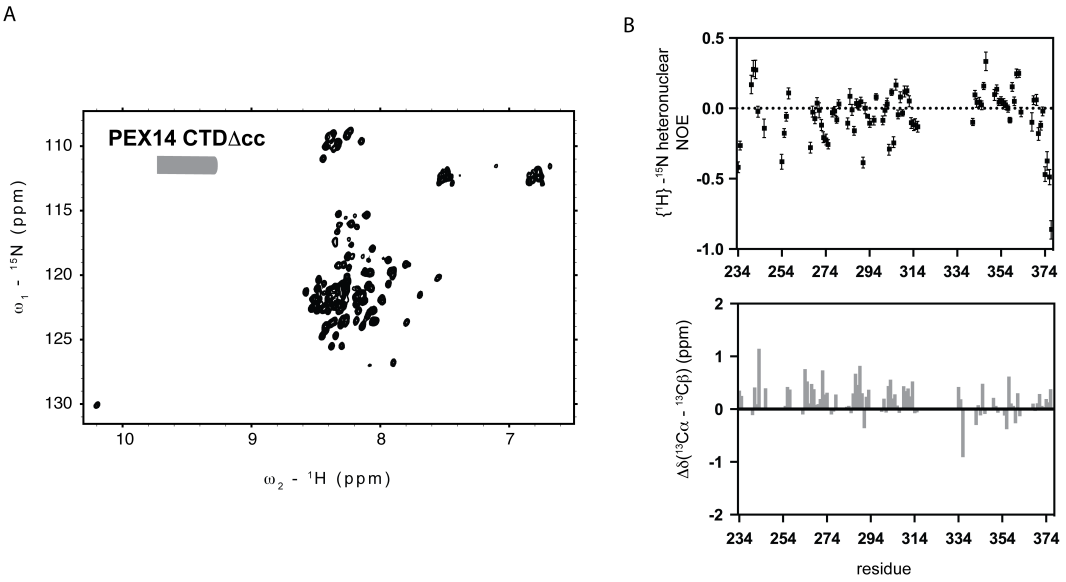


Figure 11. Pex14 CTDΔcc NMR analysis. A) ^{1}H,^{15}N HSQC spectrum of 1 mM protein in NMR buffer. B) Heteronuclear NOE plot and backbone $\Delta\delta(^{13}C_{\alpha} - ^{13}C_{\beta})$ is plotted in top and bottom graph respectively. Gaps in the graphs are due to signal overlap or missing backbone assignment.

main. In agreement with initial observation from ^{1}H,^{15}N HSQC spectrum, both heteronuclear NOE experiment and $\Delta\delta(^{13}C_{\alpha} - ^{13}C_{\beta})$ indicate the absence of any structure (Fig. 11B). Heteronuclear NOEs, which report on the dynamics of the protein, are found to be on the order of zero, indicating a high degree of local structural fluctuations (Fig. 11C).

3.1.3. Pex14 CTDΔcc binds Pex5 TPR in vitro

To investigate whether Pex14 CTDΔcc selectively interacts with Pex5, we compared ^{1}H,^{15}N spectra of ^{15}N labeled Pex14 CTDΔcc protein in the presence and absence of Pex5 full length unlabeled protein. Since the second molecule is unlabeled, its presence is not detectable by this 2D correlation NMR experiment.

We observed that when full length Pex5 was added, certain signals in the spectrum disappeared (Fig. 12A). This line broadening of peaks indicates that their chemical environment and dynamic are altered upon addition of Pex5. Such behavior could be explained by an increase of the molecular weight of Pex14 CTDΔcc in solution if it tumbles in complex with Pex5.

To further validate this finding and locate the interaction interface we repeated the same experiment using different fragments of Pex5. Two different constructs, composed of either the N-terminal disordered region or the C-terminal TPR domain, were mixed separately with Pex14 CTDΔcc. Strikingly, while in the case of Pex5 NTD no detectable changes in the spectrum of Pex14 were observed, the presence of the TPR domain caused the same peaks to disappear as upon addition of full length Pex5 (Fig. 12B,C). These experiments show that there is a second interaction interface between Pex14 and Pex5 involving the C-terminal parts of both proteins. Since backbone assignments of the construct was available, we identified the residues, which are affected upon addition of the TPR domain. Interestingly, a stretch of residues containing a Trp-Gln-Ile sequence (WQI) in the Pex14 CTDΔcc residues was identified (Fig. 12D).

3.1.4. Backbone assignment of Pex5 TPR

Mapping of the TPR-WQI peptide interface on Pex5 was achieved by assigning backbone amides of the TPR domain. As anticipated for a 35 kDa protein, deuteration was mandatory to proceed with backbone assignment (Fig. 13). Experiments described in 2.4.6 were recorded employing a ^{2}H,^{15}N and ^{13}C uniformly labeled sample. Despite the high alpha helical content of the domain, 73% of the backbones amides were unambiguously assigned using deuteration and TROSY type backbone assignment experiments (Fig. 14).

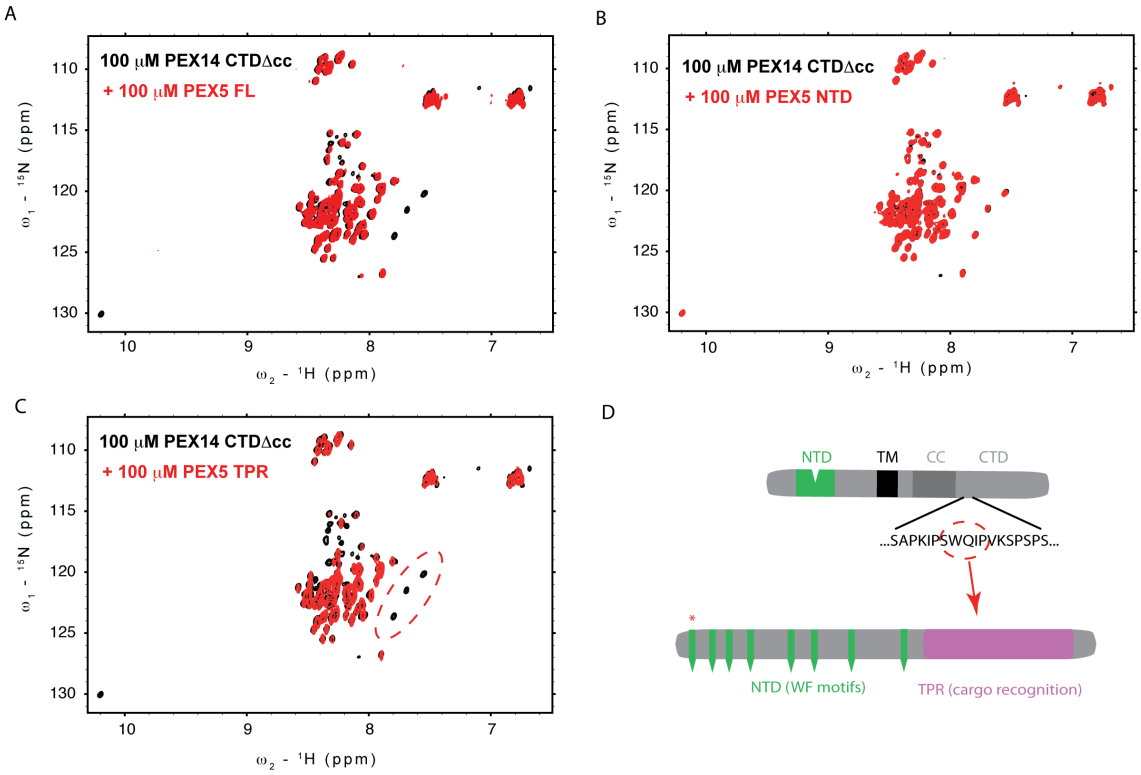


Figure 12. NMR titrations of Pex14 CTDΔcc with Pex5 protein. A) ^{1}H,^{15}N HSQC spectrum of 100 µM Pex14 CTDΔcc (black) superimposed with a spectrum of Pex14 supplemented with an equimolar amount of full length Pex5 (red). (B and C) Overlay of the ^{1}H,^{15}N HSQC spectra of 100 µM Pex14 CTDΔcc (black) and with an equimolar amount of Pex5 NTD (B) and TPR (C). The resonances mostly affected upon TPR addition are marked with a circle. D) Schematic overview of the Pex14 and Pex5 C-terminal interaction. The amino acids which correspond to the most affected residues highlighted in (C) are shown.

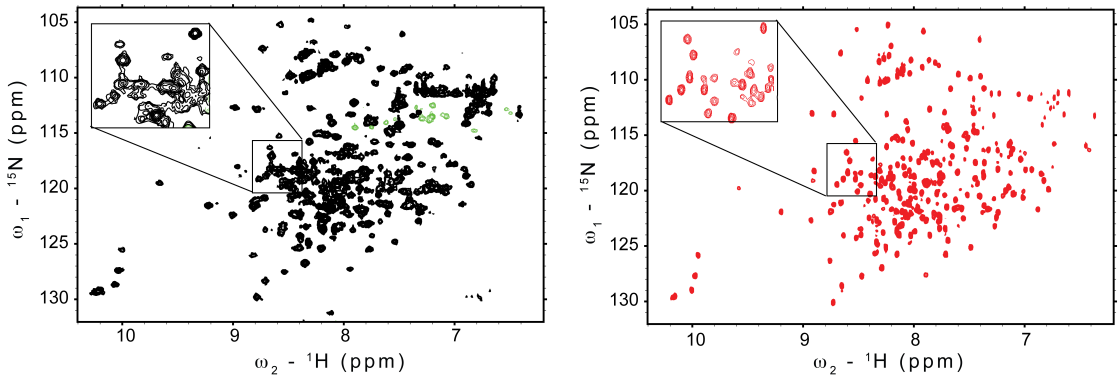


Figure 13. Deuteration of Pex5 TPR. Comparison of a ^{1}H,^{15}N HSQC spectrum of protonated (left) and a ^{1}H,^{15}N TROSY-HSQC of deuterated Pex5 TPR domain (right).

3.1.5. Mapping of the Pex5/Pex14 interaction

Unlabeled peptide (sequence: SAPKIPSWQIPVKSPSPS) was used in NMR titrations with Pex5 TPR (Fig. 15A). In these experiments only the Pex5 amides were ^{1}H,^{15}N labeled, while the rest of the protein was deuterated. The titrations confirm that this peptide indeed interacts with the Pex5 TPR domain. The majority of affected peaks show fast chemical exchange indicating a relatively weak interaction (Fig. 15B,C). Calculating the chemical shift perturbations from these titrations revealed that the region of Pex5 TPR is most affected upon addition of the WQI (Fig. 16A). These residues are localized exclusively on TPR 4 and 5 of the domain, in specific residues 418 to 517. Notably, this region is not related to the PTS1 binding site of the Pex5 TPR (Fig. 16B,C).

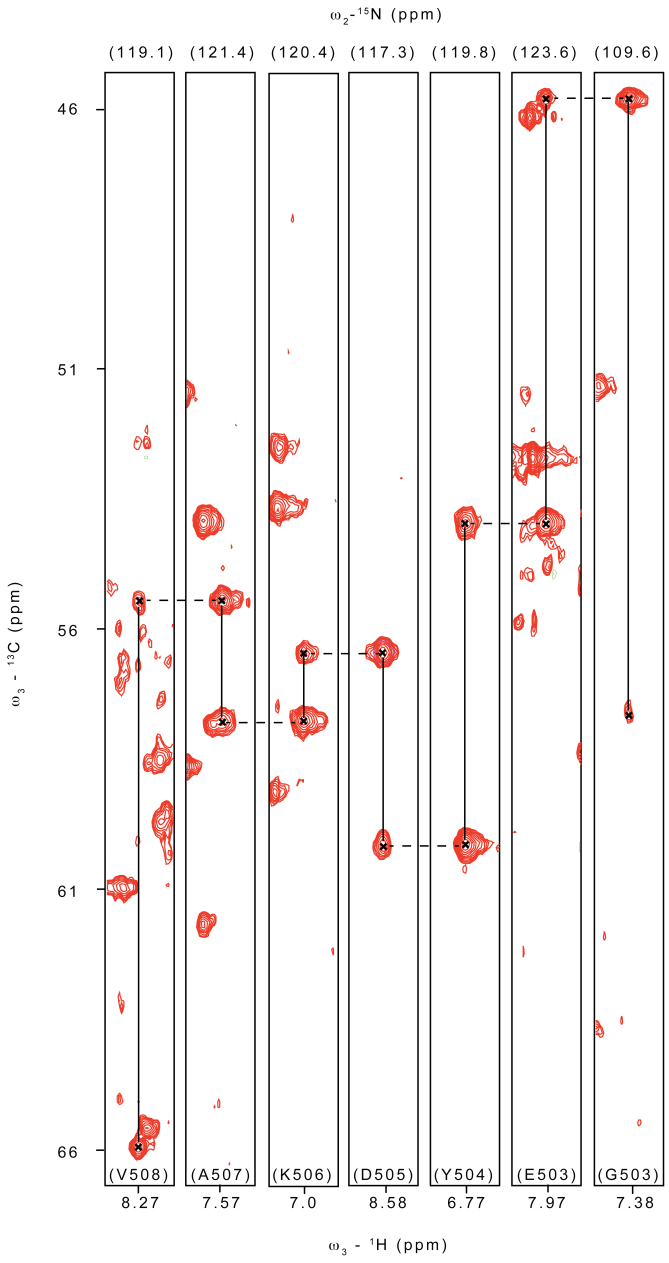


Figure 14. TPR backbone assignment. Slices of a 3D TROSY-HNCA spectrum showing C_α and $C_{\alpha-1}$ carbon resonances connecting residues V508-A507-K506-D505-Y504-E503-G502. In each strip the stronger peak corresponds to the intra residual connectivity to C_α, while the weaker peak to the inter residual $C_{\alpha-1}$. The sequential walk is shown with solid and dashed lines.

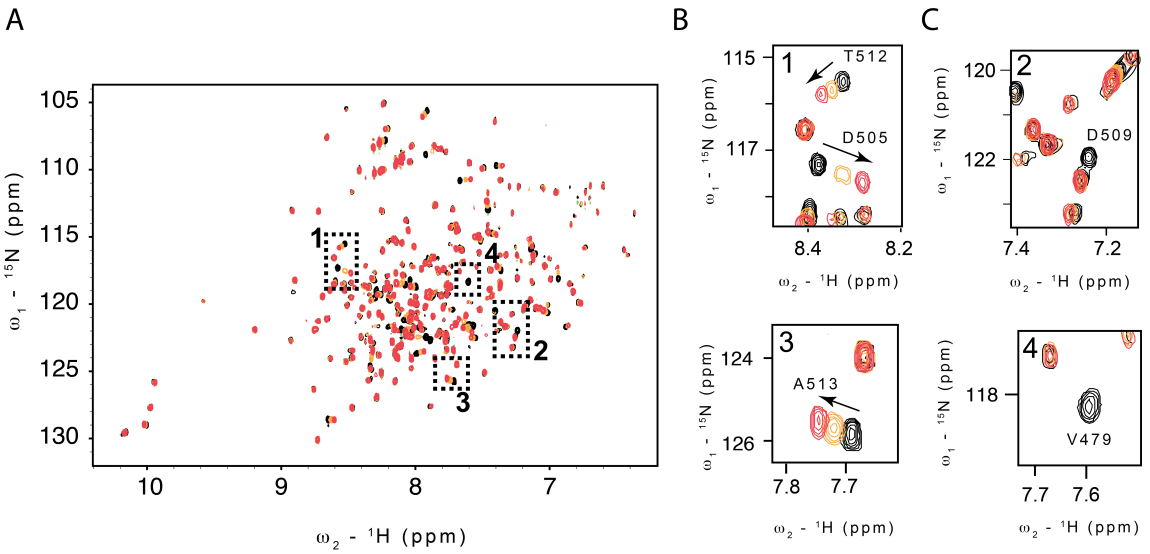


Figure 15. NMR titration of TPR with the WQI peptide. A) Overlay of 150 µM TPR spectra (black) with 150 µM TPR in the presence of 300 µM (orange) and 1500 µM (red) WQI respectively. Numbered regions in rectangles represent areas shown in B (fast exchange) and C (slow exchange).

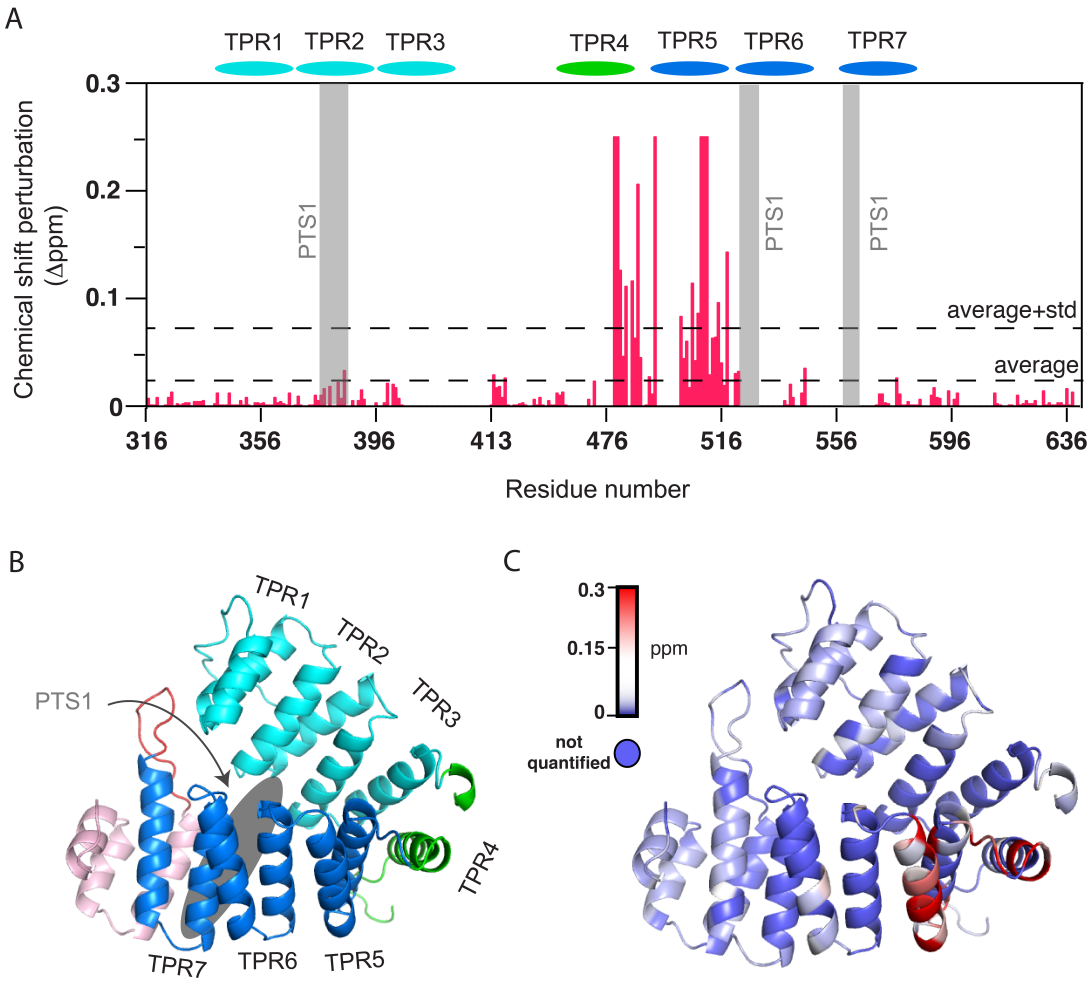


Figure 16. Mapping of the TPR-WQI peptide interaction. A) Chemical shift perturbations of Pex5 TPR amides upon addition of WQI peptide. Line broadened residues beyond detection have set to maximum perturbation of 0.25. TPR repeats in the sequence of the protein are shown on top of the graph. PTS1 binding regions are highlighted with grey boxes. B) TPR repeats as revealed from crystal structure of Pex5 TPR domain (PDB : 2C0M) (Stanley et al., 2006). The binding site of PTS1, is indicated as a grey oval shape. C) Cartoon representation of the Pex5 TPR domain colored depending on the magnitude of chemical shift perturbation. Regions colored in red correspond to most affected regions upon WQI peptide binding, while light blue least affected. Residues with non-quantified perturbations are colored with blue.

