TUM School of Natural Sciences

Structural biology of the human peroxisomal translocation machinery

Stefan Ferdinand Rudolf Gaussmann

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzende/er: Prof. Dr. Franz Hagn

Prüfende/-r der Dissertation: 1. Prof. Dr. Michael Sattler

2. Prof. Dr. Kathrin Lang

3. Prof. Dr. Ralf Erdmann

Die Dissertation wurde am 12.07.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 16.10.2022 angenommen.

Structural biology of the human peroxisomal translocation machinery

Technical University of Munich

TUM School of Natural Sciences
Chair of Biomolecular NMR Spectroscopy
Bavarian NMR Center

Institute of Structural Biology
Helmholtz Center Munich

Doctoral Thesis

Stefan Ferdinand Rudolf Gaussmann

Munich 2022

Summary

The complexity of eukaryotic cells is reflected by the presence of numerous cellular compartments, i.e. membrane-surrounded or membrane-less organelles, that exhibit a defined microenvironment for specialized biochemical activities. Peroxisomes are cellular organelles that accommodate various essential steps in lipid metabolism including fatty acid degradation and the detoxification of the byproduct hydrogen peroxide. Malfunction of peroxisomal processing in early stage development often leads to severe phenotypes with short life expectancies. Peroxisomal proteins are translated in the cytosol and then targeted to the peroxisomal lumen or the peroxisomal membrane depending on a peroxisomal targeting signal (PTS1 or PTS2) or membrane targeting signal (mPTS), respectively. These proteins are recognized by two distinct import machineries, which depend on peroxisome related proteins called peroxins (PEX). Peroxisomal targeting of peroxisomal membrane proteins (PMPs) is mediated by PEX19, PEX3 and PEX16 where PEX19 is the cytosolic receptor and transport factor that recognizes the mPTS, while PEX3 and PEX16 are membrane-associated proteins responsible for docking and insertion. The majority of enzymes targeted to the peroxisomal lumen carry a PTS1 signal, which is recognized by the cytosolic receptor PEX5. The receptor tethers a peroxisome-targeted protein as cargo to the peroxisomal membrane, where it docks onto the membrane-associated proteins PEX13 and/or PEX14. Subsequent to docking, a transient pore is formed and the cargo translocated into the lumen.

In this thesis, the key peroxins in PTS1-mediated import, PEX5, PEX13 and PEX14 are studied using integrative structural biology combining nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, with biophysical techniques, including isothermal titration calorimetry (ITC), static light scattering (SLS) and circular dichroism (CD) spectroscopy. Structural biology is supported by biochemical and functional studies in cells.

Chapter 1 provides an introduction and overview into peroxisome biogenesis. Chapter 2 introduces the NMR methods and experimental approaches, while Chapter 3 outlines the aims and scope of this thesis. In Chapter 4 the methods and experimental procedures as well as the material used are described.

The results of the thesis are presented in Chapter 5-7. In Chapter 5 the membrane interactions of the N-terminal domain (NTD) of the soluble receptor PEX5 and of the PEX14 NTD are studied. The intrinsically disordered PEX5-NTD harbors multiple (di)aromatic peptide motifs (WxxxF/Y or related) that are recognized by the PEX14-NTD, which represents a key step in peroxisomal matrix import. Interactions of PEX5-NTD and PEX14-NTD with membrane-mimicking bicelles and nanodiscs were analyzed using NMR spectroscopy and isothermal titration calorimetry. PEX14-NTD is found to weakly interact with membrane-mimicking bicelles with a positively charged surface that partially overlaps with the WxxxF/Y binding site. The PEX5-NTD harbors multiple membrane interaction sites that involve a number of amphipathic α-helical regions, which partially overlap with some of the WxxxF/Y-motifs. The preformed helical conformation of these regions is stabilized in the presence of bicelles. ITC data show that the interaction between the PEX5 and PEX14-NTDs is largely unaffected by the presence of the membrane reflected by similar free binding enthalpies, where reduced binding enthalpy in the presence of bicelles is compensated by a reduced loss of entropy. This demonstrates that docking of PEX5 to PEX14 at the membrane does not reduce the overall binding affinity between the two proteins, providing insights into the initial phase of PEX5-PEX14 docking in the assembly of the peroxisome translocon.

In Chapter 6, a comprehensive biochemical and computational analysis of (di)aromatic peptide motifs from PEX5-NTD binding to PEX14-NTD is presented. Here, a combination of biophysical (ITC, NMR, CD), biochemical and computational methods was used to compare binding affinities and thermodynamics, which identified key features of the (di)aromatic peptide motifs for PEX14-NTD binding. Human PEX5 possesses eight of these conserved motifs distributed within its 320-residue disordered N-terminal region. All motifs exhibit distinct affinities and energetic contributions for the interaction with the PEX14-NTD. Molecular dynamics (MD) simulations of the docked peptides and analysis of binding energies identifies the specific amino acids features that stabilize a helical conformation of the peptide ligands and mediate important contact sites. The discovered key features were cross-validated in a pep-scan assay with a large number of PEX5 WxxxF/Y motifs from other organisms. The results suggest a refined consensus binding motif of **WΦxE(F/Y)Φ** for the PEX14-NTD.

Chapter 7 presents a structural analysis of the C-terminal domain (CTD) of the human peroxisomal membrane protein PEX13 and its interactions with the peroxins PEX5 and PEX14. Peroxisomal

matrix enzymes are recognized by the PEX5 C-terminal TPR domain and guided to the peroxisomal membrane, where PEX5 interacts with PEX13 and PEX14 to mediate docking and protein translocation. Docking of the cargo-loaded receptor is thought to be mediated by binding of the intrinsically disordered N-terminal domain of PEX5 to the NTD of the membrane-associated protein PEX14 via (di)aromatic peptide motifs. PEX13 was shown to be essential for PTS import, but molecular and structural details have so far remained elusive. Structural analysis of the PEX13-CTD revealed an unexpected autoinhibition of the PEX13-SH3 domain by an internal FxxxF motif, located C-terminally to the SH3 domain. Binding studies using NMR and ITC reveal interactions of the PEX13-CTD with PEX14 and of the PEX13-SH3 domain with the PEX5-NTD, which is modulated by the intrinsic FxxxF motif of PEX13. We show that the PEX13 FxxxF motif affects PTS1 import in a cellular assay indicating a novel regulatory interaction that fine-tunes peroxisome biogenesis. These results provide unexpected insight into the molecular interactions between PEX13, PEX14 and PEX5 and serve as basis for further studies to analyze sequential steps of the peroxisome pore formation and cargo translocation.

The results of this thesis are discussed in Chapter 8 presenting conclusions and an outlook for future studies to unravel the molecular basis of peroxisome biogenesis.

Acknowledgements

With this, I like to thank my supervisor Prof. Dr. Michael Sattler for providing me with the opportunity to study in a great scientific environment. He supported me with discussions and options to collaborate with other groups, which helped me to learn and to do the science in an independent way, the most important skill of a scientist.

Thanks to the members of my advisory committee, Prof. Dr. Kathrin Lang and Prof. Dr. Franz Hagn for the time they spent to review my work and valuable discussions and advices.

To my mentors Dr. Andreas Schlundt and Dr. Florent Delhommel who were always open for discussions and help with experimental setup as well as data analysis.

To my lab mates, students, faculty and staff members of the Bavarian NMR Center (BNMRZ) and the Helmholtz Center Munich (HMGU) for creating a very pleasant and productive working environment. Special thanks to Dr. Sam Asami, Dr. Gerd Gemmecker and Dr. Mark Bostock for support with NMR experiments and pulse sequences. Last but not least I thank my friend Maximilian Fottner for all the valuable talks, drinks etc. under the Greek sun/rain/snow at day/night.

To my colleagues from Ralf Erdmann's group at the Ruhr University Bochum for the ongoing successful partnership.

To my parents who made it possible to create and follow my passion for science free and independently.

Greatest thanks to my better half, *Sabrina*, for her patience, support, and understanding during the second half my study, which was the most productive and successful time. I greatly appreciate you supported me with workload reduction and understanding in stressfull times.

… and to all the people who supported me and I forgot to mention here

List of publications

Results reported in manuscripts 3, 6 and 8 are described in this thesis.
* Equal contribution

Manuscripts in preparation

8. Stefan Gaussmann, Julia Ott, Krzysztof Zak, Florent Delhommel, Grzegorz Popowicz, Wolfgang Schliebs, Ralf Erdmann, Michael Sattler, *Intramolecular autoinhibition of HsPEX13 modulates protein-protein interactions in peroxisomal import*

- To be submitted

7. Leonidas Emmanouilidis*, Jessica Sehr*, Katharina Reglinski*, David Goricanec, Jonathan Kordon, **Stefan Gaussmann**, Dominic Waithe, Philip Hublitz, Verian Bader, Konstanze F. Winklhofer, Martin Jung, Wolfgang Schliebs, Christian Eggeling, Ralf Erdmann, Michael Sattler, *A novel PEX14/PEX5 interface links peroxisomal protein import and receptor recycling*

- Submitted to *Nature Communications*

6. Mohanraj Gopalswamy*, Chen Zheng*, **Stefan Gaussmann***, Hamed Kooshapur, Eva Hambruch, Wolfgang Schliebs, Ralf Erdmann, Iris Antes, Michael Sattler, *Distinct conformational and energetic features define the specific recognition of (di)aromatic peptide motifs by PEX14*

- Submitted to *Biological Chemistry,* in revision

Journal publications (peer reviewed)

5. Mengqiao Li, **Stefan Gaussmann**, Bettina Tippler, Julia Ott, G Popowicz, Wolfgang Schliebs, Michael Sattler, Ralf Erdmann, V Kalel, *Novel Trypanocidal Inhibitors that Block Glycosome Biogenesis by Targeting PEX3–PEX19 Interaction, Frontiers in Cell and Developmental Biology, 2021, Vol. 9 Pages 737159, DOI: 10.3389/fcell.2021.737159*

4. Maximilian Fottner, Maria Weyh, **Stefan Gaussmann**, Dominic Schwarz, Michael Sattler, Kathrin Lang, *A modular toolbox to generate complex polymeric ubiquitin architectures using orthogonal sortase enzymes, Nature Communications, 2021, Vol. 12 Issue 1, DOI: 10.1038/s41467-021-26812-9*

3. Stefan Gaussmann*, Mohanraj Gopalswamy*, Maike Eberhardt, Maren Reuter, Peijian Zou, Wolfgang Schliebs, Ralf Erdmann, Michael Sattler, *Membrane Interactions of the Peroxisomal Proteins PEX5 and PEX14, Frontiers in Cell and Developmental Biology, 2021, Vol. 9 Issue 743 Pages 651449, DOI:10.3389/fcell.2021.651449*

2. Maren Reuter, Hamed Kooshapur, Jeff-Gordian Suda, **Stefan Gaussmann**, Alexander Neuhaus, Lena Brühl, Pratima Bharti, Martin Jung, Wolfgang Schliebs, Michael Sattler, Ralf Erdmann, *Competitive Microtubule Binding of PEX14 Coordinates Peroxisomal Protein Import and Motility, Journal of Molecular Biology, 2021, Vol. 433 Issue 5 Pages 166765, DOI: 10.1016/j.jmb.2020.166765*

1. Vishal C Kalel, Mengqiao Li, **Stefan Gaussmann**, Florent Delhommel, Ann-Britt Schaefer, Bettina Tippler, Martin Jung, Renate Maier, Silke Oeljeklaus, Wolfgang Schliebs, Bettina Warscheid, Michael Sattler, Ralf Erdmann, *Evolutionary divergent PEX3 is essential for glycosome biogenesis and survival of trypanosomatid parasites, Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2019, Vol. 1866 Issue 12 Pages 118520, DOI:10.1016/j.bbamcr.2019.07.015*

Table of Contents

Chapter 1:

Biological background

Biological background

The cell is the center and the smallest unit of any life on earth. Depending on the evolutionary and developmental states cells differ in architecture and complexity. The highest degree of complexity is represented by eukaryotic cells harboring separated organelles with defined microenvironments to ensure perfect functioning, e.g. of metabolic pathways.

The Peroxisome

Peroxisomes are single lipid bilayer membrane-encapsulated organelles ubiquitous in eukaryotic cells, which were first described as 0.2-1 µm small microbodies by Rhodin (1954). One decade later, Christian de Duve and Pierre Baudhuin characterized these microbodies as organelles with a function in hydrogen peroxide breakdown by the enzyme catalase and named them peroxisomes (De Duve & Baudhuin, 1966; Duve, 1969). Studies of the past 30 years corroborated their indispensability for human health and development.

Peroxisomes show variations in number and size depending on the metabolic demand and are specialized by organism and cell type. Subclasses of Peroxisomes with specialized functions are glycosomes in trypanosoma (Opperdoes & Borst, 1977), glyoxysomes in plants (Tolbert & Essner, 1981) and Woronin bodies in fungi (Jedd & Chua, 2000). Within a diverse range of functions found in different tissues and organisms, H_2O_2 detoxification and β-oxidation of fatty acids are well conserved among them.

In mammalian cells, peroxisomes are heavily involved in fatty acid metabolism. The breakdown of fatty acids in peroxisomes is classified as β- and α-oxidation. Briefly, β-oxidation is described for very-long chain fatty acids (VLCFAs), dicarboxylic fatty acids, prostanoids and the toxic bile acid intermediates di-/tri-hydroxycholestanioc acid (DHCA, THCA) (Hashimoto *et al*, 1999; Mannaerts & Van Veldhoven, 1993; Reddy & Mannaerts, 1994; Wanders, 2004) while α-oxidation is reported for 3-methyl- branched fatty acids (Wanders *et al*, 2001). Peroxisomes are further known to mediate the synthesis of ether lipids such as plasmalogen, cholesterol and other isoprenoids (Wanders *et al*, 2010). It is estimated that over 50 enzymes, which catalyze metabolic reactions, are found in the peroxisomal matrix (lumen) (Wanders & Waterham, 2006). One of those enzymes is the glyoxylate aminotransferase, which catalyzes the reaction from glyoxylate into glycine. In absence of this enzyme, glyoxylate is converted into the toxic metabolite oxalate by the enzyme lactate dehydrogenase, which can cause severe consequences for the cell (Jonassen *et al*, 2005; Wanders *et al.*, 2010).

Peroxisome biogenesis

Peroxisomes are very dynamic organelles, which have the ability to multiply (proliferate) in response to external environmental stimuli. Thus, peroxisome formation (biogenesis) and degradation (pexophagy) are dynamic processes of the cell as well. The molecular mechanism of biogenesis is still an actively debated topic. To date, two possible mechanisms have been described: (i) The growth-and-division model, in which peroxisomes derive from pre-existing ones (Lazarow & Fujiki, 1985; Purdue & Lazarow, 2001; Schrader *et al*, 2012) (**Figure 1A**). (ii) The *de novo* formation model where peroxisomes derive from a preperoxisomal compartment (P-ER), i.e. a subdomain of the endoplasmatic reticulum (ER) (Hoepfner *et al*, 2005; Kim *et al*, 2006; Kragt *et al*, 2005; Matsuzono *et al*, 1999; Tam *et al*, 2005) (**Figure 1.1B**). Present evidence supports both models suggesting a co-existence of these pathways with discrete regulations (Agrawal & Subramani, 2016; Nuttall *et al*, 2011)

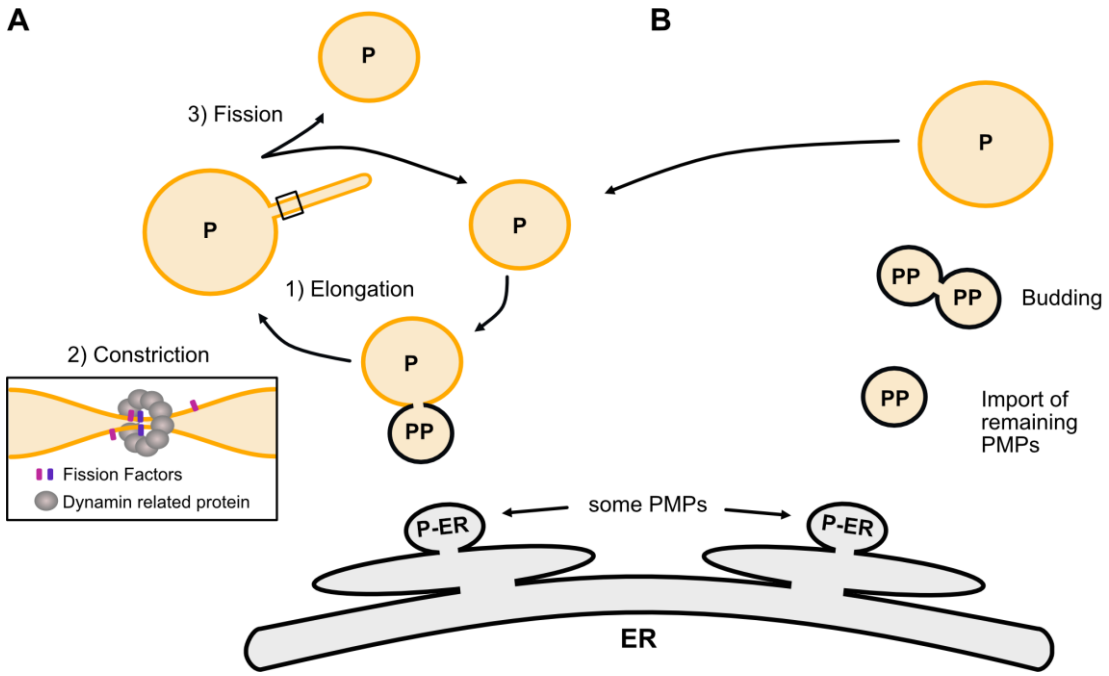

Figure 1.1: Models for peroxisomal formation and biogenesis of peroxisomal membrane proteins (PMPs). (A) Growth-and-division model. Peroxisomes (P) grow by budding with pre-peroxisomes (PP) derived from a specialized subdomain of the ER, which contain a small set of PMPs. Larger peroxisomes can then multiply by division. (B) Peroxisome de novo biogenesis. The pre-peroxisomes bud from the subdomain of the ER and fuse homotypically to form larger vesicles. At the same time, additional PMPs are imported as part of the maturation process, which is completed once the import machinery has been inserted.

Early observations in peroxisome proliferation (Lazarow & Fujiki, 1985) and peroxisome membrane protein (PMP) biogenesis on free ribosomes (Fujiki *et al*, 1984; Kindl & Kruse) gave rise to the development of the multi-step growth-and-division model. These steps are described as: (1) Elongation; describing the growth and remodeling of certain membrane structures in a tubular matter. Lipids are synthetized in the ER and transported to the peroxisome. (2) Constriction; where PMPs and mitochondrial fission factors (Fis1 and Mff in mammals) concentrate in the tubular extension at the point of division to recruit dynamin related protein (DLP1 in human), a mitochondrial and peroxisomal fission factor. DLP1 oligomers form a large ring-like structure wrapping around the membrane by GTP consumption. (3) Fission; the membranes are forced to form contact sites, which may lead to spontaneous division of the spherical parent and the newly formed peroxisome (**Figure 1.1A**). Mitochondrial fission factors may play a regulatory role (Delille *et al*, 2010; Koch *et al*, 2010; Schrader *et al.*, 2012; Schrader & Fahimi, 2006). Although the involved factors and mechanistic details may vary between different species, proteins of the Pex11 family (PEX11α, β, γ) were shown to regulate peroxisome abundance (Erdmann & Blobel, 1995; Marshall *et al*, 1996) and pre-peroxisomal membrane topologies. In brief, the expression of various Pex11 variants blocked peroxisomal fission, which led to formation of elongated pre-peroxisomal membrane structures (TPAs) and large membrane stacks (JEPs) (Koch *et al.*, 2010; Schrader *et al.*, 2012).

The growth and division model was challenged when ΔPex3p/ΔPex19p yeast cells lacking peroxisomal structures were shown to synthesize peroxisomes *de novo* upon reintroduction of Pex3p and Pex19p (Hettema *et al*, 2000). Similar results were obtained from human Zellweger syndrome patients with PEX3 and PEX19 dysfunction (Matsuzono *et al.*, 1999; Shimozawa, 2000). Thus, Pex3 and Pex19 were identified to be essential for peroxisome *de novo* biogenesis. It was later shown that Pex3 is routed through the ER where pre-peroxisomal structures bud off in a Pex19 dependent manner (Hoepfner *et al.*, 2005). These pre-peroxisomes grow by homotypical fusion and mature by the import of remaining PMPs to functional peroxisomes (Hoepfner *et al.*, 2005; Kim *et al.*, 2006; Kragt *et al.*, 2005; Matsuzono *et al.*, 1999; Tam *et al.*, 2005) (**Figure 1.1B**). It was further demonstrated that peroxisome formation in mammals relies on Pex16, which is co-translationally targeted to the ER. It is speculated that Pex16 is a receptor for Pex3, since it was shown to target overexpressed Pex3 to the ER in peroxisome deficient cells (Kim *et al.*, 2006). However, with the *de novo* biogenesis model, the PMPs were classified as either Pex19 dependent PMPs (class I) or Pex19 independent PMPs (class II) such as Pex3 and Pex16 (Fang *et al*, 2004; Fujiki *et al*, 2006).

The majority of PMPs belong to the class I, which harbor a membrane peroxisome targeting signal (mPTS): Those PMPs are synthesized in the cytosol and directly targeted to the peroxisomal membrane via mPTS recognizing Pex19 **(Figure 1.2)**.

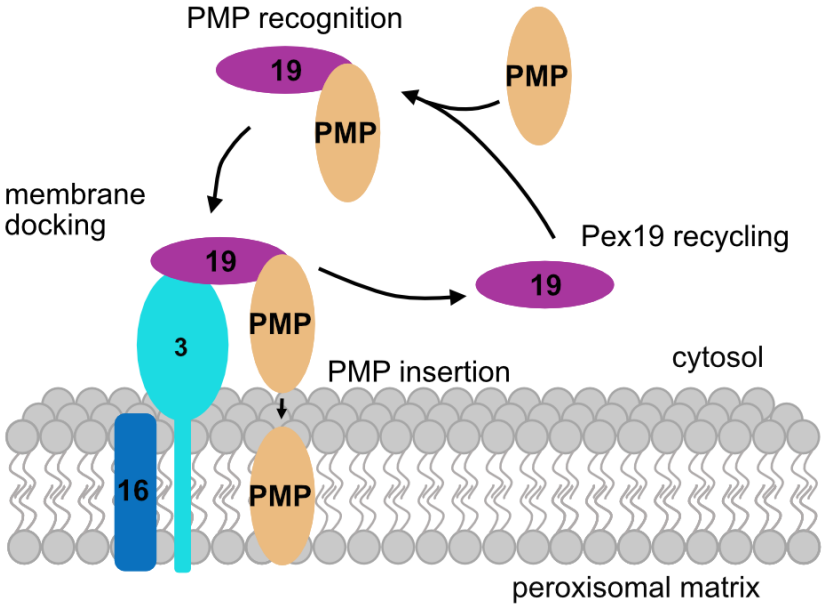

Figure 1.2: Pex19 mediated peroxisomal membrane protein of class I PMPs. Cytosolic Pex19 binds class I PMPs and tethers them to membrane bound Pex3 (and Pex16 in mammals). After docking, the PMPs are inserted into the membrane and Pex19 is recycled to the cytosol.

Peroxisome disorders

The first cases of peroxisome disorders were clinically described by Bowen, Zellweger and Lindberg in 1964 and Smith et al. in 1965 (Bowen *et al*, 1964; Smith *et al*, 1965). Later studies in 1967 and 1969 further described the phenotypic features of this rare disease leading to the given name hepato-renal syndrome. The name was then changed to Zellweger syndrome after the pediatrician Hans Zellweger who described two of the first patients (Passarge & McAdams, 1967). Although the phenotype was well-defined, the causal link between Zellweger Syndrome and peroxisomes was not made until Goldfischer *et al* (1973) discovered the absence of peroxisomal structures in hepatocytes and proximal tubules.

As human peroxisomes are essential for diverse metabolic pathways, dysfunction of one or more proteins can cause a range of clinical phenotypes depending on the affected pathway. Peroxisomal disorders are thus classified in two major groups: (i) peroxisome biogenesis disorders (PBDs), which are caused by one or more mutations within 14 different *PEX* genes essential for peroxisome

biogenesis, peroxisome proliferation or matrix protein import (**Table 1.1**) (Waterham & Ebberink, 2012) and (ii) single peroxisomal enzyme / transporter deficiencies (PEDs), which affect lipid and H_2O_2 metabolic pathways as well as glyoxylate detoxification (**Table 1.2**) (Wanders & Waterham, 2006).

PBDs are described as two distinct subtypes: Zellweger Syndrome Spectrum (ZSS) disorders, which cover a broad range of phenotypic presentations with similar biochemical abnormalities including the described phenotypes of Zellweger Syndrome (ZS), neonatal adrenoleukodystrophy (NALD), infantile Refsum disease (IRD), the recently discovered Heimler Syndrome (HS) (Kelley & Moser, 1984; Poulos *et al*, 1984; Ratbi *et al*, 2015) and rhizomelic chondrodysplasia punctata (RCDP) type 1 and type 5 (Waterham *et al*, 2016). The patients of ZSS are characterized based on the deficiency of functional peroxisomes in their cells. In severe cases, the cells entirely lack functional peroxisomes. Cells from milder phenotypes can still contain lower amount of intact organelles or cells with and without functional peroxisomes depending on their type (Gootjes *et al*, 2004). The majority of peroxisomal enzymes are unstable or inactive in the cytosol, which leads to their degradation and thus accumulation or shortage of substrates or end products, which are usually metabolized by the peroxisomes. Frequently accumulated substrates are VLCFAs, pristanic acid, phytanic acid, DHCA and THCA while plasmalogens, cholic and chenodeoxycholic acid as well as docosahexaenoic acid are not sufficiently produced (Steinberg *et al*, 2006; Wanders & Waterham, 2006; Waterham & Ebberink, 2012). RCDP type 1 and type 5 patients suffer from mutations or a frameshift in *PEX7* and *PEX5* respectively. Both are cytosolic receptor proteins, which recognize newly synthetized peroxisomal matrix proteins comprising a peroxisomal targeting signal type 2 (PTS2, PEX7) or type 1 (PTS1, PEX5). The mutations in PEX7 (RCDP type 1) as well as the loss of PEX5L, the long isoform of PEX5, caused by a frame shift affect only import of PTS2 enzymes because PTS1 enzymes are still imported by a shorter isoform of PEX5 (PEX5S). However, RCDP type 1 patients can suffer from severe phenotypes (Barøy *et al*, 2015; Waterham *et al.*, 2016).

Table 1.1:Frequency in distribution of PEX gene defects of patients with peroxisome biogenesis disorders including Zellweger syndrome spectrum disorders reported by **Waterham & Ebberink (2012).**

***PEX* gene**	**Frequency (%) among PBDs**	**Frequency (%) among ZSS disorders**
PEX1	48.5	58.9
PEX2	3.2	3.9
PEX3	0.5	0.5
PEX5	1.4	1.7
PEX6	13.1	15.9
PEX7	17.7	-
PEX10	3.4	4.2
PEX11β	<0.1	<0.1
PEX12	5.9	7.1
PEX13	1.4	1.7
PEX14	0.2	0.3
PEX16	1	1.2
PEX19	0.3	0.4
PEX26	3.4	4.2

Single peroxisomal enzyme / transporter deficiencies (PEDs) are defined as defects in peroxisomal enzymes as well as peroxisomal membrane proteins involved in metabolite transport. A brief summary of the genetic causes and the phenotypes known to date is described below.

Biogenesis disorders of ether phospholipid such as plasmalogens are found to be caused by either dihydroxyacetone-phosphate acyltransferase (DHAPAT) or alkyl-dihydroxyaceton-phosphate synthase (ADHAPS) deficiency causing the phenotype of RCDP type 2 or type 3 respectively (**Table 1.2, Nr.1**). As peroxisomal β-oxidation mediates the breakdown of several different species of fatty acids (eg. VLCFAs, Pristanic acid, DHCA, THCA etc.) multiple enzymes and enzyme-complexes are involved. To date, five different defects including X-linked adrenoleukodystrophy (ALDP), Acyl-CoA oxidase 1 (ACOX1) deficiency, D-bi-functional protein (DBP) deficiency, 2-methylacyl-CoA racemase (AMACR) deficiency and sterol carrier protein X (SCPx) deficiency have been described (**Table 1.2, Nr.2**). In humans, phytanoyl-CoA hydroxylase is encoded by the *PHYH* (*aka PAHX*) gene is essential for α-oxidation of 3-methylated fatty acids in the peroxisome. Mutations in the *PHYH* gene lead to impaired α-oxidation, which represents the phenotype of Refsum disease (**Table 1.2, Nr.3**). Another crucial function of peroxisomes is the detoxification of glyoxylate, which is catalyzed by the enzyme alanine glyoxylate aminotransferase (AGT). Dysfunction of AGT leads to primary hyperoxaluria type 1 (**Table 1.2, Nr.4**). The best-studied function of peroxisomes is the breakdown of H_2O_2 to H_2O and O_2 by the enzyme catalase (CAT),

which causes acatalasemia in a loss-of-function mutant setup (**Table 1.2, Nr.5**) (Wanders & Waterham, 2006; Waterham *et al.*, 2016)

Table 1.2: List of single enzyme / transporter deficiencies summarized by Wanders & Waterham (2006).

Nr.	Affected peroxisomal pathway	Enzyme defect	Gene
1	Ether phospholipid synthesis	DHAPAT	*GNPAT*
		ADHAPS	*AGPS*
2	Peroxisomal beta-oxidation	ALDP	*ABCD1*
		ACOX1	*ACOX1*
		DBP	*HSD17B*
		AMACR	*AMACR*
		SCPx	*SCP2*
3	Peroxisomal alpha-oxidation	PHYH/PAHX	*PHYH/PAHX*
4	Glyoxylate detoxification	AGT	*AGXT*
5	H_2O_2 metabolism	CAT	*CAT*

Peroxisomal matrix protein import

Peroxisomes are simple membrane encapsulated organelles lacking genetic material. Hence, all enzymes acting inside the peroxisomal lumen must be translocated post-translationally. Matrix enzymes are not routed through the ER but are cytosolic expressed and imported into mature peroxisomes (**see Peroxisome biogenesis**). This import pathway is mediated by a machinery of specialized peroxins and exhibits some unique features. In contrast to import pathways into mitochondria for example, proteins can be imported in their folded and even in oligomeric states (Glover *et al*, 1994). Other experiments showed that even artificial protein complexes of several hundred kDa and gold particles of 9 nm size were imported into the peroxisome (**Figure 1.3A**). Hence, proving the dynamic and the omnivorous character of the import mechanism (Walton *et al*, 1995; Yang *et al*, 2018). Although not all peroxins are conserved among different species the general cascade of interactions is found in all eukaryotes (**Figure 1.3B**) (Erdmann & Schliebs; Gould *et al*, 1990; Hu *et al*, 2012).

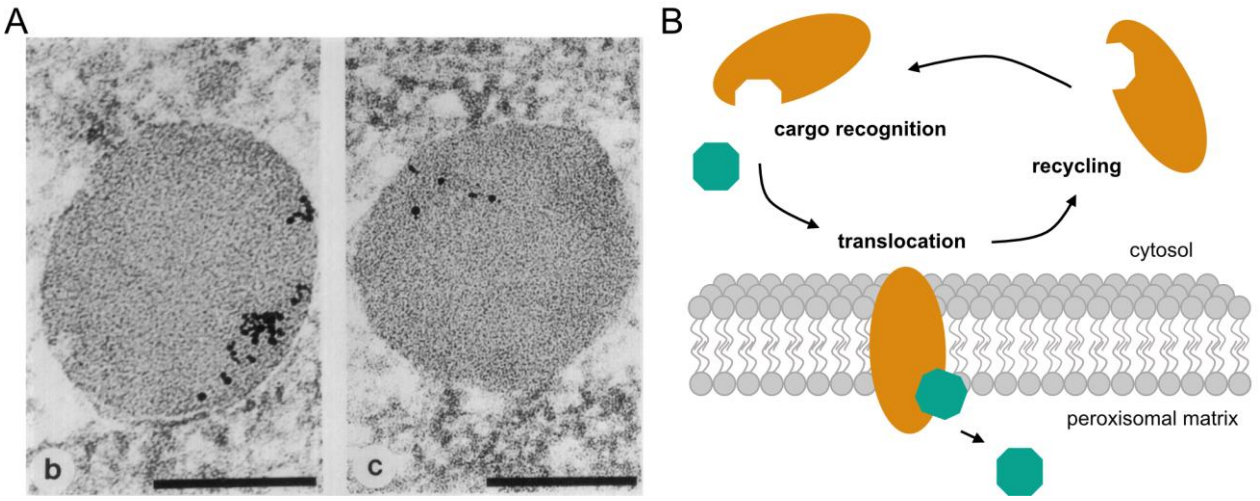

Figure 1.3: Peroxisomal matrix protein transport. (A) Peroxisomal import of gold particles (4-9 nm diameter) coated with human serum albumin modified with a C-terminal peroxisomal targeting sequence (adapted from Walton *et al.* (1995). (B) Schematic representation of peroxisomal matrix protein translocation: Cargo proteins are tethered to the organelle via a receptor, which associates with docking proteins and is recycled back to the cytosol after cargo release**.**

Peroxisomal matrix protein import can be schematically described in three steps; (i) A cargo protein destined for transfer to the peroxisome contains a signal sequence is bound by a cargo receptor. (ii) The cargo-receptor complex is tethered the organelle where it associates with one or more docking proteins. (iii) A transient pore is formed and the cargo is shuttled through the membrane into the peroxisomal lumen. After release of the cargo the receptor is recycled to the cytosol and is available for another round of cargo shuttling (**Figure 1.3B).**

Yeast models have proven to be particularly useful to study the mechanism behind peroxisomal matrix import. Key features of the findings in yeast have often been successfully transferred to other organisms. However, differences in the involved peroxins and their interactions have also been observed. In the following two sections, the import mechanisms in yeast and mammals are described to point out differences and unknowns of the mammalian system.

Matrix import pathways in yeast

Newly synthetized matrix proteins bearing a peroxisomal targeting signal sequence located either on their C - (PTS1) or N-terminus (PTS2) are shuttled to the peroxisomal membrane via distinct receptors. the PTS1 pathway is far more frequently used then PTS2 simply due to the larger number of proteins encompassing it. The PTS1 signal was originally described as a dodecapeptide from

monkey catalase with the sequence LIKAKKGGKSKL (Gould *et al*, 1987). A consensus sequence was later specified to SKL, which was shown to universal to different organisms (Gould *et al.*, 1990; Sommer *et al*, 1992). However, the consensus sequence was subsequently broadly expanded depending on the organism, which seems to be the result from mutations in Pex5 (Amery *et al*, 1998; Reumann, 2004)**.** Further studies show that the recognition sequence for specific substrates can be larger as the tripeptide. (Brocard & Hartig, 2006; Hagen *et al*, 2015). However, cytosolic PTS1 cargos are recognized by Pex5p and can be directly shuttled to the membrane **(Figure 1.4A)**. PTS2 proteins possesses a C-terminal nona-peptide with a very broad consensus sequence (Kunze, 2020; Petriv *et al*, 2004). Those cargo proteins are directed via a bipartite complex of Pex7p associated with the co-factors Pex18p or Pex21p **(Figure 1.4B) (Lazarow, 2006)**. A third option for cargo proteins lacking a PTS1 or PTS2 signal is being transported as complexes with a PTS-containing protein called "piggyback" mechanism (Saryi *et al*, 2017).

Cargo-loaded receptors dock at the peroxisomal membrane on either Pex14p or Pex13p, which are the central components of the docking complex **(Figure 1.4)**. Albeit Pex13p is essential for docking, it has not been found to be part of the translocon (Meinecke *et al*, 2010). For long, the PTS1 and PTS2 import mechanism was thought to converge at the peroxisomal membrane and utilize the same pore. However, recent studies show strong evidence that both pathways utilize distinct translocon compositions **(Figure 1.4B)**. PTS1 and PTS2 cargo proteins shuttle with Pex5p and Pex18p/Pex7p complex through the membrane respectively. Both receptors, Pex5p and Pex18p, are essential parts of the pores themselves and cycle between a soluble and a membrane bound state (Meinecke *et al.*, 2010; Montilla-Martinez *et al*, 2015). The cargo release as well as the interplay between import and the export machineries is linked to Pex8p, which was proposed based on the observation that ΔPex8p cells show mislocalization of PTS1 and PTS2 matrix proteins (Rehling *et al*, 2000). Although the pore components are well-characterized, the structural mechanism of the cargo translocation and release process remains unclear.

After the cargo translocation, the receptors need to be recycled to be available for further import rounds. For this, Pex5p is monoubiquitinated by the Pex22p-anchored ubiquitin-conjugating enzyme Pex4p (E2-type) and the ubiquitin ligase Pex12p (E3-type), which is a component of the RING-finger complex containing additional Pex2p and Pex10p. The ubiquitin signal recruits the Pex15p conjugated AAA ATPase complex Pex1p/Pex6p, which pulls the receptor out of the membrane in an ATP-dependent manner (Platta *et al*, 2013). There is evidence that the receptor

Pex18p takes a similar recycling route via monoubiquitination via Pex4p/Pex22 (E2) and Pex12p/Pex10p (E3) ligases and subsequent delocalization via Pex1p /Pex6p. It is thought that the receptor Pex7p shuttles with the co-factor in and out of the peroxisome (El Magraoui *et al*, 2013; Nair *et al*, 2004). Deubiquitination in the cytosol makes the receptors available for import again (Platta *et al.*, 2013).

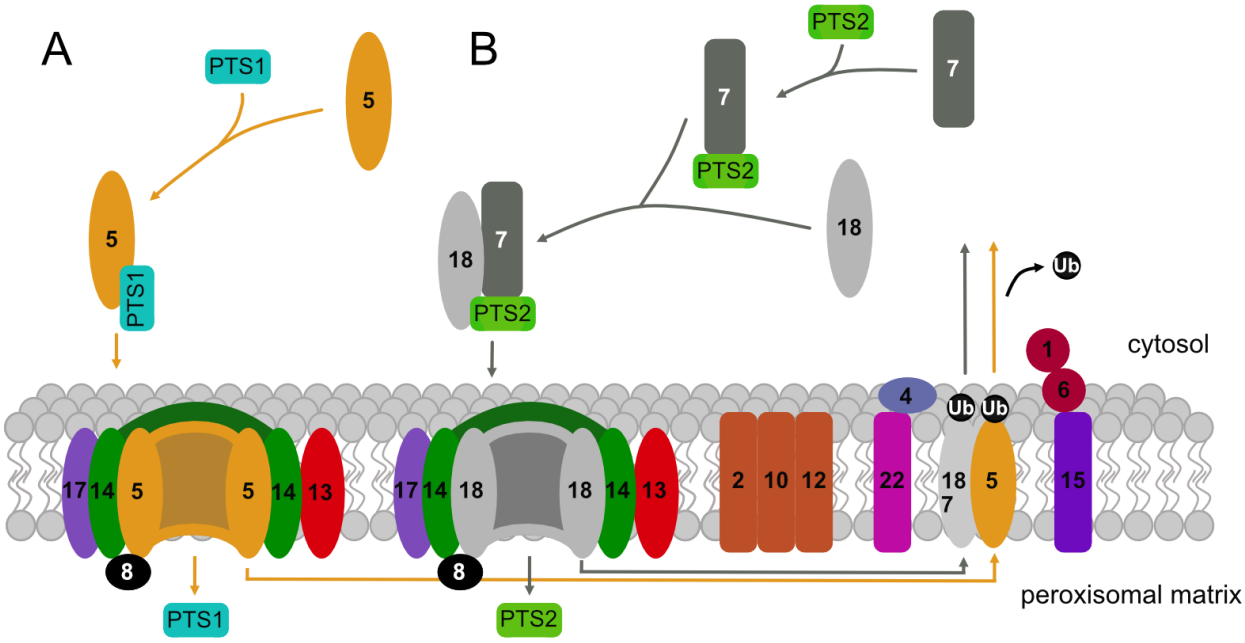

Figure 1.4: Peroxisomal matrix import pathways in yeast. (A) Cytosolic PTS1 cargo is recognized by Pex5p and shuttled to the peroxisomal membrane where it docks onto Pex14p or Pex13p. After formation of the docking complex a dynamic pore of Pex5p and Pex14p is formed and the cargo translocated into the lumen. Cargo release and receptor recycling is triggered by Pex8p. Receptor recycling is mediated by monoubiquitination via an ubiquitin ligase cascade including Pex4p and the RING finger complex Pex12p/Pex10p/Pex2p and delocalization via the AAA ATPase Pex1p/Pex6p. (B) The PTS2 pathway relies on the cyclic Pex7p and the co-factors Pex18p or Pex21p (not illustrated), which resemble the docking capabilities of Pex5p. The receptor complex docks at the membrane in the same manner as PTS1. In the translocation pore Pex5p is replaces with Pex18p. After cargo release the Pex7p/Pex18p takes a similar recycling route as Pex5p.

Matrix import pathways in mammals

The targeting signal sequences PTS1 and PTS2 are conserved among yeast and mammals, which was shown by the ability to import PTS cargos of one organism into peroxisomes of a different species (Gould *et al.*, 1990). Studies identified hereby the same PTS receptors PEX5 and PEX7.

Opposite to yeast, mammalian cells do not express homologues of Pex18p and Pex21p but two isoforms of PEX5, which differ in length and are therefore called PEX5S (short) and PEX5L (long). The PEX5S isoform has a deletion of amino acid 215-251 (**Figure 1.6A)**. However, both versions are able to mediate PTS1 import but solely PEX5L can interact with PEX7-PTS2 **(Figure 1.5)** replacing the co-factors in yeast (Braverman *et al*, 1998; Otera *et al*, 2000). The PTS independent "piggybag" pathway was shown to be existent in mammals as well (Islinger *et al*, 2009).

Both PTS1 and PTS2 cargo complexes dock at the peroxisome on the N-terminal domain of PEX14 (**Figure 1.5E)** (Neufeld *et al*, 2009; Neuhaus *et al*, 2014). It was shown that not only PEX5 but also cargo loaded PEX7 can directly interact with PEX14 (Shimizu *et al*, 1999). The role of PEX13 is not understood so far and a docking complex as known from yeast has not been reported yet. Either way, knockout of PEX13 is linked to peroxisomal biogenesis disorders with PTS import defects (Liu *et al*, 1999; Maxwell *et al*, 2003) and was further reported to be essential for the import of catalase, a non-canonical PTS1 cargo, but not for PTS1 import (Otera *et al.*, 2000). Thus clearly indicating the involvement in peroxisomal matrix import.

After the docking event, a transient pore is formed and the cargo shuttled through the membrane **(Figure 1.5)**. The actual mode of the import process and the composition of all components is still elusive. It has been suggested that the major constituents of the pore may be PEX14, PEX13 and the PTS receptor PEX5. At this point, PEX5 was found to act as an integral membrane protein migrating with PEX14 on a blue native gel and sucrose gradient (Gouveia *et al*, 2000; Reguenga *et al*, 2001). PEX13, on the other hand, does not seem to migrate with PEX5/PEX14 but forms complexes of higher mass containing only PEX13 (Reguenga *et al.*, 2001). Nevertheless, these studies do not exclude PEX13 to be involved in the translocation event due to their insensitivity. Moreover, pull down experiments showed an interaction of the PEX13 N-terminal region with PEX5 (Otera *et al*, 2002).

How the cargo is released into the peroxisomal lumen remains another unsolved question **(Figure 1.5)**. This task is assumed to be accomplished by Pex8p in yeast. But a mammalian homologue has yet to be identified. A study from Freitas *et al* (2011) suggests the involvement of the PEX14 N-terminal domain in cargo release; at least for catalase, which harbors a non-canonical PTS1 sequence KANL.

The mechanism of receptor recycling via monoubiquitination on a conserved cysteine with subsequent dislocation by the AAA ATPase complex PEX1/PEX6, on the other hand seems to be

conserved from yeast to mammals (Okumoto *et al*, 2011). Admittedly, mammalian cells lack clear Pex4p- and Pex22p-orthologs but functional –related isoforms of the UbcH5 were shown to mediate ubiquitination *in vitro* **(Figure 1.5)** (Grou *et al*, 2008). The mechanism of ubiquitination *in vivo* has not been described yet. Clearly the action of the RING-finger complex is needed. Mutations of the RING-finger complex components PEX12, PEX10 and PEX2 lead to phenotypes of Zellweger syndrome (Krause *et al*, 2006; Okumoto, 1998; Okumoto *et al*, 1998; Shimozawa *et al*, 1999). PEX12 and PEX2 have been observed to be subunits of a protein assembly together with PEX14 and PEX5 (Reguenga *et al*., 2001). The monoubiquitinated receptors PEX5S or PEX5L are then extracted from the peroxisomal membrane by the PEX26 recruited AAA- ATPase complex consistent of PEX1 and PEX6 (Matsumoto *et al*, 2003).

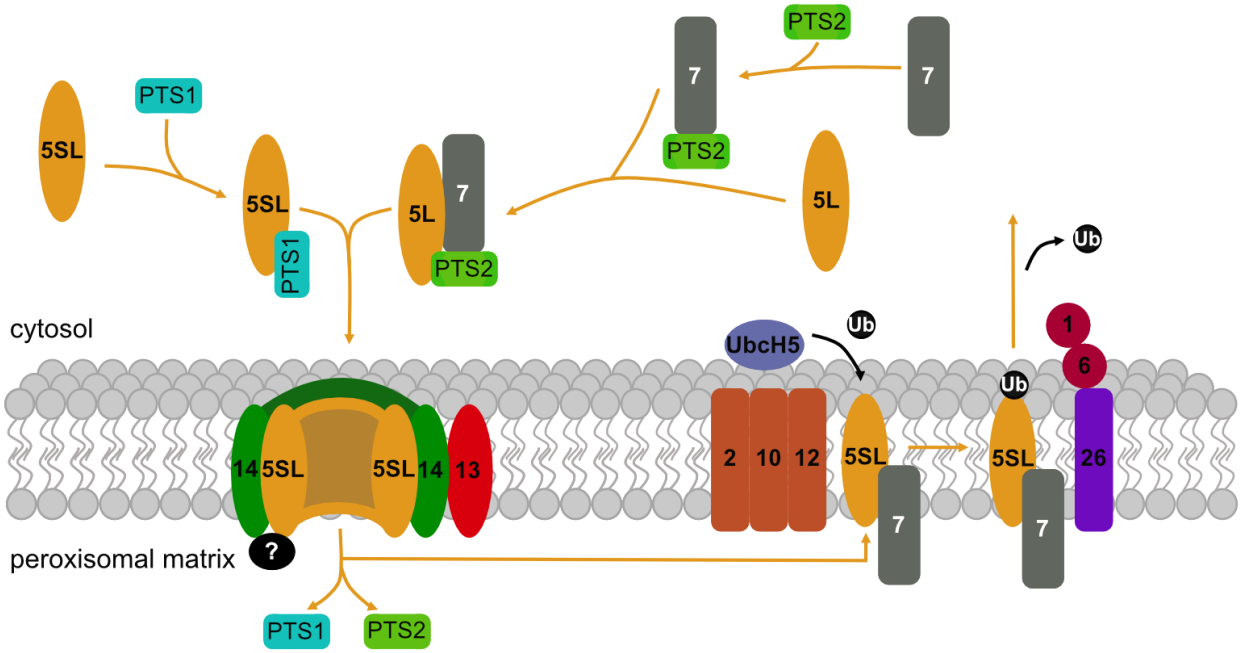

Figure 1.5: Peroxisomal matrix import pathways in mammals. Cytosolic PTS1 and PTS2 cargo is recognized by PEX5S and PEX7/PEX5L respectively. Cargo complexes are shuttled to the peroxisomal membrane where it docks onto PEX14. After formation of the docking complex a dynamic pore of PEX5, PEX14 and eventually PEX13 is formed and the cargo translocated into the lumen. What the cargo release triggers is not known at this point. Receptor recycling is mediated by monoubiquitination via an ubiquitin ligase cascade including UbcH5 proteins and the RING finger complex Pex12p/Pex10p/Pex2p. and delocalization via the PEX26 anchored AAA ATPase PEX1/PEX6.

The peroxins PEX5, PEX14 and PEX13

As discussed above the PEX5 receptor is essential for peroxisomal matrix import shuttling between a cytosolic an integral membrane state. It was shown that PEX5 can spontaneously insert into artificial membranes but is not able to form a pore by itself (Kerssen *et al*, 2006). Its function is mediated by two distinct domains located either at the N-terminus or C-terminus. The N-terminal domains of PEX5L (residues 1-315) and PEX5S (residues 215-251 missing) are intrinsically disordered and harbor seven or six (di)aromatic pentapeptide motifs with the consensus sequence WxxxF/Y (W-motifs), respectively (**Figure 1.6A**) (Otera *et al.*, 2002). The residues 215 to 251 that harbor the W5 motif were identified to mediate PEX7-PTS2 binding. However, W motifs and another N-terminal located LVAEEF (W-like) motif (**Figure 1.6A**) are essential for PEX14 binding and peroxisomal membrane targeting (Neuhaus *et al.*, 2014; Schliebs *et al*, 1999). All (di)aromatic peptide motifs except W4 bind the PEX14-NTD with nanomolar affinity (Saidowsky *et al*, 2001; Schliebs *et al.*, 1999). Structures of PEX5 W0 and W1 in complex with PEX14-NTD show α-helical conformations of the motifs, which are either preformed or induced upon binding (**Figure 1.6B, C**) (Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014). PTS1 cargo recognition is mediated by its C-terminal TPR domain (residues 315-639), which consists of 7 tetratricopeptide repeat (TPR) motifs arranged into two clusters: The first one consists of TPRs 1-3 and is connected to the second cluster, composed of TPRs 5-7, *via* a hinge loop formed by TPR 4 (**Figure 1.6D**). Moreover, the C-terminus harbors another ~ 50 residues forming an additional bundle of three helices (**Figure 1.6D**). Those helices provide an extra binding site for cargos like SCP2. However, high affinity PTS1 binding per se is mediated by large parts of TPR 5-7 and TPR 1-3 (Dodt *et al*, 1995; Gatto *et al*, 2000; Stanley *et al*, 2006).

Besides PEX5, PEX14 is the second most studied component of the docking and translocation machinery. The 377 amino acids long protein contains a N-terminally located small α-helical domain followed by a hydrophobic –putative transmembrane- domain (annotated as MD), a coiled coil region and an intrinsically disordered C-terminal region (**Figure 1.6E**) (Fransen *et al*, 1998; Shimozawa *et al*, 2004). Although PEX14 was clearly located at the peroxisomal membrane, its topology is still under debate. Some studies propose a N- in C- out membrane topology (Barros-Barbosa *et al*, 2019; Will *et al*, 1999), which is in conflict with other observations proposing a N-out, C- in topology (Bharti *et al*, 2011; Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014). In addition, even a C-, N- out topology has been proposed (Shimizu *et al.*, 1999). It is likely that PEX14 can

adopt multiple topologies depending on the peroxisomal state and metabolic demand. The NTD is the only structured and best characterized domain of PEX14. Besides binding of PEX5 (di)aromatic peptide motifs, which tethers PEX5 to the peroxisomal membrane (**Figure 1.6B, C, E**), it was shown to interact with the PEX19 FxxxF motif involved in PMP import. PEX14 was further shown to mediate microtubule-based peroxisome motility via its N-terminus (Bharti *et al.*, 2011; Reuter *et al*, 2021). To date, direct interactions with PEX13 have not been reported.

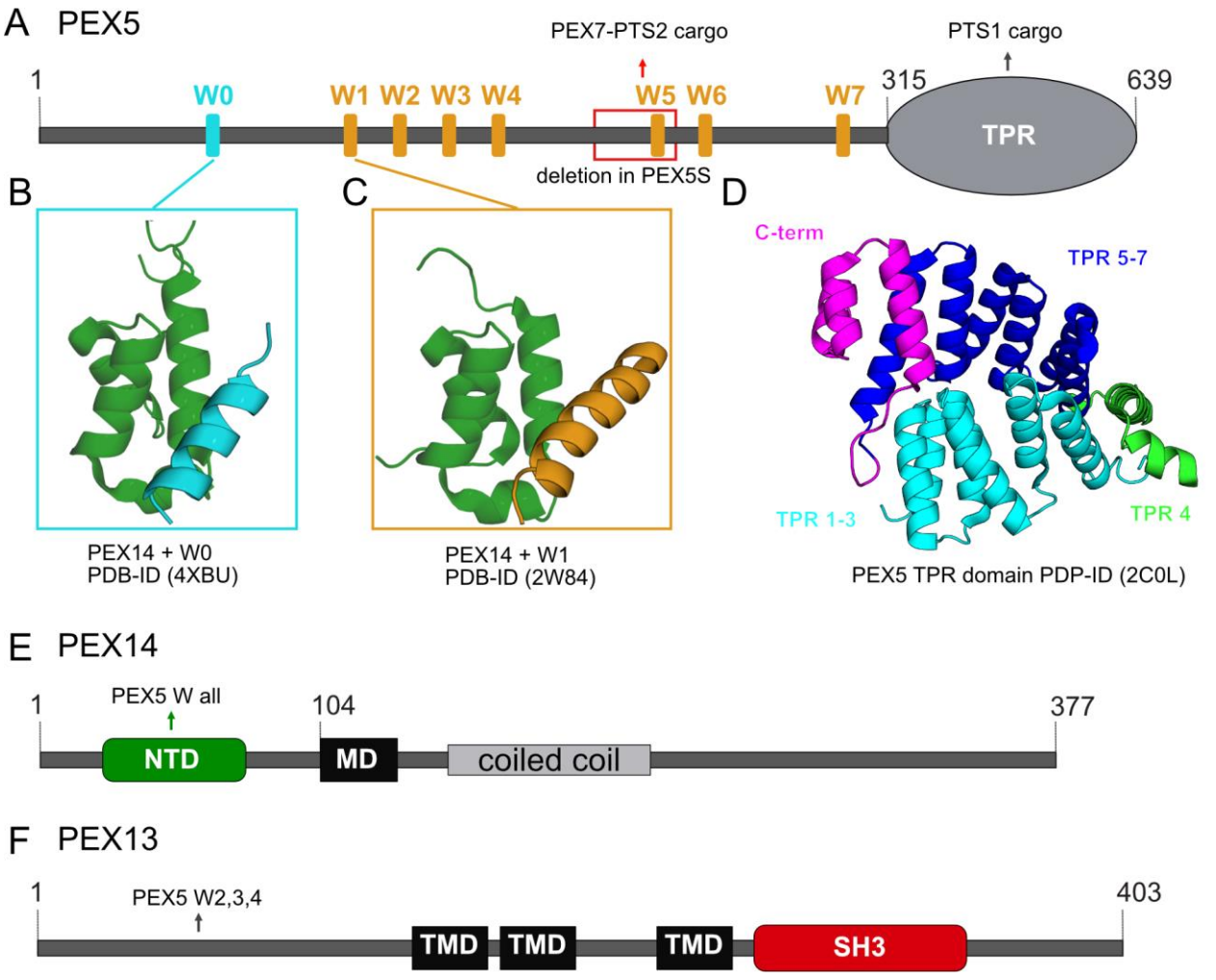

Figure 1.6: Domain architecture of the docking and translocation complex components PEX5, PEX14, PEX13. (A) PEX5 consists of an N-terminal disordered domain and a C-terminal TPR domain. The NTD harbors multiple WxxxF/Y motifs and a LVAEEF motif (W0-W7) indicated in blue and orange colors. The shorter version PEX5S lack amino acid 215 to 251 including W5 (red box). (B) NMR structure of PEX14-NTD (green) with PEX5 W0 (cyan). (C) NMR structure of PEX14-NTD (green) with PEX5 W1 (orange). (D) X-ray structure of human PEX5 TPR (TPR1-3 cyan, TPR4 green, TPR 5-7 blue) and helical C-terminus (magenta). (E) The

membrane associated PEX14 consists of a N-terminal, -helical domain (NTD) followed by a membrane, a coiled coil domain and a disordered C-terminus. PEX14-NTD is able to bind all W (-like) motifs from PEX5. (F) The integral membrane protein PEX13 possesses a disordered N-terminus, three transmembrane domains and a C-terminal region, which includes an SH3 domain followed by an intrinsic disordered region.

Human PEX13 is an integral peroxisomal membrane protein with an intrinsically disordered N-terminal domain followed by a transmembrane region of likely three spans and a mostly unstructured C-terminal region harboring a SH3 domain. (**Figure 1.6F**). Analogous to PEX14, two different topologies of PEX13 have been reported (Barros-Barbosa *et al.*, 2019; Gould *et al*, 1996a). PEX13 was first identified and studied in the context of peroxisomal import (Gould *et al.*, 1996a) and Zellweger spectrum disorder (Liu *et al.*, 1999). Clear indications for its function in peroxisomal import are limited. PEX13 was reported to be essential for catalase with a non-canonical PTS1 peptide but not for PTS1 import in general. The same study showed that PEX13 interacts with PEX5 via its N-terminus in GST pulldown experiments (**Figure 1.6F**), which is the only reported interaction to PEX5 to date (Otera *et al.*, 2002). Later the Zellweger mutation W313G located in the PEX13-SH3 domain was demonstrated to disrupt PTS1, but not PTS2 import while interaction with PEX14 was still intact (Krause *et al.*, 2006; Krause *et al*, 2013).

Role of (peroxisome-related) glycosomes in trypanosomatidea

Parasitic infections with the Trypanosoma genus are responsible for several diseases. Those pathogenic unicellular flagellates are transmitted by insect-vectors (e.g. Tsetse fly) in tropical and subtropical regions of Latin America, Africa and Asia. The most common diseases to human are Human Afrikan trypanosomiasis (HAT) caused by *Trypanosoma brucei* , Chagas Disease caused by *Trypanosoma cruzi* and leishmaniasis induced by *Leishmania* (Desjeux, 2004; Poltera *et al*, 1977; Rassi *et al*, 2010). Besides infecting humans, the parasites infect wild and domestic animals causing mortality and severe economic losses. However, these diseases are also present in non-endemic regions, including North-America, Europe or in the Western-Pacific region due to population mobility (Coura & Viñas, 2010). Different types of trypanosomiases are not very well treatable since available drugs show major side effects and lack efficacy against different stages of the disease (Castro *et al*, 2006). Moreover, there is no vaccination available.

Glycosomes - a potential drug target

Glycosomes are peroxisome-related organelles found in the human pathogenic trypanosomes *T. brucei, T. cruzi* and *Leishmania*. In Trypanosomatids, the glycolytic and peroxisomal function is located in the glycosome. Glycosomal activity is essential for the survival of parasites, as glycolysis is the sole source of adenosine-triphosphate (ATP) in the bloodstream stage. Deficiency in glycosome biogenesis leads to accumulation of glycolytic enzymes in the cytosol where their unregulated enzyme activities deplete cellular ATP levels by runaway phosphorylation of glucose, which accumulates up to the toxic levels and leads to death of the bloodstream parasite (Bakker *et al*, 2000; Furuya *et al*, 2002; Haanstra *et al*, 2016; Kalel *et al*, 2017). This behavior makes glycosome biogenesis an interesting drug target. Notably, the glycosome/peroxisome biogenesis including glycosomal import of enzymes is mechanistically conserved from higher eukaryotes to trypanosomatids. This led to the successful development of small-molecule inhibitors that block glycosomal matrix protein import by inhibiting the TbPEX14-TbPEX5 interaction leading to death of *Trypanosoma* parasites (Dawidowski *et al*, 2017; Dawidowski *et al*, 2020). We recently identified trypanosomal PEX3, which is required for glycosome biogenesis and thus trypanosomatid survival (Kalel *et al*, 2019). We furthermore identified small molecule inhibitors that target TbPEX3-TbPEX19 interaction and block glycosome biogenesis (Li *et al*, 2021). Those molecules are potential precursors for drug development against trypanosomiases.

Chapter 2:

Nuclear Magnetic Resonance (NMR) spectroscopy

Introduction of NMR spectroscopy

NMR spectroscopy is uniquely suited to study the structure, molecular interactions and dynamics of biomolecules in solution. The physics behind this method was discovered by a couple of scientists along with Isidor Isaac Rabi, who was awarded "for his resonance method for recording the magnetic properties of atomic nuclei". Further development was pushed by Felix Bloch and Edward Mills Purcell in the 1950, which were awarded the Nobel prize in 1952 "for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith." In 1952 the first NMR spectrometer HR-30 (30 MHz) was launched by Varian. NMR spectroscopy is, together with X-ray crystallography and cryo-electron microscopy, one of the three major techniques in structural biology. Every technique has its limitations; so has NMR, but it comes with a big advantage, which is the ability to not only determine structures but map molecular interactions and study conformational dynamics of biomolecules in solution. This allows us to detect molecular motions, weak interactions and low populated states.

Basic principles

The principle of NMR spectroscopy is based on the nuclear spin angular momentum μ of a given nucleus, which is represented by the characteristic spin quantum number (I). The spin quantum number (I) depends on the number of unpaired protons and neutrons (N_p, N_n) in the nucleus and can adopt the values of zero if N_p and N_n are even (e.g. ^{12}C, ^{16}O), integer numbers if N_p and N_n are odd (e.g. ^{2}H) or half-integer numbers if one of the numbers is odd and the other even (e.g. ^{1}H, ^{13}C, ^{15}N). The magnitude of the magnetic moment is directly proportional to the angular momentum via a specific factor, the gyromagnetic ratio γ, which depends on the fine structure of the nucleus (**Eq. 1**). Only isotopes with a non-zero spin are NMR active, and nuclei with spin ½ are most relevant for biomolecular NMR. As spin ½ carbon and nitrogen nuclei are only present with <1% at natural abundance isotope-labeling is important for NMR studies of biomolecules, which mostly involves ^{1}H, ^{13}C and ^{15}N nuclei (**Table 2.1**).

Table 2.1: Table of isotopes commonly used in biomolecular NMR spectroscopy.

Isotope	Natural Abundance (%)	Nuclear Spin (I)	γ (MHz/T)
^{1}H	99.985	1/2	42.58
^{13}C	1.108	1/2	10.71
^{15}N	0.365	1/2	-4.32
^{19}F	100	1/2	40.04
^{31}P	100	1/2	17.23

In the presence of an external magnetic field, the spins align with it (**Figure 2.1A**) and yielding two states +1/2 (α-state) and -1/2 (β-state) (Zeeman levels) with distinct energy levels, which are proportional to the magnetic field B_0 (**Eq. 2**) **(Figure 2.1B**). The energy difference between both states depends on the nuclei specific gyromagnetic ratio γ and can expressed in a frequency term, which corresponds to the Larmor frequency (**Eq. 3**).

Eq. 1 $$\mu = \gamma \vec{I}$$

with μ = nuclear spin angular momentum; γ = gyromagnetic ratio; I = spin quantum number

Eq. 2 $$E = -m\hbar\gamma B_0$$

with E = energy in Joule; m = quantum number (eigenvalues of I_z); ħ = reduced Planck constant; B_0 = magnetic field strength

Eq. 3 $$\Delta E = E_\alpha - E_\beta = +\frac{1}{2}\hbar\gamma B_0 - (-\frac{1}{2}\hbar\gamma B_0) = \hbar\gamma B_0 = \hbar\omega_0$$

with ΔE=energy difference; $m_{1/2}$= +1/2, -1/2; ħ=reduced Planck constant; γ=gyromagnetic ratio; B_0 magnetic field strength; Lamor frequency ω_0=-γB_0

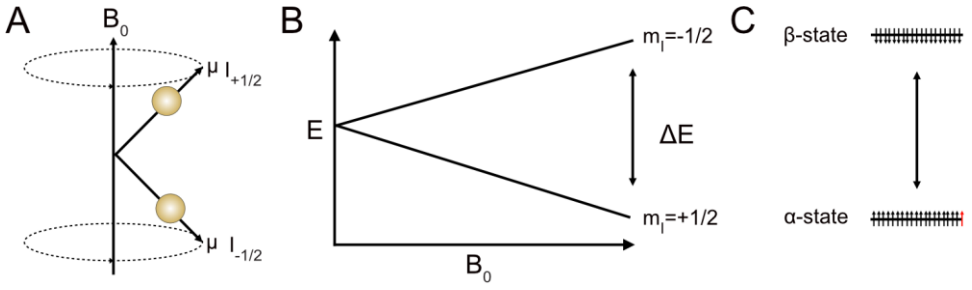

Figure 2.1: Behavior of nuclear spins with half-integer spin quantum number in the presence of an external magnetic field. (A) Nuclear spins orient parallel or anti parallel towards B_0 field. (B) The two states adopt distinct energy (Zeeman) levels, which are proportional to the external

magnetic field strength. (C) The lower energy α-state is slightly more populated than the β-state. The distribution is described by Boltzmann.

The spin distribution of the two states is described by the Boltzmann equation (**Eq. 4**).

Eq. 4

$$\frac{N_\alpha}{N_\beta} = e^{\frac{-\gamma\hbar B_0}{k_B T}}$$

With N_α and N_β = number of spins in the α- and β-state; γ = gyromagnetic ratio; ħ = reduced Planck constant; B_0 = magnetic field strength; k_B = Boltzmann constant and T = temperature

Solving the Boltzmann equation for protons (^{1}H) at 600 MHz shows an almost equal distribution of the spins in the α- and β-state with 1.00001 α and 1 β, which is a reason for low sensitivity of NMR spectroscopy (**Figure 2.1C)**. Thus technically, a way to gain sensitivity is to increase the magnetic field or lower the temperature and generate a more favorable spin state distribution.

In reality, the spins are not strictly aligned with the external magnetic field and thus divided in α- and β-states but would adopt all different combinations. Still the average over the large assembly of spins also described as "bulk magnetization" or "net magnetic moment" is aligned in the direction of the magnetic field M_Z (**Figure 2.1A**). The behavior of the bulk magnetization M_Z can be described in a vector model in a Cartesian coordinate system where the external magnetic field B_0 corresponds to the *z* axis. In equilibrium, the spins are aligned with the magnetic field with net magnetization corresponding to the difference of N_α and N_β (**Eq. 4**). The angular momentum of the spins causes a precession around the *z* axis corresponding to the Larmor frequency (ω_0=-γB_0) (**Figure 2.2B**). Randomly distributed *x*- and *y*- components in the transverse plane average out to zero.

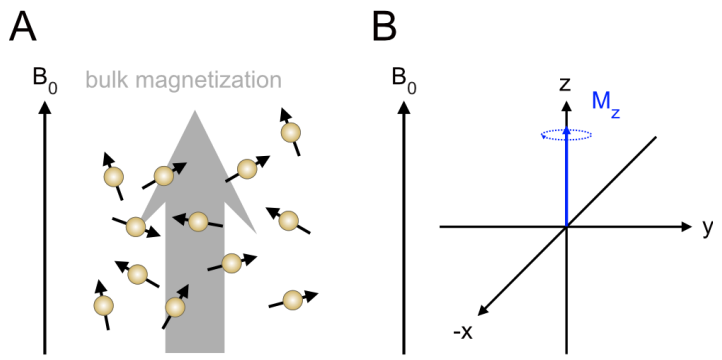

Figure 2.2: (A) Bulk magnetization arises over the large assemble of spins in thermal equilibrium. (B) In the vector model, the bulk magnetization is aligned with *z* rotating with the Larmor frequency (ω_0=-γB_0).

Signal detection

Before the experiment starts, bulk magnetization is aligned with the external magnetic field along the *z* axis (**Figure 2.3A**). In order to detect a NMR signal, spin states have to be changed. This is achieved by an on-resonance radio frequency (rf) pulse perpendicular to the *z* axis (on *x* or *y*), which induces a second magnetic field B_1. On resonance means the pulse oscillates with the Larmor frequency of the observed nucleus. For detection of a signal, the ideal pulse length causes a 90° flip of the magnetization into the transversal *x*-*y* plane ($M_{x,y}$) where the detection coil of the spectrometer sits (**Figure 2.3B)**. Now the magnetization precedes in the *x*-*y* plane before it naturally returns into equilibrium state along the *z* axis. The magnetization oscillating in the transverse plane induces a decaying signal, the so called free induction decay (FID) in the receiver coil, which is recorded and digitized (**Figure 2.3C)**. This time-domain signal is a superposition of all frequencies of different spins present in the sample. Ultimately, the time-domain signal is converted to frequency-domain by Fourier Transformation (FT) (**Figure 2.3D)**.

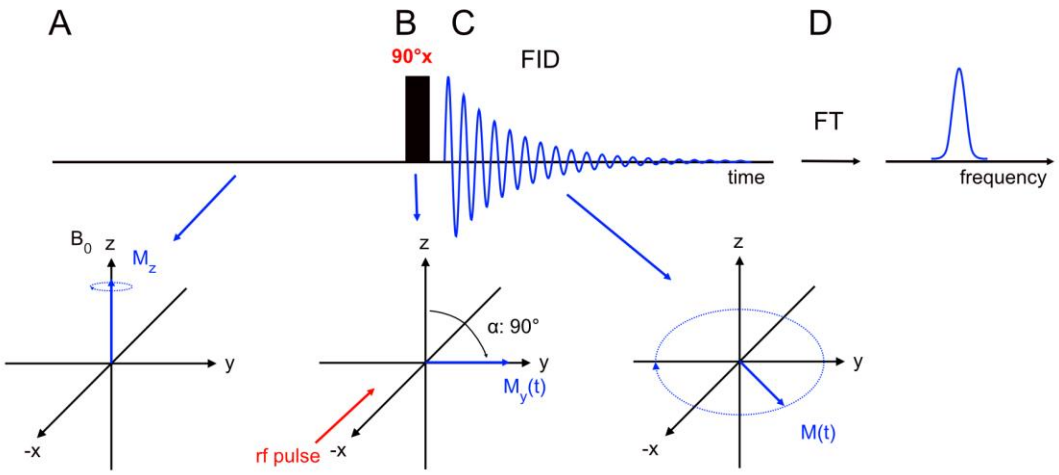

Figure 2.3: Schematic description of a 1D NMR Experiment. (A) Before the experiment starts the spins are in thermal equilibrium aligned on z. (B) A radio frequency (rf) pulse with a distinct amplitude is applied along *–x* to flip the bulk magnetization M_0 at 90° to the *y*-axis. (C) The angular momentum of the spins forces the precession in the *x*-*y* plane, which induces an oscillating current in the receiver coil of the spectrometer. This decaying signal, called FID, is monitored until the magnetization is decayed in the *x*-*y* plane, while recovery of *z* magnetization towards equilibrium (driven by T_1 relaxation) is required before the next scan can be recorded. (D) The superposition of signals in the time domain is then transformed into the frequency domain via Fourier Transformation.

The Larmor frequencies depend on the gyromagnetic ratios of the observed nuclei and the magnetic field strength. Accordingly, different rf pulses have to be applied. In theory, NMR spectra of each nucleus type should provide one identical resonance. But nuclei in molecules experience slightly different magnetic fields due to their chemical environment, which shifts the Larmor frequency. This is called chemical shift and is an important readout in biomolecular NMR. In ^{1}H NMR, chemical shifts are given in respect to the reference compound tetramethylsilane (TMS) and appear, spectrometer-specific (B_0), in Hz. Those values are converted in the spectrometer-independent unit ppm (parts per million) by dividing the shifts by the spectrometer frequency (**Eq. 5**).

Eq. 5. $$\delta ppm = \frac{\upsilon - \upsilon_{TMS}}{spectrometer\ frequency}$$

Relaxation

After flipping the bulk magnetization to the *x*-*y* plane it will return into the equilibrium state of only z-magnetization (M_z). This process is called relaxation and is mainly caused by dipole-dipole interactions and anisotropy. There are two different types of relaxation; (i) the spin-lattice relaxation (expressed by the time T_1), also known as longitudinal relaxation, where M_z is re-

established along the axis of the static applied magnetic field and (ii) the spin-spin relaxation *(T_2)*, also known as transverse relaxation, where $M_{y,x}$ decays due to the loss of phase coherence. The exponential decays of magnetization in the two relaxation phenomena are mathematically described by the Bloch equations (**Eq. 6 and 7**). Relaxation is primarily given by rates described as R_1 (R_1=1/T_1) and R_2 (R_2=1/T_2).

Eq. 6 $$\frac{dM_z(t)}{dt} = -\frac{1}{T_1}[M_z(t) - M_z^0]$$

Eq. 7 $$\frac{dM_{yx}(t)}{dt} = -\frac{1}{T_2}M_{yx}(t)$$

The relaxation depends on the rotational correlation time τ_c, which is proportional to the molecular size. Small molecules like chemical compounds with fast rotational diffusion and thus a small correlation time τ_c (pico- to nanoseconds) have relatively long relaxation times (milliseconds to seconds) with similar T_1 and T_2 (T_1~T_2). With increasing τ_c, both relaxation times become shorter. T_1 relaxation is most efficient for molecules tumbling at rates corresponding to the inverse Larmor frequency, which demonstrates the magnetic field dependency. With increasing τ_c, T_1 appears to be longer again while T_2 times further decrease. Larger biomolecules, compared to chemical compounds, such as proteins have longer correlation times (single digit to double digit nano-seconds) and thus shorter T_2 times (micro- to milliseconds) than T_1 times (milliseconds to seconds) (T_1>T_2) (**Figure 2.4**) (Gáspári & Perczel, 2010; Lee *et al*, 2006).

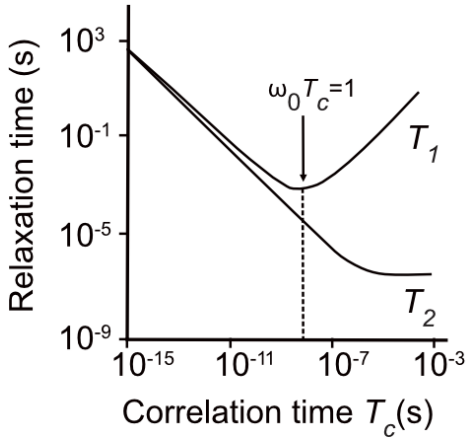

Figure 2.4: Correlation of relaxation times T_1 and T_2 and correlation time τ_c. T_2 decreases with increasing τ_c ($\propto$ molecular weight). T_1 is shortest when τ_c matches the frequency of spin state transitions. Adapted from Demangeat (2013).

Approaches in biomolecular NMR

NMR spectra of biomolecules provide not only information about chemical shifts but also about magnetic couplings of the neighboring spins, the so called scalar- or J- couplings. J-coupling arises from different spin transitions and are mediated typically over ~3 electron-bonds. The couplings are seen as peak-splitting in multiplets depending on the amount of coupling spins. The strength of the coupling depends on the distance of the spins and the bond angles between them, which is described by the Karplus curve (Karplus, 1963). The observation that J-couplings are able to transfer spin polarization between nuclei revolutionized multidimensional NMR spectroscopy. Hence, the method called 'insensitive nuclei enhanced by polarization transfer' (INEPT) where polarization is transferred from a nucleus of high gyromagnetic ratio (typically ^{1}H) to a nucleus of low gyromagnetic ratio (e.g. ^{15}N or ^{13}C) (Morris & Freeman, 1979) is commonly used in multidimensional NMR experiments. With this methodological advances and the introduction of recombinant protein expression together with isotope labelling, protein NMR became much more applicable. The enrichment of the protein sample with NMR-active nuclei such as ^{15}N and ^{13}C (**see Table 2.1**) allows us to record 2D and 3D heteronuclear experiments with reasonable resolution.

2D NMR spectra and NMR titrations

Basic and commonly used 2D spectra in protein NMR are ^{1}H,^{15}N or ^{1}H,^{13}C correlation experiments, also called "fingerprint", providing a protein-specific fingerprint of ^{1}H, ^{15}N/^{13}C resonances. A commonly recorded spectrum is the ^{1}H, ^{15}N heteronuclear single quantum coherence (HSQC) experiment where the magnetization is transferred from the proton to nitrogen and back through the ^{1}H-^{15}N bond (**Figure 2.6A**). This type of experiment is simple but the most important one in protein NMR as it provide residue-level resolution and serves as basis for a variety of applications like binding studies, relaxation- and dynamics experiments, or structure calculation.
Ligand binding is commonly explored by NMR titrations using ^{1}H-^{15}N correlation spectra. For this technique, a reference spectrum of a labeled protein is recorded before addition of the ligand. The same experiment is repeated after each stepwise ligand addition up to saturation monitoring the binding by chemical shift perturbations (CSPs) or line-broadening of the ^{1}H-^{15}N resonances, which is caused by change of the microenvironment and increasing molecular weight. The outcome of such experiments vary due to the kinetics of the complex formation with a ligand (L), which are

described by three exchange regimes (i) slow ($\Delta\nu >> k_{ex}$) **(Figure 2.5A)**, (ii) intermediate ($\Delta\nu \approx k_{ex}$) **(Figure 2.5B)** and fast exchange ($\Delta\nu << k_{ex}$) **(Figure 2.5C).** Here, $\Delta\nu$ is the frequency difference of the free and bound state and k_{ex} is the exchange rate, $k_{ex} = k_{on}[L]+k_{off}$, k_{on} and k_{off} are the on- and off-rates of the ligand binding to the protein and [L] is the ligand concentration (Göbl *et al*, 2014). In the slow exchange regime, only two sets of signals are visible, corresponding to the free and bound state respectively **(Figure 2.5A)**. The intensity of the signals correlate with to the relative population of the two states. This kind of exchange typically indicates a tight interaction in the nM range. In the case of weaker interactions in the µM to mM range, which show higher exchange rates, the observed signals correspond to the average chemical shift values of the free and bound state **(Figure 2.5C)**. Thus it becomes feasible to track chemical shift perturbations. When the exchange rate of a given interaction is in between of slow and fast exchange, a combination of line-broadening and chemical shift perturbation is observed, referred as intermediate exchange **(Figure 2.5B)**. This is the most unfavorable regime since the signals are often broadened beyond. Such a binding can be pushed towards slow or fast exchange regime by adjusting the temperature or the external magnetic field (Kleckner & Foster, 2011).

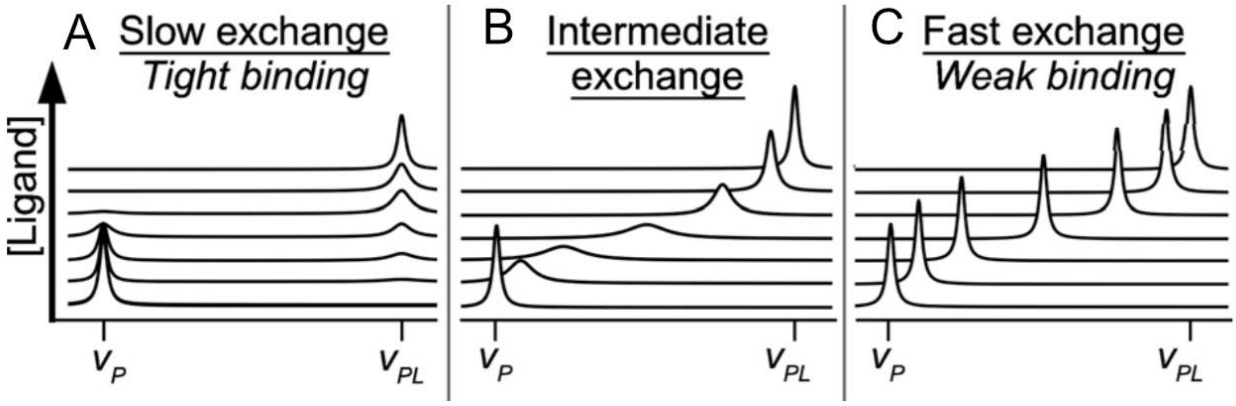

Figure 2.5: Binding regimes of complex formation. (A) slow exchange ($\Delta\nu >> k_{ex}$); (B) intermediate exchange ($\Delta\nu \approx k_{ex}$) and (C) fast exchange ($\Delta\nu << k_{ex}$) with $\Delta\nu$ = frequency difference of the free and bound state and k_{ex} = exchange rate. Adapted from Kleckner & Foster (2011).

3D NMR spectra and backbone chemical shift assignments

The ^{1}H, ^{15}N correlation spectrum is not only the basis for binding studies but also for others such as relaxation- and dynamics experiments. For the analysis of those spectra, each residue of the polypeptide chain has to be assigned to the corresponding ^{1}H-^{15}N signal, which is achieved by recording a set of 3D experiments to obtain backbone and sidechain information (mostly CO, Cα

and Cβ) of each residue. The most prominent experiment, among these, is the HNCACB where the magnetization from the proton (^{1}H) is first transferred to N_H and subsequently to Cα and Cβ nuclei before it is transferred back to H_N for detection (**Figure 2.6B**). Since the N_H of the i residue also couples over ^{2}J with the Cα of the preceding i-1 residue, this experiment is well suited for a sequential backbone resonance assignment by performing the "backbone walk". However, a typical set of 3D experiments for backbone resonance assignments consists of the HNCACB and further experiments like the CBCA(CO)NH, HNCO and HN(CA)CO. An overview of NMR experiments for protein assignments and structure determination is described by Sattler M *et al* (1999).

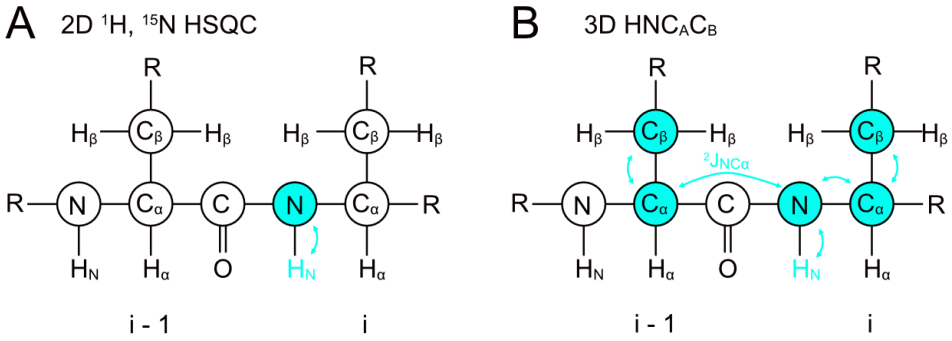

Figure 2.6: 2D and 3D heteronuclear NMR experiments. (A) In a simple ^{1}H, ^{15}N HSQC - 2D experiment, the magnetization is transferred from the proton to nitrogen and back through the ^{1}H-^{15}N bond. (B) In 3D experiments such as HNCACB, the magnetization is further transferred from the ^{15}N nucleus to Cα and Cβ nuclei, and subsequently back to H_N for detection.

Transverse relaxation-optimized spectroscopy TROSY

Larger biomolecules with slow rotational diffusion show fast transverse relaxation caused by dipole-dipole (DD) coupling and chemical shift anisotropy (CSA), which is a limiting factor for biomolecular NMR. Production of fully or partially deuterated proteins has proven useful to reduce the proton density eliminating a major source of relaxation, and thus enhance spectral quality of higher mass proteins (>25 kDa) (LeMaster, 1989; Sattler & Fesik, 1996). However, manipulating the ^{1}H, ^{15}N HSQC pulse sequence for optimal suppression of transverse relaxation using the "TROSY effect" has facilitated NMR of large proteins up to ~1MDa (Fiaux *et al*, 2002).

Basis of this technique is the use of different relaxation rates of the individual multiplet components by DD and CSA interference in a two spin system (*I, S*) such as ^{1}H, ^{15}N. A spectrum of two coupled spins gives rise to a multiplet of four, which is usually merged into one signal by respective decoupling. However, it turns out that signals of this multiplet appear with different line shapes

caused by DD and CSA interference leading to different T_2 relaxation times. TROSY observes solely the component with the narrowest linewidth experiencing transverse relaxation mainly from DD interactions with remote protons, which can be strongly reduced by perdeuteration. This specific coherence selection is archived by a couple of pulsed field gradients and phase cycling (Pervushin *et al*, 1997). This technique is successfully applied to three dimensional NMR for backbone chemical shift assignment (Loria *et al*, 1999; Salzmann *et al*, 1998). Note that perdeuteration only allows NMR experiments, which start on H_N nuclei.

Secondary structure from secondary chemical shifts

In the classical way, protein secondary structure is determined by nuclear Overhauser effect (NOE) measurements. However, this approach requires a time consuming full assignment of backbone protons. The use of carbon secondary chemical shifts is much simpler and less time consuming.
Chemical shifts are sensitive to a couple of atomic and molecular effects, which in turn are dependent on atomic composition, the solvent constitution and molecular geometry. In protein NMR, the atomic composition and solvent constitution is limited to the naturally occurring amino acids and aqueous solution, respectively. The geometry of the polypeptide backbone, on the other hand, depends on the secondary structure and is quite different between random coil, α-helices or β-sheets. Within the secondary structures of α-helices and β-sheets, the backbone torsion angles ϕ and ψ adopt approximately the same values, which leads to a specific down- or upshift behavior of the involved proton and carbon resonances such as Hα, ^{13}CO, ^{13}Cα and ^{13}Cβ relative to random coil. Those values are defined as secondary chemical shifts ($\Delta\delta = \delta_{observed} - \delta_{random\ coil}$). As a backbone chemical shift assignment is crucial for most applications in NMR, the ^{13}Cα and ^{13}Cβ are already available without performing additional experiments. Studies of protein structures and chemical shifts showed a downfield shift in helices and a upfield shift in sheets for ^{13}Cα resonances and the opposite behavior for ^{13}Cβ with respect to random coil values (Szilágyi, 1995; Wishart *et al*, 1991). Plotting the difference of $\Delta\delta^{13}$Cβ and $\Delta\delta^{13}$Cα secondary chemical shifts ($\Delta\delta(^{13}C\alpha) - \Delta\delta(^{13}C\beta)$, first used by Metzler *et al* (1993), produce a clearer-to-evaluate read out of secondary structure propensities.

^{15}N NMR relaxation to study backbone dynamics

Intramolecular motions of proteins, on a timescale ranging from picoseconds to seconds, such as the overall tumbling correlation time (nanoseconds), dynamics of the protein backbone (pico- to nanoseconds) or chemical exchange (micro- to millisecond), are important factors to determine biological activities. NMR spectroscopy is uniquely suited for measurements of relaxation rates, even for those, which are shorter than the rotational correlation time.

As mentioned earlier, relaxation is mostly driven by DD interactions and CSA. However, T_2 relaxation is also affected by other pseudo-first-order processes with a major contribution of ^{15}N conformational exchange (R_{ex}) (Bloom *et al*, 1965). ^{15}N relaxation is typically measured as a two-dimensional, ^{1}H detected inverse recovery experiment with delays introduced to measure ^{15}N relaxation in the range of milliseconds to seconds or microseconds to milliseconds for T_1 or T_2 relaxation, respectively. That means long delay times allow relaxation to evolve, which in turn produce low signal intensities. Shorter delay times on the opposite, produce higher signal intensities with a maximum at time 0. T_1 or T_2 relaxation experiments are analyzed by fitting the decay of the different signal intensities over time using a two parameter function (**Eq. 8)** (Farrow *et al*, 1994). The readout of such fitting are the relaxation rates R_1 and R_2 which can be used to determine the global isotropic correlation time by the ration of R_2/R_1 or rapid identification of the chemical exchange (R_{ex}) by the product of R_1R_2 (Kay *et al*, 1989; Kneller *et al*, 2002)

Eq. 8 $$I(t) = I_0 e^{(-\frac{t}{T_{1,2}})}$$

With I(t) = intensity after delay time *t*, I_0 = intensity at time 0, t = delay time, $T_{1,2}$ = relaxation times.

The above mentioned relaxation experiments were developed to characterize folded proteins, where overall tumbling and internal motion can be assumed to be well decoupled, and are not well suited to describe dynamics in intrinsically disordered proteins (IDPs) adopting many interchanging conformational states of similar energies at low levels. The unfolded state makes the labile backbone amid protons (H_N) accessible to the solvent providing excellent qualification for proton exchange experiments. The relaxation-based steady-state {^{1}H}-^{15}N nuclear Overhauser effect (NOE), also called {^{1}H}-^{15}N hetNOE experiment, detects motions on the pico- to nanosecond time scale and is often used to identify intrinsically disordered regions in folded proteins (or low populated secondary structures in IDPs). In this type of experiment, a steady-state

NOE between ^{1}H and ^{15}N of the protein backbone is measured in proton-saturated and -unsaturated spectra, which are normally recorded in an interleaved manner to reduce progressive sample inhomogeneity. Fluctuations of spin interactions caused by local vibrations, dihedral angle jumps between rotamers or rotational diffusion lead to cross-relaxation within a few nanoseconds to tens of nanoseconds and thus reduce signal intensities of the given ^{1}H, ^{15}N-NOE resonances in the saturated spectrum. The unsaturated spectrum shows full signal intensity and serves as control spectrum, which is used to calculate the hetNOE value (hetNOE = $I_{sat}/I_{unsat.}$). Note that relaxation experiments are sensitive to the field strength. Thus datasets to be compared need to be measured on the same magnetic field with the identical experimental setup (Bolik-Coulon *et al*, 2019; Farrow *et al.*, 1994; Stone *et al*, 1992).