3.1.6. Conformation of the WQI peptide bound to Pex5 TPR

The Pex14 C-terminal region 234-251, comprising the WQI element, was cloned into a pETM11 expressing vector. After removal of the histidine tag and purification of the peptide, uniformly ^{15}N and ^{13}C labeled peptide was obtained (Fig. 17A). Using standard backbone assignment methods, as described in 2.1.5, all peptide resonances were assigned. To determine the fold of the peptide bound to Pex5 TPR, we employed transfer NOEs. The nature of the complex, 36 kDa size and low affinity, made a transfer NOESY experiment ideal for this case.

Comparison of NOESY spectra with and without the presence of the TPR domain, revealed NOEs specific to the bound stage. Almost all these transferred NOEs were assigned to the peptide tryptophane (Trp241) aromatic ring. In particular, the aromatic ring shows contacts to the preceding proline ring protons (Pro239) and the following glutamine (Gln242) aliphatic side chain (Fig. 17B). Additional transferred NOEs were observed in the experiment, involving the Ile238 and Ile243 methyl groups. However, due to low intensity and high ambiguity in their assignment, they were not included in the structure calculation. Nevertheless, all these weak NOEs are consistent with the calculated model. Although, initial structure calculations included the complete peptide, finally only the core region of the peptide, comprising residues 238-243, was used. The reason was that on one hand in ^{1}H,^{15}N HSQC titrations amides of residues 234-236 and 245-251 were not affected upon Pex5 TPR addition (Fig. 17A), and on the other hand only intra-residue NOEs were obtained for these parts. Thus, the final fold of the region comprising residues 237-244 of the peptide was calculated from a total of 125 short and medium range intramolecular NOEs (Fig. 17C,D). The backbone RMSD of the ensemble demonstrates a fairly rigid fold. A key element of the fold is the tryptophan ring, which stacks with Pro239 and Gln242 side chains creating a small kink at the backbone (Fig. 17C).

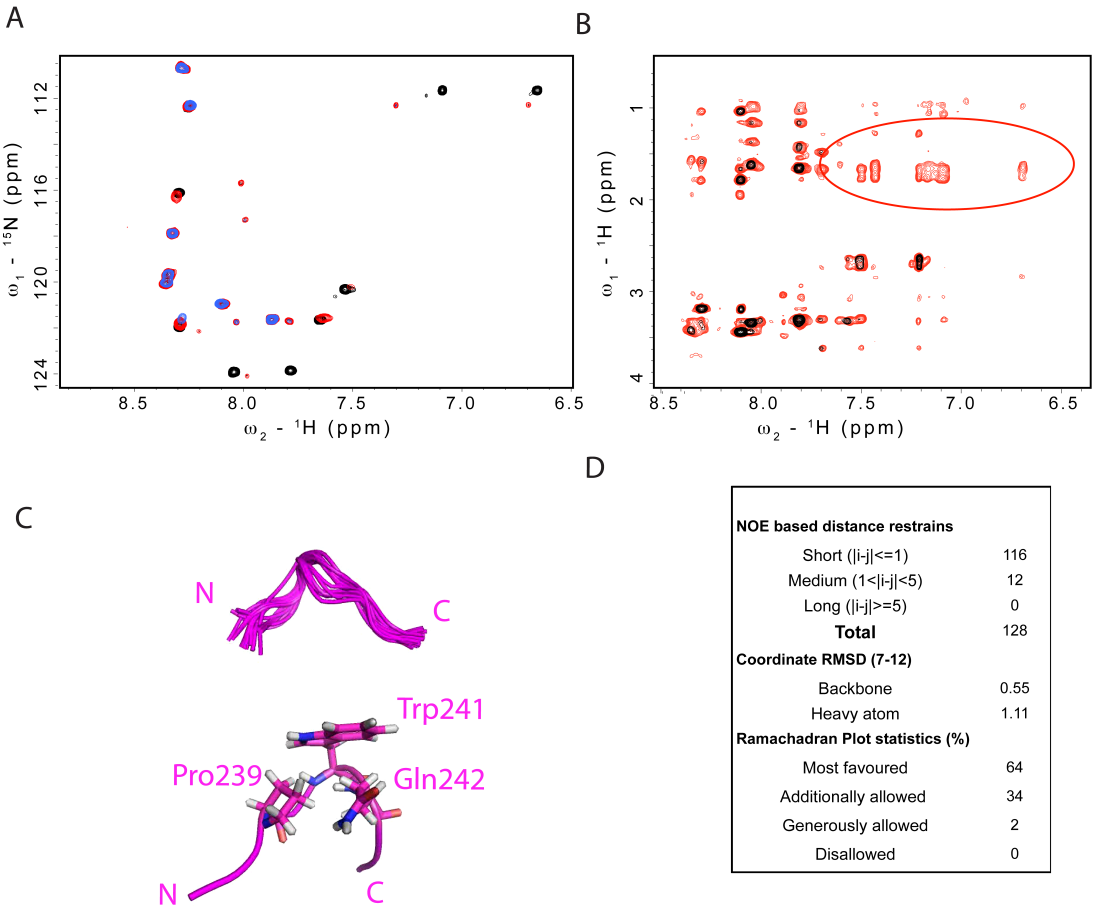


Figure 17. Conformation of WQI peptide. A) ^{1}H,^{15}N HSQC of ^{15}N labeled WQI peptide 150 µM (black) supplemented with 50 µM (blue), 150 µM (magenta) and 400 µM (red) unlabeled Pex5 TPR. B) ^{1}H,^{1}H NOESY experiment showing the amide – aliphatic region. The black spectrum was obtained from a 1 mM solution of the WQI peptide; while in red 50 mM of unlabeled Pex5 TPR is added. The additional peaks highlighted in circle correspond to transfer NOEs from the bound conformation of the peptide to the free. C) NOE based structure calculation of core of the peptide depicted as an ensemble of the 10 lowest energy structures (upper) and with key side chains shown as sticks (lower). D) Structure statistics of the WQI peptide fold.

3.1.7. Intermolecular NOEs between Pex5 TPR and WQI peptide

Intermolecular NOESY experiments were measured using the Pex5 TPR sample saturated with the WQI peptide. Since the protein contained 69 Leu, Val and Ile residues the methyl region of $^{1}H,^{13}C$ HMQC spectra showed substantial signal overlap. This was overcome employing a selectively methyl-labeled ((^{1}H-$\delta 1$ methyl)-isoleucine, (^{1}H-$\delta 1$ methyl)-leucine and (^{1}H-$\gamma 1$ methyl)-valine) sample (Tugarinov & Kay, 2003). This scheme, together with $^{1}H,^{13}C$ HMQC titrations enabled the identification of methyls affected upon WQI binding (Fig. 18A,B). From these methyls, Ile486, L476 and V479 give rise to intermolecular NOEs (Fig. 18C). Interestingly, isoleucine 486 is the only isoleucine $\delta 1$-methyl that exhibits chemical shift perturbations (Fig. 18A). All three methyls had intermolecular NOEs to protons corresponding to aliphatic protons of the WQI peptide, in particular in the region of 0.9 to 1.9 ppm (Fig. 18C), suggesting a mainly hydrophobic nature of the interaction. Additionally, the L476 $\delta 1$-methyl shows intermolecular NOEs involving aromatic protons of the peptide (Fig. 18C). This NOE was unambiguously assigned to the tryptophane aromatic protons of the peptide. Intermolecular NOEs of both I486 and V479 at a ω_1 chemical shift of 0.9 ppm were ambiguously assigned to the $\delta 1$ methyl resonances of either peptide isoleucines (Ile238 or Ile243). Similarly, the $\gamma 1$ methyls of either these peptide isoleucines were found in contact with the $\delta 1$ methyl of I486. All intermolecular NOEs observed at 1.9 ppm were originate from either Pro239 or Pro244 of the peptide. Notably, also the backbone of the peptide was found to be in close proximity to I486 $\delta 1$ methyl. In particular, the NOE observed at 4.4 ppm matches the frequency of the H_α resonance from either Ile238 or Pro244 of the peptide. In summary, all intermolecular NOEs used in the complex structure calculation are shown in Table 6 in section 2.7.8.

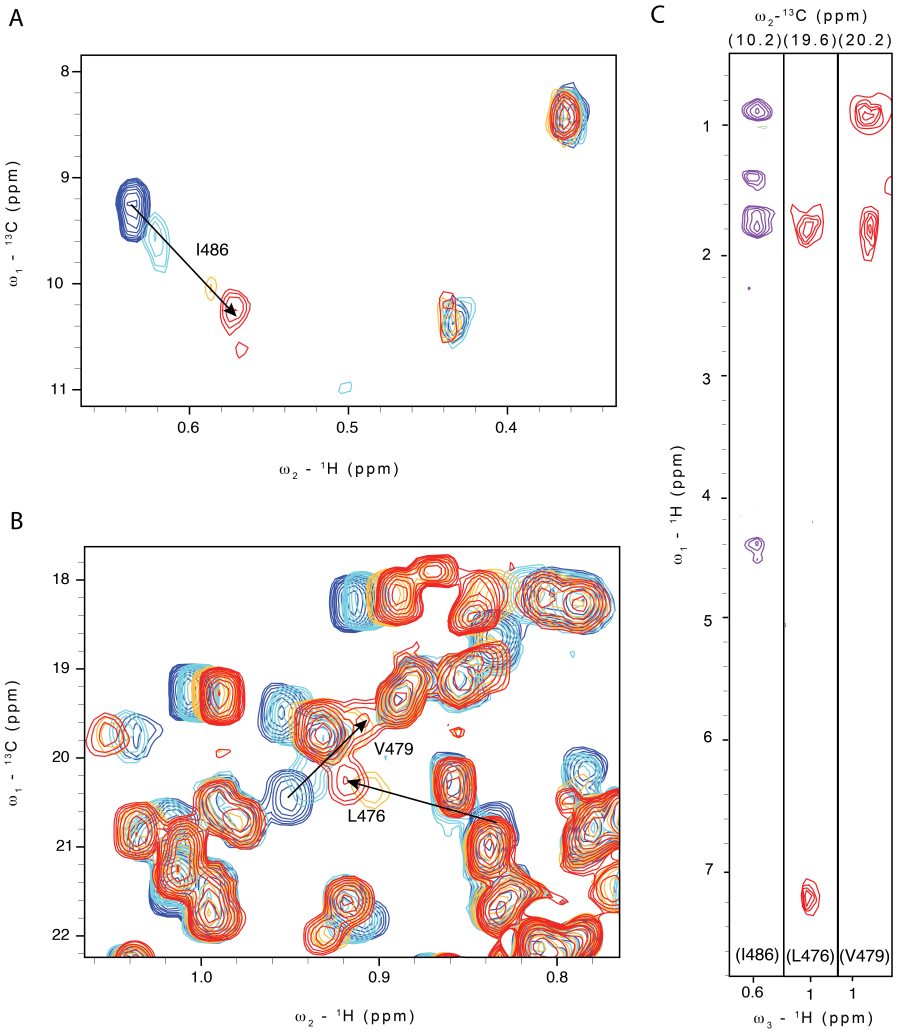


Figure 18. Intermolecular NOEs between Pex5 TPR and the WQI peptide. A) ^{1}H,^{13}C HMQC spectra overlay of Pex5 isoleucine δ1 methyls of 100 µM TPR (dark blue) with 100 µM (light blue), 300 µM (orange) and 1000 µM (red) WQI peptide. Isoleucine I486 δ1-methyl showing the largest chemical shift perturbation and intermolecular NOEs is displayed. B) Same titration as in A zoomed in the Leu δ1- and Ile γ1- methyls region. Methyls exhibiting NOEs with WQI peptide are marked. C) ω1-filtered/ω2-edited NOESY strips displaying intermolecular NOEs from Pex5 I486 δ1-methyl (purple), L476 δ1-methyl and V479 γ1-methyls to the WQI peptide.

3.1.8. Structure calculation of the Pex5 TPR - WQI complex

The High Ambiguity Driven protein-protein DOCKing (HADDOCK) web server was used to dock the WQI peptide on the Pex5 TPR structure based on the chemical shifts perturbations and intermolecular NOEs (see 2.7.8). Using a 3 Å RMSD clustering criterion, HADDOCK generated 2 clusters with a total of 199 complex structures. While in both clusters the WQI peptide was found at the same interface, they differed slightly in the orientation of the peptide. The overall lower energy, interface and ligand RMSD made the least populated cluster 2 solution of choice (Table 7).

The peptide in both clusters is located on the interface between TPR 4 and 5 of the domain (Fig. 19A). A shallow groove on the surface of these helices accommodates Pro238, Trp241 and Gln242 of the peptide (Fig. 19B,C). The majority of the intermolecular interactions involve hydrophobic groups highlighting the nature of the binding. Interestingly, two hydrogen bonds are present in the best cluster models (Fig. 19B). The glutamine of the peptide appears to form a hydrogen bond with the Pex5 Q491 side chain of Pex5 TPR. Moreover, the peptide's tryptophane indole formed the second hydrogen bond with D505 of the protein (Fig. 19C).

3.1.9. Affinity measurement between WQI peptide and Pex5 TPR

To measure the affinity of the peptide for the TPR domain, a N-terminally fluorescein isothiocyanate (FITC) labeled peptide was used in the fluorescence polarization experiment. The peptide exhibits low affinity for the wild type Pex5 TPR, at a concentration of 250 µM. Based on a HADDOCK complex model, the Pex5 TPR double mutant Q491A/L494A was produced and tested for in vitro binding to the WQI peptide (Fig. 20A). The binding constant of the peptide for the mutant is 4-5 times weaker, and falls into the millimolar range (Fig. 20B). Unfortunately, an accurate estimate was not possible due to non-saturating conditions.

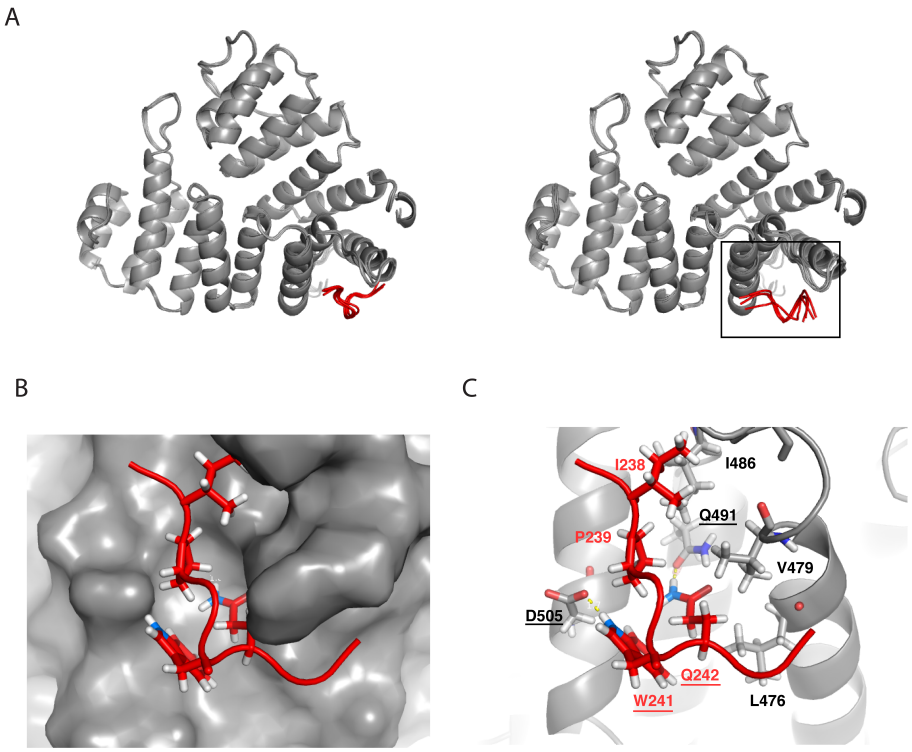


Figure 19. HADDOCK modeling. Ensemble of 4 lowest energy models from each cluster. left cluster1, right cluster2. The peptide is shown in red. B) Surface representation of lowest energy model from cluster 1 around WQI peptide. C) Interface of interaction displaying important residues. Marked residues had intermolecular NOEs. Underlined residues are in optimal geometry for hydrogen bond formation.

Table 7. HADDOCK modeling statistics

HADDOCK statistics		
Parameter	Cluster1	Cluster2
HADDOCK score	-102.8	-104.7
Cluster size	185	14
i-RMSD[a]	2.1 Å	0.5 Å
l-RMSD[a]	5.9 Å	1.0 Å
fnat[c]	0.85	0.47
VdW[d]	-35.2 kcal/mol	-39.2 kcal/mol
Elec[e]	-92.7 kcal/mol	-114.9 kcal/mol
Violation[f]	33.2 kcal/mol	34.9 kcal/mol
BSA[g]	931.7 $Å^2$	1062.7 $Å^2$

[a] interface root mean square deviation

[b] ligand root mean square deviation

[c] fraction of native contacts

[d] Van der Waals contribution to intermolecular energies

[e] electrostatic contribution to intermolecular energies

[f] restrains violation energy

[g] buried surface area

Strikingly, the presence of the PTS1 peptide has a similar effect on the binding of the WQI peptide. Likewise the double mutant, no saturation was achieved even at supra millimolar TPR concentrations (Fig. 20B). Therefore, Q491A/L494A was additionally tested for in vivo localization of PTS1 proteins, resulting in a severe mislocalization.