Chapter 3:

Scope of the thesis

Scope of the thesis

The complexity of eukaryotic cells is given by organelles, which utilize specialized functions. The majority of nuclear encoded proteins are targeted to one of those organelles to fulfill their functions. Unlike in other organelles, the function of peroxisomes relies exclusively on the import of proteins since the organelle does not contain genomic information. Still, the mechanisms of the sequential steps from docking to translocation and release remain elusive. In the human system most insight was obtained through research studying peroxisome biogenesis disorder diseases. This especially applies to the peroxin PEX13. However, the absence of structural and mechanistic data is likely linked to the transient and highly dynamic nature of the pore consisting of membrane-associated, partly disordered proteins, which is challenging for structural biology and biophysical techniques. In this thesis I combine NMR spectroscopy and X-ray crystallography with biophysical methods to address questions regarding peroxisomal matrix protein import. The thesis focusses on three relevant aspects of the peroxisomal docking complex:

I. PEX5 and PEX14 are key proteins in peroxisomal targeting and subsequent pore formation. Docking is mediated by (di)aromatic peptide motifs from the intrinsic disordered N-terminal half of PEX5 that interact with the globular N-terminal domain of membrane associated PEX14. In particular, W0 is important for initial tethering. PEX5 appears special in that it adopts both a cytosolic and membrane associated form and can spontaneously insert into artificial membranes. However, the binding mechanism and structural information of the binding interface remains unknown. Hence, part of this study aims at identification of the membrane interaction sites of PEX5-NTD. Another aim is the evaluation of possible membrane interaction sites in the PEX14-NTD.

II. Interestingly, individual (di)aromatic peptide motifs of PEX5 NTD show different binding affinity with the N-terminal domain of PEX14. Still, it is not clear which residues in the (di)aromatic peptide motifs are responsible for optimal binding. We are thus interested in the characterization of binding energetics and thermodynamics of the individual motifs, which help to identify key residues and favorable peptide conformations. The goal is to use this information to refine the consensus sequence that helps to identify good binders solely by sequence analysis.

III. Very little is known about the peroxin PEX13, which is in fact an integral peroxisomal membrane protein shown by biochemical assays. However, structural information is not available to date. PEX13 are mostly studied in the context of Zellweger spectrum disorders but its molecular

functions and role in peroxisomal import remains limited. Noteworthy, interactions between PEX13 and PEX5 as well as PEX14 are reported on a biochemical level, but seem not to be conserved between yeast and human. The aim of this project part is the first structural characterization of the PEX13 C-terminus and mapping of its interaction sites with PEX14 and PEX5 to explore functions in peroxisomal matrix import.

Chapter 4:

Material and Methods

Material and Methods

Materials

Chemicals and consumables

Table 4.1: Chemicals and Consumables. The suppliers from chemicals vary over time. Chemicals were typically purchased from VWR, Merck, Carl-Roth or Sigma-Aldrich and isotope labeled chemicals from Cambridge Isotope Laboratories or Sigma-Aldrich.

Chemical	Supplier
$^{15}NH_4Cl$	Sigma-Aldrich, Steinheim, Germany
2-propanol	Sigma-Aldrich, Steinheim, Germany
40% Acrylamide/Bis-Solution (37.5:1)	Serva, Heidelberg, Germany
99% Glycerol	VWR, Darmstadt, Germany
Acetic acid	Merck, Darmstadt, Germany
Agar	Sigma-Aldrich, Steinheim, Germany
Agarose	Merck, Darmstadt, Germany
APS	Sigma-Aldrich, Steinheim, Germany
Biotin (D-(+))	Carl-Roth, Karlsruhe, Germany
Boric acid	Carl-Roth, Karlsruhe, Germany
$CaCl_2$	Merck, Darmstadt, Germany
$CoCl_2$ x 6 H_2O	Sigma-Aldrich, Steinheim, Germany
Coomassie Briliant Blue G250	Sigma-Aldrich, Steinheim, Germany
$CuCl_2$ x 2 H_2O	Sigma-Aldrich, Steinheim, Germany
D_2O	Sigma-Aldrich, Steinheim, Germany
D7PC	Avanti Polar Lipids, Alabama, USA
D-Glucose	Merck, Darmstadt, Germany
D-Glucose (U-13C)	Cambridge isotope laborytories, Massachusetts, USA
D-Glucose (U-13C, 1, 2, 3, 4, 5, 6, 6-D7)	Sigma-Aldrich, Steinheim, Germany
DMPC	Avanti Polar Lipids, Alabama, USA
DMSO	Merck, Darmstadt, Germany
DOPC	Avanti Polar Lipids, Alabama, USA
DOPE	Avanti Polar Lipids, Alabama, USA
DPC	Avanti Polar Lipids, Alabama, USA
DTT	Sigma-Aldrich, Steinheim, Germany
EDTA	Merck, Darmstadt, Germany
EtOH	VWR, Darmstadt, Germany
$FeCl_3$ x 6 H_2O	Sigma-Aldrich, Steinheim, Germany
H_3BO_3	Carl-Roth, Karlsruhe, Germany
HCl	VWR, Darmstadt, Germany
Imidazole	Merck, Darmstadt, Germany
IPTG	Carl-Roth, Karlsruhe, Germany
Kanamycin	Carl-Roth, Karlsruhe, Germany
KH_2PO_4	Merck, Darmstadt, Germany

LiCl	Carl-Roth, Karlsruhe, Germany
MgCl2	VWR, Darmstadt, Germany
MgSO4	VWR, Darmstadt, Germany
$MnCl_2$ x 6 H_2O	Merck, Darmstadt, Germany
Na_2HPO_4	VWR, Darmstadt, Germany
NaCl	VWR, Darmstadt, Germany
$NaClO_4$	Sigma-Aldrich, Steinheim, Germany
NaH_2PO_4	VWR, Darmstadt, Germany
NaOH	Merck, Darmstadt, Germany
Ni(III)Cl_2	Merck, Darmstadt, Germany
SDS	VWR, Darmstadt, Germany
Sodium acetate	Sigma-Aldrich, Steinheim, Germany
Temed	Carl-Roth, Karlsruhe, Germany
Thiamin	Carl-Roth, Karlsruhe, Germany
Tris-aminomethane	VWR, Darmstadt, Germany
Tris-HCl	Amresco, Ohio, USA
Tryptone	Merck, Darmstadt, Germany
Urea	Carl-Roth, Karlsruhe, Germany
Yeast extract	Merck, Darmstadt, Germany
$ZnCl_2$	Sigma-Aldrich, Steinheim, Germany
β-Mercaptoethanol	Sigma-Aldrich, Steinheim, Germany
Consumable	**Manufacturer**
50x TAE buffer	Serva, Heidelberg, Germany
6x DNA loading dye	New England Biolabs, Ipswich, USA
Amicon MWCO 10k, 15mL	Sigma-Aldrich, Steinheim, Germany
Amicon MWCO 3k, 15mL	Sigma-Aldrich, Steinheim, Germany
DNA stain G	Serva, Heidelberg, Germany
DpnI	New England Biolabs, Ipswich, USA
Falcon tubes 15, 50mL	Greiner bio one, Kremsmünster, Austria
Filter 0.22/0.45µm	Sortarius, Göttingen, Germany
GeneRuler 100bp	ThermoFisherScientific, Waltham, USA
GeneRuler 1kb	ThermoFisherScientific, Waltham, USA
Lysozyme	Serva, Heidelberg, Germany
Ni-NTA resin	Qiagen, Monheim am Rhein, Germany
Pipette tips 10, 200, 1000µL	StarLab, Berlin, Germany
Q5 polymerase kit	New England Biolabs, Ipswich, USA
Reaction tubes 1.5, 2.0mL	Eppendorf, Hamburg, Germany
SnakeSkin Dialysis tubing	ThermoFisherScientific, Waltham, USA
SUMO hydrolase	Arie Geerlof, HMGU, Germany
Syringes 2, 5, 10, 20, 50ml	Braun, Melsungen, Germany
TEV protease	Arie Geerlof, HMGU, Germany
Transferpette S 10, 100, 200, 1000µL	Brand, Wertheim, Germany
Wizard® Plus SV Minipreps DNA Purification System	Promega, Madison, USA
Wizard® SV Gel and PCR Clean-Up System	Promega, Madison, USA

Vectors

All proteins and protein fragments were cloned into a pETM-11 with TEV cleavable N-terminal His_6-Tag or a with an N-terminal His_6-SUMO modified pETM13 (pETM13S) vector. For details see experimental procedures of chapter 5, 6 or 7.

Bacterial strains

Table 4.2: Bacterial strains. *E.coli Bl21* (DE3) and DH10β were used for protein expression and DNA amplification respectively.

Strain	Genotyp
E. coli BL21 (DE3)	F^- *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)
E. coli DH10β	Δ*(ara-leu) 7697 araD139 fhuA* Δ*lacX74 galK16 galE15 e14- ϕ80*d*lacZ*Δ*M15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1* Δ*(mrr-hsdRMS-mcrBC)*

Media and solutions for protein expression and DNA amplification

Table 4.3: Media components for protein expression and DNA amplification. All media were supplemented with 50 mg/l kanamycin and sterile filtrated with a 0.22 µm Steritop filter. Components used for perdeuterated protein expression were lyophilized and dissolved in 99.9 % D_2O.

Medium	Components per liter
Lysogeny broth (LB)	10g tryptone, 10g NaCl, 5g, yeast extract
M9 minimal medium	7.52g Na_2HPO_4-$2H_2O$, 3g KH_2PO_4, 0.5g NaCl, 0.5g $^{15}NH_4Cl$, 4g glucose or 2g U-[^{13}C]-D-glucose/U-[^{13}C]-D-glucose-d_7, 1 mM $MgSO_4$, 0.3 mM $CaCl_2$, 1µg Biotin, 1µg Thiamin, 10ml 100x trace elements stock solution (see below)
100x trace elements stock solution	5g EDTA, pH 7.5, 0.83g $FeCl_3$ x $6H_2O$, 84g $ZnCl_2$, 13g $CuCl_2$ x $_2$H2O, 10g $CoCl_2$ x $_6$H2O, 10g H_3BO_3, 1.6g MnCl2 x $6H_2O$

Buffers for protein purification:

Table 4.4: Buffers for protein purification.

Buffer	Components
Lysis a	50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 20mM Imidazole
Wash a	50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 20mM Imidazole
Wash a-Urea	50 mM Tris-HCl, pH 7.5, 750 mM NaCl, 20mM Imidazole, 8M Urea
Wash a-high salt	50 mM Tris-HCl, pH 7.5, 750 mM NaCl, 20mM Imidazole
Elution a	50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 500mM Imidazole
Lysis b	50 mM Tris-HCl, pH 8, 100 mM NaCl
Wash b	50 mM Tris-HCl, pH 8, 100 mM NaCl, 20mM Imidazole
Elution b	50 mM Tris-HCl, pH 8, 100 mM NaCl, 500mM Imidazole
Lysis c	50 mM Tris-HCl, pH 8, 1M NaCl, 1% (w/v) DPC
Wash c	50 mM Tris-HCl, pH 8, 0.5M NaCl, 0.2% (w/v) DPC, 20mM Imidazole
Elution c	50 mM Tris-HCl, pH 8, 0.5M NaCl, 0.2% (w/v) DPC, 500mM Imidazole

Buffers a were used for all PEX5 and PEX13 constructs while buffers b and c were used for PEX14 1-104 and PEX14 1-137 respectively. For detailed information, see experimental procedures of chapter 5, 6 or 7.

Final buffers:

Table 4.5: NMR/ITC/Xtal buffer compositions

Buffer	Components
NMR/ITC (pH 6.5)	20mM sodium phosphate (7.13 mM Na_2HPO_4, 12.85mM NaH_2PO_4, 50 or 100mM NaCl, 2 to 4 mM DTT (if needed)
NMR/ITC (pH 7.5)	20mM Tris-HCl, 50 mM NaCl
Xtal (pH 7.5)	5mM Tris-HCl, 50mM NaCl

Molecular biology

Cloning

If not stated otherwise, all proteins and protein fragments as well as mutations were cloned using Site-directed, Ligase-Independent Mutagenesis (SLIM) (Chiu, 2004) in an extended version. In this approach, two sets of two primers from, which one is short and one has a tail are designed. The tail primers (forward-tail and reverse-tail) harbor the mutation and/or just an overlapping region with the complementary tail primer **(Figure 4.1)**. This setup allows the mutation, excision and insertion of certain DNA sequences. Two PCR reactions **(Table 4.6)** with a short and tail primer each, which amplify the full vector create two sets of overlapping DNA strings **(Figure 4.1)**, which are then annealed in three cycles of a temperature gradient **(Table 4.7)**.

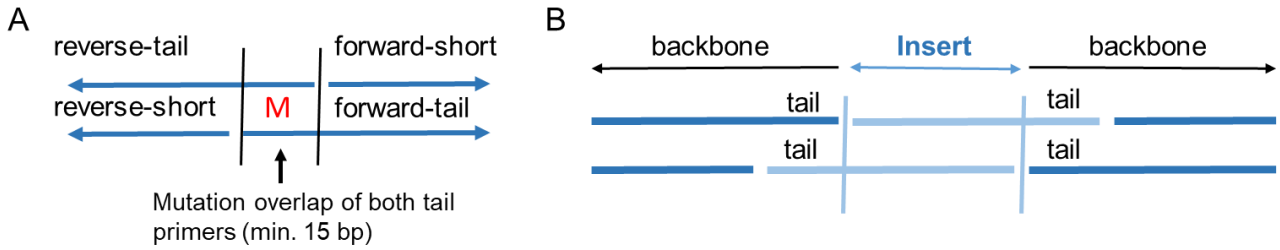

Figure 4.1: Schematic overview of SLIM primer design. (A) simple mutation, excision or insertion (B) insertion of larger inserts – amplified separately.

For larger inserts, which exceed the limits of primer synthesis, two sets of four primers have to be designed and the insert and the vector backbone have to be amplified separately **(Figure 4.1B)**, which are later mixed in a molar ration of 1:5 were 1 part is the vector backbone and 5 parts is the insert. The annealing process is equivalent to simple mutation, excision and insertion **(Table 4.6)**.

Table 4.6: Components for 25 µl PCR reactions using Q5 polymerase

Component	Volume
5x Q5 Reaction Buffer	5 µl
Q5 polymerase	0.25 µl
10 mM dNTPs	0.75 µl
10 ng DNA template	1 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
Nuclease free Water	15.5 ul

Table 4.7: Temperature and times of the annealing cycle (modified from Chui et al., 2004)

Temperature	Time	Cycles
99 °C	3 min	
65 °C	5 min	3
30 °C	40 min	

PCR products were treated with 2 µl DpnI and incubated at 37 °C for 1h before purification with Wizard® SV Gel and PCR Clean-Up System. DnpI degrades methylated DNA and thus removes the template DNA from the mixture. The pure products were used for a 20µl SLIM annealing reaction **(Table 4.7)** containing 4 µl 5x buffer H (300 mM NaCl, 50 mM Tris, pH 9.0, 20 mM EDTA, pH 8.0) from, which 10 µl were used for transformation of 50 µl chemically competent DH10β cells. The cells were grown on LB agar plates containing antibiotic. Single colonies were picked from plate and cultured in 10 ml LB medium with antibiotic in 50 ml tubes for 7h or overnight. The cells were harvested after and the plasmids extracted using the Wizard® *Plus* SV Minipreps DNA Purification System and sequenced at Eurofinsgenomic or Azenta Genewiz.

Agarose gel electrophoresis

Agarose gel electrophoresis is a DNA separation method based on the molecular weight, which corresponds to the amount of base-pairs. The DNA is mobilized through the gel due to external electric field and the charge of negatively charged phosphate backbone. Smaller molecules are migrating faster than big molecules as their movement is less hindered by the mesh of agarose molecules. The speed of migration can thus be influenced by the agarose concentration in the gel. The Bands are visualized under UV light with a DNA-intercalating dye and the size determined by comparison with a standard marker.

Agarose gels were prepared by mixing 1% (w/v) agarose with TAE buffer (40 mM **T**ris-aminomethan, 20 mM **A**cetic acid, 1 mM **E**DTA pH 8) with a subsequent heating and cooling step. The DNA-intercalating dye (DNA Stain Clear G)) was added at a temperature of 50-60 °C before cooling down to RT. Samples were mixed with 6x loading dye and loaded on the gels together with a 1kb marker. The gels run for 20 to 25 min at 120V and were then analyzed using Geldoku (BioRad).

Transformation of chemically competent *E.Coli* cells

Transformation of chemically competent *E.Coli* cells is a quick and easy way to bring genetic material coding for a protein of interest into the cell. Once the plasmid has entered the bacterial lumen it will be kept via antibiotic selection.

50 µl of chemically competent cells are thawed on ice and then mixed with few ng of the plasmid of choice. The cells were incubated with DNA for 5 min, then heated 1 min to 42 °C and cooled down for 2 min on ice. After the addition of 500 µl LB medium, the cells were incubated for 1 to 1.5h at 37 °C while shaking. Finally, the cell suspension was plated on agar plates and or transferred in 50 ml liquid LB culture.

Protein expression

All proteins used in this study are encoded on pET vectors, which enables recombinant production in *E.Coli* strains carrying a gen for T7 RNA polymerase (λDE3). The pET vectors carry genes encoding for the lac repressor (lacI), the T7 promoter - recognized by the T7 RNA polymerase, the lac operator, the f1 origin of replication and an antibiotic resistance gene. In the absence of lactose, the lac repressor binds to the lac operator and therefore sterically inhibits the T7 promoter, which leads to suppression of protein production to allow the cells to grow in a normal manner. Once lactose or the analogue Isopropyl-β-d-thiogalactopyranosid (IPTG) is present, the lac repressor is released and the protein expression initiated.

Recombinant proteins were expressed from pETM (EMBL) vector systems in Bl21(DE3) cells cultured in LB or M9 minimal medium supplemented with Kanamycin **(Table 4.3)**. M9 medium was either H_2O or D_2O based and contained $^{15}NH_4Cl$ for ^{15}N labeling scheme or $^{15}NH_4Cl$ and hydrated or deuterated [U-^{13}C]-D-glucose for ^{15}N, ^{13}C labeling scheme **(Table 4.3)**.

Protein expression was induced at an OD_{600}: of 0.6 to 0.8 with IPTG concentrations from 0.2 to 1 mM for 3 to 4 hours at 37 °C or 14 to 16 hours at 18 °C. The cells were then harvested by 20 min centrifugation at 6000 g and 4 °C. Protein expression protocols and labeling scheme differ from protein to protein. For details see experimental procedures of chapter 5, 6 or 7.

Protein purification

Immobilized metal affinity chromatography (IMAC)

All proteins of this study harbor an N-terminal six-histidine (His_6) tag, which enables the robust purification protocol via the IMAC technique called Ni-NTA affinity chromatography. The term Ni-NTA refers to a nickel^{2+} ion that is chelated by nitrilotriacetic acid (NTA), which is coupled to agarose resin. The immobilized Ni^{2+} is still able to form two more ionic bonds, which are used to chelate two histidine residues. The interaction of a His_6 tag and Ni-NTA resin is strong (K_D of 14 nM) (Knecht *et al*, 2009) while the rest of the cell extract exhibits only weak interactions. Weak unspecific binders can be removed by adding small amount of imidazole to the wash buffer. For the elution of His_6 tagged protein, a large excess of several hundred mM Imidazole is needed.
This study exclusively used gravity flow Ni-NTA affinity chromatography as first purification step. The process can be generally described in three steps; (i) binding; where the supernatant of the cell lysate is applied to the column, (ii) wash; where unspecific binding is removed, (iii) elution; where the protein of interest is eluted from the Ni-NTA beads. However, for proteins with a protease cleavable tag, an extra step was done to remove the His_6 and His_6-SUMO tag together with the His_6-tagged protease and more unspecific binders from the first column. First, a dialysis to remove imidazole from the elution has to be done. This is usually done simultaneously with the cleavage of the tag (dialysis buffer = wash buffer without imidazole). Then a reverse Ni-NTA column is performed. Reverse means that all impurities bind to the Ni-column and the clean protein comes with the flow through. All steps of the purification were monitored by SDS-PAGE.
Buffer compositions for lysis, wash, elution final buffer differ from protein to protein. For a complete list see **Table 4.4** and the experimental proceduress of chapter 5, 6 or 7. All purification steps were carried out at 4 °C either in the cold room or with pre-cooled solutions

Size exclusion chromatography

NMR spectroscopy and X-ray crystallography experiments require a protein purity of >95%. This high purity is not reached by Ni-NTA affinity chromatography alone. To further enhance purity, I used size exclusion chromatography. Here, molecules are separated by size (radius of gyration) by passing through a column filled with porous material. Small molecules are transient trapped in the material and move slower as bigger molecules. The range of molecular weight is covered by

different column types filled with smaller or larger particles. The protein trace is traced by UV light at multiple wavelengths such as 280 nm for proteins (Trp and Tyr residues) or 260 nm for RNA. Size exclusion chromatography was done using a HiLoad 16/600 Superdex S75 (Cytiva, USA) for small proteins (<40 kDa) or a HiLoad 16/600 Superdex S200 (Cytiva, USA) coupled to an ÄKTA pure system (Cytiva, USA) with multi wavelength detector and 96 well collector. In this procedure, the sample from the reverse Ni-NTA was concentrated in 3, 10 or 30 kDa (Amicon) concentrator to 2.5 ml before being loaded into a 5 ml loop. The loop was flushed with 10 ml with a flow rate of 1 ml/min onto the equilibrated column (for final buffers see **Table 4.5).** The flow rate was kept constant over the full run. The pure protein fractions were collected and analyzed using SDS-PAGE.

SDS-PAGE

Sodium dodecyl sulfate (SDS) – polyacrylamide gel electrophoresis (PAGE) is a standard method for analysis of proteins and protein mixtures. SDS denatures proteins by attaching to hydrophobic residues and introduces a negative charge to the protein at the same time. This negative charge is greater than the net charge of the protein, which enables electrophoretic separation according to molecular weight in polyacrylamide gels with different percentages.
For analysis, a few µl of protein were mixed with 4x Laemmli buffer and loaded onto 12% or 15% polyacrylamide gels. The gels ran with max. 220V and 45mA per gel for 40 to 45 min. The gels were then stained with a Coomassie Brilliant blue (R-250 or G-250) containing stain solution, which has a detection limit of ~0.1 - 0.5 µg (Brunelle & Green, 2014).

Protein concentration determination

The protein concentration was measured using a NanoDrop ND-1000 (ThermoScientific) machine. The determination is based on the UV light absorption at 280 nm and the law of Lambert Beer ($A=\varepsilon lc$), where the amino acids tryptophan ($\varepsilon=5500\ M^{-1}cm^{-1}$), phenylalanine ($\varepsilon=1490\ M^{1}cm\text{-}^{1}$) and cysteine (only disulfide bonds; $\varepsilon=125\ M^{-1}cm^{-1}$) exhibits maximum absorbance.

Biophysical techniques

NMR spectroscopy

This section describes typically used NMR techniques and procedures. For a detailed introduction in NMR or experimental setups see chapter 2 or chapters 5, 6 and 7. All NMR experiments of this study were performed on one of the spectrometers listed in **Table 4.8**.

Table 4.8: NMR spectrometers *according to proton Larmor frequency.

Field strength in Mhz*	Probes	Console
500	Cryo-TCI / ^{1}H, ^{13}C, ^{15}N, z-gradient	Bruker Avance III
600	Cryo-QCI / ^{1}H, ^{31}P, ^{13}C, ^{15}N, z-gradient	Bruker Avance III
600	TXI / ^{1}H, ^{13}C, ^{15}N, z-gradient	Bruker Avance III
800	Cryo-TCI / ^{1}H, ^{13}C, ^{15}N, z-gradient	Bruker Avance III
900	Cryo-TCI / ^{1}H, ^{13}C, ^{15}N, z-gradient	Bruker Avance III
950	Cryo-TCI / ^{1}H, ^{13}C, ^{15}N, z-gradient	Bruker Avance III+

Backbone chemical shift assignments

For the sequential backbone assignments in this study, the following triple resonance experiments were used: HNCA, HN(CO)CA, CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, or the corresponding TROSY versions (not CBCA(CO)NH) (Salzmann *et al.*, 1998; Sattler M *et al.*, 1999; Weisemann *et al*, 1993).

Secondary structure from secondary chemical shifts

For analysis of secondary structure, the difference between the experimental and random coil chemical shifts (Δδ), called secondary chemical shifts, are calculated. The final read out is the difference of $\Delta\delta^{13}C\beta$ and $\Delta\delta^{13}C\alpha$ ($\Delta\delta(^{13}C\alpha) - \Delta\delta(^{13}C\beta)$) (Kjaergaard & Poulsen, 2011; Schwarzinger *et al,* 2001). Positive values indicate α-helical structures and negative values β-strands. A useful hybrid tool to calculate this values and predict secondary structure based on NMR structures at the BMRB database under consideration of the plus and minus residue is TALOS+ (Shen *et al*, 2009).

NMR Titrations

NMR titration were typically performed with 100µM of ^{15}N-labeled protein and increasing concentration of the ligand. ^{1}H-^{15}N, HSQC spectra were recorded for the reference spectrum and after each titration step. If the ligand was at low concentration, each titration step was prepared in a separate NMR tube to avoid dilution effects. The chemical shift perturbation ($\Delta\delta_{avg}$) was calculated by using formula $\Delta\delta_{avg} = [(\Delta\delta_H)^2 + (\Delta\delta_N/6.3)^2]^{0.5}$. Dissociation constants (K_D) were fitted to the ligand concentration with a one-site specific binding model within using the equation $\Delta\delta = \Delta\delta_{max}/(2\ [P_t])*\{[L] + [P_t] + K_D - (([L] + [P_t] + K_D)^2\text{-}4[P_t][L])^{1/2}\}$, where $\Delta\delta$ is the individual and $\Delta\delta_{max}$ the maximum shift distance, P_t is the total protein concentration and [L] is the ligand concentration.

NMR relaxation to study backbone dynamics

To identify low populated secondary structure elements, {^{1}H}-^{15}N heteronuclear NOE experiments were used to sample motions of distinct N-H bond vectors. Those experiments were performed using the pulse sequence hsqcnoef3gpsi (Bruker, Avance version 12.01.11) with a 4 to 5 second interscan delay. NOE values are given by the ratio of the peak heights in the experiment with and without proton saturation (hetNOE = I_{sat}/I_{unsat}) (Renner *et al*, 2002**)**. {^{1}H}-^{15}N heteronuclear NOE values below and above the average indicate protein regions with high or low backbone dynamics respectively.

A different approach to obtain protein backbone dynamics over a wide range of time scales is the measurement of nitrogen T_1 and T_2 relaxation times. T_1 and T_2 pulse sequences are ^{15}N HSQC-based developed from Farrow *et al.* (1994) with water-control during the relaxation period in the T_1 sequence using a cosine-modulated IBURP-2 pulse (Gairí *et al*, 2015) and modifications in the T_2 sequences based on (Lakomek *et al*, 2012)**.** For both T_1 and T_2 experiments, 8 time points with delays of 80, 160, 240, 320, 400, 64, 800, 1000ms (T_1) and 14.4, 28.8, 43.2, 57.6, 72.0, 86.4, 100.8, 115.2ms (T_2) were measured respectively. The analysis of the product of both relaxation rates (R_1R_2) gives indication for chemical exchange independent of overall correlation time and motional anisotropy (Kneller *et al.*, 2002).

Software for NMR processing and analysis

NMR-Spectra were processed using Topspin (Bruker Biospin, Rheinstetten, Germany) or NMRPipe (Delaglio *et al*, 1995) and analyzed using CcpNMR Analysis 2.4.2 (Vranken *et al*, 2005). For plotting of chemical shift perturbations for dissociation constant calculations, Origin software (OriginLab Corporation, United States) was used.

X-ray crystallography

X-ray crystallography allows structure determination up to atomic resolution. In this thesis, structures of apo PEX13-SH3 domain and in complex with the peptide ligands PEX13 FxxxF and PEX5 W4 were solved. Details on the screening conditions and results can be found in chapter 7.

Protein crystallization

In order to solve a 3-dimensional protein structure, the protein must first form a crystal, which means adapting a solid-state phase where all the molecules are placed periodically in three dimensions. To reach such a state, a highly pure (>95%) protein (Geerlof *et al*, 2006) with a concentration close to its saturation limit is necessary. The solution is then brought into a state of supersaturation, a crucial step in nucleation and crystal growth, by gradual dehydration of the solvent or change of protein solubility. The protein solubility can be reduced by manipulating conditions such as temperature, pH, ionic strength and the addition of salts or organic precipitant such as polyethylene glycol (PEG). When the protein solution reaches a labile supersaturation zone, spontaneous nucleation, protein arrangement in a lattice, occurs. Crystal growth takes place at the labile and metastable supersaturation zone. When the protein or precipitant concentration too high and the high supersaturation zone is reached, the protein precipitates (See McPherson & Gavira (2014) for more information). To date, there are no means of predicting a protein's optimal crystallization conditions. It is rather a "try and error" approach, where lots of different conditions are tested systematically. Within this thesis, an automated 96 well screening approach in the sitting drop format was used. In this method, a small drop of 200 nl of buffered protein **(Table 4.5)** is placed on an elevated pedestal next to a reservoir filled with approximately 70 µl of screening buffer representing the screening condition. The protein solution is then mixed with equal amount (200 nl) of screening solution. After the automated setup of the 96 well plate, the plate including

all single wells with the screening conditions is sealed with a foil to block air exchange with the surrounding environment. The initial high concentrated protein is buffered in neutral pH, low salt and low buffer concentration. The reservoir solution on the opposite typically contains a different pH and a high concentration of salt or other precipitants. Over the time, the solvent of the diluted protein drop evaporates and diffuses towards the reservoir solution. This gradually increases the protein concentration in the drop until it reaches a state of supersaturation, which ideally initiates nucleation and crystal growth. The crystal structure is formed by random contact sites based on weak interactions such as Van der Waals forces, hydrogen bonds or hydrophobic interactions and contain typically 40-60% solvent (Matthews, 1968). Protein crystals are thus very fragile. The crystallization can be described as an equilibrium driven by minimization of free energy of a system by converting from a soluble to crystalline state (Pusey & Van Megen, 1986) meaning that crystals can dissolve once the chamber is opened and water from the air can diffuse in.

Crystal fishing and data collection

Once the protein crystals were formed, the sealing of the well was opened and the crystal fished with a nylon loop of a few µm size. To protect the crystal from being destroyed through ice crystal formation, the crystal was soaked in 20% or 25% cryoprotectant solution before flash freezing in liquid nitrogen. The cryoprotectant solutions were prepared by mixing glycerol or ethylene glycol with the screening solution from the according condition. The frozen crystals were packed in UniPucks and shipped to the synchrotron at Paul-Scherer-Institute (Villingen, CH) or EMBL (Hamburg, Hamburg, DE). Data were collected on SLS beamline X06DA (PSI) or on beamline P11 at PETRA III (EMBL).

Data processing

Collected data were processed using CCP4i2 suite (Potterton *et al*, 2018; Winn *et al*, 2011) using XDS package (Kabsch, 2010) for data reduction, aimless for data scaling, MOLREP (Vagin & Teplyakov, 2010) or PHASER (McCoy *et al*, 2007) for molecular replacement, COOT (Emsley & Cowtan, 2004) for model building and REFMAC5 (Murshudov *et al*, 2011) for refinement. Refined structures were uploaded to wwPDBdeposition using pdb_extract (Yang *et al*, 2004).

Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry is a technique to study binding affinities and energetics as well as stoichiometry between tag free molecules in solution. Here, a MicroCal PEAQ-ITC (Malvern Panalytical, Malvern, Uk) was used. The device consists of two cells, one sample cell and one reference cell. Both cells are kept at defined temperatures. The higher concentrated ligand is located in a syringe that injects fixed volumes to the sample cell in regular intervals. The binding of the ligand induces small a temperature changes in the sample, which is an increase if the reaction is exothermic or a decrease if the reaction is endothermic. In any case, the re-equilibrates the temperatures of both cells and records the required power. Saturation is reached when the temperature change per injection is small and constant, which is usually caused by dilution effects of the titrant. The power difference per injection is then plotted over the experimental time and each injection integrated to obtain a binding isotherm, which is fitted to an appropriate binding model (one or more binding sites). From this curve, the K_D and the binding enthalpy ΔH is calculated. The binding stoichiometry N is given by the molarity at the vertex point. From the determined parameters and the given temperature, the entropy ΔS and the free enthalpy ΔG can be obtained by calculation ($\Delta G = -RT \ln K = \Delta H - T\Delta S$). For details on explicit experimental setups see experimental procedures of chapter 5, 6 or 7.

Size exclusion chromatography - light scattering

Size exclusion chromatography coupled with light scattering detectors is a powerful tool to determine the molecular mass (Static Light Scattering, SLS) and radius of gyration (Dynamic Light Scattering, DLS) of given biomolecules. This is especially useful to detect oligomerization or stoichiometries of protein complexes at a given concentration (Korepanova & Matayoshi, 2012; Some *et al*, 2019). The experiments were performed on an all in one OmniSEC machine (Resolve and Reveal) from Malvern Panalytics (Malvern, Uk), which consists of an autosampler, a HPLC-SEC with exchangeable column and a set of detectors. As the protein solution passes through an analytical SEC column, it is separated by size and shape. Next, the eluted protein is analyzed by a set of detectors that measure the refractive index (concentration), UV/Vis absorption (concentration), light scattering (molecular weight) and intrinsic viscosity. Both light scattering detectors, dual (right angle 90° and low angle 7°) and multi-angle light scattering detector,

measuring the intensity of scattered light, which is directly related to the molecular weight. The right angle light scattering (RALS) detector sits perpendicular to the beam. This detector is optimal for molecules (R_g < 15 nm) because those molecules scatter light isotropically, which allows the direct measurement of the molecular weight with best signal to noise ratio. The low angle light scattering (LALS) detector sits 7° tilted to the beam. Ideally it would sit at 0°, which is not practicable since it would detect mostly the initial beam. This situation is best for molecules that do not scatter light isotropically (R_g > 15 nm) as the molecular weight can be directly measured. In contrast to the RALS and LALS detectors, the multi-angle light scattering (MALS) detector detects scattering from numerous angles, which is used to extrapolate the scattering back to 0° to calculate molecular weight.

For MW determination of the small protein fragment PEX13 CTD (~16kDa), it was feasible to use static light scattering methods with dual light scattering detector (RALS and LALS). See experimental procedures of chapter 7 for experimental details.

Circular dichroism (CD)

Circular dichroism refers to the property of chiral molecules to absorb right- and left- circularly polarised light differentially. The secondary structure elements of proteins, such as α-helices, β-strands and random coil can produce characteristic CD spectra that exhibit specific minima or maxima of ellipticity at different wavelengths in a recorded spectrum. Pure α-helical proteins show double minima at 208 and 222 nm and a stronger maximum at 191-193 nm, while β-sheet structures show a characteristic minimum at 215 nm and a maximum at 198 nm. The spectral appearance of random coil structures is essential the opposite to helices or strands with a minimum at ~198 nm and a broad maximum at ~220 nm (Kelly *et al*, 2005). These features can be used to estimate the secondary structure content of a protein and can be used to monitor protein or protein-complex unfolding during temperature exposure. In such experiment, the measured wavelength is set to a minimum (e.g. 220 nm), which is then measured during increments of increasing sample temperature. Unfolding of the protein is then monitored by a decreased signal intensity. For details on explicit experimental setups see experimental procedures of chapter 5 or 6)

Cell-based assays

All cell-based experiments were performed in collaboration with Prof. Dr. Ralf Erdmann and Dr. Wolfgang Schliebs at the Ruhr-Universität Bochum. Experimental details are described in the respective chapters.

Chapter 5:

Membrane interactions of the peroxisomal proteins PEX5 and PEX14

The work in this Chapter is published as:

Stefan Gaussmann*, Mohanraj Gopalswamy*, Maike Eberhardt, Maren Reuter, Peijian Zou, Wolfgang Schliebs, Ralf Erdmann, Michael Sattler, *Membrane Interactions of the Peroxisomal Proteins PEX5 and PEX14, Frontiers in Cell and Developmental Biology, 2021, Vol. 9 Issue 743 Pages 651449, DOI:10.3389/fcell.2021.651449*

Membrane interactions of the peroxisomal proteins PEX5 and PEX14

Author contribution: The author of this thesis, Stefan Gaussmann and Mohanraj Gopalswamy contributed equally to the original work of this article. Stefan Gaussmann carried out molecular biology and protein expression and purification, NMR and ITC experiments and analysis and preparation of the manuscript.

Abstract

Human PEX5 and PEX14 are essential components of the peroxisomal translocon, which mediates import of cargo enzymes into peroxisomes. PEX5 is a soluble receptor for cargo enzymes comprised of an N-terminal intrinsically disordered domain (NTD) and a C-terminal tetratricopeptide (TPR) domain, which recognizes peroxisomal targeting signal 1 (PTS1) peptide motif in cargo proteins. The PEX5-NTD harbors multiple WF peptide motifs (WxxxF/Y or related motifs) that are recognized by a small globular domain in the NTD of the membrane-associated protein PEX14. How the PEX5 or PEX14-NTDs bind to the peroxisomal membrane and how the interaction between the two proteins is modulated at the membrane is unknown.

Here, we characterize the membrane interactions of the PEX5-NTD and PEX14-NTD *in vitro* by membrane mimicking bicelles and nanodiscs using NMR spectroscopy and isothermal titration calorimetry. The PEX14-NTD weakly interacts with membrane mimicking bicelles with a surface that partially overlaps with the WxxxF/Y binding site. The PEX5-NTD harbors multiple interaction sites with the membrane that involve a number of amphipathic α-helical regions, which include some of the WxxxF/Y-motifs. The partially formed α-helical conformation of these regions is stabilized in the presence of bicelles. Notably, ITC data show that the interaction between the PEX5 and PEX14-NTDs is largely unaffected by the presence of the membrane. The PEX5/PEX14 interaction exhibits similar free binding enthalpies, where reduced binding enthalpy in the presence of bicelles is compensated by a reduced entropy loss. This demonstrates that docking of PEX5 to PEX14 at the membrane does not reduce the overall binding affinity between the two proteins, providing insights into the initial phase of PEX5-PEX14 docking in the assembly of the peroxisome translocon.

Introduction

Peroxisomes are ubiquitous membrane enveloped organelles of eukaryotic cells involved in various metabolic pathways, including β-fatty acid oxidation and removal of toxic oxidation products (Erdmann *et al*, 1997b; Fujiki & Lazarow, 1985; Wanders, 2004; Wanders & Waterham, 2006). Peroxisome biogenesis depends on a number of proteins, the so-called peroxins (Distel *et al*, 1996; Ma *et al*, 2011). As peroxisomes lack a protein synthesis machinery, peroxisomal matrix proteins need to be imported into the organelle post-translationally. The majority of these cargo proteins are imported via a peroxisomal targeting signal 1 (PTS1), a conserved C-terminal peptide motif, with SKL as canonical sequence (Ghosh & Berg, 2010; Gould *et al.*, 1987). The soluble peroxisomal receptor PEX5 recognizes the PTS1 motif by a C-terminal tetratricopeptide (TPR) domain (Gatto *et al.*, 2000; Stanley *et al.*, 2006). Cytosolic PEX5 shuttles the cargo protein to the peroxisomal membrane (Dammai & Subramani, 2001; Dodt & Gould, 1996; Erdmann & Schliebs, 2005; Rucktaschel *et al*, 2011). For this, its intrinsically disordered N-terminal domain (NTD) interacts with the membrane-anchored peroxins PEX14 and PEX13 (Neufeld *et al.*, 2009; Saidowsky *et al.*, 2001; Schliebs *et al.*, 1999). Subsequently, a transient pore is formed and the cargo is tunneled through the membrane (Erdmann & Schliebs, 2005). This step of membrane passaging has been characterized in *Saccharomyces cerevisiae*, where Pex5p and Pex14p are key components of the protein conducting channel (Meinecke *et al.*, 2010).

In contrast to PEX14, which is an integral membrane protein with a single transmembrane span, PEX5 does not contain a classical transmembrane domain (Emmanouilidis *et al*, 2016). However, it harbors WxxxF/Y (W1-7) and one LVAEF (W0) motif in the NTD, which bind to the PEX14-NTD and have been hypothesized to potentially mediate membrane interactions (Emmanouilidis *et al.*, 2016; Saidowsky *et al.*, 2001). PEX5 cycles between a soluble and a membrane associated state. While this suggests that PEX5 may be able to interact with the membrane, it still requires a co-factor to maintain it at the membrane (Azevedo & Schliebs, 2006). It is expected that the membrane protein PEX14 localizes PEX5 to the membrane, since the PEX14-NTD is able to bind to all eight WF-like -motifs of PEX5 (Emmanouilidis *et al.*, 2016; Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014).

Although the molecular interactions between PEX14 and PEX5 are known, the mechanism by which the cargo is translocated is still poorly understood (Emmanouilidis *et al.*, 2016). It has been proposed that the PEX14-NTD may recruit PEX5 by binding to the W0 and additional WF motifs

and thereby initiates pore formation. Recent studies have proposed that the NTD of PEX14 is located inside the peroxisomal lumen (Barros-Barbosa *et al.*, 2019; Neuhaus *et al.*, 2014), and thus would not be easily available for initiating contacts with the PEX5-NTD in the cytosol. Potentially, PEX5 might be recognized by other parts of the docking complex (such as PEX13) or may be targeted to the peroxisomal membrane by direct binding. It is also conceivable that the PEX14-NTD may be transiently exposed to the cytosol and subsequent to PEX5 binding translocate into the peroxisomal lumen. In any case, a common prerequisite for all these models is a membrane localization of PEX5 and PEX14.

Studying proteins in membrane-like environment can be challenging: For *in vitro* binding studies, the membrane mimic should represent the lipid composition of the native environment and compatible with the experimental approach. The peroxisomal membrane lipid composition of eukaryotic cells is not well characterized. Nevertheless, analysis of peroxisomes from rat liver showed a distribution of 27.5% phosphatidylethanolamine (PE), 56.6% phosphatidylcholine (PC), 4.7% phosphatidylinositol (PI), 3.7% sphingomyelin (SPM), and 3% phosphatidylserine (PS) (Hardeman *et al*, 1990). The high percentage of almost 60% of phosphatidylcholine is well feasible for studies using solution state NMR spectroscopy, since bicelles composed of 1,2-Dimyristoyl-sn-Glycero-3-Phosphocholine (DMPC) and 1,2-Diheptanoyl-sn-Glycero-3-Phosphocholine (D7PC) are well established and favorable due to the relatively low molecular weight. The bicelles assemble into discoid bilayers, where DMPC forms the planar surface of the disk while the short-chain lipids from D7PC form the curvature on the edges, which leads to the dependency of the disk size on the molar ratio of DMPC to D7PC (q). Isotropic bicelles are typically made by a molecular ratio q ranging between 0.2 and 0.5, where the size is shrinking with the value of q (Marcotte & Auger, 2005; Sommer *et al*, 2012; Warschawski *et al*, 2011). NMR spectroscopy benefits from the small molecular size of bicelles in terms of resolution derived from relatively sharp line shapes. On the other hand, the high curvature does not represent the native membrane environment very well. For proteins with transmembrane spans, a better membrane mimic can be achieved using nanodiscs, which consists of a planar lipid bilayer encircled with a membrane scaffold protein (MSP). Recent developments in this field allow the production of smaller nanodiscs with a diameter of 6–8 nm with favorable features for NMR studies (Hagn *et al,* 2018).

The present study combines NMR spectroscopy and isothermal titration calorimetry to characterize the membrane interaction of the NTDs of PEX5 and PEX14 as central components of the

peroxisomal pore and investigates the interaction between the two proteins in solution and in the presence of a membrane mimicking environment, thus providing novel insight into early steps of peroxisomal protein translocation.

Results

Conformational Analysis of the PEX5-NTD by NMR Spectroscopy

The N-terminal domain of PEX5 is intrinsically disordered (Emmanouilidis *et al.*, 2016). The NMR spectrum of ^{15}N-labeled PEX5-NTD, comprising residues 1–315 (**Figure 5.1A**) show poor spectral dispersion, with most signals showing amide proton chemical shifts between 8 and 8.5 ppm, typical for intrinsically disordered proteins (IDPs). To reduce signal overlap, we divided the PEX5-NTD into roughly 100 aa long regions with three to four amino acid residues overlapping to the next or previous subconstruct and without disrupting any of the known WF-like motifs. The 2D ^{1}H, ^{15}N correlation spectra of PEX5 comprising residues 1–113, 110–230, and 228–315 (**Figure 5.2A**) show good dispersion and very little signal overlap. Moreover, comparison of the sum of the NMR correlation spectra with the those obtained for the full PEX5-NTD shows little chemical shift differences (**Figure 5.1B**), suggesting that analysis of the three subregions faithfully reports on the properties in the context of the PEX5-NTD.

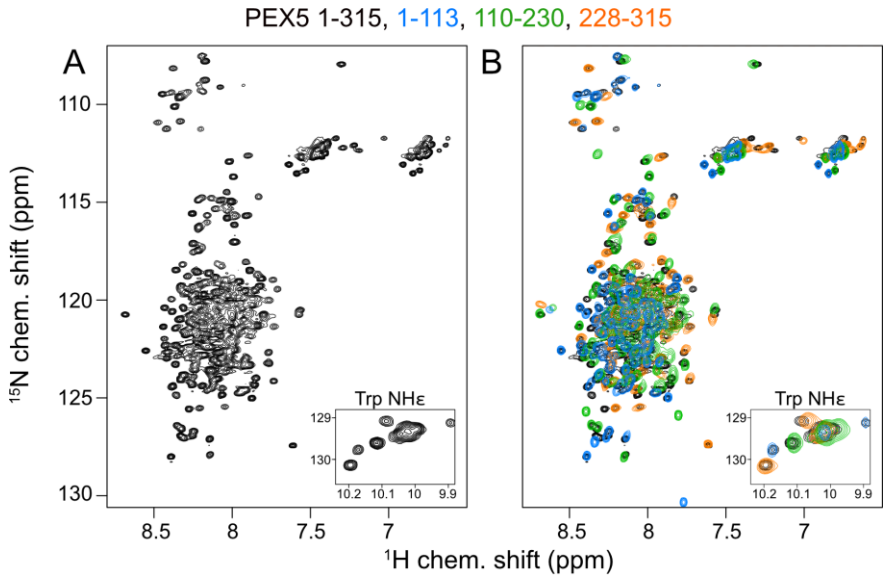

Figure 5.1: ^{1}H,^{15}N HSQC NMR spectra of (A) human PEX5-NTD residues 1-315 and (B) superposition of spectra of human PEX5 residues 1-113 (blue), 110-230 (green) and 228-315 (orange) with the complete NTD (black). The spectral region showing the tryptophan side chain indole NH signals are shown as inserts.

Next, the residue-specific backbone chemical shifts of the three regions were assigned using standard triple resonance experiments Sattler M *et al.* (1999), enabling a comprehensive NMR analysis of the PEX5-NTD. This allowed us to analyze the secondary structure and conformational flexibility of the polypeptide backbone, and to map molecular interactions with membrane mimics. First, we investigated the polypeptide backbone flexibility of the PEX5-NTD using $\{^1H\}$-^{15}N heteronuclear NOE experiments for the three NTD subregions. The conformational flexibility of the backbone is reflected by the heteronuclear NOE, with values of ~0.9 corresponding to a rigid backbone conformation as expected for a globular folded protein (Daragan & Mayo, 1997). The first 110 amino acids of PEX5, including the W0 (LVAEF) motif, shows a highlight flexible backbone conformation, while the remaining 205 amino acids exhibit significantly reduced conformational flexibility. Most of these regions coincide with the WF motifs W1 to W6, with W5 and W6 showing the highest values. This indicates that the region comprising the W5 and W6 motifs is less flexible in solution (**Figure 5.2B**).

Next, we analyzed the secondary structure of the PEX5-NTD based on $^{13}C_\alpha$ and $^{13}C_\beta$ secondary chemical shifts, which are the difference of chemical shifts compared to those in a random coil conformation (Spera & Bax, 1991). Values around zero indicate random coil while positive and negative values correlate with α-helical and β-strand/extended conformations. Interestingly, the PEX5-NTD shows some regions with α-helical characteristics, which are not related to the WF motifs (**Figure 5.2C**). The largest positive secondary chemical shifts are found for residues 285 to 305. Notably for this region also positive heteronuclear NOE values are observed, indicating the presence of a largely formed α helix (α_4). Four additional regions, residues 81–96 (α_0), 210–220 (α_1), 237–250 (α_2), and 271–285 (α_3) exhibit transient (partially formed) helical regions indicated by secondary chemical shifts between 1 and 2 ppm (**Figure 5.2C**). All these helical regions are very well conserved among eukaryotes and have amphipathic character (**Figure 5.2F**), as is well known for the WF motifs. Thus, these regions could mediate protein-protein and/or membrane interaction (Schliebs *et al.*, 1999). Experimental structures have previously shown that W0 and W1 adopt an α-helical fold in the complex with the PEX14-NTD and for some of the PEX5 WF motifs it has been shown that they adopt a partially preformed helix in solution (Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014). Our data show that the WF motifs W5 and W6 are indeed the most helical regions. We identify four additional more transient helices (α_0–α_4) that maybe involved in additional interactions.

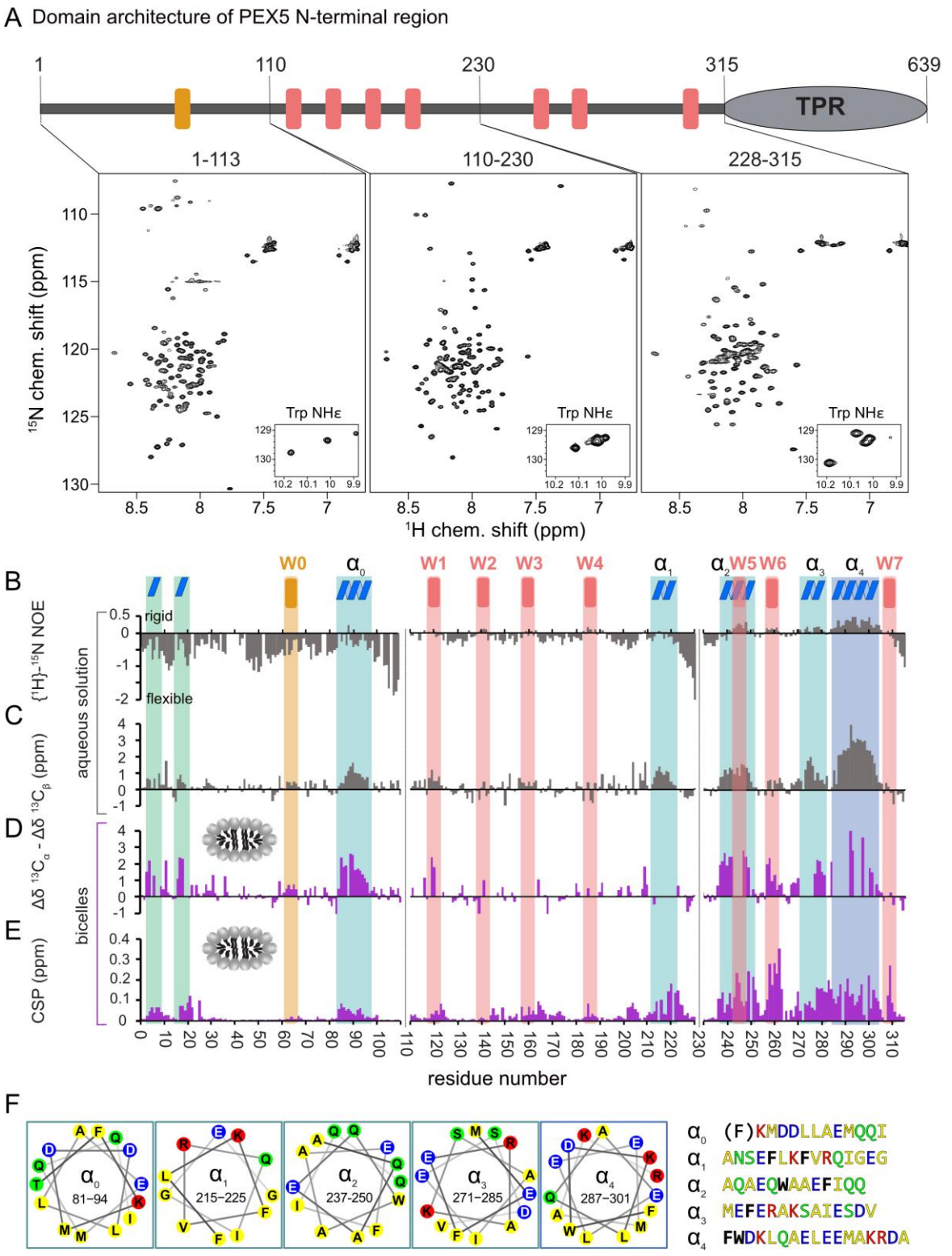

Figure 5.2: Analysis of PEX5-NTD in the absence and presence of bicelles by NMR spectroscopy. PEX5-NTD domain architecture and ^{1}H–^{15}N HSQC spectra of PEX5 constructs 1-113, 110-230 and 228-315 (A). {^{1}H}-^{15}N heteronuclear NOE experiments of the three constructs in aqueous solution (B). negative values represent a flexible and positive values rigid NH backbone. WF and the WF-like motif W0 are indicated by red or orange bars respectively. {^{1}H}-^{15}N

heteronuclear NOE experiments in aqueous solution (B) indicating less flexible regions which were classified as helices α_0 - α_4 (blue boxes) by the chemical shift index ($\Delta\delta^{13}C_\alpha$ - $\Delta\delta^{13}C_\beta$) (C). In the presence of bicelles as a membrane mimic, the helices as well as two α-turns located in the first 20 aa, are stabilized which is shown by CSI (D). Chemical shift perturbations extracted from ^{1}H-^{15}N HSQC experiments (E) demonstrate membrane binding which is mostly mediated by WF6 and the amphipathic helices α_0 - α_4 visualized as helical wheels in (F). The residues are colour coded in yellow for hydrophobic, green for polar, blue for negative charged and red for positive charged sidechains. Phe and Trp residues are shown in black and bold letters with exception of the Phe in α_0 which does not contribute to the hydrophobic face of the helix.

PEX5-NTD Interacts with Membranes

NMR membrane binding studies were performed with bicelles consisting of DMPC, a saturated C14:0 lipid. While this is a favorable and well-established membrane mimic for NMR studies it does not reflect the composition of peroxisomal membranes. To assess the validity of using DMPC mimics we performed and compared flotation analyses of PEX5 and PEX5-NTD with liposomes consisting of DMPC and of three volumes 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) and seven volumes 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC). This resembles the composition of 28% phosphoethanolamine and 57% phosphocholine, respectively, of the membranes of rat liver peroxisomes (Hardeman *et al.*, 1990). Consistent with previous observations (Kerssen *et al.*, 2006) a similar fraction of PEX5 was found in association with floated vesicles consisting of DOPE/DOPC (**Figure 5.3A**). Remarkably, the same flotation behavior was observed with liposomes constituted with DMPC only (**Figure 5.3A**). Also the PEX5-NTD co-migrated with both floated DMPC and DOPE/DOPC liposomes with comparable efficacy (**Figure 5.3B**). Taken together, the flotation analyses confirm that the lipids used for mapping the binding sites by NMR are useful and relevant proxies for the lipid binding properties of PEX proteins.

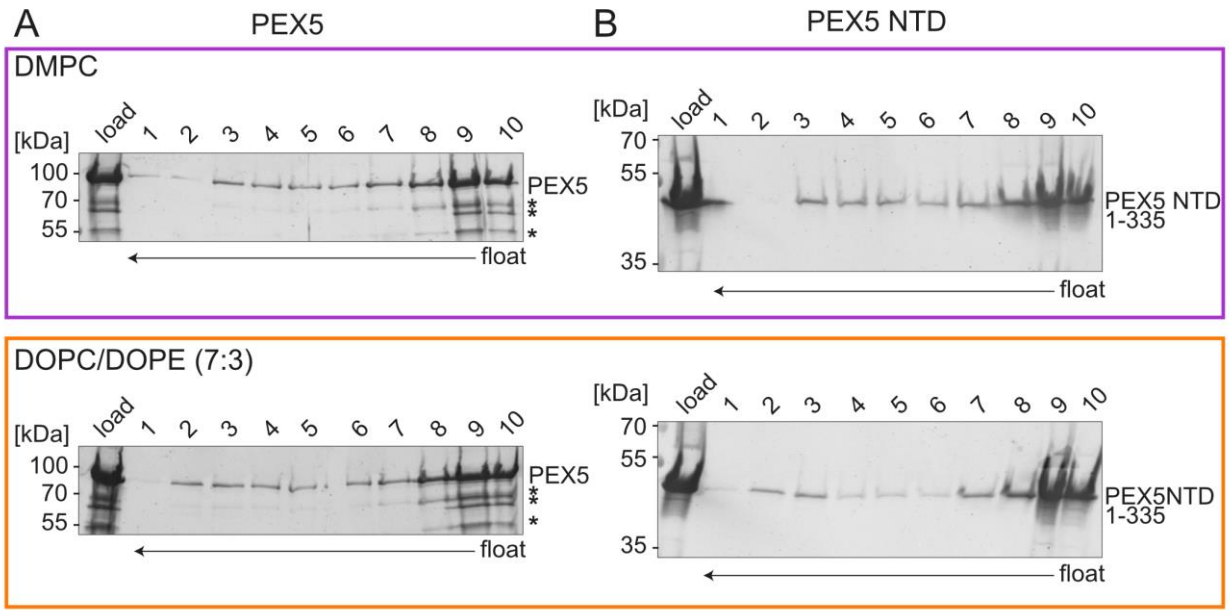

Figure 5.3: Recombinant PEX5 associates with phospholipid bilayer vesicles with different composition.1.5 nmol purified human PEX5 full length (A) and PEX5-NTD (1-335) (B) were incubated for 1h at room temperature with liposomes (ratio 1:750) either consisting of DMPC (purple box) only or of a mixture of DOPC/DOPE (7:3) (orange box). The samples were subjected to flotation gradient centrifugation (50% sucrose-0% sucrose). The gradient was collected as ten fractions from top to bottom. Equal volumes of fractions were analyzed by immunoblotting using polyclonal antibodies against human Pex5L. Asterisks indicate degradation products of PEX5, which did not co-migrate with floated liposomes.

To study potential secondary structure and map the lipid binding regions of the PEX5-NTD we performed NMR titrations with increasing bicelle concentration (**Figure 5.2D, E**). We recorded ^{15}N correlation spectra in the presence of 0.9 mM bicelles (DMPC/D7PC, q: 0.2). At these concentrations, we observe large chemical shift perturbation and substantial line broadening in the NMR spectra of the PEX5-NTD regions. To confirm and track chemical shift assignments additional triple resonance experiments were recorded in the bicelle-bound state. Comparison of NMR spectra in buffer and in the presence of bicelles reveal significant chemical shift perturbations (CSP) and changes in secondary chemical shifts exclusively for the α-helical regions. We find two small but strongly enhanced helical regions between residues 1 to 20, which show very low helical propensity in aqueous solution and were therefore not classified as preformed helical motifs (**Figure 5.2D**). These induced helical motifs are well conserved in sequence and contribute to PEX5 membrane interaction as can be judged from the chemical shift perturbations (**Figure 5.2E**). A notable increase in helical propensity is also observed for the helical regions α_0, α_2, α_3, and α_4 as

well as for the WF motifs W1, W2, W5, and W6. Unfortunately, the lack of chemical shift assignments prevents conclusions for W3, W4 and the transient helix α_1. Interestingly, the helical content of the W0 motif, which lacks a tryptophan, is not much affected by bicelle binding.

Large chemical shift perturbations (CSP) are seen for the last third of PEX5-NTD (**Figure 5.2E**, **5.4A**). The first 200 residues experience significant spectral changes only at equimolar protein/bicelle ratio, i.e., 5- fold higher bicelle concentration (**Figure 5.4B**). The strongest CSPs are mostly found in helical regions, including some of the WF-like motifs, suggesting these regions mediate the membrane interaction (**Figure 5.2E**, **5.4B**). The C-terminal region of the PEX5-NTD, encompassing W5, W6, and W7, shows the largest CSPs, suggesting the strongest membrane interaction compared to the other regions. This is reflected in the membrane affinities derived from the NMR titrations. The apparent dissociation constants for the bicelle interactions are K_D^{app} = 196 ± 16 μM, 82 ± 13 μM, and 20 ± 7 μM, for the regions comprising PEX5 residues 1–113, 110–230, and 228–315, respectively **(Table 5.1 and Figure 5.4C, D**). Interestingly, a common feature of the helical regions involved in membrane binding is the presence of one or more phenylalanine and or tryptophan residues (**Figure 5.2 F**). Taken together the NMR data reveal an unexpected extent of regions adopting an amphipathic α-helical conformation in the PEX5-NTD, which are stabilized or induced by membrane interactions.

Table 5.1: NMR-derived membrane binding affinities

Protein	**K_D DMPC (μM)**	**K_D^{app} bicelle (μM)**
PEX5 1-113	9040 ± 746	196 ± 16
PEX5 110-230	3763 ± 589	82 ± 13
PEX5 228-315	916 ± 333	20 ± 7
PEX14 16-80	79 ± 13	1.7 ± 0.3

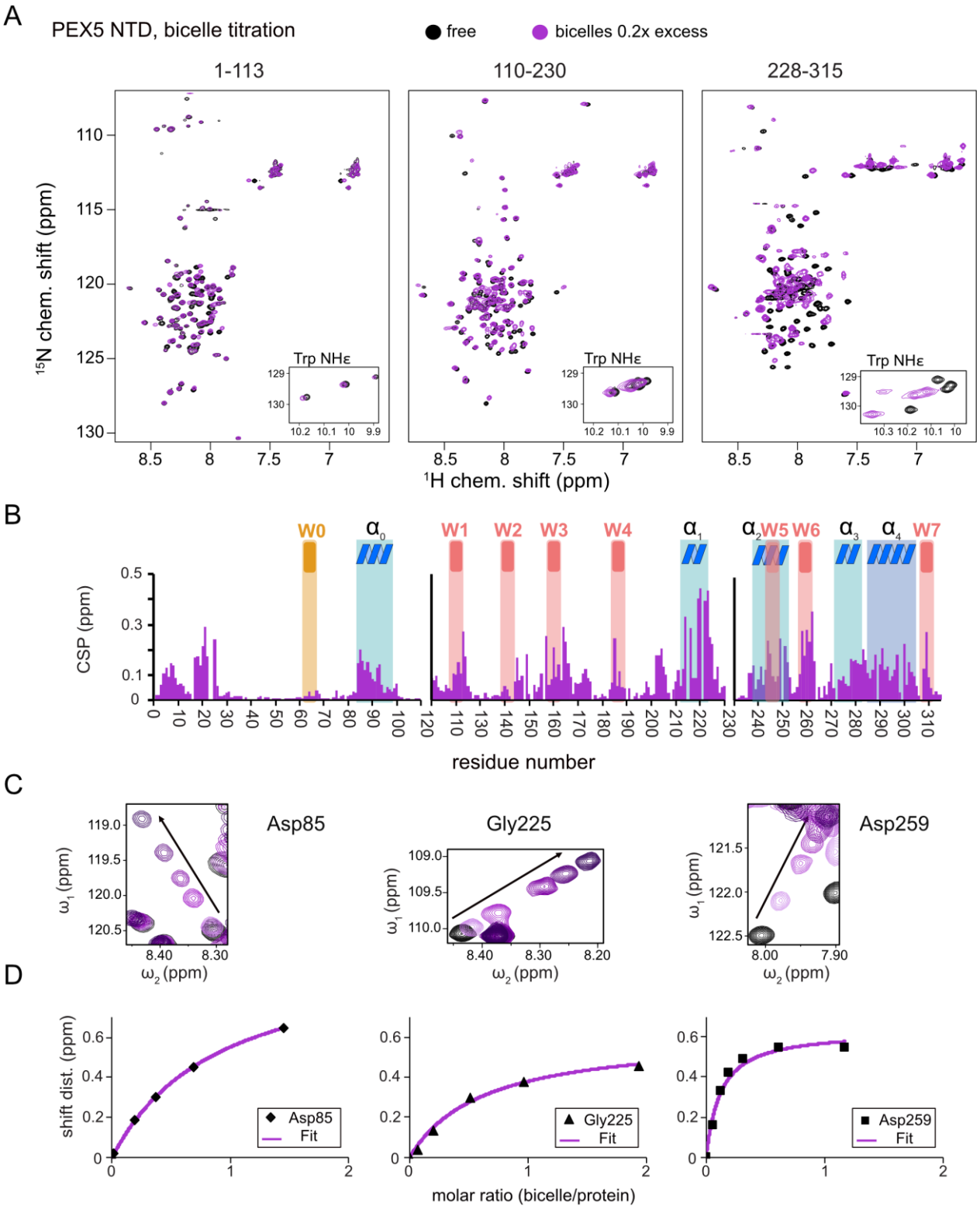

Figure 5.4: PEX5-NTD - membrane binding. ^{1}H,^{15}N HSQC NMR spectra overlay (A) of free PEX5 1-113, 110-230, and 228-315 (back) and in the presence of 0.2x molar excess of DMPC/DHepPC bicelles with a q value of 0.2. The subconstruct 228-315 is largely affected by 0.2x molar excess while the constructs 1-113 and 110-230 show comparable effects at a molar

excess of 0.7x and 0.9x (B). Tracing the chemical shift perturbations of largely affected residues Asp85, Gly225, and Asp259 of the subconstructs (C) and plotting shift distance against the molar ratio of bicelle to protein (D) Fitting of the NMR titration data (chemical shift difference to the free state) to a one-site binding model as a function of the molar bicelle:protein ratio (see experimental procedures).

PEX14-NTD Interaction with Membranes

PEX14 is embedded in the peroxisomal membrane by a transmembrane region predicted for residues 107–129. Given that the PEX14-NTD is in close proximity to this transmembrane region (**Figure 5.5A**), we wondered whether the NTD also has some intrinsic affinity to the membrane in the absence of the transmembrane region and if this could affect the interactions with PEX5. We first confirmed that the N-terminal region of PEX14 up to the transmembrane span (residues 1–104) harbors the α-helical globular domain (residues 16–80) but is otherwise unstructured (**Figure 5.5B**). This is indeed demonstrated by the virtually identical secondary chemical shifts for the region comprising the globular domain, while the flanking regions exhibit random coil chemical shifts and low heteronuclear NOE values and are thus intrinsically disordered (**Figure 5.5**B**, C**). The heteronuclear NOE data show a small increased rigidity for residues 90–97 (corresponding to the amino acid sequence QPPHLISQP), which is often observed in P-rich regions **(Chen *et al*, 2013).**

We therefore studied the potential membrane interaction of the PEX14-NTD focusing on the globular domain (residues 16–80). For this, we performed NMR titration experiments of a ^{15}N-labeled sample of the globular α-helical fold in the PEX14-NTD (residues 16–80) (Neufeld *et al.*, 2009) with preformed bicelles up to full saturation (**Figure 5.7A**, **5.6A, B**). Substantial chemical shift perturbations and some line-broadening are observed for many residues indicating a significant interaction of the PEX14-NTD with the bicellar surface. The secondary chemical shifts and heteronuclear NOE values in the absence and presence of bicelles are very similar (**Figure 5.5C**) and the overall spectral dispersion of the NMR signals is not affected (**Figure 5.5A**), demonstrating that the globular fold of the domain is still intact. This is also supported by the very similar circular dichroism spectra and melting temperatures in the absence and presence of bicelles (**Figure 5.6C, D**). To confirm the significance of the spectral changes in the presence of bicelles we performed a control experiment with an unrelated RNA binding domain, which is expected to not interact with membranes. Here, virtually no spectral changes are seen upon addition of bicelles (**Figure 5.6E**).

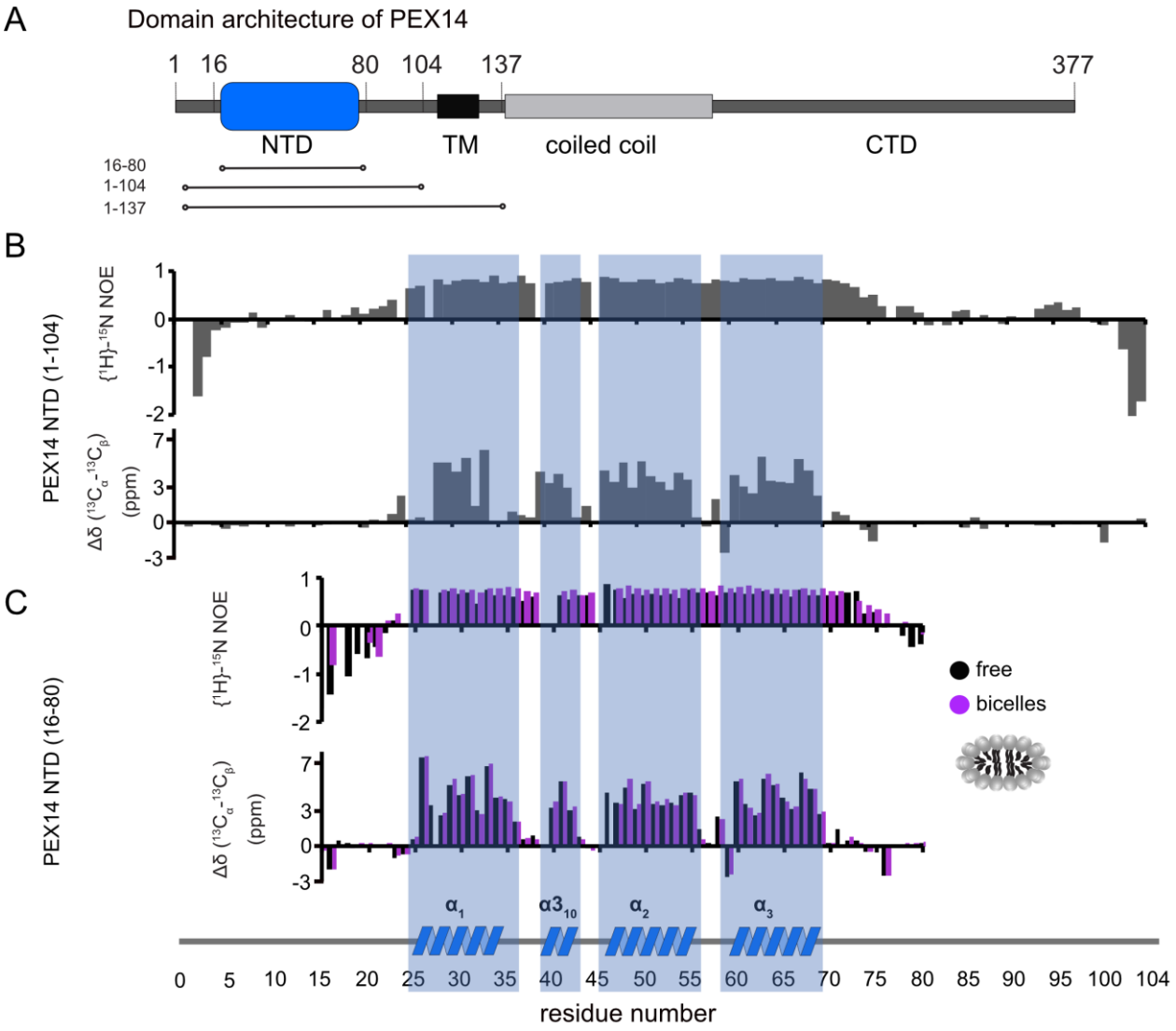

Figure 5.5: Human PEX14-NTD secondary structure. (A) Human PEX14 comprises a globular N-terminal domain (NTD) and a short transmembrane span which is followed by a coiled coil region and an unstructured C-terminal domain. (B, C) $\{^1H\}$-^{15}N heteronuclear NOE (top) and ^{13}C secondary chemical shifts (bottom) of (B) the PEX14-NTD (residues 1-104) free in solution and (C) of the globular domain (residues 16-80) in solution (black boxes) and in the presence of bicelles (magenta boxes).

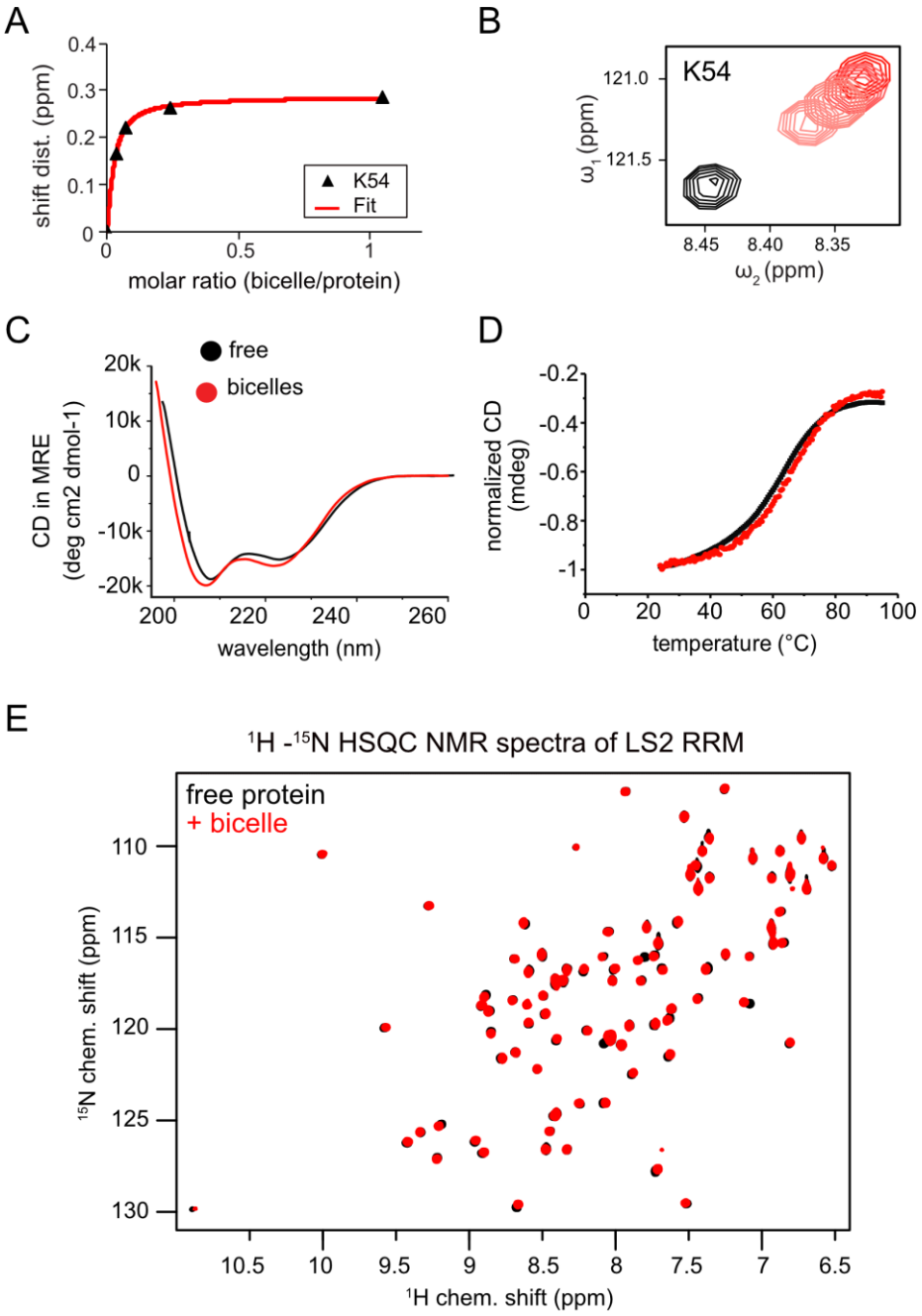

Figure 5.6: Biophysical PEX14 free and in the presence of bicelles. (A) Fit of chemical shift distance vs. bicelle concentration for PEX14 L53 and K54 for K_D calculations. (B) Zoomed spectral regions for of ^{1}H,^{15}N HSQC NMR spectra for amides of L53 and K54. (C) CD spectra of free and bicelle-bound PEX14-NTD (residues 16-80) (D) Melting curves of PEX14-NTD in buffer and in the presence of bicelles, recorded at 222nm. (E) Control titration experiment of the RRM of the *D. melanogaster* LS2 RNA binding protein with bicelles.

We next wanted to explore whether the membrane anchoring of the PEX14-NTD affects the membrane interaction of the globular domain. For this we compared NMR spectra of the PEX14-

NTD (1–104) with a construct that additionally includes the transmembrane region, PEX14 TM (residues 1–137), which was assembled into nanodiscs (**Figure 5.7B**). Notably, mapping the CSPs onto the sequence of PEX14 reveals very similar chemical shift perturbations (**Figure 5.7C**). This shows that the PEX14-NTD has an intrinsic membrane affinity, that is independent of being anchored to the membrane via the TM region. The strongest CSPs (above a threshold of 0.08 ppm) are seen in α-helices 1 and 2 and the $\alpha 3_{10}$ helix (**Figure 5.7C**). Mapping the CSPs onto the structure of the PEX14-NTD (**Figure 5.7D, E**) reveals three hotspots for the interaction, located around Ala32 (helix α1), Arg42 ($\alpha 3_{10}$), and Lys54 (helix α2). Arginine and lysine residues in these three helices form a positively charged surface while the other side of the domain is mainly negatively charged (**Figure 5.7F**). Interestingly, the positively charged surface partially overlaps with the binding interface for WF motifs (**Figure 5.7G**), where PEX14 Lys56 is reported to form an important salt bridge with PEX5 Glu121 (Neufeld *et al.*, 2009). This suggests an at least partially competitive binding of PEX14-NTD to membranes and WF-like motifs in PEX5.

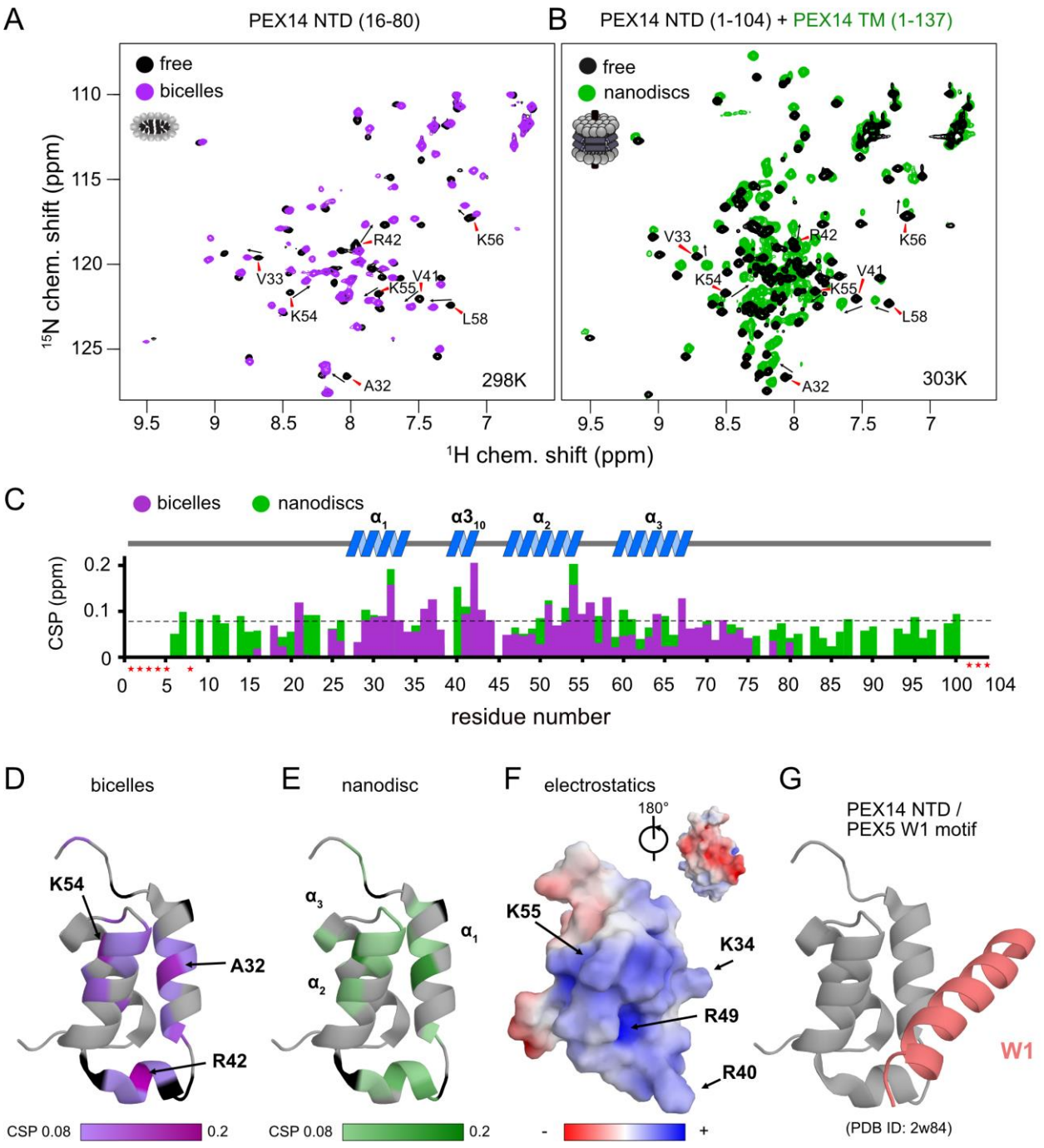

Figure 5.7: The PEX14-NTD interacts with bicelles and nanodiscs. (A) ^{1}H,^{15}N HSQC NMR spectra of the isolated NTD (16-80) in the absence (black) and presence of bicelles (1.1-fold molar excess, magenta). (B) Spectra of the extended PEX14-NTD (residues 1-104) in solution with the PEX14 TM (residues 1-137), anchored in nanodiscs. (C) Chemical shift perturbations vs. residue in the presence of bicelles (magenta) and nanodiscs (green). (D) Mapping of the CSPs shown in (C) in the presence of bicelles (D) and anchored to nanodiscs (E) onto the structure of the globular domain in the PEX14-NTD. Hotspots cluster around Ala32, Arg42, and Lys54. (F) The membrane binding interface is strongly positive charged due to the presence of Arg and Lys residues, as seen

by electrostatic surface rendering. (G) The membrane binding surface partially overlaps with binding site of the PEX14-NTD with WxxxF/H motifs from the PEX5-NTD.

PEX5 PEX14 Interaction in the Presence of Lipids

To investigate the potential competitive binding of PEX14-NTD to membranes and WF-like motifs in PEX5, we performed isothermal titration calorimetry (ITC) with the PEX5-NTD (residues 1–315) and PEX14-NTD (1–104) in aqueous solution (**Figure 5.8A, 5.9A**) or in the presence of bicelles (**Figure 5.8B, 5.9B**). The bicelle concentration used correspond to a 0.9-fold and 1.5-fold molar excess for PEX14 and PEX5, respectively, to ensure saturated membrane binding. Since PEX5-NTD harbors eight possible binding sites for PEX14, the number of sites was fitted to 1/8. This results in a dissociation constant, K_D of 147 ± 16 nM, for the interaction of the PEX5 and PEX14-NTDs in aqueous solution (**Figure 5.8C and Table 5.2**). The free Gibb's energy $\Delta G = -9.3$ kcal/mol is composed of a binding enthalpy, $\Delta H = -147.0$ kcal/mol and $-T\Delta S = 137.7$ kcal/mol, indicating that the interaction is enthalpy driven with a negative change of entropy ΔS. In the presence of bicelles $K_D = 260 \pm 26$ nM, with $\Delta G = -8.9$ kcal/mol, and thus in the same range as without bicelles. Interestingly, both the binding enthalpy and the entropic contributions are reduced by about 60% to $\Delta H = -92.0$ kcal/mol and $-T\Delta S = 83.0$ kcal/mol, respectively, compared to the interaction between the two proteins in the absence of membrane. The reduced binding enthalpy is consistent with a partially competitive binding between the two proteins toward each other and the bicelles. However, the enthalpy reduction is compensated by a reduced entropy loss associated with the PEX5-PEX14 interaction in the presence of a membrane environment. This may result from the increased helical conformation observed for WF motifs in the presence of membrane, which thereby can reduce the entropic cost associated with the formation of a full helical conformation when bound to the PEX14-NTD.

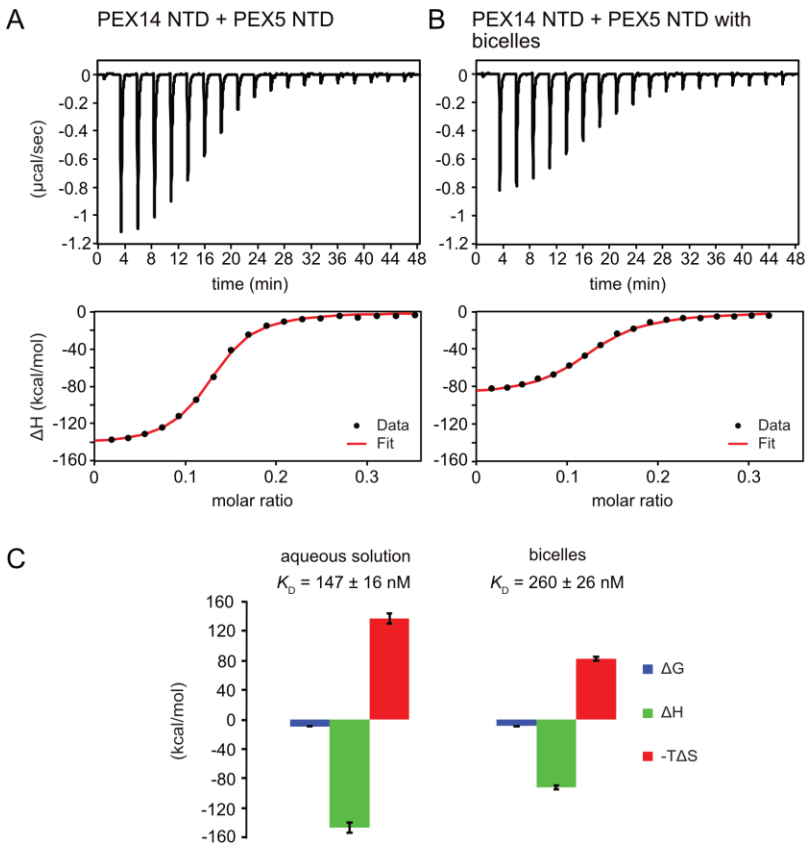

Figure 5.8: PEX5 - PEX14 interaction by isothermal titration calorimetry. ITC experiments were performed as triplicates. (A) Titration of 50 µM PEX5 in aqueous buffer into a 30 µM aqueous solution of PEX14. (B) The same experiment in the presence of 44 µM bicelle corresponding to full saturation of PEX5 and PEX14 with 0.9- fold and 1.5- fold molar excess, respectively. (C) The titration experiments in membrane-like environment show reduced binding enthalpies ΔH and entropic contribution –TΔS and slightly reduced affinity compared to experiments in aqueous solution.