3.1.10. Refinement of TPR structure

Residual dipolar couplings for the Pex5 TPR domain were measured in the presence of the WQI peptide. For that purpose, TROSY based HNCO experiments for $^{1}H^{N}$-^{15}N, $^{13}C_{\alpha}$-^{13}CO and ^{15}N-^{13}CO couplings as described in 2.7.7 were

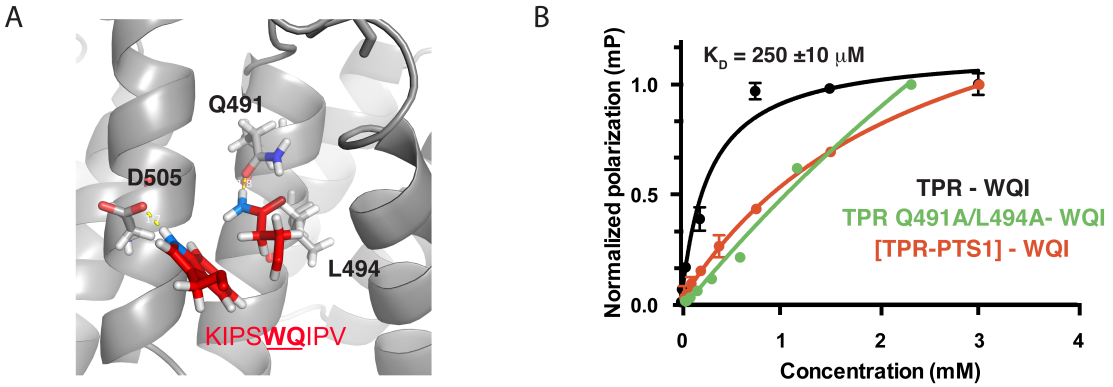


Figure 20. Affinity measurements of WQI peptide for TPR. A) Cartoon representation of HADDOCK model of Pex5-WQI peptide complex. In sticks are depicted the sidechains selected for alanine mutations. Tryptophane and glutamine side chain of the peptide, which interact with TPR domain, are shown in red sticks. B) Affinity measurement of WQI peptide for Pex5 TPR with fluorescence polarization.

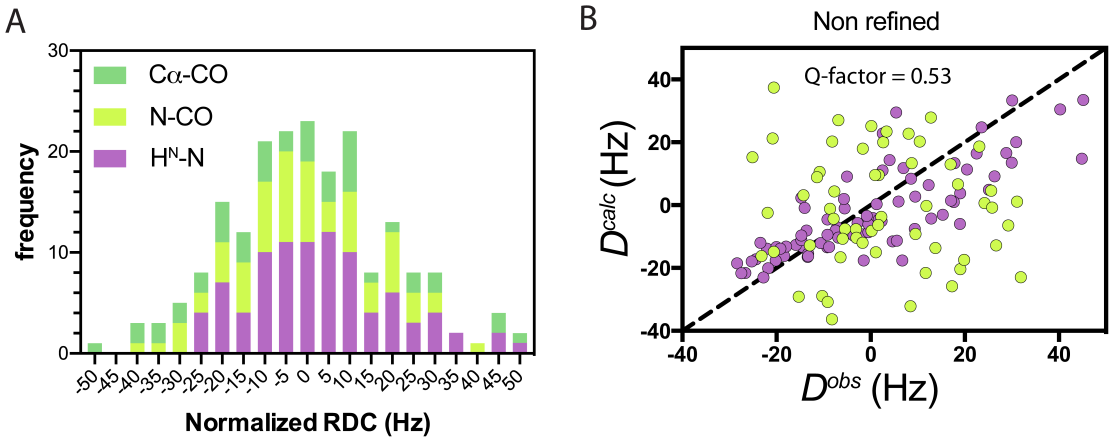


Figure 21. Residual dipolar couplings analysis with histogram method. A) Histogram distribution of all normalized RDCs color-coded. B) Correlation of observed RDC values (NH purple, NC green) with back-calculated values from the crystal structure

recorded (Fig. 22). Analysis of a total 95 $^{1}H^{N}$-^{15}N, 75 ^{15}N-^{13}CO and 51 $^{13}C_{\alpha}$-^{13}CO couplings revealed a Q_{factor} of 0.53 indicating that they can be used to refine the free TPR structure (Fig. 21).

3.1.11. Discussion

Peroxisomal translocation is one of the least understood process of the cell, due to the relative recent discovery of the organelle. Roles and structures of peroxins Pex5 and Pex14, although key components of the pathway, till now remain largely unknown. With the results collected in this project, we proved that the interplay between them is more complex than previously thought. Both proteins seem to share the same bipartite domain organization, only reversed. While Pex5 bears a mostly unstructured N-terminal and an α-helical C-terminal domain, Pex14 on the other hand is structured in the N-terminal region while C-terminal is disordered. Moreover our data demonstrate that the two peroxins have at least two distant binding sites. Since both molecules are part of the same pathway, such tight interaction is not surprise. On the other hand the nature of the interaction reveals to a large extent the dynamic nature of the peroxisomal cargo import. In both interactions mentioned, the disordered regions of the proteins may heritage the system the necessary flexibility to adapt to various cargo sizes and perform the different steps of the translocation. It would be difficult to have a cargo-adaptive machinery if all components involved are rigid.

The C-terminal region of Pex14 found to interact with TPR domain of Pex5 exhibits some very interesting features. Firstly, the conformation of the peptide derived from that region, WQI peptide, revealed that even bound to TPR does not adopt any secondary structure. This is the opposite behavior to the N-terminal interaction, where WF motifs of Pex5 become helical upon binding to the Pex14 NTD. Secondly, the very weak affinity of the interaction indicates that it is involved in an essential dynamic step of the import. It is intriguing that the low

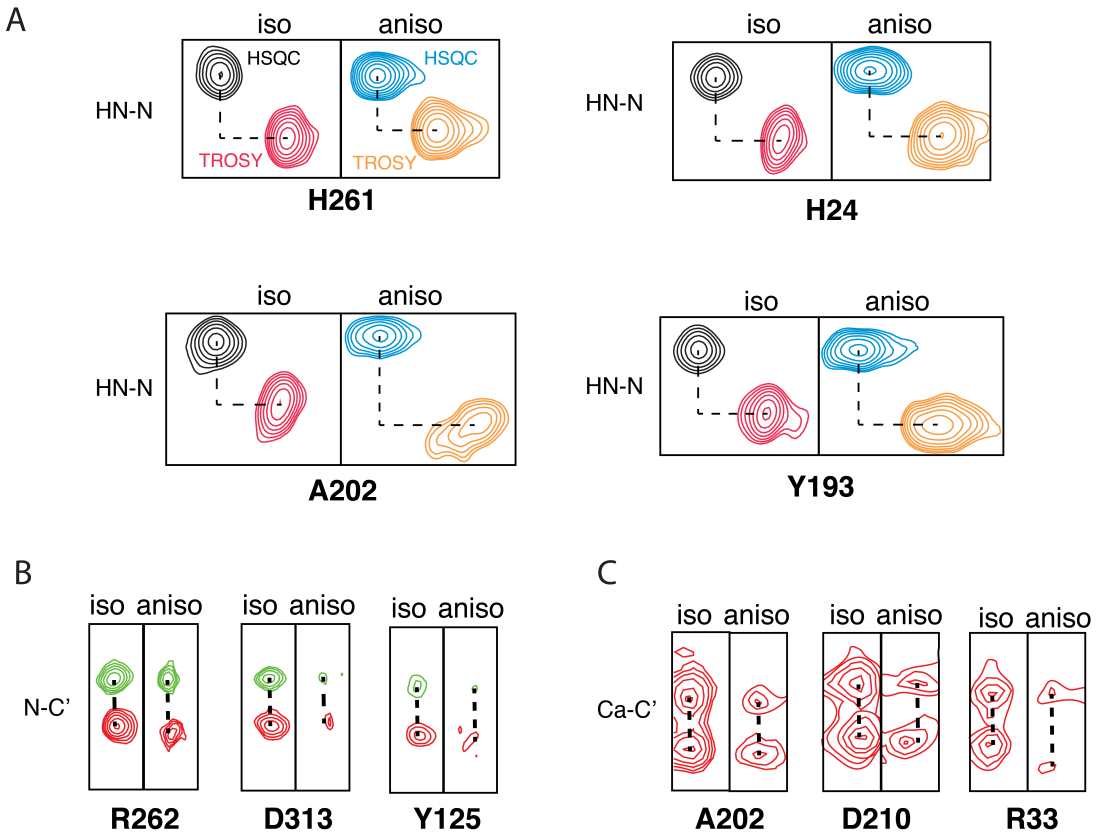


Figure 22. Residual dipolar couplings measurements of Pex5 TPR. A) HSQC and and TROSY component of selected amides in isotropic (iso) and anisotropic solution (aniso) containing 15 mg/mL Pf1 phages. B) Amides of selected residues recorded from TROSY-HNCO with splitting in nitrogen dimension due to N-C' coupling. C) Carbonyls of selected residues recorded from TROSY-HNCO, splitting in carbon dimension is caused by Ca-C' coupling.

nanomolar interaction of the WF motifs for the Pex14 NTD is dispensable for the cell (Neuhaus et al., 2014), while the high micromolar WQI-TPR interaction is absolutely necessary. Such weak interactions are usually found in nods of regulation by the cell, while very strong interactions keep the constituents together. In fact, careful sequence observation reveals the presence of many serine residues in close proximity to WQI peptide. Already in the peptide used, Ser240 is in the middle of the binding site. The fact that the serine-rich flanking region of WQI is conserved, points to a potential mechanism how cell would regulate the specific binding by phosphorylation of these serines. Lately, a mass spectrometry study identified in Hansenula polymorpha two in vivo phosphorylation sites that correspond to the WQI region in human (Tanaka et al., 2013). Although exact position of the modification might differ by some residues among species, indicates that the interaction described here might be regulated in cells.

Furthermore, in the study by Stanley and coworkers (Stanley et al., 2007) it is reported that TPR undergoes an interesting structure rearrangement upon cargo binding. According to their findings the ring-like formation of the TPR repeats becomes tighter as 1-3 and 4-7 TPR repeats come closer when bound to cargo. They identified the fourth TPR to be the pivot point of this movement. Interestingly, WQI peptide binds at this region between TPR 4 and 5. Pex14 might recognize structural changes on that region via the WQI interaction.

Up to date the only role that has been assigned to TPR domain of Pex5 was the cargo recognition. Data presented here support a more complex function of this domain, since it also interacts with the C-terminal region of Pex14. Unfortunately, it was not possible to pin down the role of this interaction, yet interesting result emerged. Localization experiments have previously shown that while the NTD of Pex14 is not solvent accessible; the CTD faces the cytoplasm (Azevedo & Schliebs, 2006). Therefore, one hypothesis could be that cytoplasmic WQI region would be the first to sense a cargo-loaded receptor and subsequently the tight N-terminal interaction will take place as soon as Pex5 approaches in the

membrane. However, the fluorescence polarization data show that loaded receptor binds weaker to WQI region than the empty receptor. Hence, the Pex14 CTD binds tighter to an empty receptor. One could rationalize this effect, by imaging that the Pex14 CTD can distinguish receptor that has completed translocation and cargo-receptor on the way to translocate. Such differentiation would be important for the cell; to avoid that Pex5 recycles back in the cytosol prior to cargo release. Although more data are required to test this scenario, it would make sense from functional point of view. Therefore, it seems more likely that this novel C-terminal interaction plays crucial role at later stage of the cargo translocation by coupling the receptor with export components. In alignment to that, published data support this possibility in the case of Leishmania donovani organism (Madrid et al., 2004).

Mainly two aspects highlight the importance of this interaction. A previous study by Niederhoff et al already showed that Pex14 possesses two distinct functional Pex5 binding sites (Niederhoff et al., 2005). With multiple Pex14 variants and yeast-to-hybrid assay they were able to identify 20-amino acid region in the Pex14 CTD to be essential for growth and Pex5 interaction. Thus, the importance of this C-terminal interaction between these proteins is emphasized by its conservation across species. Furthermore, till now no Pex5 mutant has been identified with a similar strong import deficiency as Q491A/L494A described here. The N-terminal mutations in single or dual WF motifs yield no significant phenotype in vivo. In fact, two of these motifs together with a recently described WF-like motif have to be mutated to alanines in order to achieve the same import deficiency as Q491A/L494A mutant (Neuhaus et al., 2014).

In summary, the results presented here add one more piece to the puzzle of the peroxisomal import pathway. The previously neglected CTD of Pex14 is biochemically and structurally characterized, revealing a flexible region. A 5-residue stretch within the CTD was identified to mediate an important interaction with the TPR domain of Pex5. The structural model obtained and the enhanced affinity for an "empty" receptor suggests possible implication in later stages of the transloca-

tion. New questions are raised regarding the in vivo role of this new interaction and cargo release regulation, further research should be conducted to gain much better overview of this machinery.

3.2. Structure-based drug design of a trypanosomal Pex14/Pex5 inhibitor

The parasites of the genus Trypanosoma enclose in the same organelle peroxisomal reactions together with many glycolytic enzymes. Hence the name of this special organelle in Trypanosomes is glycosome. It has been suggested that impaired activity of these organelles will lead to parasite ATP-starvation and death. In this project we exploited previous reports about the biogenesis of the glycosome to target its function and design an trypanocidal compound as new tool for therapy.

3.2.1. NMR structure of free tbPex14 NTD

The same peroxins that govern peroxisome biogenesis are found in *Trypanosoma brucei* (tb). In particular, for the PTS1 import pathway, cargoes in the cytoplasm are targeted to the membrane-anchored tbPex14, via initial tbPex5 binding. The last contains several di-aromatic motifs (WF) that are accommodated in hydrophobic pockets of tbPex14 NTD. Although structures from both rat and human Pex14 NTD are available, for a parasite specific drug the tbPex14 NTD structure is required as template for small molecule design. Therefore the solution structure of tbPex14 NTD was determined (Fig. 23A).

The construct used for this purpose was somewhat different from the previously published rat and human. Since no structural information was available we included in both termini of the protein additional amino acids not to miss important contacts or structure elements. Hence the tbPex14 (1-84) was uniformly ^{15}N and ^{13}C labeled and standard backbone and side chains assignment spectra were recorded and analyzed as described in 2.7.1. Calculation of the solution NMR structure was accomplished as described in 2.7.8 and revealed a three helical bundle domain very similar to previously published structures (backbone coordinate

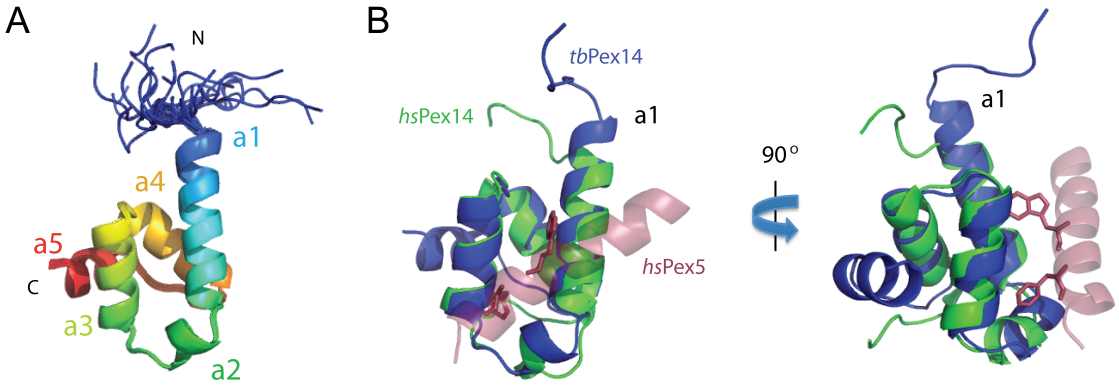


Figure 23. Solution NMR structure of tbPex14 NTD. A) Cartoon representation of the tbPex14 NTD ensemble. 20 lowest energy models are depicted. The disordered N-terminal and the additional α5 helix are visible, in blue and red respectively. Backbone RMSD = 0.11, Heavy atom RMSD = 0.74. B) Comparison of human (green) and T.brucei (blue) of Pex14 NTD. The human protein is complexed with Pex5 WF motif, which is shown in transparent red cartoon. The aromatic residues of WF motif are displayed as sticks. Rotation by 90^o of the structures reveals the missing C-terminal helix of the human construct.

RMSD = 0.5 Å) (Fig. 23B). Interestingly, the additional C-terminal residues present form an α-helix which folds on to the back side of the WF binding interface (Fig. 23B). Contrary, the initial 19 residues are completely flexible and do not form any structural element or contacts with the folded domain.

3.2.2. Pharmacophore model and initial compounds

The NMR structure of the free tbPex14 NTD described in 3.2.1 together with the previously published structure of the hsPex14 complex with WF motif determined the key features absolutely required for this interaction. Two $\pi - \pi$ stacked aromatic phenylalanines (35 and 52) exist at the core of NTD fold and shape two highly hydrophobic pockets exposed to the surface. Since the accommodation of these pockets by the WF peptide, is responsible for most of the binding

free energy, small molecule design was based on this.

A 3D pharmacophore model mimicking the binding mode of the aromatic residues of WF repeats to the tbPex14 NTD pockets (Fig. 24A). Pharmacophore-based in silico screening, perfomed by Dr. G. Popowicz, of >20,000,000 compounds was followed by 3D docking and visual inspection of the highest scoring compounds. The ten most promising compounds were selected to confirm in-vitro binding to tbPex14 NTD by NMR. From ^{1}H,^{15}N HSQC titration of ^{15}N labeled protein with candidate compounds we confirmed that binding affects residues on the WF interface (Fig. 24C). This leads to the identification of "drug-like" scaffold of pyrazolo[4,3-c]pyridine (Fig. 24B), which was easy to chemically decorate, by chemist Dr. M. Dawidowski, and achieve high enough affinity for competing WF peptide. As expected, at this early stage most compounds tested in NMR titrations exhibited chemical shift perturbations in fast exchange indicating moderate affinity (Fig. 24C). The ability to disrupt the tbPex14-tbPex5 complex formation was evaluated using AlphaScreen assay. Indeed the experiment revealed that already these initial compounds were able to compete with Ki = 58 µM (see 2.6.2).