Table 5.2: ITC titration with PEX5-NTD (1-315) on PEX14-NTD (1-104)

Buffer conditions	**N**	**K_D (nM)**	**ΔG (kcal/mol)**	**ΔH (kcal/mol)**	**-TΔS (kcal/mol)**
Buffer	0.125	147 ± 16	-9.3 ± 0.07	-147.0 ± 7	137.7 ± 7
Bicelles	0.125	260 ± 26	-8.9 ± 0.06	-92.0 ± 3	83.0 ± 3

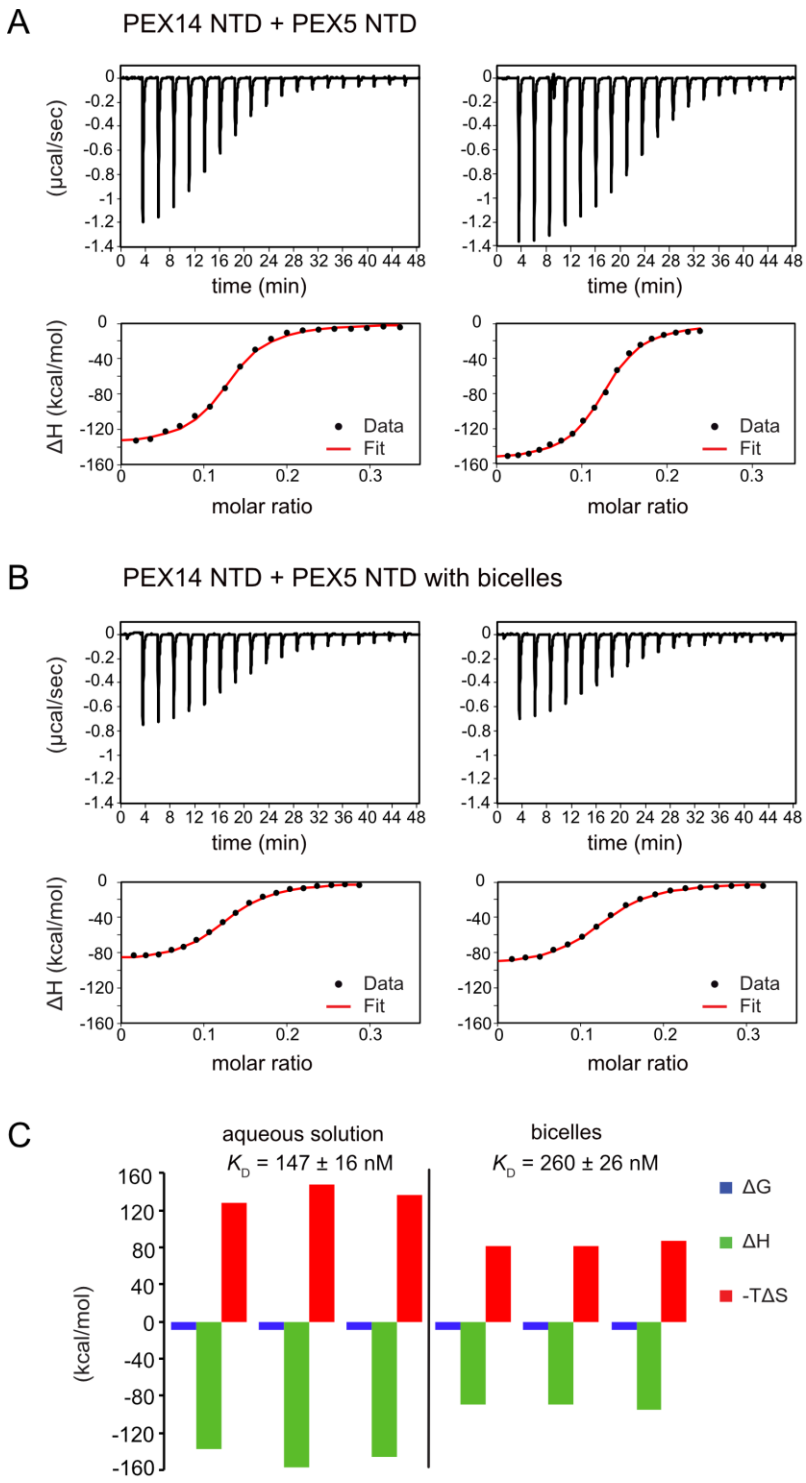

Figure 5.9: ITC experiments for the PEX5-NTD and PEX14-NTD interaction. PEX5-NTD (50µM) in aqueous buffer was titrated to 30 µM PEX14 (A) in aqueous solution or (B) in bicelles containing buffer. (C) Thermodynamic parameters for the titration experiments in solution and in the presence of bicelles. The titration experiments in membrane like environment show similar Gibb's free energies ΔG (blue) but significantly reduced energies ΔH (green) and entropic

contribution –TΔS (red) with a slightly reduced affinity compared to experiments in aqueous solution.

Discussion

Here we present a comprehensive NMR and biophysical analysis of the membrane-associated N-terminal regions of the peroxisomal targeting receptor PEX5 and its binding partner PEX14, which play critical roles for initial steps of the assembly of the peroxisome translocon.

The PEX5-NTD contains eight WF-like binding motifs (W0-W7) that are recognized by the PEX14-NTD. Although this region is overall largely disordered our study reveals that some of the WF-motifs exhibit partial helical conformation, with the highest propensity for W5. The preformed helical conformation may reduce the entropic loss associated with the formation of helix upon complex formation with PEX14. This is consistent with the fact that the W5 motif has the highest binding affinity amongst the eight WF-like binding motifs (Gopalswamy et al., in preparation). Surprisingly, we identified five conserved, amphipathic helical regions α_0 (residues 81–96), α_1 (210–220), α_2 (237–250), α_3 (271–285), and α_4 (287–301), where helices α_0 to α_3 are transient and partially formed, while helix α_4 seems almost fully formed in solution (**Figure 5.2C**). Interaction with membrane-mimicking bicelles stabilizes the α-helical character of these helices including the WF-motifs, with the exception of the W0 motif, which lacks a tryptophan as second aromatic residue. The stabilization of the amphipathic helices might rise from electrostatic interactions from charged amino acids which are often present on the hydrophilic surface (Gimenez-Andres *et al*, 2018), as especially seen for helices α_3 and α_4 of PEX5 (**Figure 5.2F**).

The NMR titrations with bicelles revealed that residues 210–310 in the PEX5-NTD represent the most important interaction site for membrane binding (**Figure 5.2E**). This region harbors the helices α_1 to α_4, which share as a common feature the exposure of phenylalanine and/or tryptophan side chains on the hydrophobic face of the amphipathic helix (**Figure 5.2F**). Interestingly, helix α_0, which lacks an aromatic residue on its hydrophobic face does not show significant spectral changes upon bicelle binding. This suggests that the aromatic residues are important as anchors for the membrane interaction, as is commonly observed for amphipathic helices (Cornell & Taneva, 2006; de Jesus & Allen, 2013; Gimenez-Andres *et al.*, 2018). Amphipathic helices found in the PEX5-NTD are 10 to 15 residues in length exposing four to seven aliphatic residues, which corresponds to two to four helical turns. Such short helices can be found in membrane channels, while the

average length helices in *bona fide* transmembrane proteins (TMPs) is 17.3 residues with a length of about 26 Å (Hildebrand *et al*, 2004). Our NMR titrations show significant chemical shift perturbations in the presence of bicelles (**Figure 5.4**), consistent with a micromolar binding to the bicelle surface. Note, that this experimental setup does not allow us to make conclusions about a potential transmembrane spanning of these regions by PEX5. Thus, we conclude that the PEX5-NTD has significant affinity to the membrane, which may play a role in the formation of the peroxisome translocon.

Our NMR data show that the secondary structure and overall fold of the PEX14-NTD (16–80) is not altered in the presence of membrane-mimicking bicelles (**Figure 5.5C, 5.6C, D**). Nevertheless, significant CSPs are observed upon bicelle binding for amide signals in helices α1 and α2 and the short helix $\alpha3_{10}$ (Ala32, Arg42, and Lys54), which highlights this region as membrane interaction surface (**Figure 5.7C, D**). Notably, NMR spectral changes seen for the PEX14-NTD when inserted into a phospholipid nanodisc by the presence of the transmembrane region, identify the same membrane interaction surface of the globular helical domain in the PEX14-NTD (**Figure 5.7E**). Additional line broadening is seen for some of the lysine and arginine residues in the binding surface presumably reflecting a stronger interaction due to the anchoring of the protein in the nanodisc. The NMR data demonstrate that the small helical fold in the PEX14-NTD has an intrinsic although weak affinity to the membrane surface, independent of the presence of the membrane-spanning helix. The membrane binding helices represent a positively charged surface (**Figure 5.7F**) harboring numerous Arg and Lys residues.

Surprisingly, the PEX14-NTD membrane interaction surface partially overlaps with the interface for the bi-aromatic WF-motifs in the PEX5-NTD (**Figure 5.7D–G**). This suggests at least a partial competition in the binding to the PEX14-NTD. Indeed, our ITC data for the PEX14/PEX5 interaction in the absence and presence of bicelles show a minor decrease of affinity from 150 to 250 nM in the presence bicelles, respectively (**Figure 5.8 and Table 5.2**). In both titration series, we observe negative binding enthalpy and entropy, which demonstrates that the interaction is driven by enthalpy. The negative entropy can be explained by the loss of conformational flexibility upon formation of an α-helical conformation of the WF peptides upon binding to the PEX14-NTD. Interestingly, in the presence of bicelles both ΔH and $-T\Delta S$ are reduced by ≈60%. This is consistent with a competitive binding of the PEX5 WF-motif and the bicelle to the PEX14-NTD. However, the free binding enthalpy remains very similar as a result of enthalpy/entropy compensation. The

partial competition for the PEX14 binding surface reduces the binding enthalpy ΔH, but is likely compensated by the fact that both binding partners are preferentially localize at the membrane and that the membrane interaction increases the pre-existing helical conformation of the WF-motifs, such that loss of conformational entropy from disordered to helical conformation of the WF peptides is reduced, compared to an interaction in the absence of membrane-mimicking bicelles.

Conclusion

We show that the PEX5-NTD, while being overall unstructured, exhibits a number of weakly populated, transient helical regions, which have amphipathic character. Notably the helical propensity is stabilized by a weak micromolar interaction with the membrane (**Figure 5.4**). As judged from the NMR chemical shift perturbations the largest contribution to the membrane binding can be mapped to residues 210–310 in the PEX5-NTD, which comprises the two WF motifs W5 and W6. The other WF motifs and the pre-existing helix α_0 which lacks an aromatic residue, are much less affected in the presence of the membrane. This supports the hypothesis that the WF-like motifs in the N-terminal region of the PEX5-NTD can initiate PEX14 binding (Neuhaus *et al.*, 2014), while the region comprising residues 210–310 help to stabilize PEX5 at the membrane. The PEX14-NTD itself is weakly membrane-associated with the membrane with micromolar affinity (**Figure 5.6**), but this interaction is readily competed out by the PEX5 WF-like motifs, which bind with significantly stronger binding affinity. The weak membrane interactions of the PEX14-NTDs may provide a proof-reading mechanism to avoid random binding events with unspecific targets from the cytosolic compartment.

Experimental procedures

Molecular cloning

The full length genes of human PEX5 (UniProtKB no. P50542) and human PEX14 (UniProtKB no. O75381) were optimized according to the codon usage of *Escherichia coli* and synthesized by IDT (IDT Europe GmbH, Germany). These sequences were used as templates to generate PEX5 (1–113), PEX5 (110–230), PEX5 (228–315), PEX5 (1–315), PEX14 (1–104), and PEX14 (1–137) constructs by polymerase chain reaction (PCR) amplification using the following primers:

PEX5 1–110: F: aaaccatggcgatgcgcgaac

		R: aaaggtaccttacgccagatcggcaacacc
PEX5	110–230:	F: aaaccatggccgatctggcgttatcg
		R: aaaggtaccttactctaaactgacctggccttc
PEX5	228–315	F: aaaccatggagagtttagagtctggtgccggatc
		R: aaaggtaccttagaggtcatcatag
PEX5	1–335	F: gatcccatggcaatgcgggagctggtggag
		R: gatcgcggccgctagtgatcagccaaggggttctcc
PEX14	1–104	F: aaaccatggctagcagcgaacagg
		R: aaaggtaccttaactacccgccggagaatacg
PEX14	1–137	F: aaaccatggctagcagcgaacagg
		R: aaaggtaccttaacctaagatcagcggaaggagg

where F/R refers to forward and reverse primers, respectively.

PEX5 fragments PEX5 (1–113), PEX5 (110–230), PEX5 (228–315), and PEX5 (1–315) were cloned into the bacterial expression pETM10 vector with a non-cleavable N-terminal His_6-tag and PEX14 fragments was cloned into pETM11 vector with His_6-tagged followed by a tobacco etch virus (TEV) cleavage site (EMBL, G. Stier) using *Nco*I and *Kpn*I restriction sites. PEX5 (1–335) was PCR-amplified using pET9d-His-TEV-PEX5L (Schliebs *et al.*, 1999) as a template and subcloned into *Nco*I/*Not*I-digested pET9d expression plasmid.

Protein Sample Preparation

PEX constructs were transformed into *E. coli* BL21 (DE3) cells and expressed in LB or isotope-enriched M9 minimal medium. Single colonies were picked randomly and cultured in the medium with 50 μg/ml kanamycin overnight at 37°C. Overnight cultures were grown at 37°C, diluted 50-fold, and grown up to an optical density of 0.4–0.6 at 600 nm. Protein expression was induced at 37°C with 0.1 mM IPTG for PEX14 (1–104), 0.5 mM IPTG for PEX14 (1–137), and 1 mM IPTG for PEX5 constructs. While PEX5 and PEX14 (1–104) were expressed for 18 h at 18°C, PEX14 (1–137) remained for 4 h at 37°C before being harvested. The cells were harvested by centrifugation at 5,000 rpm for 20 min at 4°C. For protein purification the cell pellets were resuspended in the different binding buffers described below and lysed by pulsed sonication (5 min, 40% power, large probe, Fisher Scientific model 550) followed by centrifugation at 15,000 rpm for 1 h. All proteins were purified using gravity flow Ni-NTA (Qiagen, Monheim am Rhein,

Germany) affinity chromatography which can be described in three steps. First a binding step where the supernatant of the cell lysate is applied to the column. Second, a wash step where endogenous proteins where removed and a third step where the protein of interest is eluted from the column. However, the different natures of the PEX5/PEX14 constructs bring the need of different buffer compositions.

Intrinsically disordered PEX5 protein constructs were lysed in buffer containing 50 mM sodium phosphate (NaP) buffer, pH 8, 300 mM NaCl, 10 mM imidazole, and 8 M Urea. The urea denatures all proteins and prevents binding of contaminants. After binding to the column, urea was removed in an extensive washing step using urea-free buffer. Then, the proteins were eluted in a high imidazole buffer (50 mM NaP, pH 8, 300 mM NaCl, 500 mM imidazole) before final purification via size exclusion chromatography (Superdex S75, 16/600, GE Healthcare, Rosenberg, Sweden) in NMR buffer (see below) and lyophilized for long term storage.

The PEX14 constructs comprising residues 1–104 and 1–137 were lysed in 50 mM Tris/HCl, pH 8.0, 100 mM NaCl, and protease inhibitor mix (Serva, Heidelberg, Germany). After lysis, the PEX14 1–137 suspension was adjusted to 1 M NaCl and 1% (w/v) dodecylphosphocholine (DPC). After binding to the column a wash step with additional 20 mM imidazole was performed. While salt concentration was kept constant for PEX14 1–104, NaCl and DPC concentrations for PEX14 1–137 were lowered to 500 mM and 0.2% (w/v), respectively. The protein was eluted by increasing imidazole to 500 mM. PEX14 1–137 was subsequently further purified for nanodisc assembly including a buffer exchange to 50 mM Tris, pH 8.0, 100 mM NaCl, 0.1% (w/v) DPC using a Superdex S200 (16/600, GE Healthcare). PEX14 (1–104) was final purified after TEV cleavage running a reverse Ni^{2+} column and a Superdex S75 (16/600), where the buffer was changed to NMR buffer containing 50 mM NaP pH 6.5 and 100 mM NaCl. The PEX14 16–80W construct (with a C-terminal Trp) was expressed and purified as described previously (Neufeld *et al.*, 2009). Protein expression and purification of full-length PEX5 and PEX5 (1–335) was carried out as described (Schliebs *et al.*, 1999). Protein purification and quality was confirmed by SDS PAGE (**Figure 5.10A, B, D**).

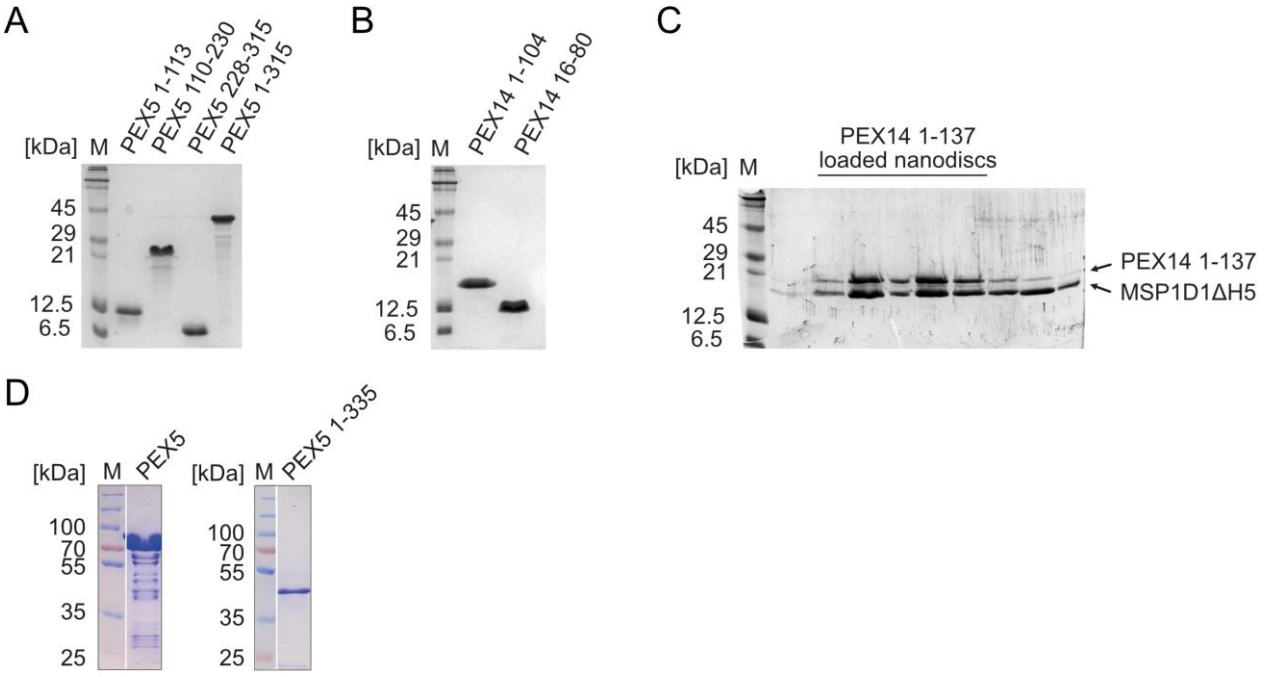

Figure 5.10: SDS PAGE analysis of purified samples for (A) PEX5 1-113, 110-230, 228-315, 1-315 and (B) PEX14 1-104, 16-80. (C) SDS PAGE analysis of the size exclusion peak corresponding to PEX14 loaded nanodiscs. (D) SDS PAGE analysis of recombinant full-length PEX5 and PEX5 1-335 used for flotation assays. Note, that the degradation bands seen for PEX5 full-length do not bind to lipid vesicles in the flotation assay.

Uniformly ^{15}N or ^{15}N, ^{13}C labeled proteins were expressed in H_2O or D_2O M9 minimal medium supplemented with 50 µg/ml kanamycin, 1 g/liter ^{15}N NH_4Cl and 2 g/liter hydrated or deuterated [U-^{13}C] glucose as the sole sources of nitrogen and carbon, respectively

Bicelles and Nanodiscs Preparation

Lipids 1,2-Diheptanoyl-*sn*-glycero-3-phosphocholine (D7PC) and 1,2-Dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) were purchased from Avanti Polar Lipids (Alabaster, United States). Bicelles were prepared according to the established protocols (Sommer *et al.*, 2012). Briefly, water free D7PC and DMPC were dissolved in chloroform to generate stock solutions of 500 mM and 100 mM, respectively. The lipids were mixed in a ratio of 1 to 1, dried under vacuum and rehydrated in 20 mM NaP, pH 6.5, 100 mM NaCl, 0.02% (w/v) NaN_3 to generate a 240 mM lipid stock. The bicelles with $q = 0.2$ were formed by several freeze and thaw cycles in liquid N_2 yielding a clear, viscous bicelle solution of 870 µM concentration. The bicelle concentration was calculated based on the number of DMPC molecules in one bicelle. The radius *R* of the bilayer region of the

bicelle (for q = 0.2) was calculated to be 2.04 nm using the formula $R = 1/2rq[\pi + (\pi^2 + 8/q)^{1/2}]$ assuming a bilayer thickness of 4 nm with a radius r = 2 nm (Klöpfer & Hagn, 2019). Thus the calculated surface area of the bicelle is 1307 Å^2, which corresponds to 46 DMPC molecules (given a surface area of 57 Å^2 per DMPC molecule).

Nanodisc assembly with freshly purified PEX14 (1–137) was performed with a lipid mixture of 75% deuterated DMPC, 25% deuterated DMPG (FB Reagents, Cambridge, United States) and the 19.5 kDa scaffold protein MSP1D1Δ5 as described (Hagn *et al.*, 2018). Buffer exchange to NMR buffer was done via size exclusion chromatography on a Superdex S200 (16/600). To confirm successful reconstitution of PEX14 (1–137) peak fractions were analyzed by SDS-PAGE (**Figure 5.10C**). Protein concentration of 150 mM of 250 µl sample was then transferred to a Shigemi (Shigemi Inc., Allison Park, United States) tube for NMR experiments.

Liposome Preparation and Flotation Assay

1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) were purchased from Avanti Polar Lipids, Inc. (United States). DOPC/DOPE lipids were mixed with a ratio of 7 to 3 in 50 mM NaCl, 20 mM Tris, pH 7.4 as described (Kerssen *et al.*, 2006). In addition, 10 mM DPMC, purchased from Sigma-Aldrich (Germany), was resuspended in 50 mM NaCl, 20 mM Tris, pH 7.4. Small unilamellar vesicles (SUVs) were obtained by sonication of the multilamellar vesicle (MLV) suspension using an ultrasonic bath (Sonorex RK 52, Bandelin) followed by 10 cycles of freezing and thawing (Pick, 1981).

Liposomes, either DOPC/DOPE or DMPC, were incubated with 1.5 nmol purified human PEX5 or PEX5 (1–335) for 1 h at room temperature (ratio protein/lipid: 1/750). Incubation of liposomes with protein and all following steps were performed in 50 mM NaCl, 20 mM Tris–HCl, pH 7.4. The loading samples were adjusted to a sucrose concentration of 45% (w/v) using 65% (w/v) sucrose solution and 0.4 ml of each sample were layered onto 0.52 ml 50% sucrose cushion at the bottom of 11 ml ultracentrifuge tubes. 1.3 ml 40% (w/v) sucrose, 5.1 ml 25% (w/v) sucrose 2.6 ml of buffer without sucrose were stepwise added. After ultracentrifugation for 4 h at 175,000 × *g* at 4°C in a swing-out rotor, the linear gradient (0 to 50% (w/v) sucrose) was collected as ten 1 ml fractions from top to bottom. The fractions were separated by SDS-PAGE and analyzed by immunoblotting using polyclonal rabbit anti PEX5 antibodies.

NMR Spectroscopy

NMR data were collected on Bruker Avance III spectrometers operating at 500, 600, 800, 900 or 950 MHz, equipped with cryogenic probes. The sequential assignment of backbone resonances for PEX5 fragments and PEX14 (1–104) were performed based on heteronuclear experiments such as ^{1}H-^{15}N-HSQC, HNCA, HN(CO)CA, CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, HN(CA)NNH and H(NCA)NNH (Sattler M *et al.*, 1999; Weisemann *et al.*, 1993). {^{1}H}-^{15}N heteronuclear NOE (hetNOE) experiments (Farrow *et al.*, 1994) were performed using the pulse sequence hsqcnoef3gpsi (Bruker, Avance version 12.01.11) with a 5 s interscan delay. NOE values are given simply by the ratio of the peak heights in the experiment with and without proton saturation (hetNOE = I_{sat}/I_0) (Renner *et al.*, 2002). NMR-Spectra were processed using Topspin (Bruker Biospin, Rheinstetten, Germany) or NMRPipe (Delaglio *et al.*, 1995) and analyzed using CcpNMR Analysis 2.4.2 (Vranken *et al.*, 2005).

All NMR experiments with PEX5 were recorded in 20 mM NaP pH 6.5, 100 mM NaCl, 0.02% (v/v) NaN_3, and 2 mM DTT at 298°K at 600 MHz for triple resonance experiments or at 500 MHz for hetNOE experiments. Triple resonance experiments of free PEX5 1–113, 110–230, and 228–315 were performed at concentrations of 225 µM, 200 µM, and 200 µM, respectively. Assignment experiments for PEX5 1–113, 110–230, and 228–315 in the presence of 870 µM bicelles (q = 0.2) were recorded at 350 µM to 430 µM. {^{1}H}-^{15}N heteronuclear NOE experiments were performed by dissolving lyophilized ^{15}N-labeled PEX5 1–113, PEX5 110–230 and PEX5 228–315 in a buffer containing 20 mM NaP pH 6.5, 100 mM NaCl, 2 mM DTT to a final concentration of 190 µM, 90 µM, and 165 µM, respectively.

Experiments with all PEX14 constructs were recorded in 20 mM NaP pH 6.5, 100 mM NaCl. While spectra of PEX14 (16-80W) were recorded at 298°K at 600 MHz proton Larmor frequency, NMR experiments for PEX14 1–104 and 1–137 without and with nanodiscs were collected at 298°K or 303°K at 800, 900, or 950 MHz proton Larmor frequency, respectively. Backbone assignments and hetNOE experiments of PEX14 16–80W in the presence of bicelles were recorded at 950 µM in 150 µM bicelle solution. Triple resonance experiments for PEX14 1–104 were recorded on 950 MHz at a concentration of 750 µM and 298°K. Nanodiscs of 6 nm size with PEX14 1–137 were assembled and purified as described above, 2D ^{1}H-^{15}N HSQC experiments were measured at a final concentration of 150 µM at 303°K and 900 MHz.

Titration experiments of preformed bicelles to ^{15}N-labeled PEX5 1–113, 110–230, and 228–315 at 280 μM, 150 μM, and 150 μM and PEX14 16-80W at 110 μM were performed with increasing bicelle concentration up to 1.5-fold, 2-fold, and 1-fold molar excess, respectively. The chemical shift perturbation ($\Delta\delta_{avg}$) was calculated by using formula $\Delta\delta_{avg} = [(\Delta\delta_H)^2 + (\Delta\delta_N/6.3)^2]^{0.5}$. Dissociation constants (K_D) were fitted to DMPC concentration with a one-site specific binding model within Origin software (OriginLab Corporation, United States). The equation used for the fitting is $\Delta\delta = \Delta\delta_{max}/(2\ [P_t])*\{[L] + [P_t] + K_D - (([L] + [P_t] + K_D)^2\text{-}4[P_t][L])^{1/2}\}$, where $\Delta\delta$ is the individual and $\Delta\delta_{max}$ the maximum shift distance, P_t is the total protein concentration and [L] the DMPC ligand concentration. The K_D of ∼20 representative residues of each construct was fitted to the DMPC concentration assuming that all DMPC molecules are associated with bicelles, whereas partially water soluble D7PC molecules may exist in an equilibrium between solution and bicelle-bound. Bicelle concentrations were derived from the DPMC concentration by scaling with a factor of 1/46 (according to the number of DPMC molecules per bicelle, see above) to obtain apparent dissociation constants K_D^{app} for the bicelle interaction.

Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) measurements were performed as triplicates at 25°C using a MicroCal PEAQ-ITC (Malvern Instruments Ltd., United Kingdom) calorimeter. Buffer conditions were 20 mM NaP pH 6.5, 100 mM NaCl containing none or 44 μM bicelles. Pex5 (1–315) at a concentration of 50 μM was injected in the cell containing Pex14 (1–104) at a concentration of 30 μM. The concentration of PEX14 was corrected with the fit, since it cannot be accurate measured at 280 nm owing to the extinction coefficient of only 1490. The dilution effect of PEX5 as a control experiment was subtracted before the data were fitted to a one-site binding model using the Malvern Analysis software.

Circular Dichroism (CD)

Far-ultraviolet circular dichroism (Far-UV CD) and thermal unfolding measurements were carried out using a Jasco J-810 spectropolarimeter equipped with a peltier thermal controller in a 0.1 cm path length quarts cuvette. Measurements were performed between 10 or 15°C to 95°C with 1°C/min scanning speed. Far UV-CD data of PEX14 (16-80W) at concentration of 25 μM in bicelle free or 44 μM bicelle (q = 2) in 10 mM sodium phosphate, 50 mM sodium chloride and pH 6.5

were collected at 25°C in the range of 190–260 nm wavelength. Protein-bicelle complexes were incubated for 2–3 h prior to the experiment. Spectra were collected in 10 accumulations and subtracted from the spectrum of the buffer control. Thermal unfolding spectra were collected at 222 nm. The midpoint of the folding and unfolding (T_m) of the complex was derived from the raw data by fitting to the equations for the sigmoidal curve, $Y = A2 + (A1-A2)/(1 + \exp((x-x_0)/dx))$. Where A1 and A2 are the folding and unfolding intercept, respectively. x is the midpoint of the cure and dx is the slope of the curve. All curves were fitted by using Origin software (OriginLab Corporation, United States).

Author Contributions

SG and MG performed and analyzed NMR and ITC experiments with bicelles. SG, MG, PZ, and ME performed cloning and protein expression and purification. ME performed NMR experiments with nanodiscs. MR, WS, and RE performed and analyzed flotation experiments. MS and MG designed the study. SG and MS wrote the manuscript. All authors contributed and approved the manuscript.

Chapter 6:

Distinct conformational and energetic features define the specific recognition of (di)aromatic peptide motifs by PEX14

The work in this Chapter is based on the manuscript:

Mohanraj Gopalswamy*, Chen Zheng*, Stefan Gaussmann*, Hamed Kooshapur, Eva Hambruch, Wolfgang Schliebs, Ralf Erdmann, Iris Antes, Michael Sattler, *Distinct conformational and energetic features define the specific recognition of (di)aromatic peptide motifs by PEX14*

Submitted to *Biological Chemistry*, in revision

Distinct conformational and energetic features define the specific recognition of (di)aromatic peptide motifs by PEX14

Author contribution: The author of this thesis, Stefan Gaussmann along with Chen Zheng and Mohanraj Gopalswamy contributed equally to the original work of this article. Stefan Gaussmann carried out molecular biology and protein expression and purification, NMR experiments and analysis and preparation of the manuscript.

Abstract

The cycling import receptor PEX5 and its membrane-located binding partner PEX14 are key constituents of the peroxisomal import machinery. Upon recognition of newly synthesized cargo proteins carrying a peroxisomal targeting signal type 1 (PTS1) in the cytosol, the PEX5/cargo complex docks at the peroxisomal membrane by binding to PEX14. The PEX14 N-terminal domain (NTD) recognizes (di)aromatic peptides, mostly corresponding to Wxxx(F/Y)-motifs, with nano- to micromolar affinity. Human PEX5 possesses eight of these conserved motifs distributed within its 320-residue disordered N-terminal region. Here, we combine biophysical (ITC, NMR, CD), biochemical and computational methods to characterize the recognition of these (di)aromatic peptides motifs and identify key features that are recognized by PEX14. Notably, the eight motifs present in human PEX5 exhibit distinct affinities and energetic contributions for the interaction with the PEX14-NTD. Computational docking and analysis of the interactions of the (di)aromatic motifs identify the specific amino acids features that stabilize a helical conformation of the peptide ligands and mediate interactions with PEX14-NTD. We propose a refined consensus motif **WΦxE(F/Y)Φ** for binding to the PEX14-NTD and discuss conservation of the (di)aromatic peptide recognition by PEX14 in other species.

Introduction

Peroxisomes are ubiquitous organelles with varying metabolic capacities dependent on species, tissues and environmental changes (for an overview of function of peroxisomes see Deb & Nagotu (2017). Peroxisomal proteins are nuclear-encoded and need to be imported into the organelle post-translationally (Emmanouilidis *et al.*, 2016; Erdmann & Schliebs; Fujiki & Lazarow, 1985; Giannopoulou *et al*, 2016; Meinecke *et al.*, 2010). Import of peroxisomal matrix proteins depends

on the recognition of cargo proteins harboring peroxisomal transport signal (PTS) peptide motifs. However, cargo proteins lacking a PTS sequence can also be transported into peroxisomes by interacting with PTS containing proteins by "piggyback" mechanism (Effelsberg *et al*, 2015; Yang *et al*, 2001). The main pathway of protein import into peroxisomes depends on the cycling import receptor PEX5, which recognizes cargo proteins with a peroxisomal targeting signal 1 (PTS1) in the cytosol (Gould *et al*, 1989; Gould *et al.*, 1987). Receptor-cargo complexes are then docked to the peroxisomal membrane by binding to the membrane-associated protein PEX14 (Brocard *et al*, 1997; Jansen *et al*, 2021; Will *et al.*, 1999).

The PEX14 N-terminal domain (NTD) forms a small globular helical fold, while the C-terminal region of PEX14 is largely unstructured (Emmanouilidis *et al.*, 2016; Gaussmann *et al*, 2021). This domain binds to (di)aromatic peptide motifs present in peroxins (Neufeld *et al.*, 2009; Otera *et al.*, 2002) and unrelated proteins as it has been shown for β-tubulin (Reuter *et al.*, 2021). The intrinsically unstructured NTD of human PEX5 (residues 1-320) harbors eight conserved peptide motifs, seven comprising a Wxxx(F/Y)-motif and one non-canonical LVxEF motif (**Figure 6.1**) (Neuhaus *et al.*, 2014; Saidowsky *et al.*, 2001). These peptide motifs bind to the conserved PEX14-NTD (Figure 6.2) (Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014; Su *et al*, 2009; Watanabe *et al*, 2016). Significant differences are observed for the binding affinity and kinetics of the individual motifs (Gaussmann *et al.*, 2021; Neuhaus *et al.*, 2014; Saidowsky *et al.*, 2001) and higher order interactions for regions comprising multiple motifs have been reported (Shiozawa *et al*, 2009). In ITC experiments a stoichiometry of 1:8 has been determined, consistent with the presence of eight (di)aromatic motifs in the PEX5 N-terminal domain (Neuhaus *et al*, 2016).

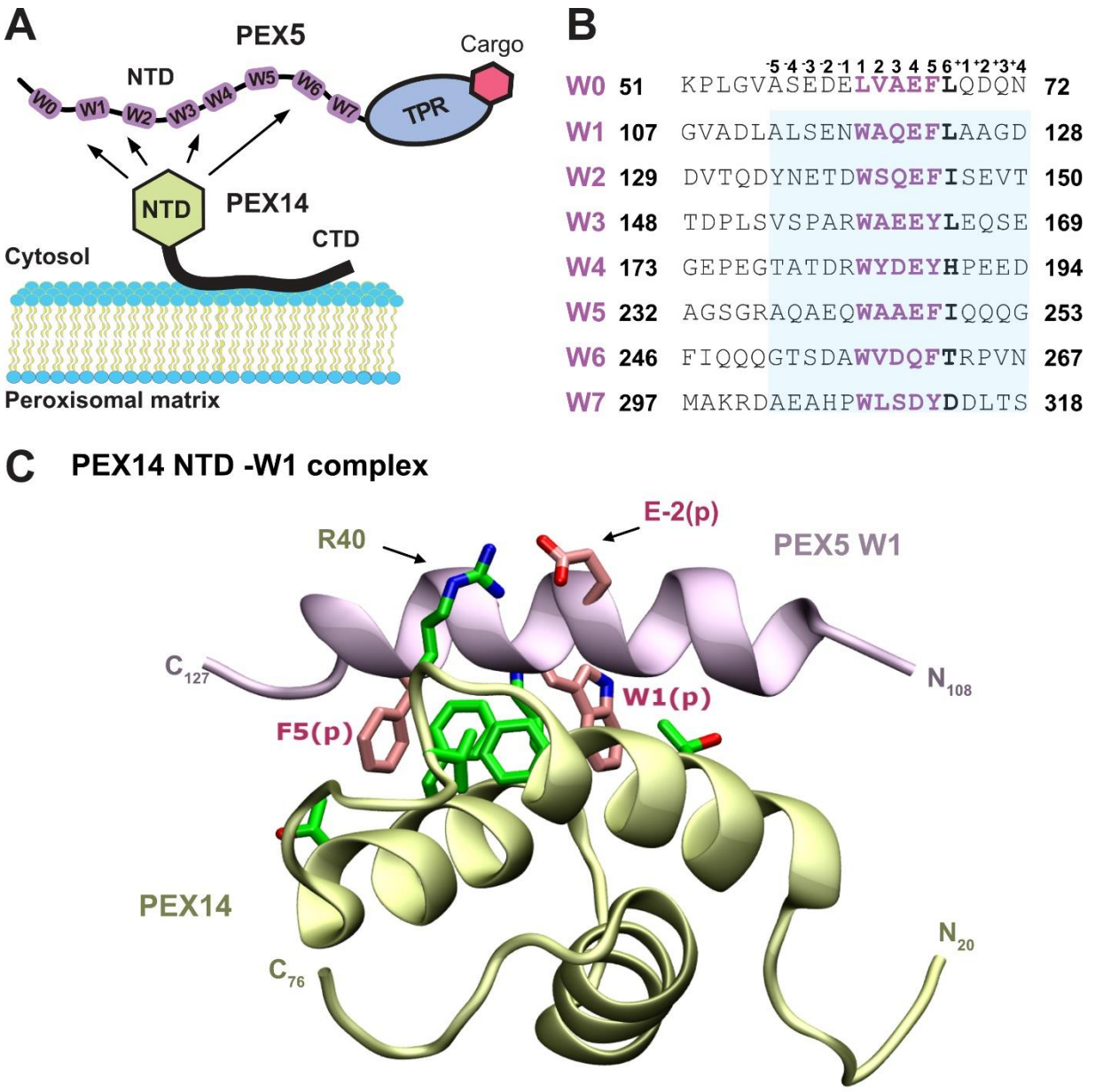

Figure 6.1: Interactions between PEX14-NTD and PEX (di)aromatic motifs. (A) The PEX14-NTD, which is located at the peroxisomal membrane, recognizes the W0-W7 motifs in the N-terminal region of PEX5. **(B)** 20-mer peptide comprising the W0-W7 motifs found in human PEX5 (UniPort ID: P50542). The 20-mer peptides were used in computational studies. Blue shaded are the 15-mer peptides (W1-W7) used for ITC experiment. For W0 ITC PEX (1-113) was used. The central 6-residues harboring the core motifs are highlighted in bold. **(C)** Structure of the PEX14-NTD/PEX5 complex (PDB-ID: 2W84). PEX14-NTD is shown in yellow/green and PEX5 in pink. Important residues are given in licorice representation. The F5(p) and W1(p) residues correspond to the Wxxx(F/Y) motif.

A potential functional relevance of these distinct thermodynamic and kinetic binding parameters was suggested based on mutational analysis. Substitution of the LVxEF motif with the W1 Wxxx(F/Y) motif impaired protein import into peroxisomes (Neuhaus *et al.*, 2014). These data suggest that the presence of multiple PEX14-binding motifs and differential interactions with the PEX14-NTD are functionally important for processing of the PTS1 receptor at the peroxisomal membrane. It has been speculated that the most N-terminally located LVxEF motif may represent an initial tethering site of PEX5, from which the cargo-loaded receptor is further processed in a sequential manner by "handing" over Wxxx(F/Y) motifs to the PEX14 at the membrane. A non-mutually exclusive function of the presence of eight PEX14 binding motifs may involve avidity effects for the PEX5-PEX14 interaction. Hence, the presence of multiple binding motifs may enhance the PEX14 interaction by an increased local concentration (Emmanouilidis *et al.*, 2016). Interestingly, the roles of Wxxx(F/Y) motifs to mediate protein interactions are conserved in yeast, Leishmania and trypanosoma, although specific contributions of individual motifs may vary (Cyr *et al*, 2008; Hojjat & Jardim, 2015; Watanabe *et al.*, 2016).

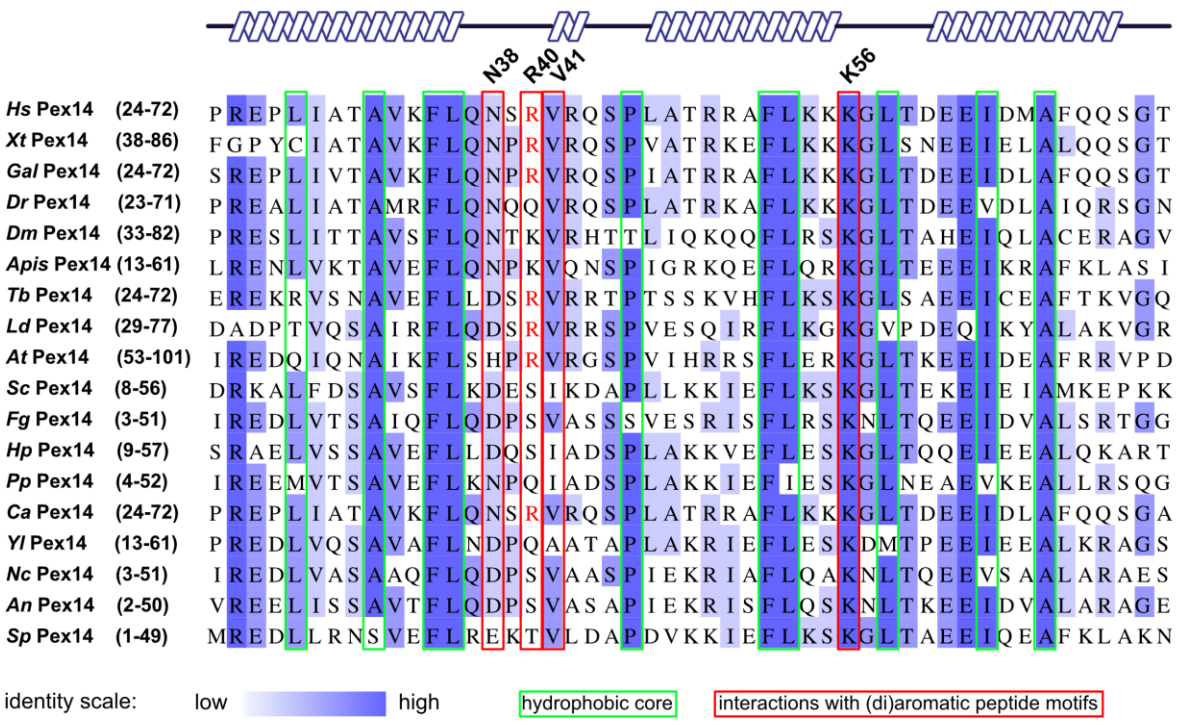

Figure 6.2: Multiple sequence alignment of PEX14-NTD from different species. Primary sequence alignment of PEX14 using Clustal Omega (EMBL-EBI webservices) from *Hs: Homo sapiens, Xt: Xenopuslaevis, Gal: Gallus gallus, Dr: Daniorerio, Ce: Caenorhabditiselegans, Ag: Ashbyagossypii, Dm: Drosophila melanogaster, Apis: Apisapis, Tb: Trypanosomabrucei, Ld: Leishmaniadonovani, At: Arabidopsis thaliana, Sc: Saccharomyces cerevisiae, Fg:*

Fusariumgraminearum, Hp: Hansenulapolymorpha, Pp: Pichiapastoris, Ca:Caviaporcellus, Yl: Yarrowialipolytica, Nc: Neurosporacrassa, An: Aspergillusnidulans, Sp:Schizosaccharomycespombe. All primary sequences are taken from Uniport database. Conservation of the residues are indicated by the identity scale showing higher degree of conservation with increasing blue shade. Key residues for peptide binding are highlighted by red boxes, including Asn38, Arg40, Val 41, and Lys56. Residues of the hydrophobic core are indicated with green boxes. Schematic representation of the secondary structure of human PEX14-NTD (PDB 2W84) is shown on top.

NMR-derived structures have been reported for the human PEX14-NTD in complex with the first Wxxx(F/Y) motif in PEX5 (W1, PEX5 residues 108-127) (Neufeld *et al.*, 2009) (**Figure 6.1B**) and with the N-terminal LVxEF motif (W0, PEX5 residues 57-71) (Neuhaus *et al.*, 2014) These structures show that all (di)aromatic motifs bind to the PEX14-NTD in an α-helical conformation utilizing two hydrophobic binding pockets in the PEX14-NTD fold to recognize aromatic and/or aliphatic side chains, suggesting a broad consensus motif. To address the different contributions of amino acids in the diverse (di)aromatic motifs, a better understanding of their interactions and binding energies with the PEX14-NTD is important. In this context a mutational analysis of the *Trypanosoma brucei* PEX14-NTD/PEX5 interaction suggested that position 4 in the Wxxx(F/Y) motif is essential for binding (Watanabe *et al.*, 2016). However, a systematic analysis of the energetic and conformational features of the eight motifs present in human PEX5 is not available.

Here, we present a comprehensive analysis of the PEX14 binding motif features combining experimental affinities and thermodynamic parameters obtained from ITC and peptide overlay binding assays, circular dichroism data, and computational analysis of the energetic contributions. Our results show that the eight motifs present in human PEX5 exhibit a broad range of affinities for binding to the PEX14-NTD. Computational analysis of the binding interfaces indicates that, in addition to the two hydrophobic (aromatic/aliphatic) side chains, other amino acid types in different positions in the motifs play important roles to stabilize a helical conformation and to mediate high affinity binding to the PEX14-NTD. Based on our analysis we propose a refined peptide consensus motif, **WΦxE(F/Y)Φ**, for recognition by PEX14 (where Φ, denotes a hydrophobic residue, x any amino acid). We analyze the conservation of this motif and its recognition by PEX14 in other species. Our integrated approach, combining experimental data and computational simulations highlights the role of the PEX14-NTD as a conserved domain for the recognition of helical (di)aromatic peptides with a broad consensus but identifies unique contributions of specific amino acids for high affinity binding.