3.2.3. Structure-based inhibitor optimization

Subsequent optimization of the aromatic moieties addressing the two pockets gave **2** and **3** with an increased potency in biochemical assays as well as enhancement in trypanocidal activity (Fig. 24D). Atomic-resolution X-ray structures of both compounds complexed with tbPex14 (Fig. 25A) (Table 8) confirmed the anticipated binding mode. The central element (II of pharmacophore model) of both compounds shields efficiently the partially exposed Phe35 and Phe52, while each of the two aromatic groups occupies its respective hydrophobic pocket. Notably, the indole ring of **2** addresses the phenylalanine pocket (III), whereas the tryptophan pocket is filled with the naphthyl ring (I). Similarly, in Pex14-3 complex the phenylalanine pocket is filled by the metoxynaphthyl moiety with the central ele-

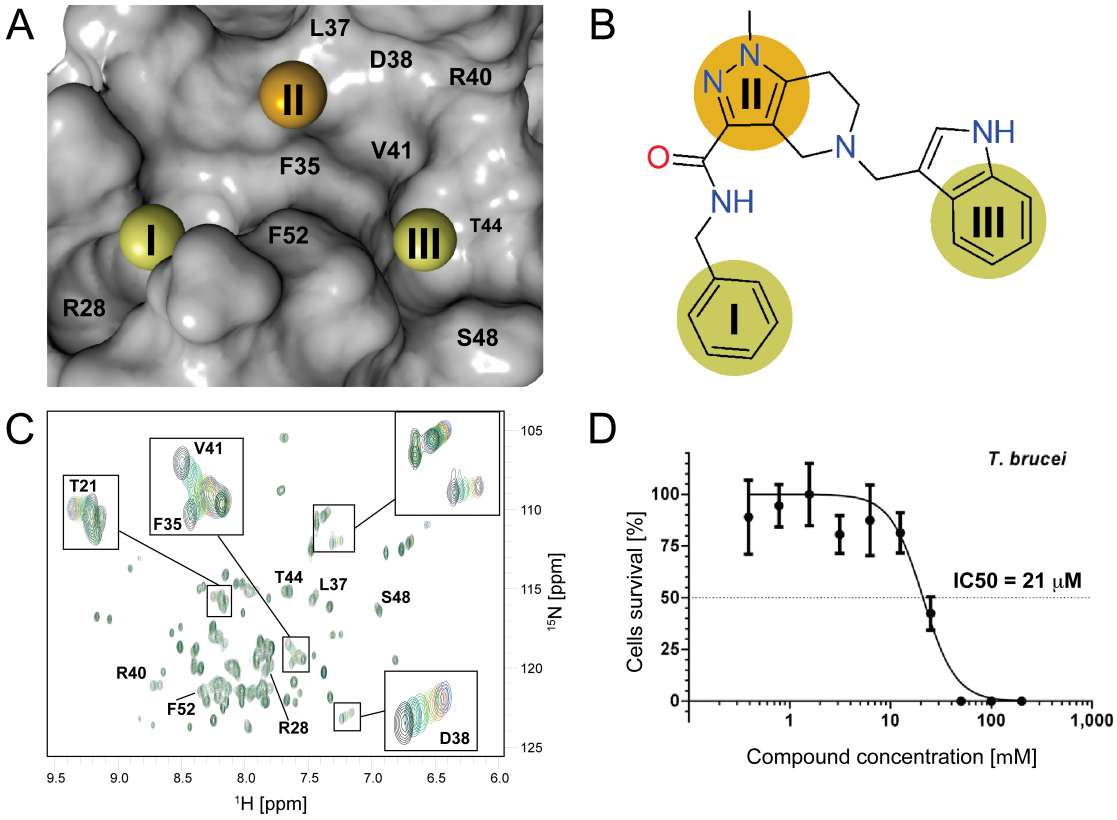


Figure 24. Pharmacophore model with initial confirmed Pex14-Pex5 inhibitors. A) Surface of tbPex14 NTD showing the WF motif binding site. The residues which establish contacts with the motif are marked. Yellow spheres locate crucial binding areas which inhibitor should bind. Positions I and III correspond to the two hydrophobic pockets of Pex14 NTD, where aromatic residues of WF are accommodated. Position II highlight the exposed phenylalanine rings (F35 and F52) which link the two pockets. B) Chemical formula of initial hit compound 5. With yellow circles are marked the elements which correspond to the desired features based on the pharmacophore model. C) ^{1}H, ^{1}N HSQC spectra of 50 µM tbPex14 NTD (black) with increasing concentrations of the compound **4** up to 500 µM (blue). Residues exhibiting the biggest chemical shifts are noted. D) Treatment of blood-stream form culture of T.brucei with increasing concentration of compound 4 indicate that 50% of the cells are dead with 21 µM of compound. Data by V. Kalel.

ment remains in a virtually identical position as in **2**. The phenyl moiety of **3** fills the tryptophane pocket.

Although the compounds interact with Pex14 almost exclusively via hydrophobic interactions, one water molecule forms two important hydrogen bonds. Remarkably, this molecule is found in all complex structures of the inhibitors tested, suggesting the importance of this interaction. Indeed, the water molecule bridges the amide groups of both compounds with the Asn31 side chain of the protein. Interestingly, this asparagine is one of the few residues (Asn31, Asp38 and Ser48) in the interface that is not conserved between human and trypanosome Pex14. In the human protein Asn31 is replaced by Thr31, which does not allow the formation of a hydrogen bond with the inhibitor at this position. Consequently, this tightly coordinated water molecule reveals a means to distinguish two almost identical interfaces.

Both **2** and **3** show significantly enhanced hydrophobic interface when compared to **1**. This is in agreement with their improved binding inhibition constant. Notably, lower Ki lead to significant improvement of trypanocidal activity, with EC_{50} being lower than Ki as was the case for **1**.

Since for all compounds solubility is low, further optimization aimed at improvement of solubility. The replacement of the hydroxyl group at position II to amine resulted compound **4,** a much more soluble and potent Pex14-Pex5 interaction inhibitor (Ki = 223 nM).

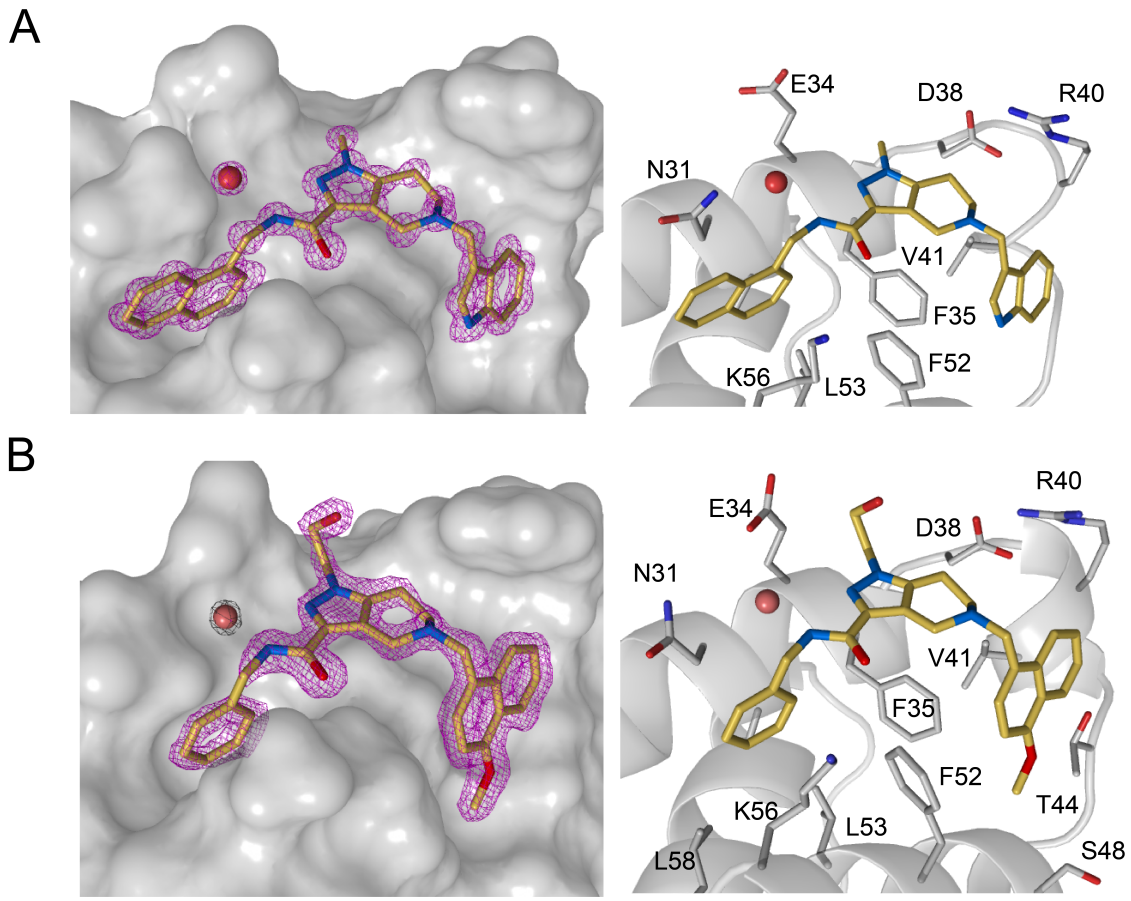


Figure 25. Crystal structures of tbPex14 in complex with compounds. Protein is colored as grey surface or cartoon and ligands as yellow sticks. The electron density as calculated from the structure is also shown in magenta. Water molecule participating in hydrogen bonding between the protein and the ligand is displayed as sphere. A) and B) correspond to **2** and **3** compounds respectively. The residues, which are direct, contact with the labeled on the cartoon representation and marked as sticks. Data analysis performed by G. Popowicz.

Table 8. Data collection and refinement statistics (Molecular replacement)

	tbPex14-2	tbPex14-3
Data collection		
Space group	C2	P1
Cell dimensions		
a, b, c (Å)	43.03, 37.26, 42.46	36.1, 36.28, 55.31
α, β, γ (°)	90.0, 117.9, 90.0	94.1, 105.9, 100.5
Resolution (Å)	37-0.86(0.88-086)	31.3-1.57(1.66-1.57)
Rmerge	4.6 (76.8)	10.9 (60.3)
$I/\sigma I$	18.2(1.62)	7.5 (2.2)
Completeness (%)	95.6(61.8)	92.7 (87.4)
Redundancy	4.0(1.63)	3.7 (3.4)
Refinement		
Resolution (Å)	37.5-0.86	27.7-1.57
No. reflections	44871	32107
Rwork/ Rfree	0.13/0.15	0.22/0.25
No. atoms		
Protein	997	1958
Ligand/ion	72	149/1
Water	136	279
B-factors		
Protein	11.3	19.1
Ligand/ion	21.3	18
Water	26.7	31.7
R.m.s deviations		
Bond lengths (Å)	0.03	0.02
Bond angles (°)	2.4	2.3

3.2.4. Mislocalization of PTS1 enzymes in parasite cultures (in collaboration with R.Erdmann lab)

All compounds were tested for trypanocidal activity. The blood-stream form of the parasite was grown in cultures and treated with various compounds. Visual inspection of such parasites with optical microscopy reveals the survival rate upon treatment. Remarkably, almost all compounds tested displayed cellular potency against *T. brucei*, compound **4** for example showed EC_{50} of 186 nM. The clinically relevant *T. brucei rhodesiense* subspecies had proved to be more susceptible for the same inhibitor (EC_{50}= 20 nM). The subspecies, which cause Chagas disease, *T. cruzi* was also affected displaying 2 times higher cytotoxicity (EC_{50}= 750 nM) than the reference drug benznidazole.

To validate that tbPex14-tbPex5 was indeed disrupted in vivo, localization of cargo proteins were performed. Immunofluorescence microscopy using antibodies against glycosomal enzymes, from both PTS1 and PTS2 families, treated with inhibitors exhibited severe mislocalization in the cytoplasm of enzymes belonging in both pathways. Glycosomal enzymes Hexokinase (HK), phosphofructokinase (PFK), aldolase (ALD) and dehydrogenase (GAPDH) were found almost exclusively in the cytoplasm (Fig. 26).

3.2.5. Pex14/Pex5 inhibitor discussion

In this project we targeted the T.brucei glycosome import pathway to develop an antiparasitic compound. Structural data from the essential Pex5-Pex14 interaction together with solution structure of tbPex14 NTD, enabled us to design small molecules mimicking the WF motifs. The small molecules developed compete and inhibit the crucial interaction between the cytosolic receptor Pex5 and its membranic counterpart on glycosome Pex14. We proved that the previously suggested strategy to eliminate the parasites by targeting the glycosome

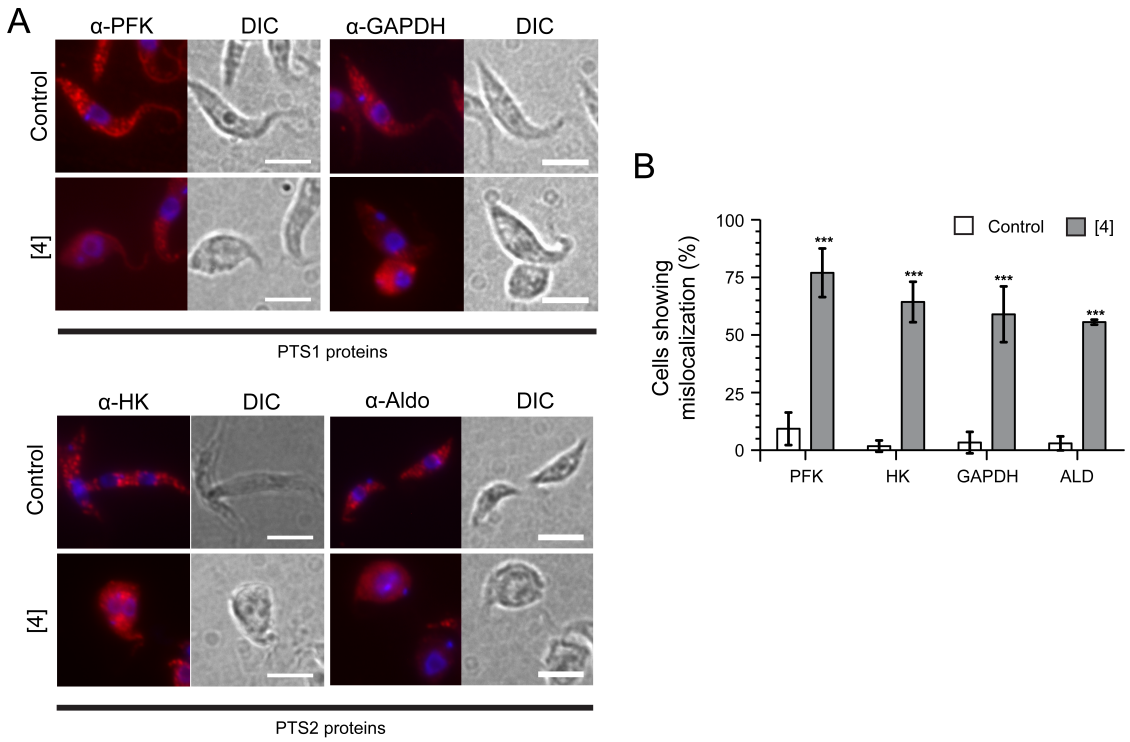


Figure 26. In-vivo target validation. A) Immunofluorescence microscopy analysis of subcellular localization of glycosomal glycolytic enzymes PFK, HK, GAPDH and ALD. All enzymes tested loose the distinct punctuate pattern of peroxisomal localization upon inhibitor treatment. DIC, Differential Interference Contrast. Scale bars, 5 µm. Statistical analysis of percentage of cells showing mislocalization of enzymes. Data shown ± s.d. from two independent biological replicates. ***, $p < 0.001$. Data by V. Kalel.

biogenesis is valid and applicable, since the developed molecules selectively kill the trypanosomes. The observed mislocalization of glycosomal enzymes in the cytoplasm in the presence of the inhibitors, confirmed that we targeted indeed the import pathway.

It is intriguing that the parasites' energy source, glucose, becomes lethal in the absence of glycosomes. Our approach was to target this weakness of trypanosomes by inhibiting protein import in the organelle. Notably, after relatively simple compound optimization based on crystal structures of Pex14-inhibitor complexes, inhibitors show trypanocidal activity comparable or better to currently approved therapeutics in cell-based assays.

It is intriguing that the trypanosomes' essential carbon source; glucose becomes trypanocidal in the absence of the glycosome. Here we exploited this phenomenon to attack trypanosomes by inhibition of glycosomal protein import.

3.3. Role of Pex19 farnesylation

The transport of peroxisomal membrane proteins (PMPs) to the peroxisome requires the soluble Pex19 protein as chaperone and import receptor. Pex19 interacts with PMPs via the C-terminal domain. Previous crystal structure of the domain shows that the domains form an α-helical fold, with indication that the first helix is involved with PMP recognition (Schueller et al., 2010). While this study revealed the structure of Pex19 CTD, it was incomplete as the important farnesylation site was not included in that construct. It is known that Pex19 farnesylation at a conserved cysteine residue in its C-terminal CaaX motif (C=Cys, a=aliphatic, X=any amino acid) enhances the PMP binding. The three-dimensional structure of the farnesylated protein (Schuetz, 2012) shows that the farnesyl moiety is buried in a hydrophobic cavity formed by aliphatic side chains of the Pex19 CTD forming a hydrophobic area on the surface of the protein. (Fig. 27A,B). Comparison of farnesylated and non-farnesylated protein shows distinct changes in helix α1 (Fig. 27C). In this project we used the farnesylated structure of the Pex19 CTD to explain the role of the modification in PMP recognition.

3.3.1. Mapping of PMP interaction

The nature of PMP proteins, and specifically the Pex19 binding site, is very hydrophobic, as expected, for membrane proteins, thus not very soluble in aqueous buffers. Based on previous data (Schuetz, 2012) less hydrophobic peptides from Pex13 and ALDP proteins were purchased (see 2.7.5). Both were prepared as described in 2.9.5 for NMR titrations. Proton 1D spectra of both peptides show that they are partially soluble (Fig. 28A). Although both exhibited much higher solubility than previously tested peptides, in case of Pex13 titration severe line broadening was observed. This effect, indicative for precipitation, aggregation or too tight binding, was much more pronounced for the farnesylated Pex19 CTD

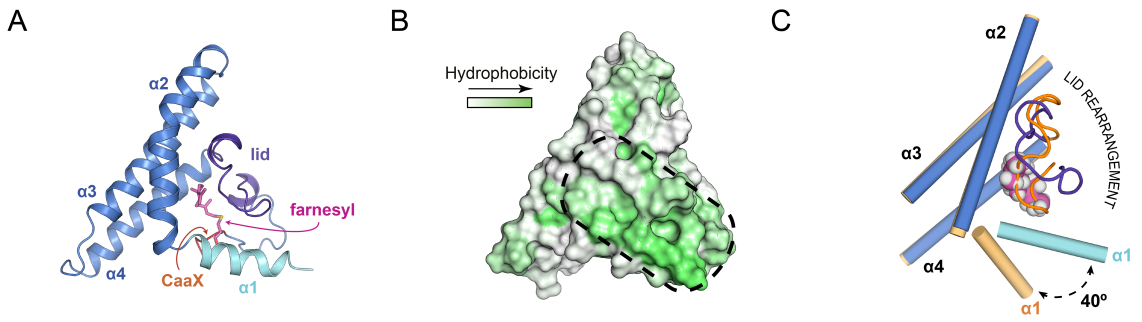


Figure 27. Solution structure of farnesylated Pex19 CTD. A) Cartoon representation of the structure. Farnesyl group shown as purple sticks. B) White-green colored gradient of protein surface depending on the hydrophobicity of residues (Kite-Doolittle scale). C) Schematic comparison of crystal non-farnesylated (orange) (PDB : 2WL8) and NMR farnesylated (blue/cyan) of the structure. The different position of lid and helix $\alpha 1$ is highlighted.

(Fig. 28B,C). Contrary, the ALDP peptide does not show the same behavior. NMR titrations with this peptide results in chemical shift perturbations (CSPs) (Fig. 28B) in the fast to intermediate exchange regime. The ^{15}N (Fig. 28A right) and ^{13}C CSPs (Fig. 29A) were quantified and plotted against protein residues (Fig. 29B), revealing the most affected residues upon interaction. The mapping of the interface on the structure of the Pex19 CTD indicates that mainly helix $\alpha 1$ and the lid region where affected (Fig. 29C). Since no other region of the protein showed significant CSP, helix $\alpha 1$ and the lid must be directly affected by the presence of the peptide.

3.3.2. Intermolecular NOEs between PMP and Pex19 CTD$_{farn}$

Further insights about the structure of the Pex19 CTD$_{farn}$-ALDP is gained by determining intermolecular contacts. A number of intermolecular NOEs were recorded. A ω_1 ^{13}C-edited, ω_3 ^{13}C,^{15}N filtered NOESY spectrum revealed that the peptide contacts the protein with mainly hydrophobic methyl-, methylene-

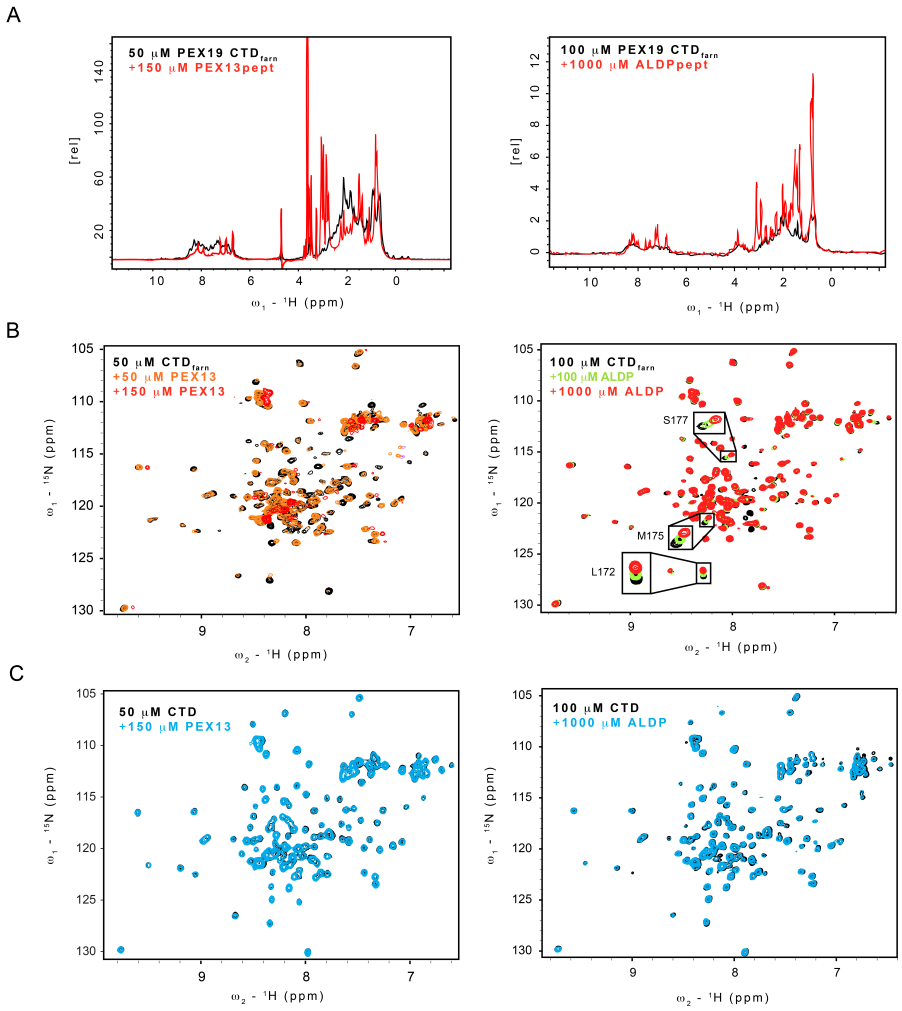


Figure 28. NMR titrations of PMP to Pex19 CTD. A) 1D proton spectra of Pex19 CTD$_{farn}$ supplemented with Pex13 peptide (left) or ALDP peptide (right). ALDP peptide was possible to reach much higher concentration in comparison to Pex13, due to better solubility. B) 2D version of the titrations as in A. Rapid line broadening in the case of Pex13 (left) titration leads to almost completely disappearing of the signals. Fast exchange chemical shifts perturbations are shown in ALDP (right) titration. Residues with non-overlapping peaks and significant perturbations are depicted in separate zoomed-in rectangles. C) NMR titrations of non-farnesylated protein with Pex13 peptide (left) and ALDP peptide (right) does not exhibit any observable chemical shift perturbation.

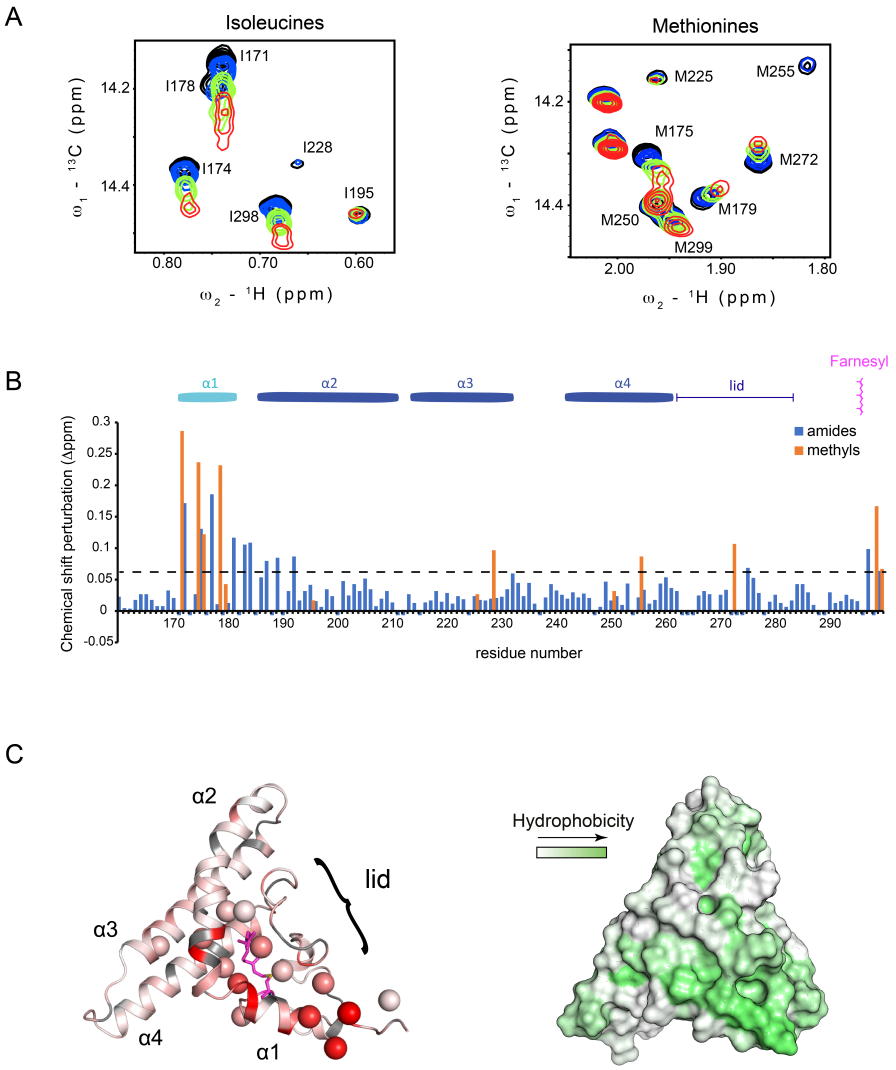


Figure 29. Mapping of the interaction of Pex19 CTD$_{farn}$ surface. A) ^{1}H,^{13}C HSQC titration of deuterated Ile,Met methyl labeled protein with ALDP peptide. Two spectra correspond to regions where labeled methyls of the protein reside. Black 100 mM Pex19 CTD$_{farn}$, blue, green, red +100 µM, +300 µM and 1000 µM respectively. B) Chemical shift perturbations of amides and methyls of the protein upon addition of ten times excess of ALDP peptide plotted against the residue numbers. Secondary structure and farnesylation site is shown above the plot. C) (left) Cartoon of Pex19 CTD$_{farn}$ colored with red residues with the largest CSPs. Ile and Met methyls are depicted as spheres. (right) Same orientation as C of the protein colored based on hydrophobicity. The PMP interface mapping correlates with an increased hydrophobicity region of the protein consisted of helix α1 and the lid.

groups. In addition, unique resonances corresponding to aromatic protons are also involved in the interaction. These resonances could be unambiguously assigned to the phenylalanine aromatic ring of the ALDP peptide. Unfortunately, assignment of the strong aliphatic NOEs is impossible due to signal overlap of the methyl-containing residues of the peptide. Nevertheless, all NOEs confirm a number of contacts involving aliphatic and aromatic protons of the PMP, that establish a specific complex with Pex19 CTD$_{farn}$. Selectively Leu and Val methyl labeling of the protein in deuterated background, considerably decreased signal overlap in the methyl region of ω1 ^{13}C-edited, ω3 ^{13}C,^{15}N filtered NOESY-HSQC spectrum. Hence, intermolecular NOEs, which originated from the ALDP phenylalanine, could be unambiguously assigned to Leu182 or Leu183 of protein.

3.3.3. Conformation of ALDP peptide

In order to determine the bound conformation of the peptide, we repeated the above experiments with excess of protein. A ω1-, ω2- filtered NOESY experiment was recorded to observe only the bound saturated form of the peptide. Unfortunately, due to instability of the complex for more than one day, it was impossible to obtain reasonable quality spectra. Therefore, we examined the conformation of the peptide in the free form. All proton and carbon frequencies were unambiguously assigned using combination of TOCSY and NOESY spectra (Fig. 30C), except leucine residues. ^{13}C secondary chemical shifts of the free peptide indicate preformed α-helical conformation (Fig. 31A). Despite the ambiguity in leucine residues assignments, due to severe signal overlap, 5 out of 8 leucine-candidate resonances also infer presence of α-helix (Fig. 31B). Therefore, this was used to build a model of ALDP (performed by Dr. K. Tripsianes) and dock it to Pex19 CTD$_{farn}$.

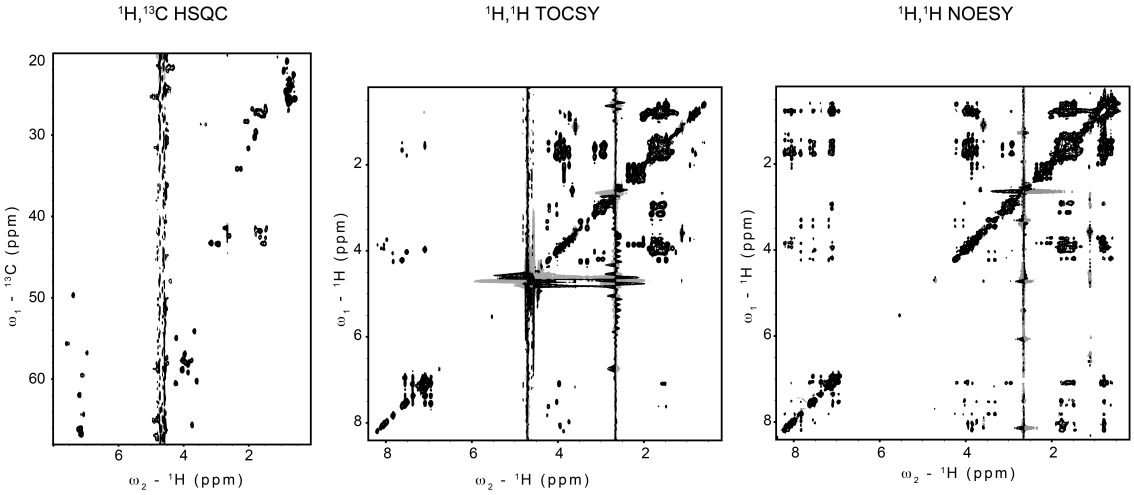


Figure 30. NMR assignments of the ALDP peptide. Natural abundance ^{1}H,^{13}C HSQC, ^{1}H,^{1}H TOCSY and ^{1}H,^{1}H NOESY spectra were recorded on a 1 mM ALDP peptide in NMR buffer. Peaks originated from intra-residue proton-proton correlations where unambiguously assigned by the TOCSY experiment, while through space correlations by NOESY. Carbon resonances where then assigned in carbon HSQC using previously known proton resonances.

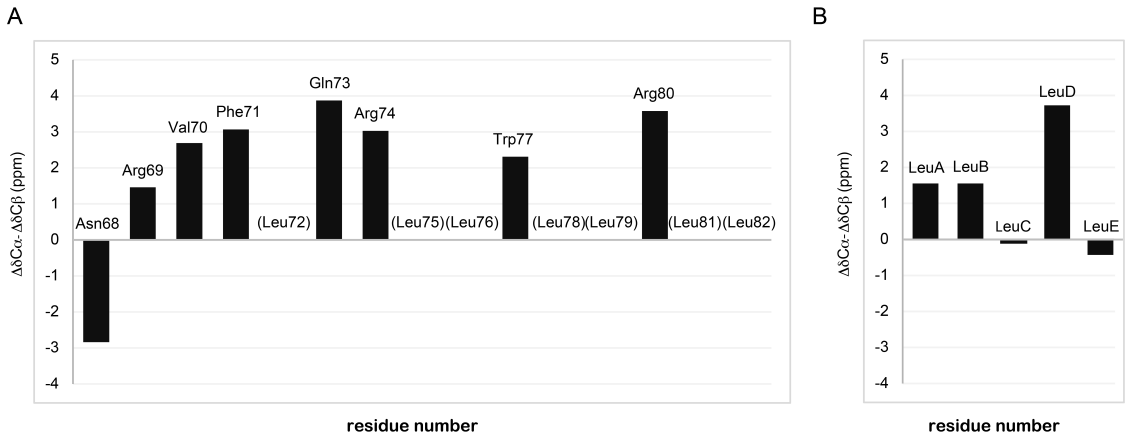


Figure 31. Secondary structure of ALDP peptide. A) Chemical shift difference between Cα and Cβ plotted for each residue. B) Same plot for potential 5 leucine spin systems. Due to severe signal overlap, the spin systems were impossible to unambiguously assign to specific residues.

3.3.4. Docking of the ALDP peptide to Pex19 CTD$_{farn}$

All intermolecular NOEs and CSPs were used to restrain docking of ALDP on protein by HADDOCK software. Peptide's helical model and the Pex19 CTD$_{farn}$ structures were used as input structures. The docking yielded two clusters of structures. The two clusters differ mainly in the orientation of the peptide relative to the protein, since the phenylalanine ring is the only anchor point to the protein. HADDOCK resulted in 95% of the water-refined structures in two clusters, indicating successful converging. In the two best scoring clusters, 2 and 5, ALDP Phe71 ring was placed in adjacent positions between helix α1 and the lid. As both clusters satisfy the ambiguous intermolecular NOEs to Pro273 or Pro274 and Leu183 both solution were found plausible. The experimental data cannot exclude cluster 5 solution, despite cluster 2 better statistics (Table 9). In fact, in the context of much larger PMP, where more hydrophobic residues are present, two hydrophobic side chains spaced by one helical turn can be accommodated in both cavities (Fig. 32).

Table 9. Pex19-ALDPpeptide HADDOCK model statistics

HADDOCK statistics		
Parameter	Cluster2	Cluster5
HADDOCK score	-38.8	-33.7
Cluster size	19	7
i-RMSD[a]	0.6 Å	4.15 Å
l-RMSD[a]	3 Å	28 Å
fnat[c]	0.85	0.47
VdW[d]	-23.2 kcal/mol	-20.9 kcal/mol
Elec[e]	-111.7 kcal/mol	-131.9 kcal/mol
Violation[f]	5.2 kcal/mol	11.7 kcal/mol
BSA[g]	725.5 Å^2	716.1 Å^2

[a] interface root mean square deviation

[b] ligand root mean square deviation

[c] fraction of native contacts

[d] Van der Waals contribution to intermolecular energies

[e] electrostatic contribution to intermolecular energies

[f] restrains violation energy

[g] buried surface area

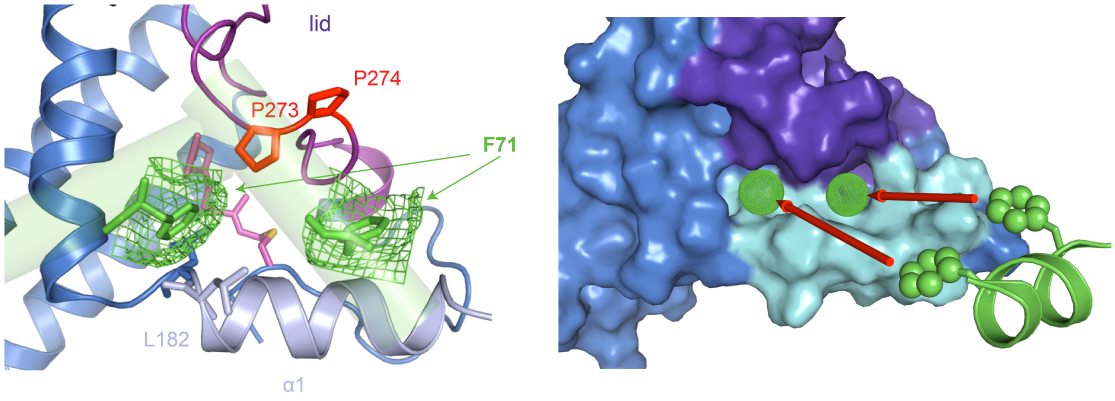


Figure 32. HADDOCK model. (left) Cartoon representation of Pex19 CTD$_{farn}$. The docking result with two different positions of green ALDP peptide is visible. Phenylalanine ring (F71) in both clusters is shown in sticks and the occupied area in mesh. Protein side chains with intermolecular NOEs are also marked and highlighted in sticks.