Results

Thermodynamic parameters of the PEX14-PEX5 peptide interactions

We first examined the interaction and thermodynamics of the recognition of the eight (di)aromatic binding motifs in the PEX5-NTD with the PEX14-NTD using isothermal titration calorimetry (ITC). For this we titrated 15-mer peptides, comprising the seven Wxxx(F/Y) motifs flanked by five residues, as well as one construct comprising PEX5 residues 1-113, which harbors the W0 (LVxEF) peptide **(Figure 6.1B)**. The binding processes represent a single transition with exothermic binding enthalpies with dissociation constants ranging from 60 nM to 6 µM (**Figure 6.3, Table 6.1**).

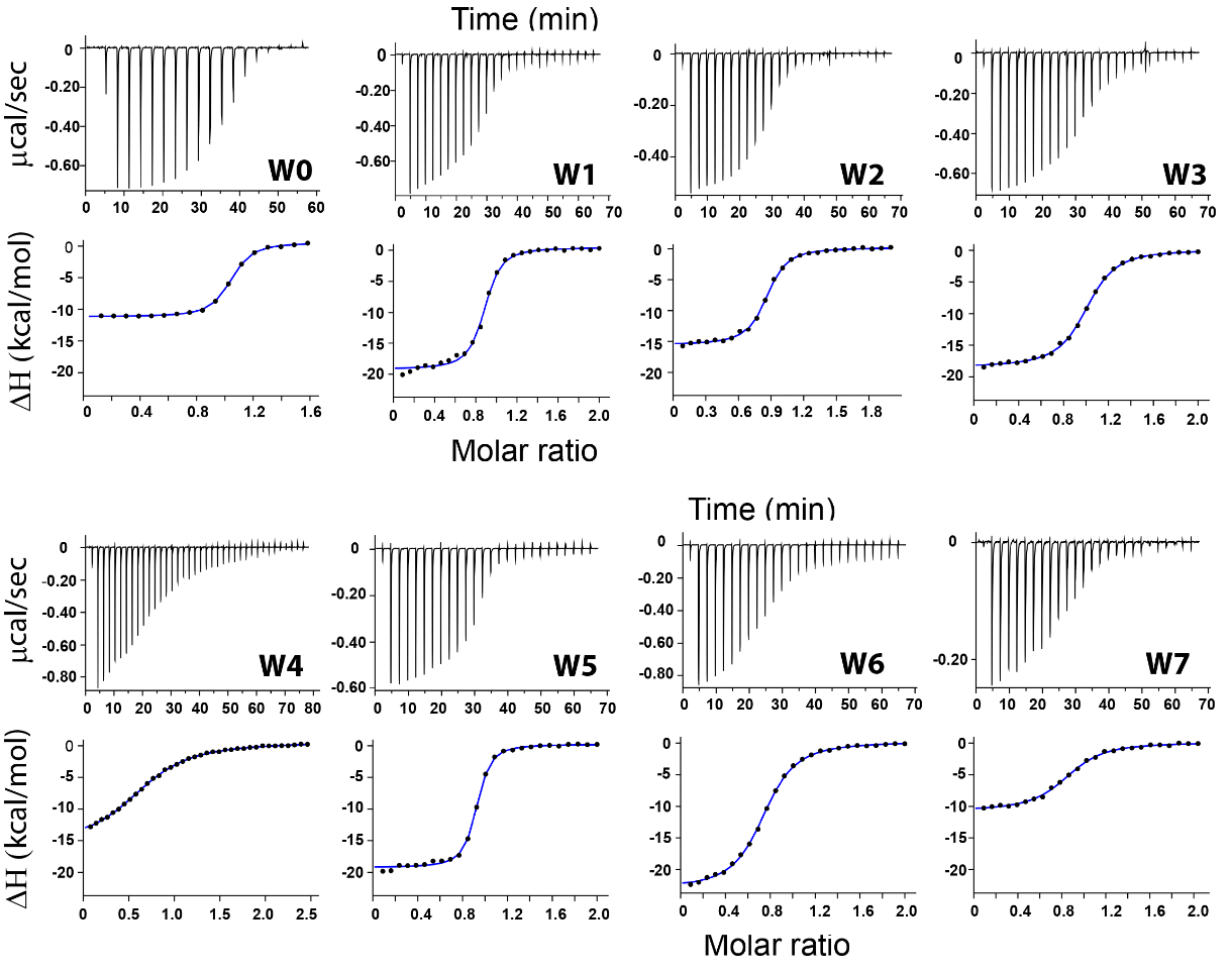

Figure 6.3: ITC profiles of the interaction of human PEX14-NTD with PEX5-derived peptides. For each titration, the upper graph indicates the injection of 1.5 µL of 200 µM of peptide titrated into 20 µM of PEX14-NTD. The lower graph shows the fit to the ITC data according to a single-site binding model. The temperature was set to 25 ºC for all experiments.

Table 6.1: Isothermal titration calorimetry of PEX5 W0-W7 peptide binding to PEX14-NTD. All titrations were performed at 25 °C. ITC data were fitted in to 1:1 binding model using Microcal Origin software. Stoichiometry for all the titration is 1:1 (protein: peptide). W0 and W1 ITC data are published previously (Neuhaus *et al.*, 2014; Shiozawa *et al.*, 2009) and showed for the comparison. The error values are obtained from the curve fit.

peptide	K_D (nM)	ΔH (kcal/mol)	ΔG (kcal/mol)	(-TΔS) (kcal/mol)
W0	173 ± 9	-11.7 ± 0.07	-9.23	2.52
W1	139 ± 20	-19.6 ± 0.26	-9.36	10.2
W2	209 ± 12	-15.7 ± 0.11	-9.12	6.62
W3	344 ± 18	-18.6 ± 0.14	-8.82	9.75
W4	6310 ± 234	-14.2 ± 0.18	-7.1	7.12
W5	60 ± 6	-19.5 ± 0.15	-9.86	9.62
W6	575 ± 26	-23.3 ± 0.17	-8.52	14.7
W7	727 ± 57	-10.9 ± 0.15	-8.38	2.57

The relative differences are comparable to previously reported fluorescence polarization studies (Saidowsky *et al.*, 2001), some minor differences variations may reflect distinct buffer conditions and temperatures used. The significant variations in the binding affinities for the different (di)aromatic peptide motifs in the PEX5-NTD suggest that – in addition to the two conserved aromatic residues – further amino acids contribute to the interaction. Interestingly, ITC experiments with yeast PEX14-NTD and PEX5-NTD show only binding to a reverse WxxxF (W3) motif with μM affinity (**Figure 5.3**), consistent with previous reports (Kerssen et al., 2006), suggesting further variations in the binding interface.

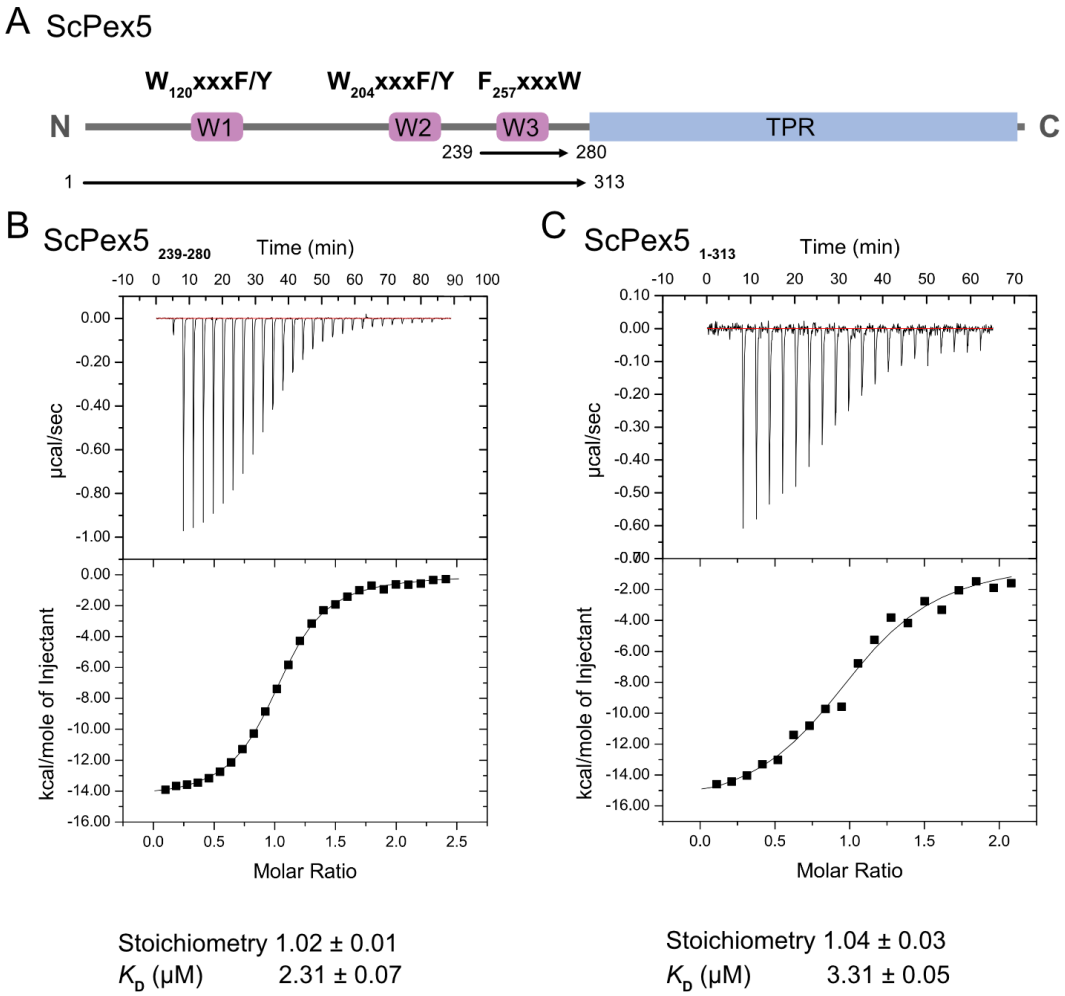

Figure 6.3: ITC data for the interaction of ScPex14-NTD and the ScPex5-NTD. (A) Schematic representation of ScPex5. (B) ITC profile of the titration of 680 µM ScPex5 (239–280) into 57 µM ScPex14-NTD. (C) ITC profile of the titration of 320 µM ScPex14-NTD into 32 µM ScPex5(1–313). The W3 peptide has a binding affinity comparable to the full PEX5-NTD, suggesting that W3 mediates the interaction, while W1 and W2 have minor contributions.

The energetics of binding are notably different for each motif (**Figure 6.4A**). Under all measurement conditions, binding between the peptides and PEX14-NTD is exclusively enthalpy-driven ($\Delta H<0$) with an unfavorable entropic contribution ($-T\Delta S>0$). The enthalpy changes for the binding of W1, W3, W5 and W6 motifs are larger (about -19 kcal/mol) compared to W0, W7 peptides (-11 kcal/mol). Moderate enthalpy values are observed for the W2 and W4 peptides (-15 kcal/mol). Noteworthy, binding of W0, which lacks the first aromatic residue, and W7 show the lowest entropic penalty. We next characterized the thermal stability of apo PEX14-NTD and in complex with the W0-W7 ligands using circular dichroism (CD) measurements **(Figure 6.4B)**. Temperature- dependent measurements allow the determination of the transition midpoint (T_m) for thermal denaturation of the free PEX14-NTD and

when bound to the W0-W7 ligands. The T_m values correspond to the temperature at which 50% of protein is unfolded and are directly correlated to the stability of protein complex (Rees & Robertson, 2001; Tol *et al*, 2013). The T_m is 61°C for the apo PEX14-NTD and ranges from 65°C to 77°C for the PEX14-NTD/ligand complexes in the presence of two-fold excess of peptide ligands **(Figure 6.4B, Figure 6.5)**. As expected, ligand binding did not cause any unfolding transition but stabilized the fold. Notably, the highest T_m value of 77 °C is observed for the complex with the W5 peptide, which also has the highest binding affinity (ΔG=-9.86 kcal/mol), while the lowest stability (T_m = 64°C) was observed for W4, which features the lowest binding affinity (ΔG=-7.10 kcal/mol). Overall, the melting temperatures and thermodynamic stabilities of the peptide complexes show a clear correlation with the free energies calculated from the ITC data (**Figure 6.4B**).

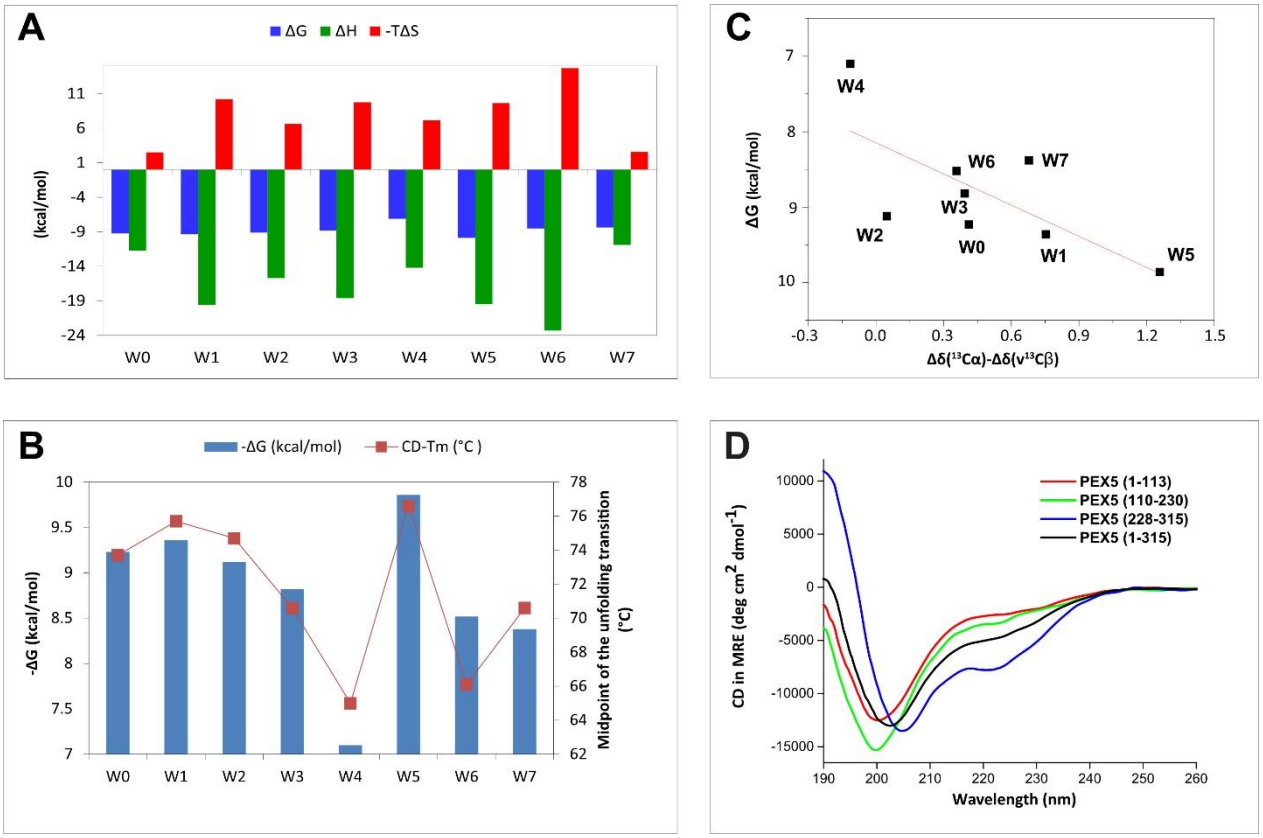

Figure 6.4: Thermodynamic and conformational features of the PEX14-NTD peptide interaction. **(A)** Bar Graph comparing thermodynamic parameters for the PEX14-NTD interaction with different (di)aromatic peptide ligands. Color codes are ΔG in blue, ΔH in green and -TΔS in red. **(B)** Correlation between free binding energy (-ΔG), calculated from the measured ITC data at 298 K (blue boxes) and thermal stability for the PEX14-NTD/peptide interactions are shown. Thermal unfolding was measured using CD at 222 nm, the transition midpoints (T_m) are shown as red squares. **(C)** Correlation between free binding energy (-ΔG) and average ^{13}C NMR secondary chemical shifts for the peptide motifs. The extend of positive NMR secondary chemical shift indicates increasing helical propensity. The correlation line is shown in red: y= 1.37x+8.14 with R^2 = 0.42. (D) CD spectra (mean residue ellipticities) of different regions and the complete PEX5-

NTD. A strong negative band below 200 nm indicates unstructured regions. The three 100-residue regions show different minor extent of helical conformations. The CD spectrum of the PEX5-NTD (1-315) shows some negative minima at 208 nm and 222 nm, consistent with the presence of partial helical folding.

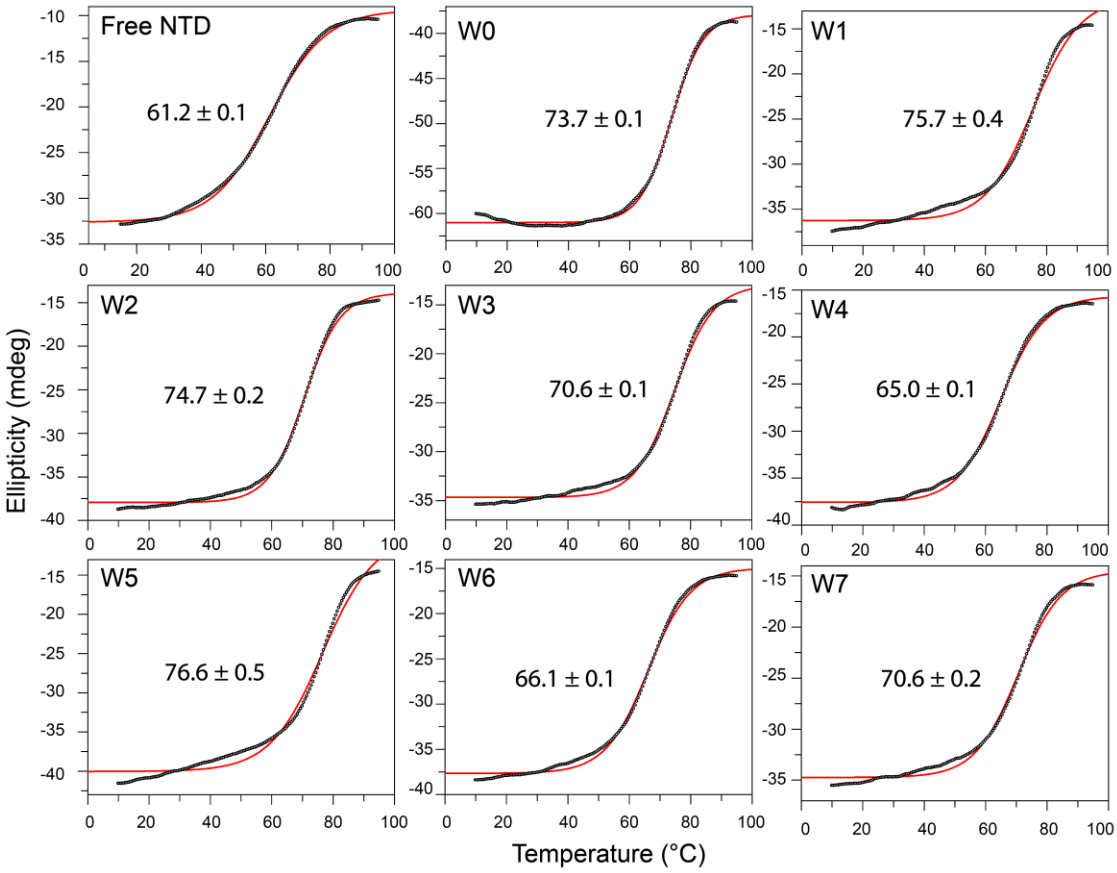

Figure 6.5: Circular dichroism (CD) thermal transition curves of the PEX14-NTD free and in complex with WxxxF/Y peptides. Ellipticity changes were measured at 222 nm from 10 ºC to 95 ºC at 1 ºC/min (black) with a concentration of 30 µM of NTD-PEX14 (free) and 60 µM of peptide motif (W0 corresponds to PEX5(1-113), W1 toW7 are 15-mer peptides, as indicated in **Figure 7.9.4B**). The complex was incubated an hour before the experiment. The solid red line is a sigmoidal fit to the collected CD data. The melting points (T_m) are indicated in °C along with the fitting error.

To assess conformational features of the PEX5 (di)aromatic motifs we used solution NMR and CD spectroscopy. Our CD spectra of the full PEX5-NTD (1-315) and the regions comprising residues 1-113, 110-230, 228-315 indicate a mostly disordered region with some α-helical propensity (**Figure 6.4D)**. These results are in agreement with our previous reported NMR analysis, where we identified defined □-helical propensities within the PEX5-NTD by analysis of ^{13}C secondary chemical shift (**Figure 6.6**) (Gaussmann *et al.*, 2021). Since these experiments did not cover the flanking sequence of W7 very well, we analyzed ^{13}C secondary shifts of a larger construct for this study

(**Figure 6.6G**). Notably, in the context of the full PEX5-NTD W5 motif exhibits the largest extent of α-helical conformation (**Table 6.2)**. There is clear correlation of the free energy of association (-ΔG) determined by ITC and helical propensity of the (di)aromatic motifs from average ^{13}C secondary chemical shift values **(Figure 6.4C)**. The weaker correlation observed for W6 and W7 may reflect additional contributions by neighboring residues outside the core motif. The good correlation between helical propensity observed by NMR and thermodynamics of binding determined by ITC also matches the stabilization effect of peptide binding to the PEX14-NTD indicated by thermal stability measurements (T_m) by CD.

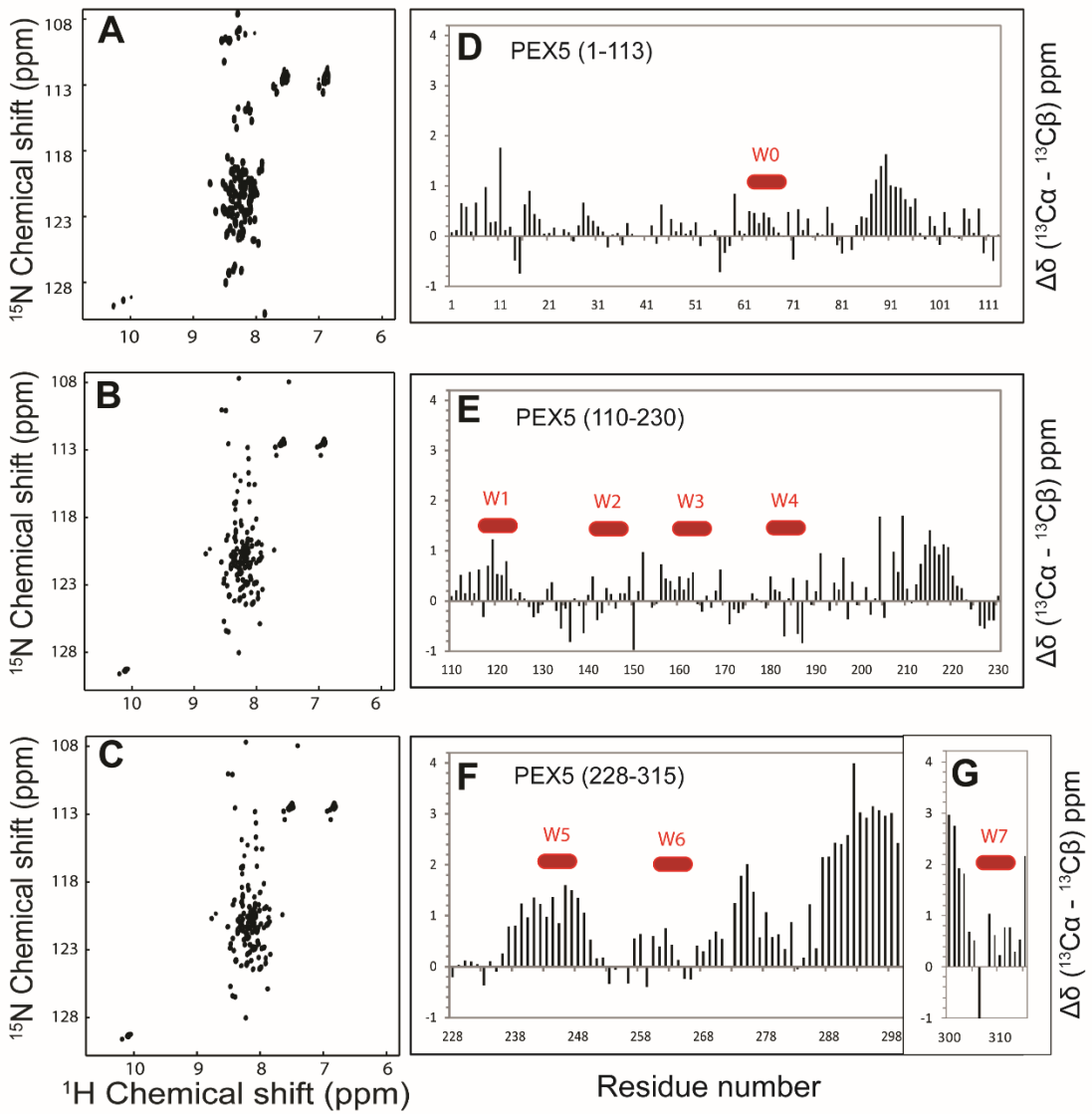

Figure 6.6: 2D ^{1}H-^{15}N-HSQC NMR spectra and secondary structure prediction of N-terminal 315 amino acids of PEX5 based on NMR chemical shift data. ^{1}H-^{15}N-HSQCs of PEX5 (1-113), PEX5 (110-230), PEX5 (228-315) shown in A, B, C respectively. Sharp and limited dispersion between 7-8.5 ppm indication of intrinsically disorder protein with residual secondary structure.

The difference between observed and random coil chemical shift values of ^{13}Ca and ^{13}Cb are plotted against amino acid sequence of Pex5 fragments (D-G). Higher positive value for residues 287 to 301 and 237 to 249 (W5 peptide region) indicates α helical structure. All other regions are designated as random coil or flexible regions of the polypeptide chain are shown on the right). Terminal amino acids are flexible in NMR. So to study about W7 motif (Plot G), residues 300 to 315 chemical shift values were measured from PEX5 (281-639) construct (unpublished data) and plotted.

Table 6.2: Comparison of experimental binding enthalpies and calculated interaction energies [a] Interaction energies were calculated for the core motif (6 residues), highlighted in bold. [b] Average ^{13}C secondary chemical shift values $\Delta\delta(^{13}C\alpha) - \Delta\delta(^{13}C\beta)$ for the 6 residues in the core motif.

Peptide	Sequence[a]	ΔH (kcal/mol)	Interaction Energy (kcal/mol)	Δδ ppm [b]
W1	ALSEN**WAQEFL**AAGD	-19.6	-198.98	0.754
W3	VSPAR**WAEEYL**EQSE	-18.6	-209.53	0.393
W5	AQAEQ**WAAEFI**QQQG	-19.5	-191.59	1.258
W6	GTSDA**WVDQFT**RPVN	-23.3	-281.28	0.357
W2	YNETD**WSQEFI**SEVT	-15.7	-157.04	0.048
W4	TATDR**WYDEYH**PEED	-14.2	-148.69	-0.114
W0	ASEDE**LVAEFL**QDQN	-11.7	-138.55	0.412
W7	AEAHP**WLSDYD**DLTS	-10.9	-104.22	0.678
W1_E4A	ALSEN**WAQAFL**AAGD	N/A	-132.92	N/A
W1_E4L	ALSEN**WAQLFL**AAGD	N/A	-174.97	N/A
W2_S2A	YNETD**WAQEFI**SEVT	N/A	-181.56	N/A
W2_S2L	YNETD**WLQEFI**SEVT	N/A	-180.60	N/A
W4_Y2L	TATDR**WLDEYH**PEED	N/A	-159.54	N/A
W7_D4E	AEAHP**WLSEYD**DLTS	N/A	-176.70	N/A
W7_D6L	AEAHP**WLSDYL**DLTS	N/A	-204.84	N/A

Computational analysis of the PEX14-NTD/PEX5 W0-W7 peptide interaction

To understand the contribution of structural features for the PEX14-NTD/PEX5 (di)aromatic peptide interactions, we performed 50 ns molecular dynamics simulations of 20-mer peptides comprising the W0 to W7 motifs bound to the PEX14-NTD. The simulations are based on the experimental structure

of the PEX14-W1 peptide complex. For the simulation of other peptides side chains were replaced by the corresponding residue with IRECS (Hartmann *et al*, 2007, 2009). The eight peptides bind to PEX14-NTD with different affinities, ranging from ΔG values of -9.89 kcal/mol to -7.10 kcal/mol determined by ITC (**Table 6.1**). To assess the importance of specific PEX14-NTD interaction pattern contributing to these values, we focused on the analysis of the enthalpic contribution ΔH (**Table 6.1**), as these values can directly be correlated to specific interactions and interaction energies. We then calculated the interaction energies for the five residues of the Wxxx(F/Y) core motif plus the following additional residue of the peptides based on optimized structures extracted from the last 10ns of MD simulations using the Dynadock program (Antes, 2010). For computational analysis of binding energies, only the five residues of the core motif plus the following residue were considered (**Figure 6.1B**). Flanking residues are included to consider structural features obtained from the molecular dynamics simulation. We hence denote residue numbers of the core peptide (p) motifs with the number A1(p)-A6(p) and the flanking residues with a number of the position relative to the core motif (…-2(p), -1(p), A1(p)-A6(p), A+1(p), …), i.e. the the W and F residues in the W1 peptide (WAQEFL) correspond to W1(p) and F5(p), respectively, (**Figure 6.1B**). Residues within PEX14 are annotated according to the human protein sequence.

Based on the binding enthalpies, ΔH, the eight peptides can be classified into three groups: W1, W3, W5, and W6 have strong, W2 and W4 intermediate, and W0 and W7 weak enthalpic contributions. Notably, the experimental binding enthalpies (ΔH) and calculated interaction energies correlate very well (**Table 6.2**). The W0 motif, which lacks the first aromatic residue shows small experimental binding enthalpy and calculated interaction energy. The W6 peptide exhibits the most favorable experimental enthalpy and consistently shows the best calculated interaction energy. Peptides with intermediate experimental enthalpy values have interaction energies around -150 kcal/mol and the calculated energies for the weakest binder W7 was determined to -100 kcal/mol. These results show that our computational simulation provides realistic peptide-PEX14 structures, which can be used for an in depth analysis of the binding modes.

A comprehensive analysis of the PEX14-peptide interaction (**Tables 6.3, 6.4, 6.5**) allows us to identify three crucial features important for binding to the PEX14-NTD, which are described as (i) a central hydrophobic core (**Figure 6.7B**), (ii) electrostatic interactions with K56 (**Figure 6.7C**) and (iii) electrostatic interactions with R40 and N38 (**Figure 6.7D**) as illustrated for the PEX14-W1 complex in **Figure 5.7**. In brief, the overall binding pattern is characterized by a central hydrophobic core region consisting of the bottom of the peptide binding groove in PEX14 and the corresponding counterpart residues in the bound peptide (**Figure 6.7A,B**). If this stable hydrophobic interaction pattern exists, a

strong intra-peptide backbone hydrogen bond between the flanking residue E-2(p) and residue Q3(p) of the core motif can be observed (**Figure 6.7D**), which stabilizes the helical conformation of the peptide (not shown). This hydrophobic core area of PEX14 is flanked at either side by two crucial charged residues, namely arginine 40 (R40) and lysine 56 (K56) (**Figure 6.7C,D**). These residues establish strong hydrophilic interactions between their positively charged side chains and negatively charged groups of neighboring peptide residues (**Figure 6.7C,D**) which effectively lock the peptide in the binding site. Moreover, R40 and K56 additionally contribute to the hydrophobic core with their aliphatic side chain. The three binding features are summarized in **Figure 6.7A** in a simplified cartoon representation. Structural details highlighting these interactions are shown in **Figure 6.7B-D**. **Figure 6.7B** illustrates the hydrophobic core region belonging either to the PEX14 binding groove (yellow) or the peptide (pink). It should be noted that both K56 and R40 contribute to this region via their aliphatic groups forming conserved hydrophobic clusters with W1(p) and L6(p)/F5(p), respectively. Thus, the hydrophobic area in PEX14 stretches over the whole binding site and forms complementary hydrophobic pockets for the corresponding peptide residues. **Figure 6.7C** shows the hydrophilic interaction pattern of K56, which forms a salt bridge with E4(p). The hydrophobic/hydrophilic interaction pattern of K56 is conserved in all peptides containing the W1(p) and E4(p) residues. The third important binding feature (**Figure 6.7D**) involves R40, which establishes hydrophilic interactions with E-2(p) and N38 through its terminal guanidino group, thus featuring a similar interaction pattern as K56. Overall, the two flanking hydrophilic interactions additionally stabilize the central hydrophobic region such that the peptide is effectively locked in its bound position. This may explain the large enthalpy contributions: after formation of the central hydrophobic core strong electrostatic interaction are established, which further stabilize the complexes by specific contacts.

Detailed analyses of the binding modes of the individual PEX14-peptide complexes are given in **Tables 6.3, 6.4** and **6.5**. In addition, representative structures from the last 10 ns of simulation (see experimental procedures) are shown in **Figure 6.8** and **Figure 6.9**. We find that the strong binding peptides, i.e. W1, W3, W5 and W6 exhibit all three interaction features (**Table 6.3**): (i) a stable central hydrophobic core, strong hydrophobic and hydrophilic interactions between the peptide and residues (ii) K56 and (iii) R40, respectively, leading to a strong intra-ligand -2(p)/3(p) backbone hydrogen bond (**Table 6.5**) and thus a very stable helical conformation of the bound peptide. In addition, R40 forms hydrogen bond networks with the core region of the peptide (W1, W5) or alternatively with the C-terminal residues (W6, W3), as shown in **Figure 6.7** and **Figure 6.9**.

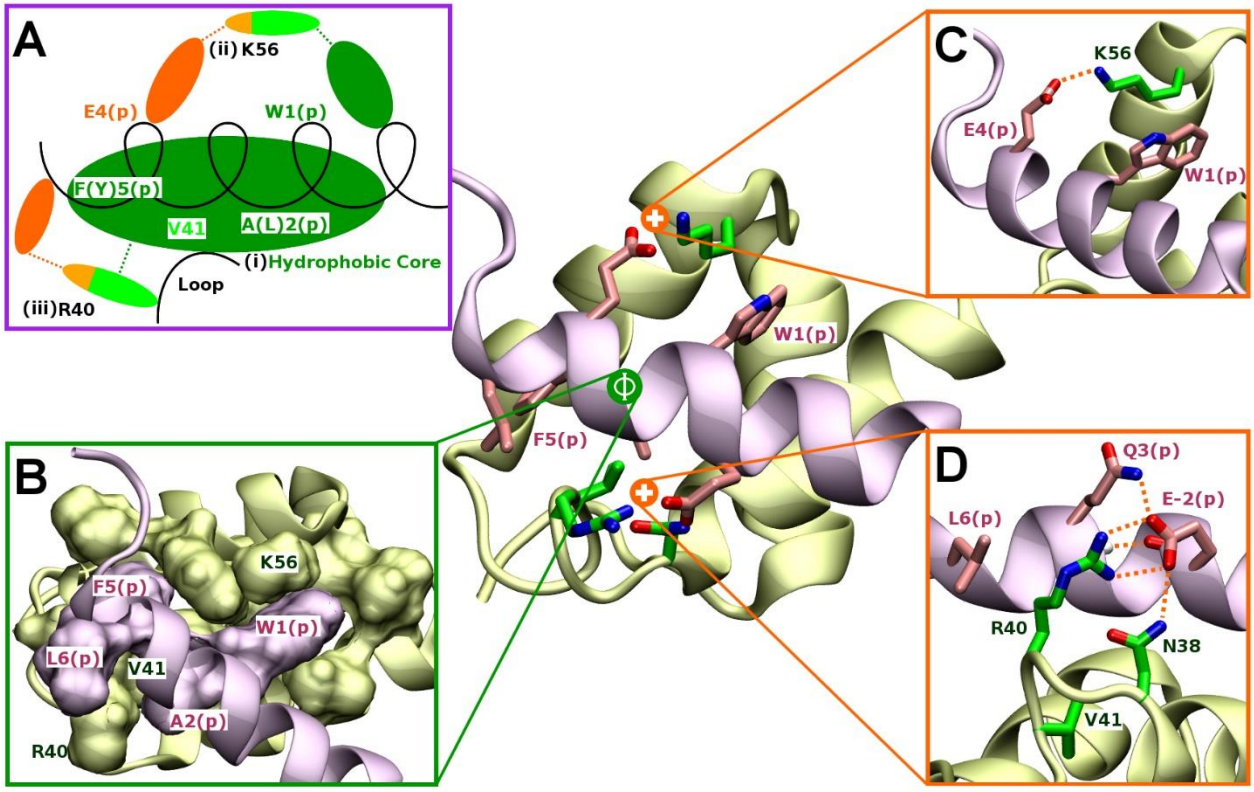

Figure 6.7: Binding details of the PEX14-NTD/PEX5 W1-peptide complex. The PEX14-NTD is shown in yellow/green, the PEX5 peptide in pink. Important residues are given in licorice representation. The central picture shows the most prominent structure of W1 during the last 10ns of MD simulation. **(A)** Schematic sketch of the binding pattern of the PEX14-NTD/PEX5 complex illustrating the three binding features (i), (ii) and (iii). **(B)** Surface representation of the residues contributing to the hydrophobic core. **(C)** Hydrophilic interactions of K56. **(D)** Hydrophilic interactions of R40.

Analysis of the complexes with intermediate and weak binding peptides (**Table 6.3** and **Figure 6.8**) reveals that some of the features observed for the strongly binding peptides are missing. In the W2 and W4 peptide complexes the central hydrophobic core is disrupted by residue 2(p), either S or Y, substituting for the hydrophobic residues A or L. This residue is located right at the center of the hydrophobic core (**Tables 6.3, 6.4**; **Figure 6.8A, B**) and thus crucial for optimal packing. Residues in the flanking regions can also affect the binding affinity. The -2(p) mutation to T in W2 abolishes the favored R40 and N38 interactions seen with W1 **(Figure 6.7D)** while the H mutation at position 6(p) in W4 causes a charge clash with R40 which destabilizes the binding (**Figure 6.8B**).

Table 6.3: Binding site features. [a] X= the interaction exists, if the residue is not E4(p) or W1(p), respectively, the residue type is given in parenthesis. Bold letters/numbers: variation from the optimal binding pattern.

[b] Number of residues with which R40 forms hydrophilic (Hphil) or hydrophobic (Hphob) interactions, as weak, fluctuating interactions with F/Y5(p) are observed in all simulations, they are not considered in the table.

	Interactions of K56 [a]		Interactions of R40 [b]		Central hydrophobic cluster	Residue +1	Interaction energy (kcal/mol)
Peptide	E4(p)-K56 salt bridge	W1(p)/K56 Hphob/arom	H phil	H phob	# residues	Amino acid	
W1	x	x	2	2	3	L	-198.98
W3	x	x	1	1	3	L	-209.53
W5	x	x	1	1	3	I	-191.59
W6	x (Q)	x	1	2	3	T	-281.28
W2	x	x	-	1	2	I	-157.04
W4	x	x	2	2	2	H	-148.69
W0	x	x (L)	-	2	3	L	-138.55
W7	x (D)	x	1	-	3	D	-104.22
W1_E4A	-	x	2	1	3	L	-132.92
W1_E4L	-	x	2	1	3	L	-174.97
W2_S2A	x	x	-	1	3	I	-181.56
W2_S2L	x	x	-	2	3	I	-180.60
W4_Y2L	x	x	2	2	3	H	-159.54
W7_D4E	x	x	2	-	3	D	-176.70
W7_D6L	x (D)	x	2	1	3	L	-204.84

The effects of these amino acid substitutions on the overall stability of the complexes are summarized in **Table 6.5.** The percentage of simulation time within the last 10 ns of the MD simulations during which either the 4(p)/K56 salt bridge or the -2(p)/3(p) backbone hydrogen bond exist, serves as a measure for the stability of the corresponding interactions and, as explained above, the stability of the helical peptide conformation in the complex. The E4(p)/K56 salt bridge exists during 90% of the simulation time (i.e. value of 0.9) for the W1 peptide, but only during 38% of the time for peptide W2. This is correlated to the presence of the 2(p)/3(p) backbone hydrogen bond, which is significantly

weakened in W2 (0.58) compared to the strong binding peptides (>0.8). In addition, the peptide interactions of R40 are weak for peptide W2. This is due to strong conformational fluctuations of the guanidino-group, which in W2 can alternatively interact with the backbone carbonyl oxygen of residue 2(p). These fluctuations lead to an additional disturbance of the already weakened central hydrophobic core (**Figure 6.8A**). The same trend holds for W4, due to the disruption of the hydrophobic core by Tyr. As this residue is also too large to fit into its binding pocket, the bulky tyrosine side chain is turned towards to solvent and clashes with the R40 side chain (**Figure 6.8B**). In addition, the side chain of R40 is stabilized by π-stacking interactions with Y2(p) and H6(p). Although this should stabilize the bound complex, it leads to weaker hydrophobic interactions due to imperfect shape complementary (i.e. a gap is introduced between peptide and protein surfaces, data not shown), as well as extra electrostatic repulsions between H6(p) and R40. Therefore, in both the W2 and W4 complexes, the central hydrophobic binding core is disrupted, which leads to an additional destabilization of the R40 peptide interactions.

The W0 motif (LVxEF) is distinct from W1-W7 in that it lacks the W1(p) tryptophan (the most conserved residue of the motif), which is replaced by a leucine. Nevertheless, W0 is one of the strongest binders in terms of ΔG but not considering ΔH (**Table 6.1**). The less favorable ΔH likely reflects that the replacement of W by L reduces the hydrophobic contact surface. Yet, most of the key features required for a strong interaction as described above are present in W0. Hydrophobic interactions with R40 and K56 as well as the E4(p)/K56 salt bridge and -2(p)/3(p) H-bond exist (**Table 6.3**). In the last 10ns of MD simulation the E4(p)/K56 salt bridge and -2(p)/3(p) H-bond are present for 99% and 89% of the time respectively (**Table 6.5**). These results are comparable with W1 showing similar affinities (**Table 6.1**). When bound to the PEX14-NTD the W0 peptide undergoes conformational fluctuations as it is less well packed, consistent with the reduced helicity observed for the W0 peptide featuring only two helical turns (other ligands have four) and the lack of the bulky tryptophan side chain and incomplete electrostatic clamping, suggesting conformational entropy compensation. This is likely also reflected in the fast off-rate (k_{off}) observed in SPR experiments (Neuhaus *et al.*, 2014).

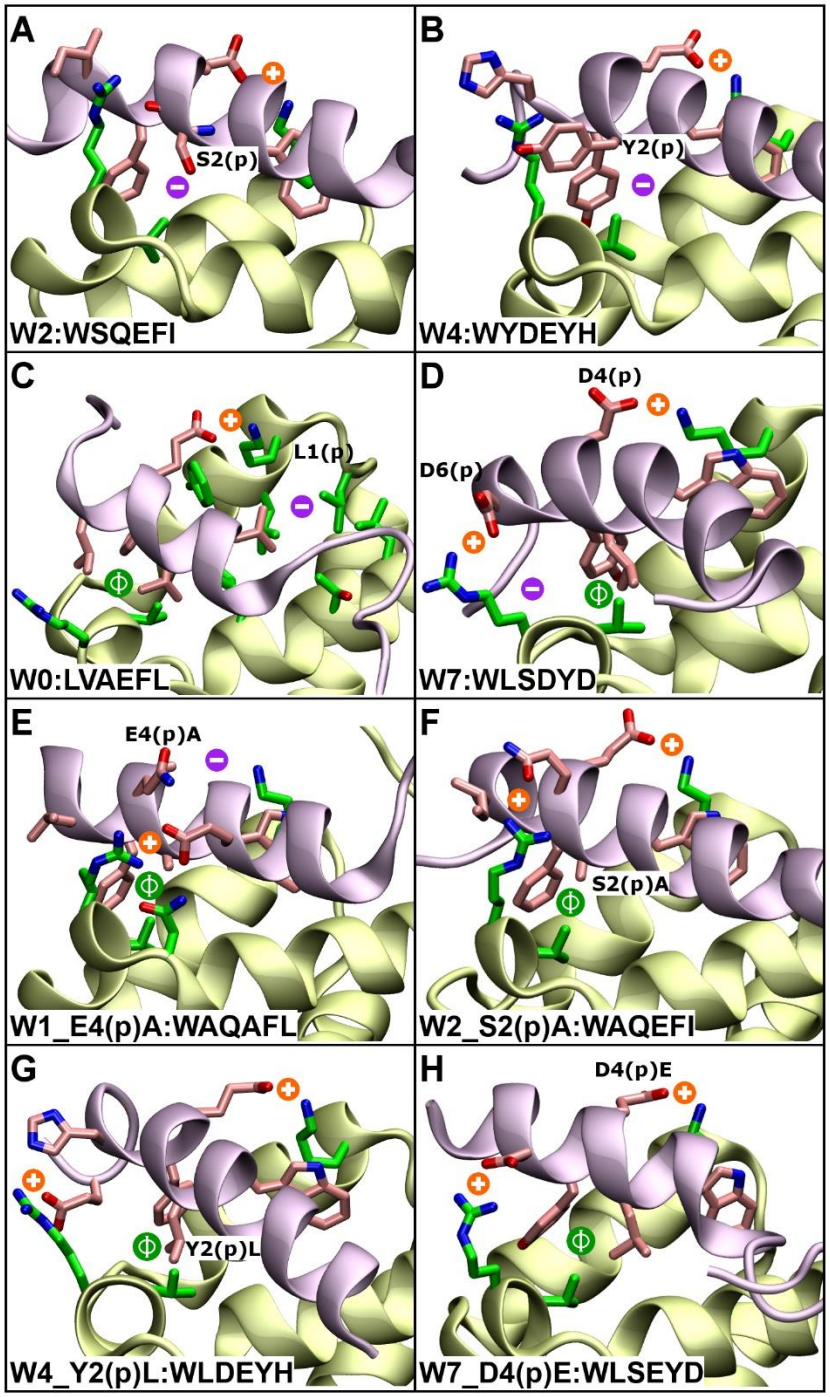

Figure 6.8. Molecular dynamics simulations. Most prominent structures observed during the last 10 ns of the MD simulations of wild type and mutated PEX14-NTD/PEX5 peptide complexes are shown. The PEX14-NTD and PEX5 peptides are shown in yellow/green and pink, respectively. Hydrophobic interactions are highlighted by green circles, charged and polar interactions are circled in orange. Weak salt bridges are indicated with dotted lines. Interaction features that are missing are indicated by a "–" sign with purple background. (A) W2, (B) W4, (C) W0, (D) W7, (E) W1_E14(p)A, (F) W2_S2(p)A, (G) W4_Y2(p)L, (H) W7_D4(p)E.

In the W7 peptide E4(p), which normally forms the essential salt bridge to K56 is mutated to D4(p). The shorter side chain does not provide an optimal length for hydrogen bond formation with K56 and leads to strong fluctuating movements of the K56 side chain as it tries to adapt to the larger distance towards D4(p). This effect is enhanced by D6(p), as the terminal guanidino group of R40 tries to form hydrogen bonds with D6(p), which is sterically not possible, thus leading to flipping R40 conformations (**Figure 6.8D**). During the MD simulations, both movements of R40 as well as K56 cause strong fluctuations in the PEX14 backbone in the neighboring binding site region, which are not observed for the other peptides (data not shown). Through these fluctuations the central hydrophobic core is weakened, as seen by the lower percentage of occurrence of the -2(p)/3(p) H-bond in W7 (0.71) (**Table 6.5**).

Table 6.4: Binding site features. Detailed binding site features, showing the amino acids presented in Table 3.

Peptide	Interactions of K56		Interactions of R40/N38		Central hydrophobic cluster
	E4(p)-K56 salt bridge	W1(p)/K56 Hphob/ aromatic	Hphil	Hphob	Residues
W6	Q4(p)	W1(p)	P+2(p), T6(p)	T6(p)	F5(p), V41, V2(p)
W1	E4(p)	W1(p)	E-2(p), N38	A2(p), L6(p)	F5(p), V41, A2(p)
W5	E4(p)	W1(p)	E-2(p)	I6(p)	F5(p), V41, A2(p)
W3	E4(p)	W1(p)	E+4(p)	L6(p)	Y5(p), V41, A2(p)
W2	E4(p)	W1(p)	-	I16(p)	F5(p), V41
W4	E4(p)	W1(p)	Y2(p), D+4(p)	Y2(p), H6(p)	Y5(p), V41
W0	E4(p)	L1(p)	-	V2(p), L6(p)	F1(p), V41, V1(p)
W7	D4(p)	W1(p)	D6(p)	-	Y1(p), V41, L2(p)
W1_E4A	-	W1(p)	E-2(p), N38	L6(p)	F5(p), V41, A2(p)
W1_E4L	-	W1(p)	E-2(p), N38	L6(p)	F5(p), V41, A2(p)
W2_S2A	E4(p)	W1(p)	-	I6(p)	F5(p), V41, A2(p)
W2_S2L	E4(p)	W1(p)	-	L2(p), I6(p)	F5(p), V41, L2(p)
W4_Y2L	E4(p)	W1(p)	E+2(p), E+3(p)	H6(p),L2(p), Y5(p)	Y5(p), V41, L2(p)
W7_D4E	E4(p)	W1(p)	D6(p), T+3(p)	-	Y5(p), V41, L2(p)
W7_D6L	D4(p)	W1(p)	T+3(p),S+4(p)	L6(p)	Y5(p), V41, L2(p)

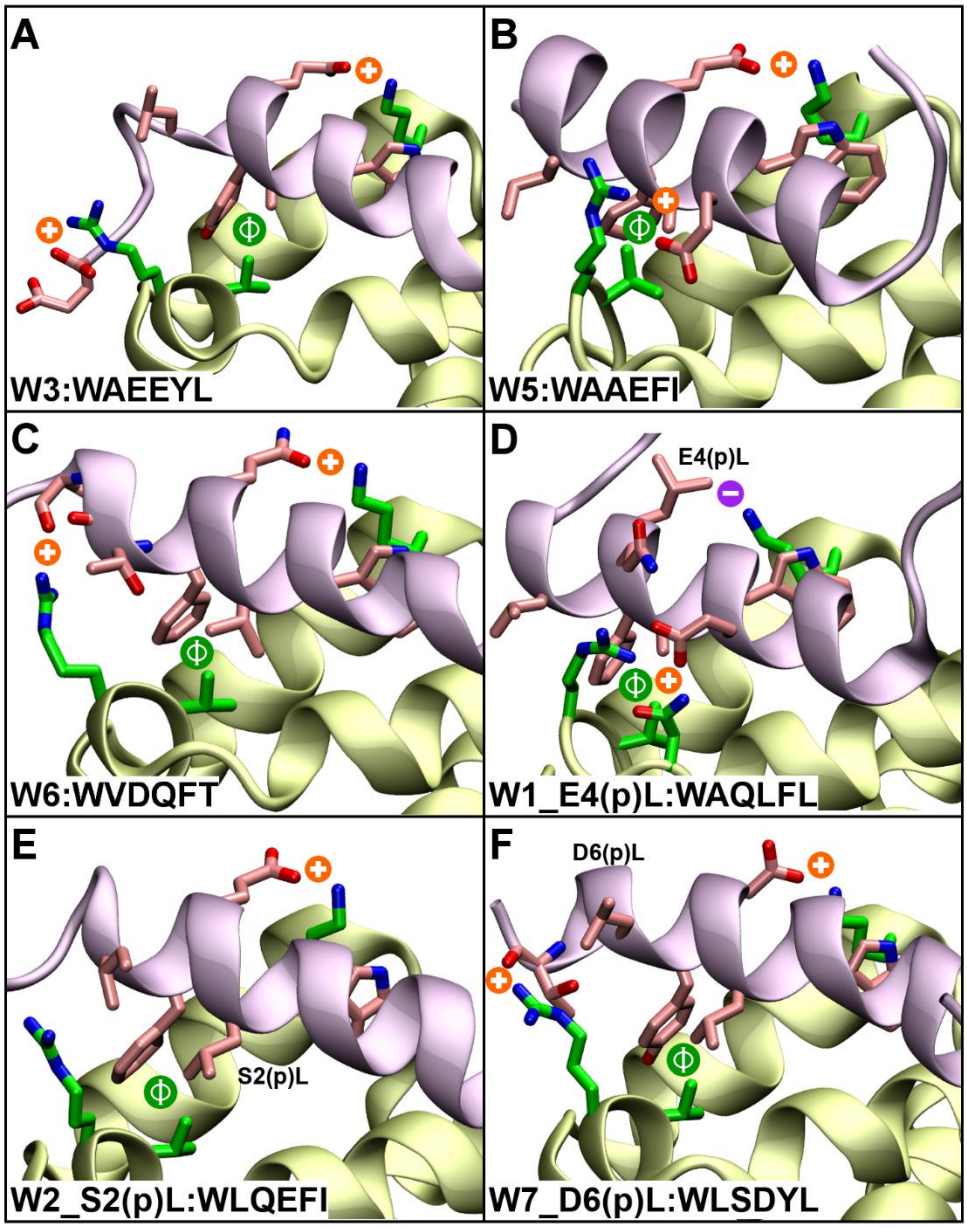

Figure 6.9: Representative structures observed during the last 10ns of MD simulations of wild type and mutated PEX14-NTD/PEX5 peptide complexes. The PEX14-NTD is shown in yellow/green and PEX5 in pink. Important residues are given in licorice representation. Hydrophilic interactions are highlighted by a red circle and hydrophobic clusters by a blue circle. The negative sign and corresponding label mean the absence of an important structural feature. (A) W3. (B) W5. (C) W6. (D) W1_E14(p)L. (E) W2_S12(p)L. (F) W7_D16(p)L.

Table 6.5. Population of hydrogen bonds or salt bridges during the simulation. The population is calculated as the fraction of time frames during the last 10 ns of the MD simulations in which the hydrogen bond and salt bridge exist, i.e. the fraction ranging from 0 to 1 corresponds to 0-100%.

Peptide	4(p)-K56 salt bridge*	-2(p) – 3(p) H-bond*
W1	0.90	0.80
W3	0.92	0.80
W5	1.06	0.91
W6	0.73	0.85
W2	0.38	0.58
W4	1.02	0.43
W0	0.99	0.89
W7	0.63	0.71
W1_E4A	-	0.94
W1_E4L	-	0.89
W2_S2A	0.51	0.43
W2_S2L	0.49	0.74
W4_Y2L	0.79	0.45
W7_D4E	0.78	0.80
W7_D6L	0.43	0.73

In silico mutational analysis to identify sequence requirements for PEX14 binding

Based on the analysis of the simulation results we carried out *in silico* mutational studies for four peptides (W1, W2, W4, and W7), performing the same type of simulations as for the natural peptides for an overall of seven variants each featuring one single "strategic" mutation within the peptide. We designed two "failure" and five "rescue" mutations. The "failure" mutations are based on W1 and were designed to eliminate the important E4(p)/K56 salt bridge (and thus destabilize the interaction). For both variants **W1_E4(p)A** and **W1_E4(p)L**, the interaction energies indeed decrease considerably in both cases leading to much weaker binding (**Table 6.2**) and a partial opening of the binding pocket (compare **Figure 6.8E** and **Figure 6.9D** with **Figure 6.7**). These results confirm the importance of the E4(p)/K56 salt bridge.

The five "rescue" mutations were designed to improve peptide-PEX14 interactions of W2, W4 and W7. Replacing S2(p) in W2 by Ala or Leu, (**W2_S2(p)A** and **W2_S2(p)L)** restore the central hydrophobic core region. As expected, these mutations lead to stronger interaction energies (**Table 6.2**) as well as

an increase in the percentage of occurrence of the E4(p)/K56 salt bridge. In the case of **W2_S2(p)L** also the -2(p)/3(p) hydrogen bond is stabilized (**Table 6.5**). The binding pattern of R40 is altered in both mutants, and R40 now forms stable interactions with the side chains of the central peptide residues, resembling the binding pattern of W1 and W5 (**Figure 6.7, 4F** and **Figure 6.9B**). In the variant **W2_S2(p)L**, the aliphatic side chain of R40 contributes to the hydrophobic core (**Figure 6.9E**).

Next, we mutated Y2(p) to Leu in W4 generating W**4_Y2(p)L**, which also leads to a stabilization of the central hydrophobic core and thus of the E4(p)/K56 salt bridge. In addition, a more stable interaction pattern of R40 with the C-terminal end of the peptide is observed, resembling the binding pattern of W6 and W3 (**Figure 6.8G** and **Figure 6.9A,C**). The two variants of W7 replacing D4(p) by Glu (**W7_D4(p)E**) and D6(p) by Leu (**W7_D6(p)L**), respectively, show increased binding energies, a stabilizing of the E4(p)/K56 salt bridge (**Table 6.5**), and of hydrophobic interactions with R40 (**Table 6.3)**. The R40/6(p) and K56-4(p) interactions for the D4(p)E variant are shown in (**Figure 6.8H**). Due to the longer side chain of the residue at position 4(p), a stable interaction pattern with K56 can be formed (compare to **Figure 5.8D**).

Analysis of additional (di)aromatic ligands in peptide overlay binding assays

To validate the general recognition features of PEX14 binding derived from the analysis above, we synthesized 80 15-mer peptides representing Wxxx(F/Y)-containing naturally occurring fragments of PEX5 proteins of various organisms. The immobilized Wxxx(F/Y)-containing peptides were analyzed by overlay incubation with purified human PEX14-NTD followed by antibody detection (**Figure 6.10**). The intensity of the staining roughly correlates with the dissociation constants as indicated for the Wxxx(F/Y) motifs W1 to W7 of human PEX5 (**Figure 6.10A**). In accordance with results obtained by ITC (**Table 6.1**), W1 and W5 gave the strongest signals, whereas W7 and W4 were not detectable.

Taken together, among the 80 tested peptides, we identified 33 binding peptides (**Figure 6.10B, Table 6.6**), from which 12 showed a strong interaction with human PEX14-NTD (**Figure 6.10C, Table 6.6**). Interestingly, the relative number of diaromatic peptide motifs that interact with PEX14 in the peptide-overlay assay, is higher in plants and animals than in fungi and protists. For instance, 5 out of 7 motifs of human PEX5 show clear PEX14 binding, whereas all tested yeast and protists PTS1-receptors contain no more than one (di)aromatic peptide interacting with the human protein. This observation is also consistent with our ITC experiments which show binding of ScPex14 to a single motif in ScPEX5-NTD with a 1:1 stoichiometry (**Figure 6.3**) and other previous reports on Pex14-Pex5 interactions (Cyr *et al.*, 2008; Hojjat & Jardim, 2015; Watanabe *et al.*, 2016).

Table 6.6: Binding of (di)aromatic peptide motifs the human PEX14-NTD in peptide overlay binding assay, with motifs that show good binding affinity shown on the right column.

All binders		**Good binders**	
Motif	**Sequence**	**Motif**	**Sequence**
Hs1	ALSEN**WAQEFL**AAGD	Hs1	ALSEN**WAQEFL**AAGD
Hs2	YNETD**WSQEF**ISEVT	Hs5	AQAEQ**WAAEFI**QQQG
Hs3	VSPAR**WAEEY**LEQSE	Gal1	ALSEN**WTQEFL**AAAD
Hs5	AQAEQ**WAAEFI**QQQG	Gal3	VSPAK**WAEEYL**EQSE
Hs6	GTSDA**WVDQFT**RPVN	Gal4	SLADK**WYEEYQ**PEDD
Xt1	ALSEQ**WSAEFV**GAEV	Gal5	DQAEQ**WATEFM**QQQG
Xt3	SSPSH**WAEEYL**QQSE	Dr1	ALSGD**WASEFL**STAD
Xt4	ALAEK**WTEEYQ**PEDE	Dr2	PADAD**WTREFI**NEVA
Gal1	ALSEN**WTQEFL**AAAD	Dr3	ADPGR**WAEEYL**EQSE
Gal2	YNEAD**WSQEFI**AEVT	Dr5	ESAES**WVDEFA**TYGP
Gal3	VSPAK**WAEEYL**EQSE	Apis4	AVAGN**WIDEFQ**KENV
Gal4	SLADK**WYEEYQ**PEDD	At7	TAEDQ**WVNEFS**KLNV
Gal5	DQAEQ**WATEFM**QQQG		
Dr1	ALSGD**WASEFL**STAD		
Dr2	PADAD**WTREFI**NEVA		
Dr3	ADPGR**WAEEYL**EQSE		
Dr5	ESAES**WVDEFA**TYGP		
Ce3	TDAGM**WSSEYL**DTVD		
Ag4	EGKHP**WLSEFN**DFYD		
Dm2	ESLDD**WISDYQ**RSTE		
Dm3	ENEHP**WLSEYN**DNMD		
Apis2	GLGPK**WAEEYI**EHSI		
Apis4	AVAGN**WIDEFQ**KENV		
Apis5	SSKHP**WLSEYD**KFYD		
Tb3	ADVEQ**WAQEYA**QMQA		
Ld3	NDVED**WAQEYA**EMQE		
At4	SAPGE**WATEYE**QQYL		
At6	HGPEQ**WADEFA**SGRG		
At7	TAEDQ**WVNEFS**KLNV		
At8	LNVDD**WIDEFA**EGPV		
At9	SSADA**WANAYD**EFLN		
Sc1	NDISH**WSQEFQ**GSNS		
Yl3	LDEQN**WEEQFK**QMDS		

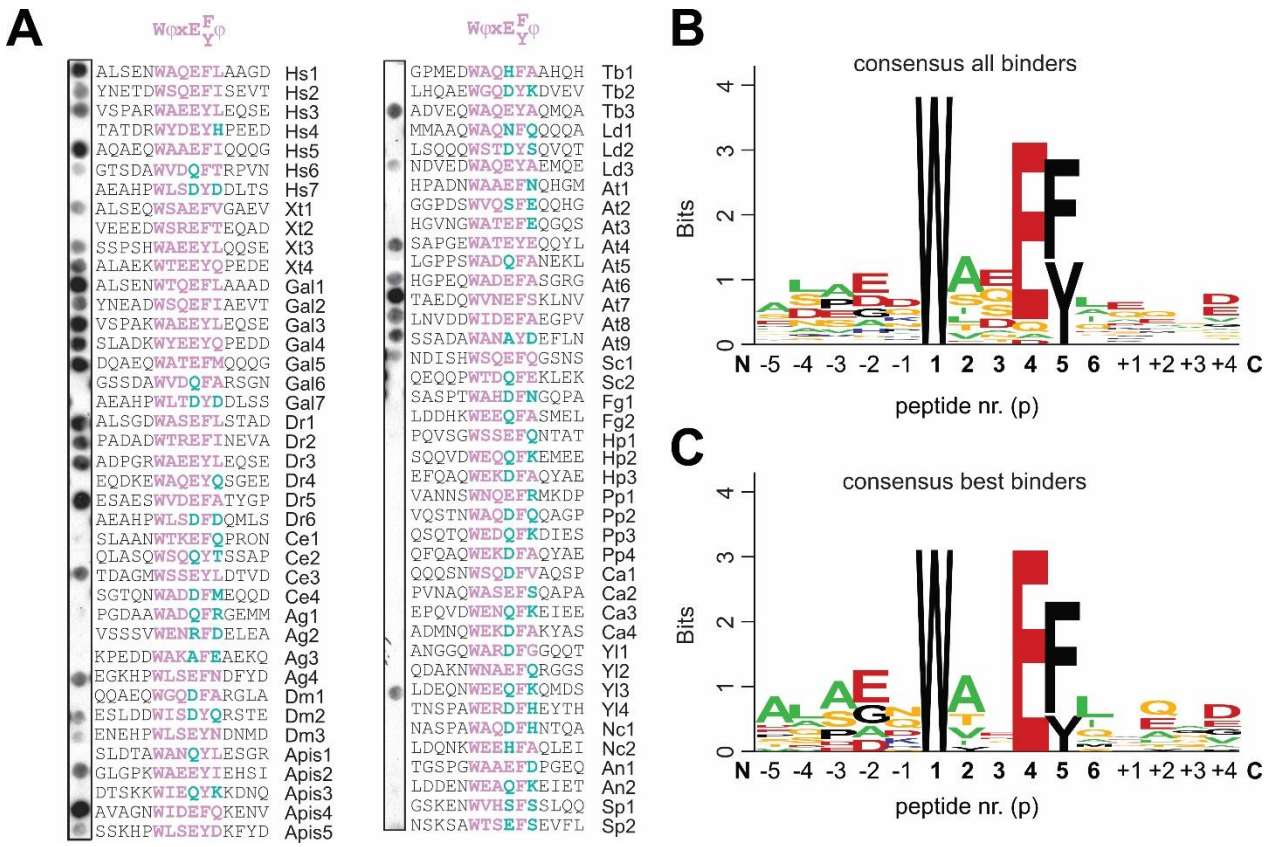

Figure 6.10: Peptide overlay binding assay with PEX14-NTD for various diaromatic peptide motifs and sequence conservation analysis of all PEX14-NTD binding motifs. **(A)** Peptide spot overlay assay with His-tagged PEX14-NTD and immobilized peptides representing Wxxx(F/Y) motifs of PEX5 of different species. Each peptide comprised 15 amino acids with a central Wxxx(F/Y) and five adjacent amino acids at each side. Bound PEX14 was visualized immunochemically with monoclonal anti-His_6 antibodies. Spots with reduced intensities indicate reduced binding affinity for PEX14. The number of Wxxx(F/Y) motifs varies within the sequences of PEX5 proteins, indicated by species name as abbreviation and position of Wxxx(F/Y) motif starting with the N-terminal W1. **(B)** Sequence logo representation of all binding and (C) the strongest PEX14-NTD binding motifs found. The total height (in Bits) of the stack indicates the degree of sequence conservation at the corresponding position and the height of each letter is proportional to its frequency at that position. The logo was generated using Berkeley's WebLogo program (Crooks *et al*, 2004). *Hs: Homo sapiens, Xt: Xenopus laevis, Gal: Gallus gallus, Dr: Danio rerio, Ce: Caenorhabditis elegans, Ag: Ashbya gossypii, Dm: Drosophila melanogaster, Apis: Apis apis, Tb: Trypanosoma brucei, Ld: Leishmania donovani, At: Arabidopsis thaliana, Sc: Saccharomyces cerevisiae, Fg: Fusarium graminearum, Hp: Hansenula polymorpha, Pp: Pichia pastoris, Ca:Cavia porcellus, Yl: Yarrowia lipolytica, Nc: Neurospora crassa, An: Aspergillus nidulans, Sp:Schizosaccharomyces pombe.*

Analysis of the peptide binding (**Figure 6.10B,C**) confirms that the predominant residue type at position 2(p) is alanine, while at position 6(p) mainly leucine is found, followed by alanine and isoleucine underlining the role of a hydrophobic residue. At position 4(p) almost all PEX14 binding peptides (n=29) exhibit a glutamic acid side-chain whereas aspartic acid at position 4(p) is the most

abundant amino acid among the non-binding peptides. Note, that the predominant residue outside of the core-motif in position -2(p) is glutamic acid. These data support the proposed recognition features that we identified in our analysis, i.e. the requirement of a stable central hydrophobic core region and stabilizing salt-bridges between the peptide and the residues K56 and R40 for high affinity binding to PEX14.

However, there are also few interacting peptides, which do not completely fulfill the criteria identified in our analysis. For example, in *Gallus gallus* (Gal) strong interactions are observed for Gal1 and Gal4. The sequence of Gal1 is almost identical with the human W1 with a mutation A2(p) to T2(p), which seems not to disturb the overall stability of the binding. Although threonine is a polar residue, the terminal of the side chain is a methyl group. Our simulations reveal that the residue at position 2(p) tends to use the terminal of the side chain to interact with V41 forming the hydrophobic core (**Figure 6.8** and **Figure 6.9**). Thus, we speculate that threonine at position 2(p) can also form the hydrophobic core and stabilize the binding compared to other polar residues, such as serine (**Figure 6B,C**). However, the analogue W4 motif Gal4 is less conserved in the flanking region. Especially, the substitution of H6(p) to Q6(p) changes charge and size specifies this position as unfavorable for large charged amino acids. In fact, none of the good binders harbors a R, K or H in the 6(p) position.

Discussion

PEX14 peptide ligands were originally defined as diaromatic pentapeptides with the consensus sequence Wxxx(F/Y) (Otera *et al.*, 2002; Saidowsky *et al.*, 2001). During the last years, various additional ligands with minor modifications of the consensus sequence were identified with a remarkable variability with respect to affinity and specificity. Here, we present a comprehensive integrated experimental and computational analysis of the (di)aromatic peptide ligands focusing on the human PEX5/PEX14 system. Our analysis allows us to refine the consensus sequence to better predict binding potential of the (di)aromatic peptide ligands.

We identified three key features that are favorable for binding to the human PEX14-NTD: (i) a stable central hydrophobic core, (ii) electrostatic interactions with K56 and (iii) interactions with residues R40 and N38 in PEX14. Based on these findings and the analysis the additional peptides from our peptide overlay binding experiments we propose a refined motif as **WΦxE(F/Y)Φ**, where Φ corresponds to a hydrophobic residue (in position 2, Φ is an aliphatic hydrophobic side chain), while x is variable. Interestingly, W0 (**LVAEFLQ**) lacks the tryptophan and thus does not fit the consensus sequence but nevertheless binds PEX14 with very high affinity (**Table 5.1**). This arises from a good fit to the PEX14-

binding surface, which still allows for a large favorable enthalpic contribution, mostly driven by hydrophobic interactions. This interaction comes with a reduced entropic binding penalty, which finally gives rise to a favorable free Gibb's energy (**Figure 6.4A).** The absence of the tryptophan side chain being recognized in a deeper binding pocket renders the complex somewhat more dynamic consistent with the fast off-rate observed by SPR (Neuhaus *et al.*, 2014). Peptide motifs similar to W0 have been reported to bind with μM affinity as discussed below.

An interesting question is whether the consensus sequence based on the interaction between human PEX14-NTD and Wxxx(F/Y) motifs allows discriminating between weak and strong binding ligands in other species. Although most important residues such as K56 are highly conserved in PEX14 from all species (**Figure 6.1**), other residues, like R40, which contribute to a minor extent to the interaction in human are not conserved. In some other animals and plants, we find a lysine instead, and in yeast and fungi usually a serine or threonine.

The PEX5 N-terminal domain of *Saccharomyces cerevisiae* harbors two regular (di)aromatic motifs: ScW1: SH**WSQEFQ**G and ScW2: QP**WTDQFE**K. Underlined residues indicate less preferred amino acids according to our analysis. Both motifs were found not to interact with ScPex14p-NTD in two-hybrid assays (**Kerssen et al., 2006**). In our peptide-scan (**Figure 6.10**), we observe weak binding of ScW1 whereas ScW2 does not interact with the human PEX14-NTD, indicating a variation of binding site features between the species. This assumption is supported by the PEX14 sequence alignment (**Figure 6.1**) showing that N38 and R40 are not conserved from yeast to human. We find a similar sequence for ScW1 compared to HsW2, which exhibits weak PEX14-NTD binding as. However, ScW2 harbors a Q in -2(p) position where an E is preferred, a glutamate in position E6(p) where an aliphatic residue is preferred, and a proline in the -1(p) position, which disfavors a helical conformation. Here, we have shown that a reverse motif with the sequence S**DFQEVW**DS in ScPex5 mediates binding to ScPEX14-NTD (**Figure 6.3)**. This motif is conserved among yeast and partly matches the consensus sequence when inverted (i.e. SD**WVEQFD**). The micromolar affinity of this interaction may reflect a non-optimal sequence missing the E in position 4(p) and a charged (instead of hydrophobic) residue in position 6(p), when considering an inverted binding directionality. However, the binding directionality has not been experimentally shown.

In *Trypanosoma brucei*, a similar situation compared to yeast is observed. TbPex5 contains three Wxxx(F/Y) motifs in the N-terminal region, of which only the third motif has been reported to bind TbPex14 with high affinity (K_D = 0.68 μM) employing pull down assays and SPR analysis (Watanabe *et al.*, 2016). This observation is again consistent with our results from the peptide-scan (**Figure 6.10A)** and the conservation of the PEX14-NTD **(Figure 6.1)**. The first motif (TbW1: ED**WAQHFA**A) has a

histidine at the E4(p) position, while the second motif (TbW2: AE**WGQDYK**D) has unfavorable residues at positions -2(p), 1(p) and 6(p) in the peptide while the third motif (TbW3: EQ**WAQEYA**Q) fulfills all stated criteria. Although, this interaction is conserved from human to trypanosomatids, the PEX5-PEX14 interaction in trypanosomal organisms is of special interest as inhibiting this interaction opens novel therapeutic concepts for drug discovery against devastating diseases, such African sleeping sickness, Chagas or leishmaniosis. As the interaction of TbPex5 and TbPex14 is essential for protein import into glycosomes, a specialized parasite-specific form of peroxisomes **(Choe *et al*, 2003)**, interfering with this interaction provides an efficient therapeutic route against trypanosomatid parasites. In fact, the validity of this approach and a proof-of-concept has been recently demonstrated using a structure-based drug discovery approach (Dawidowski *et al.*, 2017; Dawidowski *et al.*, 2020).

The revised definition of a consensus for PEX14 interaction motifs is valuable to improve the prediction of peptide ligands. However, the binding capability of PEX14-NTD is not strictly limited to motifs found in PEX5. Additional interactions are known that exhibit often much lower binding affinities in the μM range. For example, the PEX14-NTD binding motif in PEX19 with the sequence EK**FFQELF**DS has been reported to bind with a K_D of 9 μM. Interestingly, this motif binds in opposite directionality compared to human consensus motifs (Neufeld *et al.*, 2009). When the inverted motif is aligned to the consensus, the key residues **(ϕ)** in 2(p) and 6(p) positions are maintained. However, positions -2(p) and 4(p) have unfavorable residues. In addition, the helix dipole is inverted, which may contribute to the differences in binding affinity. Recent studies have demonstrated that PEX14 is responsible for peroxisomal motility by interaction with β-tubulin which was mapped to PEX14-NTD binding two (di)aromatic peptide motifs. Here, motif 1 (K**AFLHWY**TG) binding with ~280 μM exhibits unfavorable residues in position -2(p) and 4(p) whereas motif 2 (ND**LVSEYQ**Q) is more similar to human PEX5 W0 and shows higher binding affinity (K_D = 5 μM) (Bharti *et al.*, 2011; Reuter *et al.*, 2021). Apart from peroxisomes, other FxxxF-like motif interactions are known from the NFAT and mPer families towards CK1 regulating nuclear import (Okamura *et al*, 2004) suggesting that this type of motifs also plays a role biological contexts to mediate other protein-protein interactions.

There is emerging evidence that the (di)aromatic motifs play an important role in many aspects of peroxisome biogenesis, as key factors, such as PEX5, PEX13, PEX14, PEX19 are regulated by protein-protein interactions involving (di)aromatic motifs. It is tempting to speculate that distinct specificities and affinities of the individual motifs to recognition domains, such as the PEX14-NTD or the PEX13-SH3 domain could contribute a balanced and fine-tuned regulation of interactions between peroxins to enable a graduated regulation of peroxisomal import and peroxisome biogenesis.

Experimental procedures

Cloning, protein expression and purification.

Cloning, recombinant expression and purification of human PEX14 (16-80) (PEX14-NTD) (UniProtKB: O75381) and PEX5 (1-113), PEX5 (110-230), PEX5 (228-315) (UniProtKB: P50542) protein constructs were performed as described previously (Gaussmann *et al.*, 2021; Neuhaus *et al.*, 2014).

In brief, unlabeled PEX14-NTD was cloned into pETM11 vector, expressed in *Escherichia coli* BL21-(DE3) cells (Stratagene) in LB medium as a fusion protein containing His_6-tagged followed by a tobacco etch virus (TEV) cleavage site. After cell lysis (cell lysis buffer- 50 mM sodium phosphate, 300 mM sodium chloride, 10 mM Imidazole, pH 8.0), PEX14-NTD was purified using nickel-nitrilotriacetic acid-agarose (Qiagen) (elution buffer- 50 mM sodium phosphate, 300 mM sodium chloride, 500 mM Imidazole, pH 8.0), followed by TEV cleavage (TEV buffer- 50 mM sodium phosphate, 100 mM sodium chloride, 1 mM DTT, pH 8.0). His_6 tag was removed by a second Ni^{2+} affinity chromatography. The final purification was done by size exclusion chromatography on a HiLoad 16/60 Superdex75 column (GE Healthcare) in 20 mM ammonium bicarbonate buffer and pooled fractions are lyophilized.

Human PEX5 (1-113), PEX5 (110-230), PEX5 (228-315) protein constructs were expressed from pETM10 vectors with a non-cleavable N-terminal His_6 tag. Unlabeled PEX5 fragments were expressed and purified with the same protocol used for PEX14-NTD with minor change. 8M urea was used in the cell lysis buffer to avoid unspecific proteolysis cleavage during purification. Urea was removed during Ni affinity chromatography by extensive wash with cell lysis buffer. Ni-Eluted fractions were further purified by size exclusion chromatography and lyophilized.

The larger PEX5 (281-639) construct was cloned into a pETM11 vector (EMBL) with cleavable N-terminal His6 tag using site-directed ligase-independent mutagenesis (SLIM) (Chiu, 2004) with the following primers:

rv_short GGCGCCCTGAAAATAAAGATTCTCAG
fw_tail ATAGAGTCTGATGTCGATTTCTGGGACAAG
rv_tail GACATCAGACTCTATGGCGCCCTGAAAATAAAGATTCTCAG
fw_short GATTTCTGGGACAAGTTGCAGGCAG

The ScPEX14-NTD (corresponding to amino acids 1–58) (UniProtKB: P53112) and ScPex5(239–280) (UniProtKB: P35056) constructs were subcloned into a pETM30 vector (EMBL) that encodes a His_6-GST tag followed by a TEV-cleavage site. The ScPex5(1–313) construct encodes a non-cleavable His6-tag. Expression and purification of the yeast proteins followed the same procedure described above for the human variants, with the exception that for ScPEX14-NTD an additional wash step with 1M NaCl was included in the Ni-affinity step. Uniformly ^{15}N,^{13}C-labeled PEX5 (281-639) was expressed in deuterated M9 minimal medium supplemented with 1 g/liter $^{15}NH_4Cl$ (Cambridge Isotope Laboratories), 2 g/liter [U-^{13}C]-glucose-d_{12} (Cambridge Isotope Laboratories), as the sole source of nitrogen and carbon. The cell culture was induced with 0.5 mM IPTG at 18 °C for 14-16h before harvesting. PEX5 (281-639) was purified in buffer containing 50 mM Tris pH 7.5 and 300 mM NaCl using Ni-affinity chromatography. After TEV cleavage and a reverse Ni-column, size exclusion chromatography on a HiLoad Superdex 75 16/600 column (Cytiva) was performed.

The PEX5 (1-113) protein fragment was used to represent W0 for the biophysical experiments. Synthetic 15-mer peptides (W1-W7) of human PEX5 were purchased from *Peptide Specialty Laboratories* (Heidelberg, Germany). Peptides with purity of ≥ 98% were dialyzed extensively against water before the experiment.

Isothermal titration calorimetry (ITC)

ITC measurements were performed at 25 °C using PEAQ-ITC or iTC200 microcalorimeters. All proteins and peptides used for titration were dialyzed overnight in ITC buffer consisting of 20 mM sodium phosphate pH 6.5, 100 mM sodium chloride, 0.02% (w/v) sodium azide. In individual titrations, 200 µM concentration of human PEX5 peptides were loaded into the syringe and injected in 1.5 µL volumes at an interval of 150 seconds into a 280 µL cell contacting 20 µM concentration of PEX14-NTD, while stirring at 750 rpm. Calorimetric data were fitted to a single site binding model using MicroCal ITC-ORIGIN software supplied with the instrument. The binding stoichiometry (n), the dissociation constant (*K*D) and the enthalpy change (ΔH) were obtained from the fitted data. The Gibbs free energy (ΔG) and change in entropy (ΔS) were calculated from $-RT \ln K_D = \Delta G = \Delta H - T\Delta S$, where R is the gas constant and T is the absolute temperature (Rees and Robertson, 2001). To account for heat of dilution, control experiments were performed and subtracted from the corresponding data.

NMR Spectroscopy

NMR experiments for PEX5 (1-113), PEX5 (110-230), PEX5 (228-315) protein fragments were described previously (Gaussmann *et al.*, 2021). NMR of PEX5 (281-639) was performed at 298 K on a Bruker Avance II 950 MHz spectrometer equipped with cryoprobe. Buffer was exchanged to 20 mM sodium phosphate pH 6.5, 50 mM sodium chloride and 10% D2O using size exclusion chromatography. The protein was measured at 750µM in a 5mm Shigemi tube. Sequential assignment of backbone resonances was done by using TROSY versions of standard triple resonance experiments (Sattler M *et al.*, 1999; Weisemann *et al.*, 1993). NMR spectra were processed using Topspin (Bruker Biospin, Rheinstetten, Germany) or NMRPipe (Delaglio *et al.*, 1995) and analyzed using CcpNMR Analysis 2.4.2 (Vranken *et al.*, 2005). Secondary chemical shifts, $\Delta\delta$ ($^{13}C\alpha$) – $\Delta\delta$($^{13}C\beta$) were calculated by subtracting random coil chemical shifts from the observed 13Cα, 13Cβ chemical shifts (Kjaergaard & Poulsen, 2011; Schwarzinger *et al.*, 2001).

Circular dichroism

Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a peltier thermal controller. A final concentration of 30 µM of PEX14-NTD and 60 µM of peptide motif (1:2 ratio) were prepared in 10 mM sodium phosphate, 50 mM sodium chloride and pH 6.5. Thermal denaturation experiments were carried out by increasing the temperature from 10 to 95 °C at 1 °C/min in a cuvette with 0.1 cm path length and the CD spectra were collected at 222 nm. The protein-peptide complexes were incubated for one hour before initiating the unfolding experiment. The midpoint of the folding and unfolding (*T*m) is derived from raw data by fitting to the sigmoidal equation, Y = A2+(A1-A2)/(1+exp[(*x*-x0)/d*x*)]. Where A1 and A2 are the folding and unfolding intercept respectively. *x* is the midpoint of the curve and d*x* is the slope of the curve (Greenfield, 2006). The curve was fitted using Origin. Far UV-CD date were collected at 25 °C in the wavelength range of 190–260 nm. Spectra from 10 accumulations were added and the spectrum of the buffer alone was subtracted.

Peptide overlay binding assays

Each peptide spot comprises regions of 15 amino acids of PEX5 proteins with the Wxxx(F/Y) motif as a central core motif and 5 flanking amino acids on each side. Peptides were directly synthesized

on a cellulose membrane as described previously (Saidowsky *et al.*, 2001). After blocking with 3% BSA in TBS (10 mM Tris/HCl pH 7.4, 150 mM NaCl), membranes were probed overnight at 4 °C with purified 10 nM His6-tagged PEX14-NTD in TBS. Bound PEX14-NTD was immunodetected by monoclonal anti-His6 antibodies in TBS + 3% BSA, and horseradish peroxidase–coupled secondary antibodies in TBS + 10% milk powder and ECL Western Blotting Detection Reagent (GE Healthcare Amersham, ECL Western Blotting Detection Reagent). Between steps, the membranes were first thoroughly washed with TBS-TT (20 mM Tris/HCl pH 7.5, 0.5M NaCl, 0.05% (v/v) Tween20, 0.2% (v/v) Triton X-100) and at the end with TBS only.

Molecular dynamics simulations

Molecular Dynamics calculations were performed for all peptides shown in **Figure 5.2B** bound to PEX14. As initial structures the complex of PEX14-NTD with PEX5 (PDB-ID: 2W84, peptide sequence W1, **Table 5.2**) was used. The sequence of PEX5 was mutated to the corresponding sequence for all other motifs using the IRECS (Hartmann *et al.*, 2007, 2009) method as implemented in the DynaCell program (Antes, 2010). he Amber14 force field parameter set (Duan et al., 2003) was used together with the Amber14 software packages (Case *et al*, 2014). The structures were prepared for minimization with the *tleap* utility (Schafmeister *et al*, 1995) and all calculations were conducted in a neutralized, rectangular TIP3P (Jorgensen *et al*, 1983) water box extending at least 12 Å from any protein atom at each side of the box. Energy minimizations were performed with sander or pmemd.MPI (Case *et al.*, 2014). For every complex, two subsequent minimizations were conducted. First, 10100 steps of restraint minimization (100 steps with the steepest descent algorithm and 10000 steps with the conjugate gradient method) were done with the protein atoms restrained using a 50 kcal·mol-1·Å-2 force constant. Second, 100100 steps of energy minimization (100 steps with the steepest descent and 100000 steps with the conjugate gradient method) were conducted without restraints. Both minimizations were considered as converged if the root-mean-square of the Cartesian components of the energy gradient was less than 0.0001 kcal·mol-1·Å-1. The non-bonded interaction cutoff was set to 8.0 Å for both energy optimizations. Before conducting production runs, all systems were heated up by stepwise increasing the temperature over 660 ps while at the same time incrementally decreasing the number of restraint atoms as well as the force acting on them.

At each heating-up step, the systems initial velocities were randomly assigned from a Maxwell-Boltzmann distribution at the given target temperature. MD simulations were performed with 1-fs time steps. Non-bonded interactions were computed applying a cutoff of 14 Å. The Particle Mesh Ewald method was used to calculated long-range electrostatic interactions (Darden *et al*, 1993). The SHAKE algorithm (Ryckaert *et al*, 1977) was applied to constrain bonds to hydrogen atoms. The temperature was kept constant using the Berendsen thermostat (Berendsen *et al*, 1984) with a time constant of 1 ps to ensure constant temperature. The Berendsenbarostat was applied with a compressibility of 45 x 10-6 bar-1 and a pressure relaxation time of 1 ps to keep a constant target pressure of 1 bar. All MD simulations were performed by the pmemd.MPI or pmemd.cuda programs from the Amber14 software package (Case *et al.*, 2014).

After equilibration, a total of 50 ns of MD simulation was performed for each system (100 ns for W5). For the analysis of the system, all frames (4000) from the last 10 ns MD of the trajectory were minimized by DynaDock (Antes, 2010) for all minimized structures. The averaged energy values were used as final interaction energy. In order to calculate the interaction energy of the variants, the corresponding residue was mutated by IRECS (Hartmann et al., 2007) based on the representative structure of the biggest structural cluster from the corresponding wildtype MD trajectory. For each variant, 50 ns of MD simulation were performed (100 ns for W2_S2L) and the last 10 ns of the trajectories were used to calculate the interaction energy. The running conditions for these MD calculations were the same as for the wildtype peptides. The last 10 ns of the trajectories were analyzed using cpptraj in AmberTools 15 (Roe & Cheatham, 2013) for hydrogen bond analysis together with an in-house Cytoscape (Shannon *et al*, 2003) plugin allowing a network-based representation of the obtained hydrogen bonds. The hydrophobic cluster analysis was performed visually using VMD (Humphrey *et al*, 1996).

Author contributions

W.S., M.S., and I.A., conceived the study and designed the experiments. M.G., C.Z., S.G., H.K. and E.H. performed experiments and calculations. M.G., C.Z., S.G., W.S., R.E., M.S., and I.A. evaluated results and wrote the manuscript. All authors reviewed the results and approved the final version of the manuscript.

Chapter 7:

Intramolecular autoinhibition of HsPEX13 modulates protein-protein interaction in peroxisomal import

The work in this Chapter forms the basis for a manuscript to be submitted:

Stefan Gaussmann, Julia Ott, Krzysztof Zak, Florent Delhommel, Grzegorz Popowicz, Wolfgang Schliebs, Ralf Erdmann, Michael Sattler, *Intramolecular autoinhibition of HsPEX13 modulates protein-protein interactions in peroxisomal import*

Intramolecular autoinhibition of HsPEX13 modulates protein-protein interaction in peroxisomal import

Author contribution: The author of this thesis, Stefan Gaussmann carried out molecular biology and protein expression and purification, structural biology, biophysical experiments with data analysis and preparation of the manuscript.