Figure 32 *(previous page).* (right) Surface representation of Pex19 CTD$_{farn}$ with the two hydrophobic cavities depicted as green spheres. Helix α1 and lid are cyan and purple respectively. The hypothetical PMP helix containing two hydrophobic side chains, demonstrates that both can bind in the pockets at the same time.

3.3.5. Mislocalization of PMPs in vivo (in collaboration with R.Erdmann lab)

The physiological relevance of the CTD$_{farn}$-PMP complex model was tested by mutational analysis in vivo. Although helix α1 has been linked with PMP binding *in vitro* and in general peroxisome function in cells (Schueller et al., 2010), the lid region has not been investigated. Thus we investigated the contributions of two residues residing at the tip of the lid in close proximity to the helix α1, Pro273 and Pro274, in PMP binding. The two prolines were substituted to arginines or phenylalanines. Phenylalanine mutations were introduced on the hypothesis that they will occupy the hydrophobic PMP binding cavities and hinder binding in cis. As control mutations, arginines were selected as large side chain residues but with hydrophilic and flexible properties. Human fibroblast cells lacking Pex19 were transfected with wildtype or variants of Pex19 for expression. GFP-PTS1 marker protein displays cytosolic mislocalization in non functionally complemented cells and diffuse immunolabelling of the membrane protein Pex14. While 72 hours post transfection, 50% of cells with wild type Pex19 exhibit the distinct punctuate pattern of both marker proteins, P273F/P274F mutant displayed much less complementation rate of only 20% (Fig. 33A,B). The effect was not due to differences in protein expression level (Fig. 33C). As this experiment was in cells, which many PMP have to be inserted in the peroxisomal membrane via the Pex19 pathway, our findings demonstrate that the lid region is critical indeed not only for ALDP case but other PMPs too.

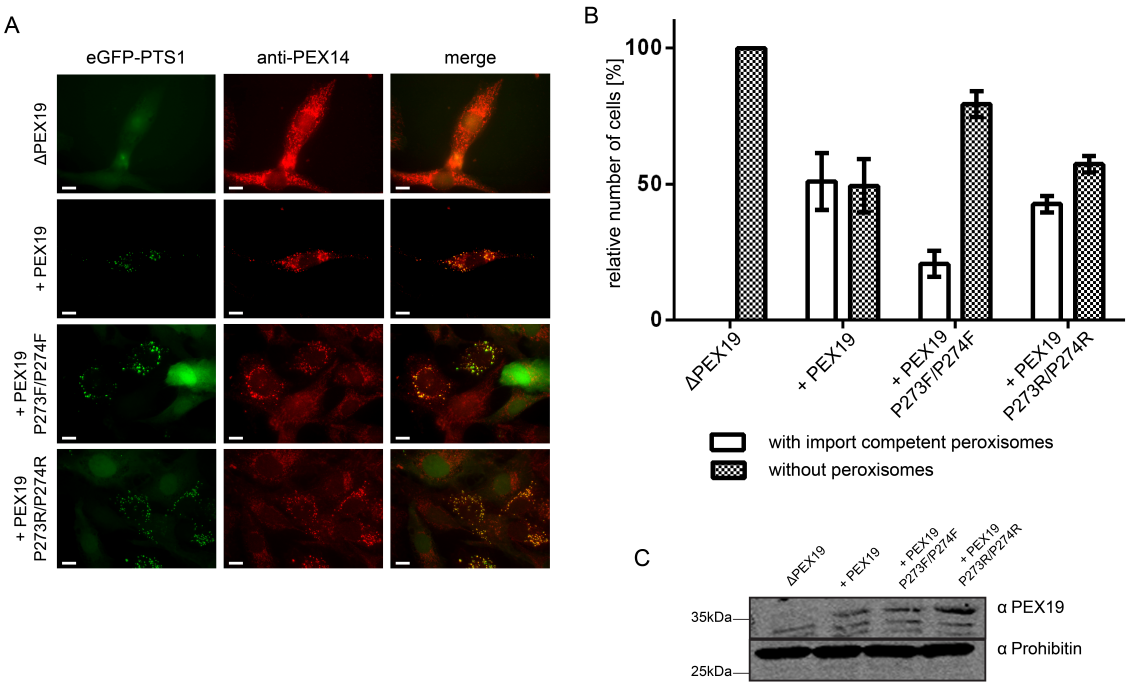


Figure 33. In vivo validation of the HADDOCK model. A) Immunofluorescence microscopy of Pex19 deficient fibroblasts transfected with bicistronic vectors expressing eGFP-PTS1 and Pex19 variants. ΔPex19 plasmid was also used as control, where diffuse staining for both eGFP-PTS1 and Pex14 is expected due to functional peroxisome absence. Punctuate pattern of both probes indicates formation of functional organelle. Scale bar: 10 µM. B) Statistical analysis of number of cells containing import competent peroxisomes. Three independent transfection experiments, 100 cells each. C) Immunoblot analysis of expression levels of all proline mutants show similar yields, similar to wildtype. Data by J. Radke.

3.3.6. Discussion: Role of farnesylation in PMP recognition

Although the farnesylation site and the modification of Pex19 is conserved across species, no study has assigned any functional role to that. In the past, the hydrophobic nature of the farnesyl group had triggered thoughts about potential role in membrane insertion (Rucktaschel et al., 2009). The solution NMR structure of the modified protein displays clearly the farnesyl group buried inside a hydrophobic cavity (Schuetz, 2012). The exposure of the group in order to become embedded in the membrane would require a huge amount of energy to overcome large conformational changes, although further experiments are required to prove this.

Comparison of the crystal (non-farnesylated) and NMR (farnesylated) structure reveals already major rearrangements in "hot spots" of the protein (Fig. 27). When the protein is unmodified, helix α1 is flexible relatively to the rest helical bundle. The exact position in the crystal structure is a result of crystal packing. In fact solution relaxation data (Schuetz, 2012) show that α1 is not preformed prior to farnesylation. The other "hot spot" of the protein, the lid, strikes out when compared between the two structures. In the case of the farnesylated protein, it creates a large hydrophobic area together with α1 helix. Combining these data, and in vivo mutational analysis, one could imagine that farnesylation is involved in PMP recognition.

Indeed the docking model presented here, based on intermolecular contacts between the farnesylated protein and a PMP-derived peptide, suggests that farnesylation regulates PMP recognition in an allosteric manner. Stabilization of elements, like helix α1 and the lid, in the correct geometry forge the perfect hydrophobic interface to accommodate the very lipophilic substrates. The functional significance of this allosteric control of PMP binding is validated by our mutational analysis in vivo. All our findings are consistent with previous reports highlighting the significance of α1 helix in PMP localization to peroxisomes (Schueller et al.,

2010). Additionally we have characterized the lid region as part of the interface of the PMP interaction.

In conclusion, the findings observed suggest that protein farnesylation could not only affect membrane localization but may also exhibit more general and novel role by allosteric regulation of ligand binding.

---

4. Conclusions and outlook

This study investigated the molecular details of processes involved in peroxisomal biogenesis. Solution NMR spectroscopy combined with other biophysical methods were applied on peroxins involved in both the peroxisomal import and the membrane insertion pathway. New interaction site between Pex14 and Pex5 is reported and it is characterized structurally. Regarding the membrane insertion pathway biochemical approaches have provided new findings on PEX19-PMP recognition. Finally, novel Pex14/Pex5 inhibitor for Trypanosoma parasite was developed based on solution NMR structure.

Even though Pex14 and Pex5 are the most well studied peroxins we still lack knowledge regarding their interplay on the peroxisomal membrane. Here one more interaction site between them is reported, yet the functional role remains unknown. Affinity measurements indicate that the new interaction is related at the later stage of the translocation, altough more research is required to proof this. Using the new Pex5 mutants should be possible in the future to study the translocon at different stages both in vivo and in vitro.

In the membrane insertion pathway, while there is indication that helix $\alpha 1$ is important for interactions with PMP cargo (Schueller et al., 2010) the exact binding site in Pex19 remained unkown. The results presented here show that the interface formed by helix $\alpha 1$ and the "lid" region is the binding site for PMPs. In agreement with previous observation that farnesylation enhances PMP binding affinity, these structure elements of the Pex19 CTD domain rearrange upon farnesylation to form a large hydrophobic cavity. This allosteric regulation of PMP

binding is important for Pex19-PMP interaction, though it is not the full picture. As farnesylation is only part of the post-translational modification of Pex19, in the future studies should aim on the role of the carboxylmethylation. Additionaly, most studies handle Pex19 as a soluble protein, including the carboxymethylated protein in the presence of the membrane protein Pex3 would reveal additional information regarding the structural and energetical interplay at the membrane.

The structure based drug design project led to new Pex14/Pex5 inhibitors for the causative agent of sleeping sickness. Although the inhibitors are not drugs, yet they exhibit remarkable antiparasitic activity, comparable to commercially available drugs. All species of the genus Trypanosoma kill thousands but as ”neglected disease” the pipeline for drug development is empty. Therefore, the new peroxisomal inihibitors should be tested in vivo and evaluated upon further solubility and bioavailability optimization.

Bibliography

Albertini, M., Rehling, P., Erdmann, R., Girzalsky, W., Kiel, J. A., Veenhuis, M., & Kunau, W. H. (1997). Pex14p, a peroxisomal membrane protein binding both receptors of the two pts-dependent import pathways. *Cell*, *89*, 83–92.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.*, *215*(3), 403–10.

Azevedo, J. E., & Schliebs, W. (2006). Pex14p, more than just a docking protein. *BBA*, *1763*, 1574–1584.

Baker, A., & Paudyal, R. (2014). The life of the peroxisome: from birth to death. *Curr. Opin. Plant Biol.*, *22*, 39–47.

Balaram, P., Bothner-By, A. A., & Dadok, J. (1972). Negative nuclear overhauser effects as probes of macromolecular structure. *J. Am. Chem. Soc.*, *94*(11), 4015–7.

Bax, A., Kontaxis, G., & Tjandra, N. (2001). Dipolar couplings in macromolecular structure determination. *Methods Enzymol.*, *339*, 127–74.

Bayer, A. M., Hunter, G. C., Gilman, R. H., Cornejo Del Carpio, J. G., Naquira, C., Bern, C., & Levy, M. Z. (2009). Chagas disease, migration and community settlement patterns in arequipa, peru. *PLoS Negl. Trop. Dis.*, *3*(12), e567.

Bharti, P., Schliebs, W., Schievelbusch, T., Neuhaus, A., David, C., Kock, K., Hermann, C., Meyer, H. E., Wiese, S., Warscheid, B., Theiss, C., & Erdmann, R.

(2011). Pex14 is required for microtubule-based peroxisome motility in human cells. *J. Cell Sci.*, *124*, 1759–1768.

Bottger, G., Barnett, P., Klein, A. T., Kragt, A., Tabak, H. F., & Distel, B. (2000). Saccharomyces cerevisiaepts1 receptor pex5p interacts with the sh3 domain of the peroxisomal membrane protein pex13p in an unconventional, non-pxxp-related manner. *Mol. Biol. Cell*, *11*, 3963–76.

Brocard, C., Lametschwandtner, G., Koudelka, R., & Hartig, A. (1997). Pex14p is a member of the protein linkage map of pex5p. *EMBO J.*, *16*, 5491–5500.

Brul, S., Westerveld, A., Strijland, A., Wanders, R. J. A., Schram, A. W., Heymans, H. S. A., Schutgens, R. B. H., Van Den Bosch, H., & Tager, J. M. (1988). Genetic heterogeneity in the cerebrohepatorenal (zellweger) syndrome and other inherited disorders with a generalized impairment of peroxisomal functions. a study using complementation analysis. *J. Clin. Invest.*, *81*, 1710–1715.

Carvalho, A. F., Costa-Rodrigues, J., Correia, I., Pessoa, J. C., Faria, T. Q., Martins, C. L., Fransen, M., Sa-Miranda, C., & Azevedo, J. E. (2006). The n-terminal half of the peroxisomal cycling receptor pex5p is a natively unfolded domain. *J. Mol. Biol.*, *356*, 864–875.

Carvalho, A. F., Pinto, M. P., Grou, C. P., Alencastre, I. S., Fransen, M., Sá-Miranda, C., & Azevedo, J. E. (2007). Ubiquitination of mammalian pex5p, the peroxisomal import receptor. *J. Biol. Chem.*, *282*, 31267–31272.

De Duve, C., & Baudhuin, P. (1966). Peroxisomes (microbodies and related particles). *Physiol. Rev.*, *46*(2), 323–57.

de Duve, D. (1969). The peroxisome: a new cytoplasmic organelle. *Proc R Soc Lond B Biol Sci*, *173*(1030), 71–83.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995). Nmrpipe: a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR*, *6*, 277–293.

Dominguez, C., Boelens, R., & Bonvin, A. M. (2003). Haddock: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, *125*(7), 1731–7.

Dosset, P., Hus, J. C., Marion, D., & Blackledge, M. (2001). A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J. Biomol. NMR*, *20*(3), 223–31.

Dujon, B., Alexandraki, D., André, B., Ansorge, W., Baladron, V., Ballesta, J. P., Banrevi, A., Bolle, P. A., Bolotin-Fukuhara, M., & Bossier, P. (1994). Complete dna sequence of yeast chromosome xi. *Nature*, *369*, 371–378.

Emmanouilidis, L., Gopalswamy, M., Passon, D. M., Wilmanns, M., & Sattler, M. (2016). Structural biology of the import pathways of peroxisomal matrix proteins. *BBA*, *1863*(5), 804–13.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of coot. *Acta Crystallogr. D*, *66*(Pt 4), 486–501.

Erdmann, R., & Schliebs, W. (2005). Peroxisomal matrix protein import: the transient pore model. *Nat. Rev. Mol. Cell. Biol.*, *6*(9), 738–42.

Evert RF, E. S. (2006). *Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development*. John Wiley Sons.

Fang, Y., Morrell, J. C., Jones, J. M., & Gould, S. J. (2004). Pex3 functions as a pex19 docking factor in the import of class i peroxisomal membrane proteins. *J. Cell Biol.*, *164*(6), 863–75.

Fransen, M., Wylin, T., Brees, C., Mannaerts, G. P., & Van Veldhoven, P. P. (2001). Human pex19p binds peroxisomal integral membrane proteins at regions distinct from their sorting sequences. *Mol. Cell. Biol.*, *21*(13), 4413–24.

Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-f., Dougherty, B. A., Bott, K. F., Hu, P.-c., Lucier, T. S., Peterson, S. N., Smith, H., Hutchison, C. A., & Venter, J. C. (1995). The minimal gene complement of myco plasma genitalium. *Science*, *270*, 397–403.

Freitas, M. O., Francisco, T., Rodrigues, T. A., Alencastre, I. S., Pinto, M. P., Grou, C. P., Carvalho, A. F., Fransen, M., Sá-Miranda, C., & Azevedo, J. E. (2011). Pex5 protein binds monomeric catalase blocking its tetramerization and releases it upon binding the n-terminal domain of pex14. *J. Biol. Chem.*, *286*, 40509–40519.

Fujiki, Y., Matsuzono, Y., Matsuzaki, T., & Fransen, M. (2006). Import of peroxisomal membrane proteins: the interplay of pex3p- and pex19p-mediated interactions. *BBA*, *1763*(12), 1639–46.

Furuya, T., Kessler, P., Jardim, A., Schnaufer, A., Crudder, C., & Parsons, M. (2002). Glucose is toxic to glycosome-deficient trypanosomes. *Proc. Natl. Acad. Sci. USA*, *99*(22), 14177–82.

Gatto, G. J., Geisbrecht, B. V., Gould, S. J., Berg, J. M., Gatto Jr., G. J., Geisbrecht, B. V., Gould, S. J., & Berg, J. M. (2000). Peroxisomal targeting signal-1 recognition by the tpr domains of human pex5. *Nat. Struct. Biol.*, *7*, 1091–1095.

Gatto, G. J., Maynard, E. L., Guerrerio, A. L., Geisbrecht, B. V., Gould, S. J., & Berg, J. M. (2003). Correlating structure and affinity for pex5:pts1 complexes. *Biochemistry*, *42*, 1660–1666.

Girzalsky, W., Saffian, D., & Erdmann, R. (2010). Peroxisomal protein translocation. *BBA*, *1803*(6), 724–31.

Glover, J. R., Andrews, D. W., & Rachubinski, R. A. (1994). Saccharomyces cerevisiae peroxisomal thiolase is imported as a dimer. *Proc. Natl. Acad. Sci. USA*, *91*, 10541–10545.

Goddard, T. D., & Kneller, D. G. (1996). Sparky 3.

Gootjes, J., Mooijer, P. A., Dekker, C., Barth, P. G., Poll-The, B. T., Waterham, H. R., & Wanders, R. J. (2002). Biochemical markers predicting survival in peroxisome biogenesis disorders. *Neurology*, *59*(11), 1746–9.

Gould, S. J., Keller, G. A., & Subramani, S. (1987). Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *J. Cell Biol.*, *105*, 2923–2931.

Guntert, P. (2009). Automated structure determination from nmr spectra. *Eur. Biophys. J.*, *38*(2), 129–43.

Haanstra, J. R., van Tuijl, A., Kessler, P., Reijnders, W., Michels, P. A., Westerhoff, H. V., Parsons, M., & Bakker, B. M. (2008). Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes. *Proc. Natl. Acad. Sci. USA*, *105*(46), 17718–23.

Halbach, A., Landgraf, C., Lorenzen, S., Rosenkranz, K., Volkmer-Engert, R., Erdmann, R., & Rottensteiner, H. (2006). Targeting of the tail-anchored peroxisomal membrane proteins pex26 and pex15 occurs through c-terminal pex19-binding sites. *J. Cell Sci.*, *119*(Pt 12), 2508–17.

Hettema, E. H., Erdmann, R., van der Klei, I., & Veenhuis, M. (2014). Evolving models for peroxisome biogenesis. *Curr. Opin. Cell Biol.*, *29*, 25–30.

Hoepfner, D., Schildknegt, D., Braakman, I., Philippsen, P., & Tabak, H. F. (2005). Contribution of the endoplasmic reticulum to peroxisome formation. *Cell*, *122*(1), 85–95.

Honsho, M., Hiroshige, T., & Fujiki, Y. (2002). The membrane biogenesis peroxin pex16p. topogenesis and functional roles in peroxisomal membrane assembly. *J. Biol. Chem.*, *277*(46), 44513–24.

Hruban, Z., Vigil, E., Slesers, A., & Hopkins, E. (1972). Microbodies: constituent organelles of animal cells. *Lab. Invest.*, *27*, 184–91.

Huhse, B., Rehling, P., Albertini, M., Blank, L., Meller, K., & Kunau, W. H. (1998). Pex17p of saccharomyces cerevisiae is a novel peroxin and component of the peroxisomal protein translocation machinery. *J. Cell Biol.*, *140*(1), 49–60.

International Human Genome Sequencing, C. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, *431*(7011), 931–45.

Jedd, G., & Chua, N. H. (2000). A new self-assembled peroxisomal vesicle required for efficient resealing of the plasma membrane. *Nat. Cell Biol.*, *2*(4), 226–31.

Johnson, M. A., Snyder, W. B., Cereghino, J. L., Veenhuis, M., Subramani, S., & Cregg, J. M. (2001). Pichia pastoris pex14p, a phosphorylated peroxisomal membrane protein, is part of a pts-receptor docking complex and interacts with many peroxins. *Yeast*, *18*, 621–641.

Jones, J. M., Morrell, J. C., & Gould, S. J. (2001). Multiple distinct targeting signals in integral peroxisomal membrane proteins. *J. Cell Biol.*, *153*(6), 1141–50.

Jones, J. M., Morrell, J. C., & Gould, S. J. (2004). Pex19 is a predominantly cytosolic chaperone and import receptor for class 1 peroxisomal membrane proteins. *J. Cell Biol.*, *164*(1), 57–67.

Jung, Y. S., & Zweckstetter, M. (2004). Mars – robust automatic backbone assignment of proteins. *J. Biomol. NMR*, *30*(1), 11–23.

Kabsch, W. (2010). Xds. *Acta Crystallogr. D*, *66*(Pt 2), 125–32.

Kadkhodaie, K., Rivas, O., Tan, M., Mohebbi, A., & Shaka, A. (1991). Broadband homonuclear cross polarization using flip-flop spectroscopy. *J. Magn. Reson.*, *91*, 437–443.

Keller, G. A., Gould, S., Deluca, M., & Subramani, S. (1987). Firefly luciferase is targeted to peroxisomes in mammalian cells. *Proc. Natl. Acad. Sci. USA*, *84*, 3264–3268.

Kemp, S., & Wanders, R. J. (2007). X-linked adrenoleukodystrophy: very long-chain fatty acid metabolism, abc half-transporters and the complicated route to treatment. *Mol. Genet. Metab.*, *90*(3), 268–76.

Kessler, P. S., & Parsons, M. (2005). Probing the role of compartmentation of glycolysis in procyclic form trypanosoma brucei: Rna interference studies of pex14, hexokinase, and phosphofructokinase. *J. Biol. Chem.*, *280*(10), 9030–6.

Kiel, J. A., Otzen, M., Veenhuis, M., & van der Klei, I. J. (2005). Obstruction of polyubiquitination affects pts1 peroxisomal matrix protein import. *BBA*, *1745*, 176–186.

Kim, P. K., & Hettema, E. H. (2015). Multiple pathways for protein transport to peroxisomes. *J. Mol. Biol.*, *427*, 1176–1190.

Kim, P. K., Mullen, R. T., Schumann, U., & Lippincott-Schwartz, J. (2006). The origin and maintenance of mammalian peroxisomes involves a de novo pex16-dependent pathway from the er. *J. Cell Biol.*, *173*(4), 521–32.

Kleckner, I. R., & Foster, M. P. (2011). An introduction to nmr-based approaches for measuring protein dynamics. *BBA*, *1814*(8), 942–68.

Kragt, A., Voorn-Brouwer, T., van den Berg, M., & Distel, B. (2005). Endoplasmic reticulum-directed pex3p routes to peroxisomes and restores peroxisome formation in a saccharomyces cerevisiae pex3delta strain. *J. Biol. Chem.*, *280*(40), 34350–7.

Kristjanson, P., Swallow, B., Rowlands, G., Kruska, R., & DeLeeuw, P. (1999). Measuring the costs of aafrican animal trypanosomaisis: the potential benefits of control and returns to research. *The Agriculture Systems*, *59*, 79–98.

Lametschwandtner, G., Brocard, C., Fransen, M., Van Veldhoven, P., Berger, J., & Hartig, A. (1998). The difference in recognition of terminal tripeptides as peroxisomal targeting signal 1 between yeast and human is due to different affinities of their receptor pex5p to the cognate signal and to residues adjacent to it. *J. Biol. Chem.*, *273*, 33635–33643.

Lanyon-Hogg, T., Warriner, S. L., & Baker, A. (2010). Getting a camel through the eye of a needle: the import of folded proteins by peroxisomes. *Biology of the cell / under the auspices of the European Cell Biology Organization*, *102*, 245–263.

Ma, C., Hagstrom, D., Polley, S. G., & Subramani, S. (2013). Redox-regulated cargo binding and release by the peroxisomal targeting signal receptor, pex5. *J. Biol. Chem.*, *288*, 27220–27231.

Madrid, K. P., De Crescenzo, G., Wang, S., & Jardim, A. (2004). Modulation of the leishmania donovani peroxin 5 quaternary structure by peroxisomal targeting signal 1 ligands. *Mol. Cell. Biol.*, *24*, 7331–7344.

Mast, F. D., Fagarasanu, A., & Rachubinski, R. (2010). The peroxisomal protein importomer: a bunch of transients with expanding waistlines. *Nat. Cell. Biol.*, *12*, 203–205.

Matsumura, T., Otera, H., & Fiyiki, Y. (2000). Disruption of the interaction of the longer isoform of pex5p, pex5pl, with pex7p abolishes peroxisome targeting sig-

nal type 2 protein import in mammals. study with a novel pex5-impaired chinese hamster ovary cell mutant. *J. Biol. Chem.*, *275*, 21715–21721.

Meinecke, M., Cizmowski, C., Schliebs, W., Krüger, V., Beck, S., Wagner, R., & Erdmann, R. (2010). The peroxisomal importomer constitutes a large and highly dynamic pore. *Nat. Cell. Biol.*, *12*, 273–277.

Motley, A. M., & Hettema, E. H. (2007). Yeast peroxisomes multiply by growth and division. *J. Cell Biol.*, *178*(3), 399–410.

Mukai, S., & Fujiki, Y. (2006). Molecular mechanisms of import of peroxisome-targeting signal type 2 (pts2) proteins by pts2 receptor pex7p and pts1 receptor pex5pl. *J. Biol. Chem.*, *281*, 37311–37320.

Neri, D., Szyperski, T., Otting, G., Senn, H., & Wuthrich, K. (1989). Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the dna-binding domain of the 434 repressor by biosynthetically directed fractional 13c labeling. *Biochemistry*, *28*(19), 7510–6.

Neufeld, C., Filipp, F. V., Simon, B., Neuhaus, A., Schuller, N., David, C., Kooshapur, H., Madl, T., Erdmann, R., Schliebs, W., Wilmanns, M., & Sattler, M. (2009). Structural basis for competitive interactions of pex14 with the import receptors pex5 and pex19. *EMBO J.*, *28*(6), 745–54.

Neuhaus, A., Kooshapur, H., Wolf, J., Helge Meyer, N., Madl, T., Saidowsky, J., Hambruch, E., Lazam, A., Jung, M., Sattler, M., Schliebs, W., Erdmann, R., Meyer, N. H., Madl, T., Saidowsky, J., Hambruch, E., Lazam, A., Jung, M., Sattler, M., Schliebs, W., & Erdmann, R. (2014). A novel pex14 protein-interacting site of human pex5 is critical for matrix protein import into peroxisomes. *J. Biol. Chem.*, *289*, 437–448.

Niederhoff, K., Meindl-Beinker, N. M., Kerssen, D., Perband, U., Schafer, A.,

Schliebs, W., & Kunau, W. H. (2005). Yeast pex14p possesses two functionally distinct pex5p and one pex7p binding sites. *J. Biol. Chem.*, *280*(42), 35571–8.

Nilges, M., Macias, M. J., O'Donoghue, S. I., & Oschkinat, H. (1997). Automated noesy interpretation with ambiguous distance restraints: the refined nmr solution structure of the pleckstrin homology domain from beta-spectrin. *J. Mol. Biol.*, *269*(3), 408–22.

Oliver, S. G., van der Aart, Q. J., Agostoni-Carbone, M. L., Aigle, M., Alberghina, L., Alexandraki, D., Antoine, G., Anwar, R., Ballesta, J. P., & Benit, P. (1992). The complete dna sequence of yeast chromosome iii. *Nature*, *357*, 38–46.

Opperdoes, F. R., & Borst, P. (1977). Localization of nine glycolytic enzymes in a microbody-like organelle in trypanosoma brucei: the glycosome. *FEBS Lett.*, *80*(2), 360–4.

Otera, H., Harano, T., Honsho, M., Ghaedi, K., Mukai, S., Tanaka, A., Kawai, A., Shimizu, N., & Fujiki, Y. (2000). The mammalian peroxin pex5pl, the longer isoform of the mobile peroxisome targeting signal (pts) type 1 transporter, translocates the pex7p.pts2 protein complex into peroxisomes via its initial docking site, pex14p. *J. Biol. Chem.*, *275*, 21703–21714.

Otera, H., Setoguchi, K., Hamasaki, M., Kumashiro, T., Shimizu, N., & Fujiki, Y. (2002). Peroxisomal targeting signal receptor pex5p interacts with cargoes and import machinery components in a spatiotemporally differentiated manner: conserved pex5p wxxxf/y motifs are critical for matrix protein import. *Mol. Cell. Biol.*, *22*, 1639–1655.

Perry, B., & Sones, K. (2007). Science for development. poverty reduction through animal health. *Science*, *315*(5810), 333–4.

Platta, H. W., Hagen, S., & Erdmann, R. (2013). The exportomer: the peroxisomal receptor export machinery. *Cell Mol. Life Sci.*, *70*, 1393–1411.

Post, C. B. (2003). Exchange-transferred noe spectroscopy and bound ligand structure determination. *Curr. Opin. Struct. Biol.*, *13*(5), 581–8.

Purdue, P. E., & Lazarow, P. B. (2001). Peroxisome biogenesis. *Annu. Rev. Cell Develop. Biol.*, *17*, 701–752.

Purdue, P. E., Yang, X., & Lazarow, P. B. (1998). Pex18p and pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the pts2 pathway. *J. Cell Biol.*, *143*, 1859–1869.

Rehling, P., Marzioch, M., Niesen, F., Wittke, E., Veenhuis, M., & Kunau, W. H. (1996). The import receptor for the peroxisomal targeting signal 2 (pts2) in saccharomyces cerevisiae is encoded by the pas7 gene. *EMBO J..*, *15*, 2901–2913.

Requena-Mendez, A., Aldasoro, E., de Lazzari, E., Sicuri, E., Brown, M., Moore, D. A., Gascon, J., & Munoz, J. (2015). Prevalence of chagas disease in latin-american migrants living in europe: a systematic review and meta-analysis. *PLoS Negl. Trop. Dis.*, *9*(2), e0003540.

Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., Linka, N., Switzenberg, R., Wilkerson, C. G., Weber, A. P. M., Olsen, L. J., & Hu, J. (2009). In-depth proteome analysis of arabidopsis leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol.*, *150*, 125–143.

Rucktaschel, R., Girzalsky, W., & Erdmann, R. (2011). Protein import machineries of peroxisomes. *BBA*, *1808*, 892–900.

Rucktaschel, R., Thoms, S., Sidorovitch, V., Halbach, A., Pechlivanis, M., Volkmer, R., Alexandrov, K., Kuhlmann, J., Rottensteiner, H., & Erdmann, R. (2009). Farnesylation of pex19p is required for its structural integrity and function in peroxisome biogenesis. *J. Biol. Chem.*, *284*(31), 20885–96.

Sacksteder, K. A., Jones, J. M., South, S. T., Li, X., Liu, Y., & Gould, S. J. (2000). Pex19 binds multiple peroxisomal membrane proteins, is predominantly cytoplasmic, and is required for peroxisome membrane synthesis. *J. Cell Biol.*, *148*(5), 931–44.

Saidowsky, J., Dodt, G., Kirchberg, K., Wegner, A., Nastainezyk, W., Kunau, W. H., & Schliebs, W. (2001). The di-aromatic pentapeptide repeats of the human peroxisome import receptor pex5 are separate high affinity binding sites for the peroxisomal membrane protein pex14. *J. Biol. Chem.*, *276*, 34524–34529.

Salido, E., Pey, A. L., Rodriguez, R., & Lorenzo, V. (2012). Primary hyperoxalurias: disorders of glyoxylate detoxification. *BBA*, *1822*(9), 1453–64.

Sato, Y., Shibata, H., Nakatsu, T., Nakano, H., Kashiwayama, Y., Imanaka, T., & Kato, H. (2010). Structural basis for docking of peroxisomal membrane protein carrier pex19p onto its receptor pex3p. *EMBO J.*, *29*(24), 4083–93.

Sattler, M., Schleucher, J., & Griesinger, C. (1999). Heteronuclear multidimensional nmr experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc.*, *34*, 93–158.

Schell-Steven, A., Stein, K., Amoros, M., Landgraf, C., Volkmer-Engert, R., Rottensteiner, H., & Erdmann, R. (2005). Identification of a novel, intraperoxisomal pex14-binding site in pex13: association of pex13 with the docking complex is essential for peroxisomal matrix protein import. *Mol. Cell. Biol.*, *25*, 3007–3018.

Schliebs, W., Saidowsky, J., Agianian, B., Dodt, G., Herberg, F. W., & Kunau, W. H. (1999). Recombinant human peroxisomal targeting signal receptor pex5. structural basis for interaction of pex5 with pex14. *J. Biol. Chem.*, *274*, 5666–5673.

Schmidt, F., Treiber, N., Zocher, G., Bjelic, S., Steinmetz, M. O., Kalbacher, H., Stehle, T., & Dodt, G. (2010). Insights into peroxisome function from the structure of pex3 in complex with a soluble fragment of pex19. *J. Biol. Chem.*, *285*(33), 25410–7.

Schrader, M., Bonekamp, N. A., & Islinger, M. (2012). Fission and proliferation of peroxisomes. *BBA*, *1822*(9), 1343–57.

Schueller, N., Holton, S. J., Fodor, K., Milewski, M., Konarev, P., Stanley, W. A., Wolf, J., Erdmann, R., Schliebs, W., Song, Y.-H., & Wilmanns, M. (2010). The peroxisomal receptor pex19p forms a helical mpts recognition domain. *EMBO J.*., *29*, 2491–2500.

Schuetz, U. (2012). Structural and functional studies of the farnesylated peroxisomal membrane protein receptor pex19. *Dissertation*.

Shen, Y., Delaglio, F., Cornilescu, G., & Bax, A. (2009). Talos+: a hybrid method for predicting protein backbone torsion angles from nmr chemical shifts. *J. Biomol. NMR*, *44*, 213–223.

Shimizu, N., Itoh, R., Hirono, Y., Otera, H., Ghaedi, K., Tateishi, K., Tamura, S., Okumoto, K., Harano, T., Mukai, S., & Fujiki, Y. (1999). The peroxin pex14p. cdna cloning by functional complementation on a chinese hamster ovary cell mutant, characterization, and functional analysis. *J. Biol. Chem.*, *274*, 12593–12604.

Shiozawa, K., Konarev, P. V., Neufeld, C., Wilmanns, M., & Svergun, D. I. (2009). Solution structure of human pex5.pex14.pts1 protein complexes obtained by small angle x-ray scattering. *J. Biol. Chem.*, *284*, 25334–25342.

Sies, H. (1974). Biochemistry of the peroxisome in the liver cell. *Angew. Chem. Int. Ed. Engl.*, *13*(11), 706–18.

Smith, J. J., & Aitchison, J. D. (2013). Peroxisomes take shape. *Nat. Rev. Mol. Cell. Biol.*, *14*, 803–817.

Stanley, W. A., Filipp, F. V., Kursula, P., Schüller, N., Erdmann, R., Schliebs, W., Sattler, M., & Wilmanns, M. (2006). Recognition of a functional peroxisome type 1 target by the dynamic import receptor pex5p. *Mol. Cell*, *24*, 653–663.

Stanley, W. A., Pursiainen, N. V., Garman, E. F., Juffer, A. H., Wilmanns, M., & Kursula, P. (2007). A previously unobserved conformation for the human pex5p receptor suggests roles for intrinsic flexibility and rigid domain motions in ligand binding. *BMC structural biology*, *7*, 24.

Stein, K., Schell-Steven, A., Erdmann, R., & Rottensteiner, H. (2002). Interactions of pex7p and pex18p/pex21p with the peroxisomal docking machinery: implications for the first steps in pts2 protein import. *Mol. Cell. Biol.*, *22*, 6056–6069.

Su, J.-R. R., Takeda, K., Tamura, S., Fujiki, Y., & Miki, K. (2009). Crystal structure of the conserved n-terminal domain of the peroxisomal matrix protein import receptor, pex14p. *Proc. Natl. Acad. Sci. USA*, *106*, 417–421.

Subramani, S. (1998). Components involved in peroxisome import, biogenesis, proliferation, turnover, and movement. *Physiol. Rev.*, *78*(1), 171–88.

Tabak, H. F., Braakman, I., & van der Zand, A. (2013). Peroxisome formation and maintenance are dependent on the endoplasmic reticulum. *Annu. Rev. Biochem.*, *82*, 723–44.

Tam, Y. Y., Fagarasanu, A., Fagarasanu, M., & Rachubinski, R. A. (2005). Pex3p initiates the formation of a preperoxisomal compartment from a subdomain of the endoplasmic reticulum in saccharomyces cerevisiae. *J. Biol. Chem.*, *280*(41), 34933–9.

Tanaka, K., Soeda, M., Hashimoto, Y., Takenaka, S., & Komori, M. (2013). Identification of phosphorylation sites in hansenula polymorpha pex14p by mass spectrometry. *FEBS Open Bio.*, *3*, 6–10.

Terlecky, S. R., Nuttley, W. M., McCollum, D., Sock, E., & Subramani, S. (1995). The pichia pastoris peroxisomal protein pas8p is the receptor for the c-terminal tripeptide peroxisomal targeting signal. *EMBO J..*, *14*, 3627–3634.

Thoms, S., Harms, I., Kalies, K. U., & Gartner, J. (2012). Peroxisome formation requires the endoplasmic reticulum channel protein sec61. *Traffic*, *13*(4), 599–609.

Titorenko, V. I., Nicaud, J. M., Wang, H., Chan, H., & Rachubinski, R. A. (2002). Acyl-coa oxidase is imported as a heteropentameric, cofactor-containing complex into peroxisomes of yarrowia lipolytica. *J. Cell Biol.*, *156*, 481–494.

Titorenko, V. I., Ogrydziak, D. M., & Rachubinski, R. A. (1997). Four distinct secretory pathways serve protein secretion, cell surface growth, and peroxisome biogenesis in the yeast yarrowia lipolytica. *Mol. Cell. Biol.*, *17*(9), 5210–26.

Tolbert, N. E., & Essner, E. (1981). Microbodies: peroxisomes and glyoxysomes. *J. Cell Biol.*, *91*(3 Pt 2), 271s–283s.

Tugarinov, V., & Kay, L. E. (2003). Ile, leu, and val methyl assignments of the 723-residue malate synthase g using a new labeling strategy and novel nmr methods. *J. Am. Chem. Soc.*, *125*(45), 13868–78.

Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., & Murshudov, G. N. (2004). Refmac5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D*, *60*(Pt 12 Pt 1), 2184–95.

van der Klei, I. J., Yurimoto, H., Sakai, Y., & Veenhuis, M. (2006). The significance of peroxisomes in methanol metabolism in methylotrophic yeast. *BBA*, *1763*(12), 1453–62.

von Wissmann, B., Machila, N., Picozzi, K., Fevre, E. M., de, C. B. B. M., Handel, I. G., & Welburn, S. C. (2011). Factors associated with acquisition of human infective and animal infective trypanosome infections in domestic livestock in western kenya. *PLoS Negl. Trop. Dis.*, *5*(1), e941.

Walton, P. A., Hill, P. E., & Subramani, S. (1995). Import of stably folded proteins into peroxisomes. *Mol. Biol. Cell*, *6*, 675–683.

Waterham, H. R., & Ebberink, M. S. (2012). Genetics and molecular basis of human peroxisome biogenesis disorders. *BBA*, *1822*(9), 1430–41.

Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, *171*(4356), 737–8.

Will, G. K., Soukupova, M., Hong, X., Erdmann, K. S., Kiel, J. a., Dodt, G., Kunau, W. H., & Erdmann, R. (1999). Identification and characterization of the human orthologue of yeast pex14p. *Mol. Cell. Biol.*, *19*, 2265–2277.

Williams, C. P., Schueller, N., Thompson, C. A., van den Berg, M., Van Haren, S. D., Erdmann, R., Bond, C. S., Distel, B., Schliebs, W., Wilmanns, M., & Stanley, W. A. (2011). The peroxisomal targeting signal 1 in sterol carrier protein 2 is autonomous and essential for receptor recognition. *BMC biochemistry*, *12*, 12.

Zhang, X., Roe, S. M., Hou, Y., Bartlam, M., Rao, Z., Pearl, L. H., & Danpure, C. J. (2003). Crystal structure of alanine:glyoxylate aminotransferase and the relationship between genotype and enzymatic phenotype in primary hyperoxaluria type 1. *J. Mol. Biol.*, *331*, 643–652.

Appendix

A. Appendix

Table 10. Pex5 TPR backbone chemical shift assignments.

	N	H	Ca	Cb		N	H	Ca	Cb
L315	122	8.066	55.739	42.516	**L459**	126.9	8.285	55.729	42.316
T316	114.4	8.051	62.079	70.636	**G460**	109.9	8.22	45.639	
S317	117.8	8.206	59.129	64.116	**S461**	115.7	8.011	58.519	64.386
A318	125.6	8.148	53.029	19.446	**L469**	117.3	7.777	58.149	41.596
T319	112.6	7.85	62.669	70.196	**E470**	119.1	7.888	59.629	29.866
K322	122.2	7.788	56.749	32.876	**V471**	119.3	8.25	67.389	31.806
G323	109	8.088	44.609		**A478**	122.6	7.887	55.359	17.866
Y324	123.9	8.47	61.269	39.596	**V479**	118.2	7.598	66.049	32.106
F326	123.9	7.85	59.339	40.406	**R480**	115.9	7.133	57.669	31.006
E327	125.7	9.939	57.019	29.446	**L481**	120	7.162	58.509	42.086
E328	130	8.722	59.289	30.656	**D482**	112.9	7.457	51.769	41.096
E329	121.9	9.199	55.029	28.326	**T484**	108.8	7.92	62.869	69.826
N330	118.7	7.742	56.219	39.656	**S485**	117.4	7.357	57.189	65.346
L332	119.9	8.034	54.469	40.776	**I486**	124.5	8.333	60.529	38.726
R333	118.4	6.847	59.259	30.756	**D487**	128.5	8.644	50.169	42.096
D334	115.4	8.092	53.289	41.126	**D489**	119.3	7.875	58.309	42.666
H335	125.1	7.477	57.209	31.616	**V490**	121.6	6.869	66.489	32.516
F339	119.4	8.471	62.299	39.756	**Q491**	115.8	7.667	58.609	27.216
E340	118.3	8.465	60.979	28.866	**S501**	110.7	7.664	58.509	63.956

E341	120.2	7.917	58.439		**G502**	109.6	7.385	45.329	
G342	111.6	8.466	48.609		**E503**	123.5	7.975	54.439	28.486
L343	120.8	7.616	58.399	41.546	**Y504**	119.8	6.773	61.319	37.236
R344	120.2	7.095	60.149	29.976	**D505**	117.3	8.581	57.199	39.106
R345	117.6	8.143	57.899	28.636	**K506**	120.6	7.013	58.639	30.316
L346	120.7	8.34	58.549	42.546	**A507**	121.5	7.571	56.089	19.396
Q347	120.2	7.785	58.909	28.426	**V508**	119.1	8.275	67.819	31.756
E348	116.3	7.321	56.409	30.236	**D509**	121.9	7.228	58.139	40.266
G349	108.4	7.8	45.889		**C510**	120.4	7.399	63.019	27.016
D350	122.3	8.046	51.709	39.956	**F511**	118.1	8.288	61.259	
L351	121.4	7.699	59.619	39.106	**T512**	115.5	8.52	67.719	69.046
N353	111.6	6.947	55.949	38.216	**A513**	125.6	7.709	55.629	17.416
A354	123.3	7.762	56.589	18.076	**A514**	121.8	8.101	56.259	19.226
V355	115.1	8.157	68.509	31.696	**L515**	118.7	7.986	57.079	43.006
L356	117.1	7.028	58.449	40.286	**S516**	115.2	7.507	61.549	63.236
L357	122.7	7.765	58.419	39.436	**Q537**	120.9	8.254	54.369	28.026
F358	121	8.65	59.979	38.516	**S538**	116.1	7.514	63.859	63.246
E359	118.2	8.896	61.029	29.546	**E539**	122.5	8.919	60.809	29.116
A360	123.6	8.143	55.619	18.736	**E540**	121.2	8.701	59.919	29.326
A361	120.3	8.391	55.839	18.646	**A541**	124.8	8.336	56.119	18.186
V362	110	7.36	63.189	31.666	**V542**	115.5	7.752	67.219	
Q363	119.6	7.139	58.689	30.126	**A543**	120	7.005	55.519	18.126
Q364	119.2	7.755	58.719	28.936	**A544**	122.5	7.552	55.519	17.406
D365	114.4	7.682	51.549	42.086	**Y545**	117.1	8.38	60.359	38.576
K367	117.3	7.949	55.899	31.226	**D521**	120.3	7.094	54.089	42.216
H368	122.2	7.76	55.659	29.686	**G570**	108.7	7.779	45.059	
M369	131.3	8.03	61.459	30.646	**A571**	127.5	7.903	50.009	16.876
E370	119.8	9.567	60.519	29.856	**H572**	119.1	6.816	60.659	31.896

A371	121.9	7.398	55.639	18.066	**R573**	119.4	8.69	60.089	29.006
W372	118.8	7.443	61.509	29.346	**V576**	115	7.533	67.749	31.276
Q373	118.8	8.185	60.619	31.026	**E577**	117.4	7.44	60.489	28.886
Y374	115.3	8.14	61.949	36.486	**H578**	117	7.647	59.769	29.876
L375	126.3	8.754	58.899	41.646	**F579**	118.4	8.196	58.929	38.606
G376	105	8.225	48.949		**L580**	118.9	8.53	58.219	41.156
T377	113	8.915	66.389	67.896	**E581**	118.8	8.204	59.489	29.656
T378	121.3	8.437	67.919	68.766	**A582**	121.9	8.162	55.659	18.066
Q379	120.8	8.404	58.539	28.026	**L583**	117.8	8.078	58.019	42.836
N382	117	8.03	52.639	40.696	**N584**	120	8.495	55.839	39.106
E383	113	7.724	58.219		**K588**	118.5	7.768	58.689	33.026
Q384	120.4	8.137		27.346	**S589**	114	7.708	59.789	64.526
L387	117	8.083	58.699	40.616	**R590**	120.6	8.022	56.489	31.256
A388	120.5	7.375			**G591**	109.4	8.026	45.329	
I389	117.7	8.302	66.099	38.546	**R593**	118.6	8.339	57.049	30.366
S390	114.2	7.717	63.109	63.796	**G594**	108.1	8.008	45.959	
A391	124.2	8.515	55.359	17.926	**E595**	121.4	8.238	57.339	30.506
D401	109.8	7.482	51.499	39.086	**G596**	110	8.408	45.879	
N402	115.7	6.719	53.539	38.596	**G597**	109.1	8.102	45.549	
Q403	127.6	8.461	60.629	29.776	**A610**	123.3	8.486	55.859	18.326
T404	116.9	7.921	66.879	69.046	**L611**	116.5	8.61	58.489	42.336
A435	116.5	7.861	54.559	18.276	**S612**	117.4	7.864	61.609	63.116
Y436	111.2	7.04	55.429	40.786	**V517**	118.2	6.924	63.049	33.336
A437	123.1	7.262	56.059	19.756	**G615**	108.3	7.871	46.169	
H438	113	8.234	57.849	29.626	**Q616**	121.4	7.358	54.539	27.386
L439	120.4	7.502	56.009	42.146	**D618**	121.9	8.752	55.919	39.406
V440	116.3	7.097	61.979	32.776	**A619**	120.8	7.873	52.439	21.516
T441	121.2	8.049	60.399	69.786	**Y620**	118	7.24	58.999	36.526

A443	125	8.247	52.869	19.496	**G621**	107.6	8.658	47.679	
E444	120.3	8.246	56.929	30.516	**A622**	125.6	8.368	55.079	
E445	122.6	8.332	57.439		**A623**	119.9	7.989	55.699	18.346
G446	110.2	8.301	45.729		**D624**	121.8	8.761	57.659	40.216
A447	124	8.021	53.049	19.446	**D627**	118.6	8.399	52.239	39.226
G448	108.3	8.309	45.829		**L628**	124.4	7.686	58.339	43.176
A450	123.9	8.127	53.049	19.436	**S629**	113	8.719	62.869	
G451	108.1	8.266	45.729		**T630**	118.7	7.364	66.769	68.156
L452	121.5	7.957	55.379	42.606	**L631**	121.6	7.333	58.999	43.196
G453	109.6	8.143	45.059		**L632**	118.9	8.634	59.579	40.506
S455	116	8.243	59.079	64.146	**T633**	114.6	7.458	66.689	69.216
K456	123.1	8.12	56.679	33.006	**M634**	122.5	8.378	59.789	33.536
R457	122.2	8.075	56.509	31.006	**F635**	114.5	7.692	59.689	38.786
I458	123.1	8.105	61.319	38.626	**G636**	109.5	7.596	47.119	
L459	126.9	8.285	55.729	42.316	**L637**	122.7	8.168	52.899	42.536
G460	109.9	8.22	45.639						

Table 11. Pex14 CTD backbone chemical shifts assignments.

	N	H		Ca	Cb
S234		6.744	**S234**	58.34665	
A235	112.2	6.881	**A235**	50.76665	18.1355
W241	112.7	7.58	**I238**	58.60665	38.3855
I243	112.9	6.904	**W241**	57.19665	29.0555
K246	113	7.505	**Q242**	55.68665	29.5655
A255	117.2	8.304	**I243**	59.18665	38.4955
V256	117.4	8.086	**K246**	56.25665	33.2355
N257	117.6	8.409	**A254**	52.58665	19.0555
D263	118.6	8.276	**A255**	52.63665	19.0955
I264	118.7	8.04	**V256**	62.47665	32.9555
S265	118.8	8.168	**N257**	54.75665	41.4355
V267	121.5	8.362	**D263**	54.75665	41.0155
S268	120	8.157	**I264**	61.10665	38.8255
N269	119.3	8.387	**S265**	57.08665	63.3855
S271	121.5	8.531	**V267**	62.57665	32.7555
G277	120.3	8.319	**S268**	58.22665	63.9655
E284	109.4	8.464	**N269**	53.57665	39.0455
G285	121.1	8.25	**S271**	58.88665	63.8055
S286	110.3	8.42	**T272**	62.02665	69.7455
T287	121.2	8.322	**S273**	58.74665	63.8255
T289	122.9	8.113	**G277**	45.46665	
Y290	118.1	8.122	**E284**	57.22665	30.2455
L292	122.9	8.126	**G285**	45.45665	
L293	121.6	8.193	**S286**	58.66665	63.9655
G294	121.7	8.242	**T287**	62.24665	69.7455
Q296	121.9	8.459	**V288**	62.78665	32.9755

G299	121.9	8.368	**T289**	61.79665	69.8755
E300	122	8.557	**Y290**	57.99665	39.1055
G301	122.1	8.189	**L292**	55.62665	42.3655
V302	122.1	8.041	**L293**	55.23665	42.7455
V303	122.1	8.426	**G294**	44.69665	
D304	124	8.199	**Q296**	55.95665	29.7455
V305	122.4	8.417	**G299**	45.52665	
K306	121.1	8.13	**E300**	56.84665	30.6055
G307	122.6	8.482	**G301**	45.53665	
V313	122.9	7.379	**V302**	62.63665	32.9655
Q314	121.9	8.051	**V303**	62.44665	33.0355
G315	122.9	8.381	**D304**	54.16665	41.4755
E316	122.9	7.199	**V305**	62.91665	32.4355
V334	120.7	8.439	**K306**	56.98665	32.6555
H336	118.6	8.309	**G307**	45.69665	
D338	127.2	8.05	**Q308**	55.94665	29.7555
D341	109.3	8.105	**R310**	56.09665	31.0355
C342	121.9	8.37	**V313**	62.49665	33.0755
L343	121	8.414	**Q314**	56.16665	29.7155
G344	120.8	8.365	**G315**	45.44665	
V345	109.2	8.308	**E316**	56.43665	33.1355
Q346	124.1	8.49	**V334**	62.54665	32.7355
G353	117.8	8.271	**S335**	58.81665	63.7755
G355	108.9	8.292	**H336**	57.14665	29.6655
Q356	121.7	7.951	**D338**	54.99665	40.8555
I357	119.8	8.18	**D341**	54.64665	41.3055
N358	119.3	8.187	**C342**	59.06665	27.8655
V361	121.4	8.437	**L343**	56.04665	42.0155

E362	120.8	7.997	**V345**	62.71665	32.8455
E368	123.7	8.425	**G277**	45.46665	
G369	121.1	8.548	**G353**	45.65665	
S371	123.7	8.103	**Q356**	56.06665	29.5955

B. Abbreviations

Table 12. Abbreviations

1D,2D,3D	one-,two-,three- dimensional	Ki	inhibitor constant
aa	amino acid	LB	lysogeny broth
CSP	chemical shift perturbation	FP	fluorescence polarization
CTD	c-terminal domain	NOE	nuclear Overhauser effect
cc	coiled-coil	NOESY	nuclear Overhauser effect spectroscopy
DMSO	Dimethyl sulfoxide	NMR	nuclear magnetic resonance
EC50	effective concentration	PCR	polymerase chain reaction
ER	endoplasmic reticulum	peroxin (pex)	peroxisomal protein
farn	farnesylated	PMP	peroxisomal membrane protein
FITC	fluorescein isothiocyanate	ppm	part per million
FID	free induction decay	RDC	residual dipolar couplings
Ftase	farnesyl transferase	rmsd	root mean square deviation
FPP	farnesyl pyrophosphate	SDS	sodium dodecyl sulfate
HADDOCK	high ambiguity driven protein-protein docking	PAGE	polyacrylamide gel electrophoresis
HSQC	heternuclear single quantum correlation	TEV	tobacco etch virus
HMQC	heteronuclear multiple quantum correlation	TOCSY	total correlation spectroscopy
kDa	kilodalton	TPR	tetratricopeptide
Kd	dissociation constant	TROSY	transverse relaxation optimized spectroscopy

C. Acknowlegdements

I would like to thank Michael Sattler for the opportunity he gave me to work in a state-of-the-art research environment. He was always supportive and gave me the freedom to pursue own ideas. His multitasking ability 24/7 and deep understanding of science was a life experience for me.

I am very thankful to the members of my thesis advisory committee Prof. Dr. Bernd Reif and PD Dr. rer. nat. Kai Hell, for their guidance and support. I would also like to thank Prof. Dr. Johannes Buchner for participation in my examining committee. I am grateful to both excellent graduate schools, Elite Netzwerk Bayern - Protein Dynamics in Health and Disease and TUM - Graduate School, which provided the funding in order to participate in international scientific conferences and soft skills courses.

Throughout the years in the lab I met bright people who really shaped me as scientist. Kostantinos Tripsianes as a unique person, helped me not only with my new life in Munich but also introduced me to the protein NMR world. Janosch Hennig is another brilliant NMR expert whose experience was invaluable. Christopher Göbl could always answer my questions and help me when I was stuck with some ”trivial” NMR analysis.

Special thanks to my dear colleagues and friends of the high-impact office, Andreas Schlundt and Jan-Niklas Tants. Our endless scientific, or not, discussions were always refreshing.

I had exceptional collaboration with Grzegorz Popowicz and Maciej Dawidowski, which i am grateful of. With them I came to know the Polish sharp sense

of humor.

I would like to thank also Gerd Gemmecker, who was always willing to answer my "quick" questions. Guelden Demiraslan is the locomotive that keeps the lab moving, many thanks for that!

During my stay in the lab I was lucky to collaborate with the students Jonathan Kordon, David Goricanec and Stefan Gaussmann who contributed a lot in my work.

I also thank Jessica, Julian, Vishal and Wolfgang from Prof. Erdmann's lab at the Ruhr-Universität Bochum who were excellent collaborators.

Loukas, Eleni, Dafni, Petros,Nikoleta and John made life in Munich so much easier and kind of felt like ... Greece.

I owe everything to my parents, Kiki and Manolis, who believed in me when I told them many many years ago that I want to do research. Their trust is a pillar I lean on and always boosted me in my efforts. I thank my sister, Valia, who always encourages me and provides me with complete new perspective of things.

To the person who showed me this way, my Koritsara.