Abstract

Peroxisomes are essential for eukaryotic life and peroxisomal malfunction is associated with severe human diseases. Organelle targeting of peroxisomal matrix proteins is guided by a peroxisomal targeting signal (PTS), which is recognized by the cytosolic cycling receptor PEX5. At the peroxisomal membrane PEX5 interacts with PEX13 and PEX14 to mediate docking and translocation, respectively. Receptor docking involves the intrinsically disordered N-terminal region of PEX5, which is recognized by a helical N-terminal domain (NTD) of the membrane-associated protein PEX14. PEX13 is crucial for PTS import and peroxisome biogenesis, and dysfunction of PEX13 is linked neonatal death. So far, molecular and structural mechanisms of PEX13 remain elusive. Here, we present a structural characterization of the PEX13 C-terminal domain (CTD) as well as intra- and intermolecular interactions with short peptide motifs by combining NMR spectroscopy, X-ray crystallography and biochemical methods. Our data reveal an unexpected autoinhibition of the PEX13 SH3 domain by a flanking FxxxF motif. Mapping of interactions between the PEX13 CTD, PEX14 and PEX5 identify a novel binding interface of the PEX13 CTD with the PEX14 NTD. The conserved interaction of the PEX13 SH3 domain with the PEX5 NTD, is modulated by the PEX13 SH3 proximal FxxxF motif. These interactions modulate PTS1 import in cell-based assays demonstrating that the PEX13 CTD contributes to peroxisome biogenesis. Our results reveal a dynamic interaction network between PEX13, PEX14 and PEX5 that regulates peroxisome biogenesis and import.

Introduction

Peroxisomes are single membrane enveloped organelles of eukaryotic cells which are essential for several metabolic pathways mainly related to lipid metabolism and to the removal of toxic oxidation products (Erdmann *et al*, 1997a; Fujiki & Lazarow, 1985; Wanders, 2004; Wanders &

Waterham, 2006). The physiological importance of these highly conserved organelles is emphasized by the diseases such as Zellweger Spectrum Disorders (ZSD) that result from defects in peroxisome biogenesis (Klouwer *et al*, 2015). Biogenesis and function rely on peroxisome-related proteins called peroxins (Distel *et al.*, 1996; Ma *et al.*, 2011) that are essential for membrane –assembly and post-translational matrix protein import (Purdue & Lazarow, 2001). Human and yeast peroxins are abbreviated as "PEX" and "Pex" respectively. Malfunction of PEX13, a peroxin crucial for peroxisomal import, leads to impaired biogenesis and neonatal death (Maxwell *et al.*, 2003). The general mechanisms of peroxisomal matrix import are evolutionary conserved. Peroxisomal cargo proteins possess a conserved peroxisomal targeting signal, at their C-terminus (PTS1) or N-terminus (PTS2) (Ghosh & Berg, 2010; Gould *et al.*, 1987). Cytosolic PTS1-cargos are recognized by the C-terminal tetratricopeptide (TPR) domain of the peroxisomal receptor PEX5 (Gatto *et al.*, 2000; Stanley *et al.*, 2006), which tethers the cargo complex to the peroxisomal membrane via its intrinsically disordered N-terminal domain (NTD) (Dammai & Subramani, 2001; Dodt & Gould, 1996; Erdmann & Schliebs, 2005; Rucktaschel *et al.*, 2011).At the peroxisomal membrane, PEX5 NTD interacts with the membrane bound components of the translocon Pex14 and Pex13 (**Figure 7.1A**) (Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014; Saidowsky *et al.*, 2001; Schliebs *et al.*, 1999). These interactions may be regulated by weak interactions of PEX5 and PEX14 with peroxisomal membranes (Gaussmann *et al.*, 2021; Kerssen *et al.*, 2006). After docking, a transient and dynamic pore is formed and the cargo is translocated into the peroxisomal matrix (Erdmann & Schliebs, 2005) (**Figure 7.1A**).

Many aspects of peroxisome biogenesis have been studied in yeast, where Pex5/cargo docking is mediated by a cytosolic complex consistent of the Pex13 SH3, the Pex14 NTD and Pex5 NTD, essential for both PTS1 and PTS2 import (Gould *et al*, 1996b). The interactions of the so-called docking complex are mediated di-aromatic peptide motifs ("W", as most contain a tryptophane residue) of Pex5 and a poly-proline (PxxP) motif of Pex14 that binds to the Pex13 SH3 domain (Barnett *et al*, 2000; Douangamath *et al*, 2002; Erdmann & Blobel, 1995). However, Pex13 appears to play a role in receptor docking as a purified peroxisome pore *in vitro* comprises only Pex5 and Pex14 (Meinecke *et al.*, 2010).

The homologous interactions in humans are less well studied. Binding between PEX5 (di)aromatic motifs and the globular N-terminal domain of PEX14 (**Figure 7.1B, 1**) has been shown to be conserved, but a docking complex with PEX13 similar to the one observed in yeast has yet to be

reported. Human PEX13 is an integral membrane protein with an intrinsically disordered N-terminal domain followed by three transmembrane spans (Barros-Barbosa *et al.*, 2019) and a mostly unstructured C-terminal region harboring a SH3 domain. (**Figure 7.1D**). PEX13 was first identified and studied in the context of peroxisomal import (Gould *et al.*, 1996a) and Zellweger spectrum disorder (Liu *et al.*, 1999). However, our knowledge of PEX13 and its role in peroxisomal import is quite limited. An early study postulated the PEX5/PEX13 interaction via PEX5 WxxxF/Y motifs W2, W3, and W4 and the PEX13 N-terminal region (**Figure 7.1B, 2**) which was shown to be essential for catalase import (Otera *et al.*, 2002). Recently, the Zellweger mutation W313G located in the PEX13 SH3 domain was demonstrated to disrupt PTS1 but not PTS2 import by abolishing homo-oligomerization of PEX13 and generating a dominant negative effect (Krause *et al.*, 2006; Krause *et al.*, 2013).

In the present study, we present the first structural characterization of the PEX13 CTD and characterize molecular interactions of PEX13 with PEX5 and PEX14 by combining biophysical methods, structural biology and functional analysis in cells. We show that a FxxxF peptide motif mediates an autoinhibitory interaction with the PEX13 SH3 domain. Our binding studies demonstrate a network of interactions of the PEX5 NTD with PEX13 and PEX14 that modulates peroxisome biogenesis in cells.

Results

Structural analysis of PEX13 CTD by NMR spectroscopy

We first assigned the backbone resonances of the soluble PEX13 CTD (261-403) construct by NMR spectroscopy. The ^{1}H-^{15}N correlation spectra shows well dispersed signals and additional signals with narrow linewidth which correspond to the globular SH3 domain and the disordered C-terminal region, respectively (**Figure 7.2A**). NMR $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ secondary chemical shifts are consistent with the secondary structure seen of SH3 domains composed of five β-strands (β1 to β5) (Saksela & Permi, 2012) **(Figure 7.1D, 7.2C)**. An additional α-helical motif located in the C-terminal region (371-383) downstream of the SH3 domain encompasses an FxxxF motif (**Figure 7.1D**) which is highly conserved across mammals (**Figure 7.2D)**. We then compared the ^{1}H-^{15}N correlation spectra of PEX13 CTD and SH3 domain. Surprisingly, significant chemical shift

differences are seen in the SH3 domain, which map to the regions of β1, β5 and in the region of β2 and an alpha turn (**Figure 7.1C, Figure 7.2A, B**).

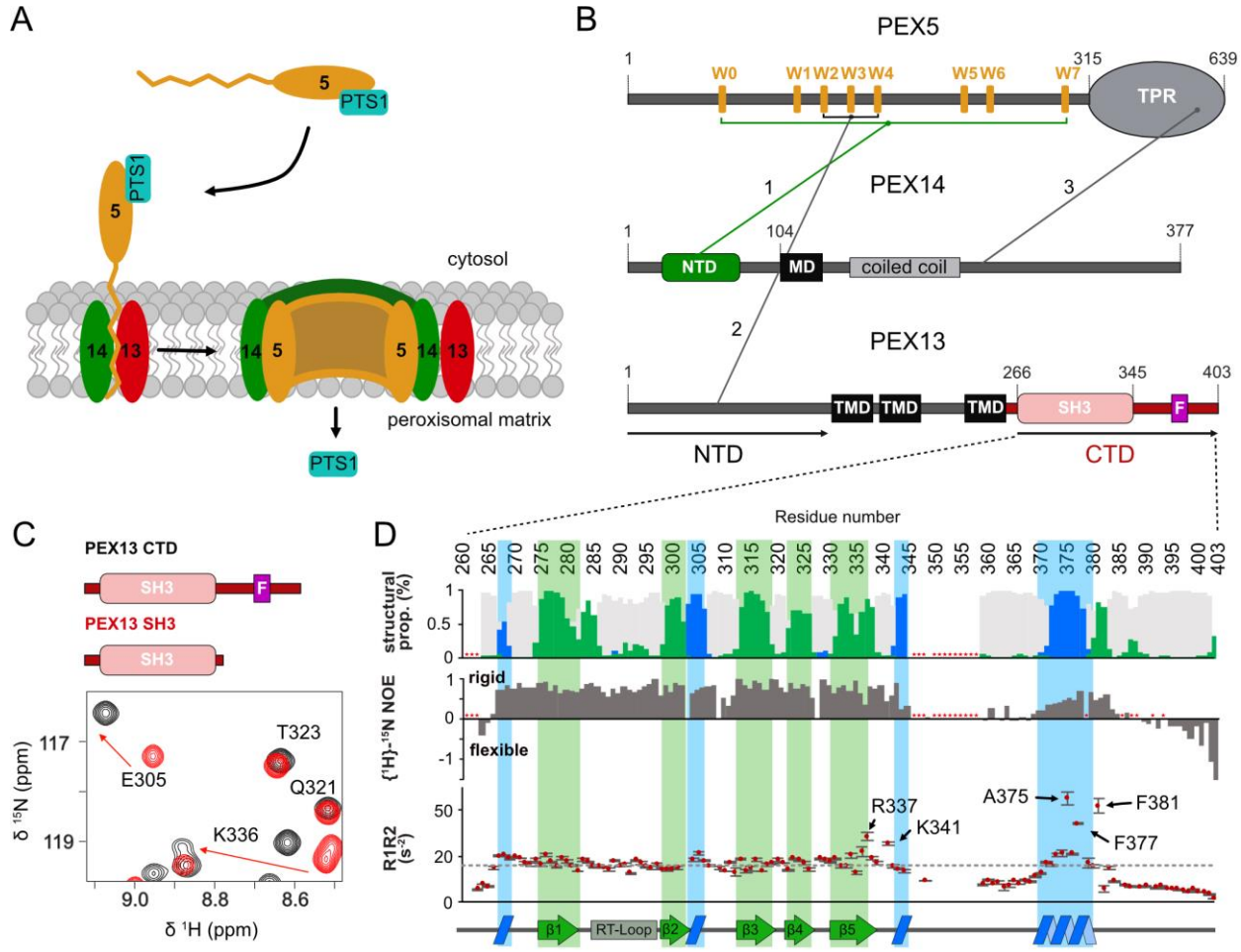

Figure 7.1. (A) Schematic overview of cargo recognition, docking and pore formation. (B) Schematic representation of domain architecture of the peroxins PEX5, PEX14 and PEX13 respectively. Lines 1, 2 (Otera *et al.*, 2002) and 3 (Emmanouilidis et al., in preparation) between the peroxins indicate binding events involving the targeted structure elements or motifs. (C) Zoom into overlaid spectra of PEX13 CTD (black) and PEX13 SH3 (red). (D) Top: TALOS+ secondary structure propensity (blue: α-helix, green: β-strand, gray: coiled coil) based on $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of PEX13 CTD (261-403). The data support the typical β-sandwich fold of the SH3 domain and the presence of a short α-helical motif comprising the FxxxF motif. Middle: elevated $\{^{1}H\}$-^{15}N heteronuclear NOE values indicate an extended SH3 fold (265-345) and a folded FxxxF motif with similar values to the SH3 domain. Asterisk indicate proline or missing assignment. Bottom: ^{15}N R_1*R_2 relaxation rates as a function of amino acid sequence. SH3 core residues (266-335) show an average of 16.6. C-terminal residues R337, K341, A375, F377 and F381 show values of 28.4 ± 1.8, 25.2 ± 0.6, 45 ± 2.6, 34,2 ± 0.6 and 42.0 ± 3.0 respectively. Values higher compared to the average in structured regions indicate the presence of conformational dynamics/and or transient interactions. Secondary structure elements are indicated at the bottom.

We then recorded {^{1}H}-^{15}N heteronuclear NOEs (hetNOE), which reflect the flexibility of the backbone at sub nanosecond timescales. Values around 0.9 are seen for amides in the SH3 domain, as expected for a rigid and well-folded domain **(Figure 7.1D)**. The reduced backbone flexibility indicated by the hetNOE data the PEX13 SH3 fold extends beyond the typical SH3 structural elements on both termini compared to the yeast Pex13 and other human SH3 domains (**Figure 7.2E**). In the C-terminal region, around residue 350, hetNOE values decrease to 0, showing the overall flexibility of the C-terminal region. However, the C-terminal FxxxF motif shows hetNOE values of ~0.8, comparable to those in the globular SH3 domain (**Figure 7.1D**), indicating conformational rigidity. The product of NMR ^{15}N R_1 and R_2 relaxation rates is indicative to discriminate rigid and flexible regions (Farrow *et al.*, 1994)**.** While similar values with an average of 16.6 s^{-2}, are seen for the rigid core domain, some C-terminal residues of the SH3 domain and the FxxxF motif show significant elevated values ranging from 25 s^{-2} to 45 s^{-2}, indicating conformational exchange (**Figure 7.1D)** (Kneller *et al.*, 2002). These suggests potential transient intramolecular interactions between the FxxxF motif and the PEX13 SH3 domain.

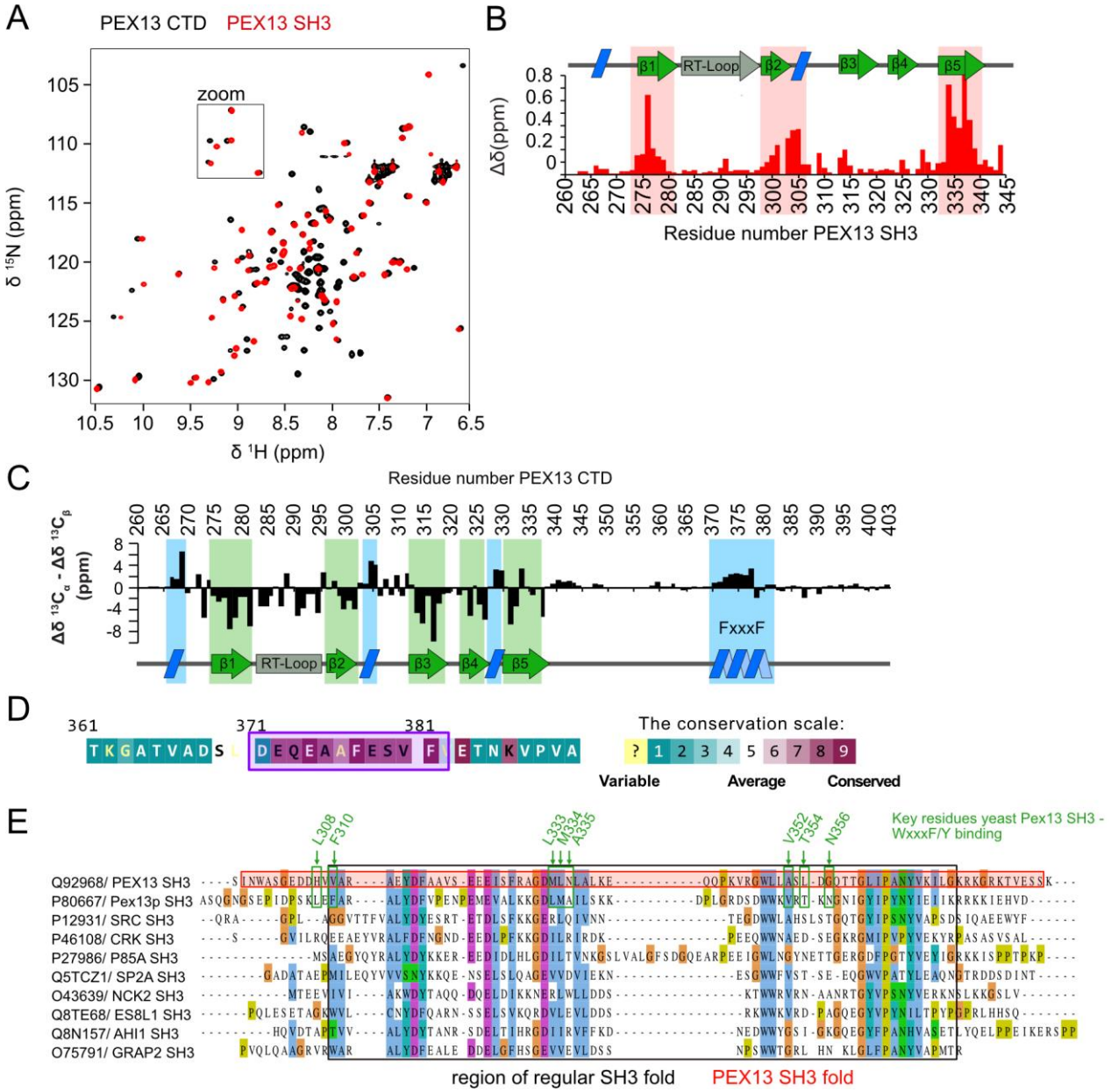

Figure 7.2. Conformation and NMR analysis of human PEX13 CTD.
(A) Overlaid 2D spectra of PEX13 CTD (black) and PEX13 SH3 (red). A zoomed-view is shown in **Figure 7.1** (B) chemical shift perturbations plotted on the PEX13 SH3 sequence with secondary structure indicated on top. (C) ^{13}C secondary chemical shifts ($\Delta\delta^{13}C_\alpha$ - $\Delta\delta^{13}C_\beta$) of PEX13 CTD (261-403) support the typical β-sandwich fold of the SH3 domain and the presence of a short α-helical motif comprising the FxxxF motif.(D) Sequence alignment with 186 mammalian PEX13 sequences plotted with ConSurf web server (Ashkenazy *et al*, 2016; Ashkenazy *et al*, 2010; Celniker *et al*, 2013) shows high conservation of the FxxxF motif (purple box) (E) Sequence alignment of human PEX13 SH3 with yeast PEX13 SH3 and SH3 domains from human SRC, CRK, AHI1, NCK2, ES8L1, GRAP2, P85A, and SPD2A (with known structures). The sequence alignment was done with the ± 10 residues flanking the SH3 fold. The black box indicates the regular SH3 fold and the red box the complete fold of human PEX13 SH3. Green arrows indicate

key residues for yeast Pex13 SH3 / WxxxF/Y binding. Note that those residues are poorly conserved from yeast to human

Autoinhibition of the SH3 domain by the internal FxxxF motif is driven by hydrophobic and electrostatic interactions

The domain boundaries defined by the NMR analysis were used to generate chimeric constructs which contain the PEX13 SH3 domain and the FxxxF motif (D_{371}EQEAAFESFV_{383}) separated by GGGGS (GS) linkers. Structures of the apo SH3 domain (**Figure 7.3A**) and SH3-2GS-FxxxF (**Figure 7.4A**) were solved by X-ray crystallography at resolutions of 1.8Å and 2.3Å, respectively (**Table 7.1**). A comparison of the apo PEX13 SH3 and the complex structure did not show significant differences upon alignment (RMSD = 0.39) (**Figure 7.3C**). Analysis of the structure of the SH3 domain shows a network of polar contacts between the N- and C-terminal regions stabilizing the β1/β5 interaction (**Figure 7.3B**), consistent with the extended domain boundaries observed by NMR.

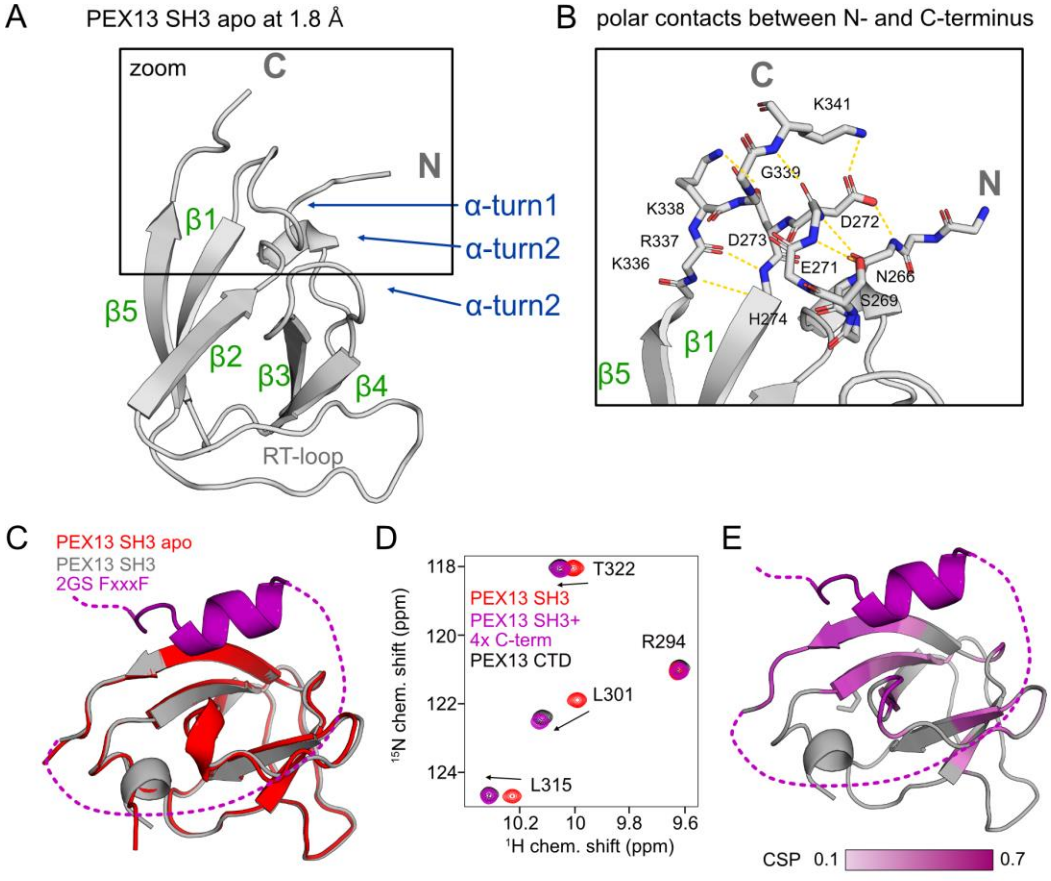

Figure 7.3. Structural features of apo PEX13 SH3 and confirmation of PEX13 SH3 – FxxxF structure in solution. (A) Apo structure of PEX13 SH3 solved with 1.8 Å resolution. (B) Zoomed-view showing a network of hydrogen bonds in the on the N – and C-terminal regions (yellow). (C) Superimposition of the structures PEX13 SH3 (red) and PEX13 SH3 (gray) in complex with FxxxF motif (purple). (D) Overlay of ^{1}H, ^{15}N correlation spectra of PEX13 SH3 (red) PEX13 CTD (black) and PEX13 SH3 titrated with FxxxF peptide (350-403) (purple). PEX13 SH3 titrated with FxxxF peptide represents the native spectrum of PEX13 CTD. (E) Chemical shift perturbations from PEX13 SH3 / FxxxF titration mapped on the PEX13 SH3-FxxxF complex structure.

The structure of the complex shows interactions of the α-helical FxxxF motif and SH3 domain mediated by hydrophobic contacts of the two phenylalanine, which clamp around β1 and β5 (**Figure 7.4A, C**), and polar interactions involving sidechain and backbone contacts. Backbone hydrogen-bonds are formed between A376 and G335 as well as S379, F381 and K336 while negatively charged sidechains E374 and E378 show electrostatic interactions with K304, E305, K336 and a water molecule (**Figure 7.4B**). Interestingly, nine out of the eleven Arg and Lys residues are located at the FxxxF binding surface, causing a highly positive charge, which is favorable for binding negative charged peptides such as the C-terminal FxxxF motif (**Figure 7.4C**). The PxxP binding site located at the other side of the SH3 domain, is on the other hand mostly negatively charged. The crystal structure was validated in solution by NMR titrations of the isolated SH3 domain with a FxxxF peptide (350-403) showing strong chemical shift perturbations in the binding site expected from the crystal structure, where the spectrum at saturated binding is very similar to the native CTD (**Figure 7.3D, E)**. Furthermore, the static light scattering (SLS) analysis of PEX13 CTD indicates a molecular weight of 15.6 ± 0.1 kDa which correlates well with the calculated mass of 15.6 kDa (**Figure 7.4D**). This confirms that the FxxxF/SH3 interaction occurs intramolecularly and does not involve oligomerization of the construct at the measured concentration. These results show that the PEX13-CTD adopts an auto-inhibited state, where the C-terminal FxxxF motif interacts with the SH3 domain.

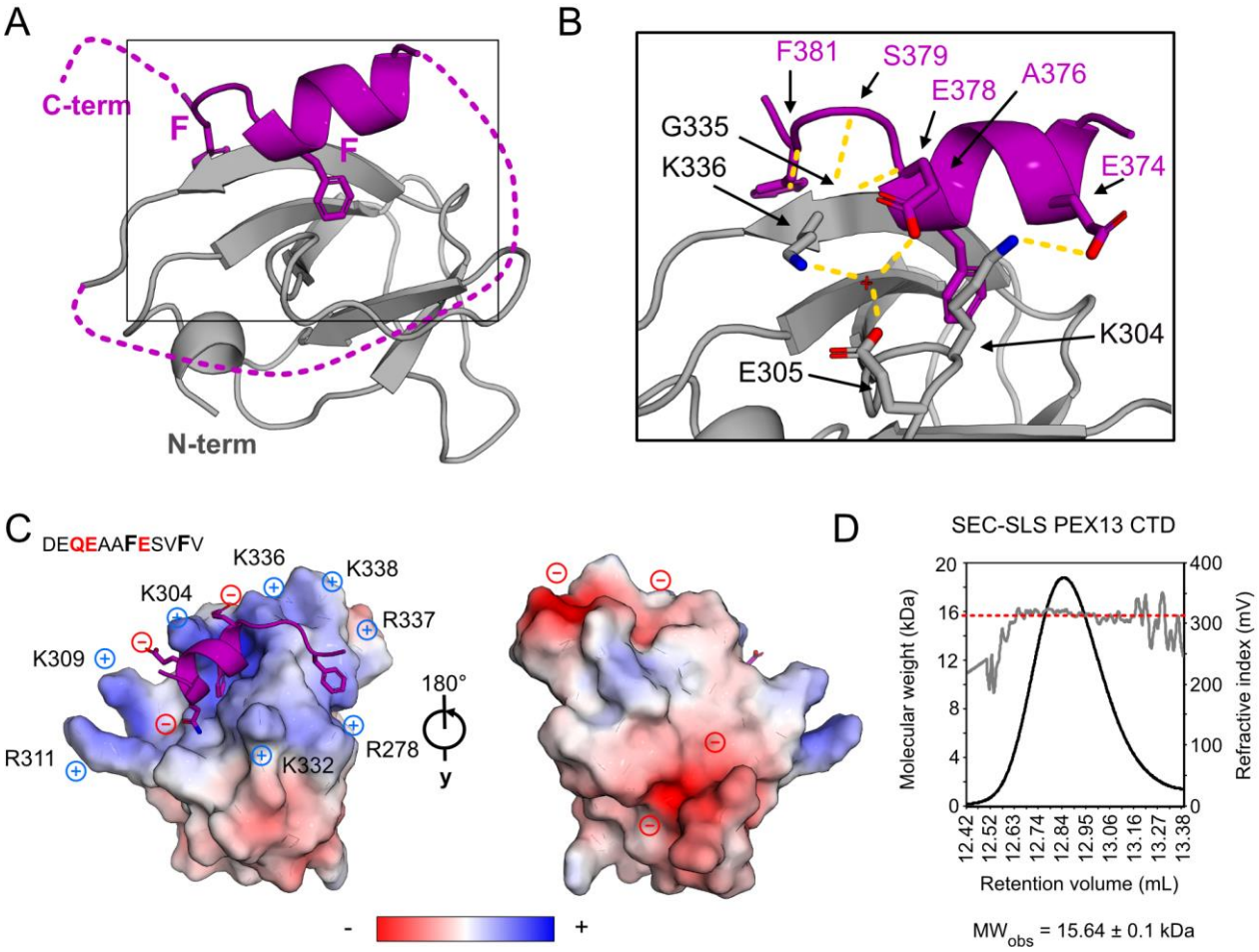

Figure 7.4. Structural analysis of PEX13 SH3 in complex with FxxxF motif (A) Crystallographic structure of PEX13 SH3 2GS FxxxF (2.3 Å resolution) showing the α-helical FxxxF motif, which clamps β1 and β5 between the two Phe residues. (B) Zoom view visualizing the hydrogen bond network between the SH3 domain and the FxxxF peptide. Polar backbone contacts are mediated byA376 and G335 as well as S379, F381 and K336 and sidechains E374 and E378 are coordinated by K304, E305, K336 and a water molecule. (C) Electrostatic surface representation showing the positively charged FxxxF binding site which is caused by seven Arg or Lys residues. The peptide in contrast, is negatively charged which is favored for the binding (Q373, E374 and E387). A 180° rotation on the Y axis reveals a negatively charged backside. (D) Static light scattering analysis of PEX13 CTD shows the molecular weight (red dashed line) of 15.64 ± 0.1 kDa versus the calculated mass of 15.56 kDa indicating a monomeric state.

Table 7.1. Crystallographic data for PEX13 SH3, PEX13 SH3 2GS FxxxF and PEX13 SH3 W4 structures.

Dataset	**PEX13 SH3**	**PEX13 SH3 2GS FxxxF**	**PEX13 SH3 GS W4**
PDB entry	**7Z0I**	**7Z0J**	**7Z0K**
Space group	I 41	P 31 2 1	P 31 2 1
Cell paramters (Å, °)	43.949 43.949 86.780, 90.00 90.00 90.00	86.808 86.808 65.436, 90.000 90.000 120.000	87.662 87.662 66.112, 90.000 90.000 120.000
Data collection			
Wavelength (Å)	0.999995	1.000029	1.0332
Temperature (K)	100	100	100
Resolution range (Å)	43.39-1.80 (1.84-1.80)	49.36-2.3 (2.38-2.3)	49.86-2.3 (2.38-2.3)
Total no. of observed reflections	104090 (6202)	260870 (25799)	468378 (25836)
Number of unique reflections	7660 (462)	12998 (1250)	13362 (1275)
R_{merge} (within I+ and I-)	0.082 (1.443)	0.282 (2.426)	0.495 (10.645)
R_{pim} (within I+ and I-)	0.034 (0.588)	0.091 (0.774)	0.118 (3.458)
CC1/2	1 (0.738)	0.997 (0.652)	0.992 (0.588)
$\langle I/\sigma(I)\rangle$	19.7 (2.1)	11.8 (1.6)	17.0 (1.8)
Completeness (%)	99.9 (99.5)	100 (100)	100 (100)
Multiplicity	13.6 (13.4)	20.1 (20.6)	35.1 (20.3)
Wilson B-factor (Å^2)	26.2	35.67	45.11
Refinement			
Rwork (%)	0.182	0.197	0.193
Rfree (%)	0.229	0.224	0.217
No. Of atoms	619	1326	1359
Ligand atoms	11	12	0
solvent atoms	51	75	61
Model quality			
RMS (Bonds)	0.0087	0.0133	0.0156
RMS (angles)	1.498	1.9	2.05
Ramachandran favored (%)	96.05	97.56	96.93
Ramachandran allowed (%)	3.95	2.44	3.07
Ramachandran outliers	0	0	0
Rotamer outliers (%)	6.35	6.62	10
Clashscore	0.8	2.63	3.71

PEX13 FxxxF motif binds to PEX14 NTD

We next evaluated interactions of PEX13 with the core components of the import machinery. First, we analyzed the PEX13 CTD / PEX14 NTD interaction by NMR titrations monitoring effects on the ^{15}N labeled PEX14 NTD upon titration of unlabeled PEX13 CTD. Significant chemical shift

perturbations are observed for amide signals (**Figure 7.5A**) and with a profile similar to the known interaction of PEX14 NTD with the PEX19 FxxxF motif (**Figure 7.5C**) (Neufeld *et al.*, 2009). A sequence comparison of the two motifs from PEX13 and PEX19 (**Figure 7.5B**) shows strong similarities with four identical residues (**Figure 7.5B**). Not surprisingly, mapping the chemical shift perturbations onto the PEX14 NTD structure highlights the involvement of key residues, that are also involved in binding of PEX5 WxxxF/Y motifs (Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014). The PEX14 NTD / PEX13 FxxxF interaction was analyzed by ITC shows binding affinities for the auto-inhibited PEX13 CTD (261-403) or a PEX13 FxxxF construct (starting at the linker region just after the SH3 domain, residues 350-403) corresponding to a K_D of 5.4 µM and 2.8 µM respectively (**Figure 7.5E, Table 7.2**). Although the dissociation constants (K_D) are in a similar range, with two-fold stronger binding for the isolated FxxxF peptide, the underlying energetics are notably different. While the interaction with PEX13 CTD profits from enthalpic and entropic effects, probably reflecting an autoinhibed bound-to- PEX14-bound transition of the FxxxF motif, the interaction with free FxxxF is comes with an entropic penalty, reflecting a free to bound transition for the FxxxF motif (**Figure 7.5F, Table 7.2)**. The ITC experiments further demonstrate a stoichiometry of one to one in both cases (**Figure 7.5E lower panel, Table 7.2).**

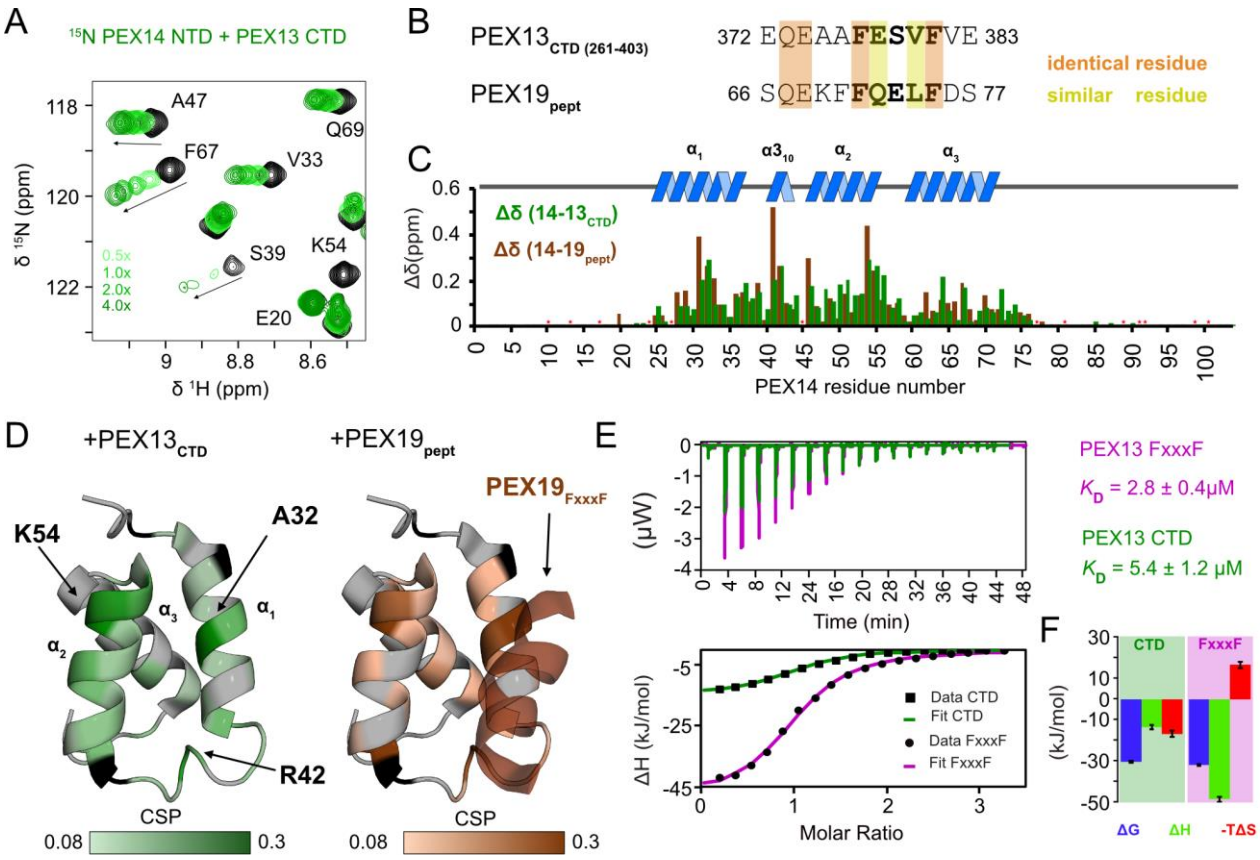

Figure 7.5. Interaction of PEX14 NTD with PEX13 CTD in comparison with PEX19 (66-77) (A) $^{1}H,^{15}N$ correlation spectra of ^{15}N-labelled recombinant PEX14(NTD) free (black), and in complex with PEX13 CTD (green scale). (B) Sequence alignment of PEX13 and PEX19 FxxxF motifs. Red and yellow boxes indicate identical and similar residues. (C) NMR chemical shift perturbations of PEX14 NTD in the presence of PEX13 CTD (green) or PEX19 peptide (brown) (Neufeld et. al.) plotted on the sequence with indicated secondary structure elements above. Asterisk indicate proline or missing assignment. (D) Chemical shift perturbations (0.08 to 0.3 ppm) of PEX13 CTD (left) and PEX19 peptide (right) mapped on the PEX14 NTD/PEX19 66-77 structure (2w85, Neufeld et al.). (E) ITC experiments of PEX14 NTD with PEX13 CTD (green) and PEX13 FxxxF (pink) showing very different energetics but the same one to one stoichiometry. (F) Energetic contribution of the PEX14 NTD interaction with PEX13 CTD (left graph) and PEX13 FxxxF peptide (right graph).

A comparison of the interaction strength of the PEX13 FxxxF and the PEX19 FxxxF motif (K_D = 9.2 µM, Neufeld *et al.* (2009)) shows a three times stronger binding of PEX13 FxxxF towards PEX14 NTD. Taken together, these data show that the auto-inhibited state of PEX13 CTD is readily released upon PEX14 binding. These findings also show that the human SH3 can interact with bi-aromatic peptide motifs on a surface opposite to the PxxP binding region, as has previously been shown for the yeast Pex13 SH3 domain (Douangamath *et al.*, 2002).

Table 7.2 Isothermal titration calorimetry of PEX14 NTD with PEX13 CTD or PEX13 C-terminal peptide (351-403).

Construct	N	K_D (µM)	ΔG (kJ/mol)	ΔH (kJ/mol)	-TΔS (kJ/mol)
PEX13 CTD	1	5.4 ± 1.2	-30.2 ± 0.50	-13.4 ± 1.15	-16.7 ± 1.47
PEX13 FxxxF	1	2.8 ± 0.4	-31.8 ± 0.36	-48.3 ± 1.3	16.5 ± 1.62

To assess whether the interaction site in PEX13 is limited to the FxxxF motif or involves additional regions we titrated unlabeled PEX14 NTD (1-104) onto ^{15}N labeled PEX13 CTD (**Figure 7.6**). Interestingly, strong chemical shift perturbations are seen not only for the FxxxF motif but also for residues in the SH3 domain **(Figure 7.6A,B, Figure 7.7A).** Notably, NMR signals of the SH3 domain shift towards their position in the SH3 domain alone (**Figure 7.6B, C yellow boxes**). Signals of the FxxxF peptide (370-386) experience large chemical shift perturbation or line-broadening and residues 378 to 383 located in the core motif show line-broadening beyond detection, indicated by negative values in the chemical shift plot (**Figure 7.6C**). We conclude that the PEX14 NTD interacts solely with the FxxxF motif, and releases the auto-inhibited conformation of the PEX13 SH3 domain (**Figure 7.6D**).

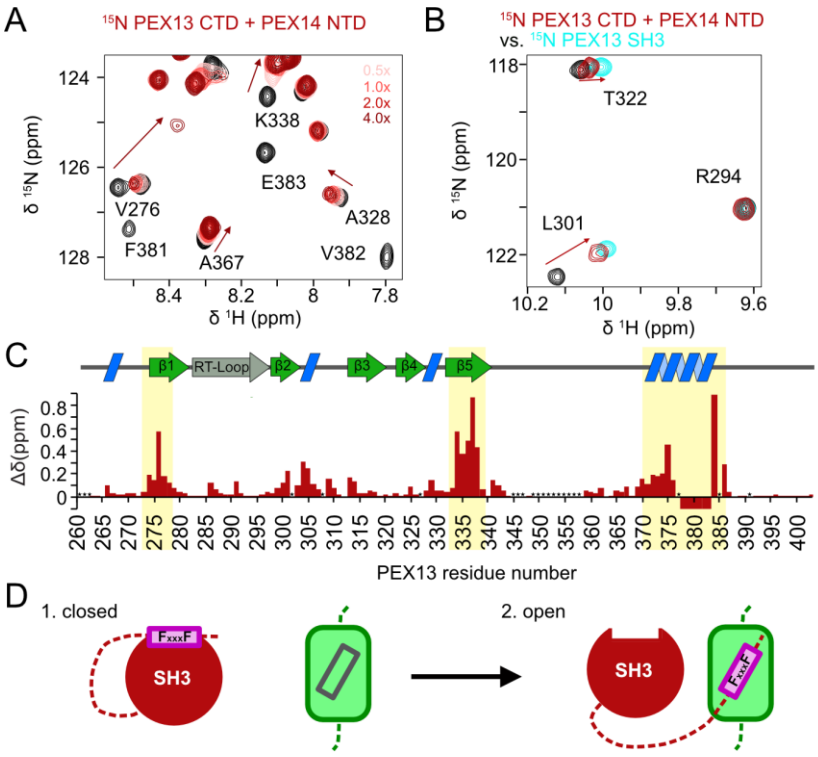

Figure 7.6. NMR titration of PEX14 NTD onto ^{15}N PEX13 CTD (A) Spectra overlay from NMR titration of unlabeled PEX14 NTD onto ^{15}N labeled PEX13 CTD where large chemical shift perturbations of resonances from the FxxxF motif were observed. (B) Spectra overlay of free PEX13 CTD (black), PEX13 CTD + 4x PEX14 NTD (dark red) and apo PEX13 SH3 (blue) showing the transfer from the closed conformation of the PEX13 CTD back to apo form of PEX13 SH3. (C) Chemical shift perturbations mapped on the sequence and structural elements (above) of PEX13 CTD visualizing the effect of the opening on the structural elements β1, β5 and the FxxxF motif. (D) Schematic representation of the opening process.

PEX13 SH3 interactions with poly-proline motifs

In yeast, the Pex14/Pex13 interaction is mediated by a poly-proline motif of Pex14, which binds to the Pex13 SH3 domain (Douangamath *et al.*, 2002). Human PEX14 also harbors a PxxP motif downstream of its NTD (residues 87-102), similar to that in yeast (residues 85-94). NMR titrations with unlabeled PEX14 NTD_{long} (1-113) onto ^{15}N labeled PEX13 SH3 were used to evaluate this potential interaction. No significant chemical shift perturbations were observed, showing that this interaction is not conserved from yeast to human (**Figure 7.7B**). Our results are in agreement with previous experiments using co-immunoprecipitation that showed an interaction of PEX14 with the PEX13 Zellweger mutant W313. This mutation is located in the SH3 domain and destabilizes the SH3 fold which affects multiple interactions of the SH3 domain. The human PEX14/PEX13 interaction was still intact while the same mutation in the yeast homologue abolished Pex14 binding

(Krause *et al.*, 2013). We thus evaluated whether human PEX13 SH3 has any ability to bind PxxP motifs. Interestingly, PxxP motifs present in the N-terminal region of PEX13 for showed significant binding in NMR titration experiments with the PEX13 SH3 domain and the PEX13 CTD. These experiments demonstrate that, in principle, the human SH3 domain can mediate PxxP interactions and that these are independent, and non-overlapping, with the FxxxF binding (**Figure 7.7C, D).** However, chemical shift perturbations mapped onto a structural model of a canonical PEX13 SH3/PxxP complex indicate that the C-terminal region of the peptide shows a distinct binding mode, suggesting that the PEX13 SH3 domain shows non-canonical interactions with PxxP peptides, distinct from the yeast orthologue (**Figure. 7.6F).**

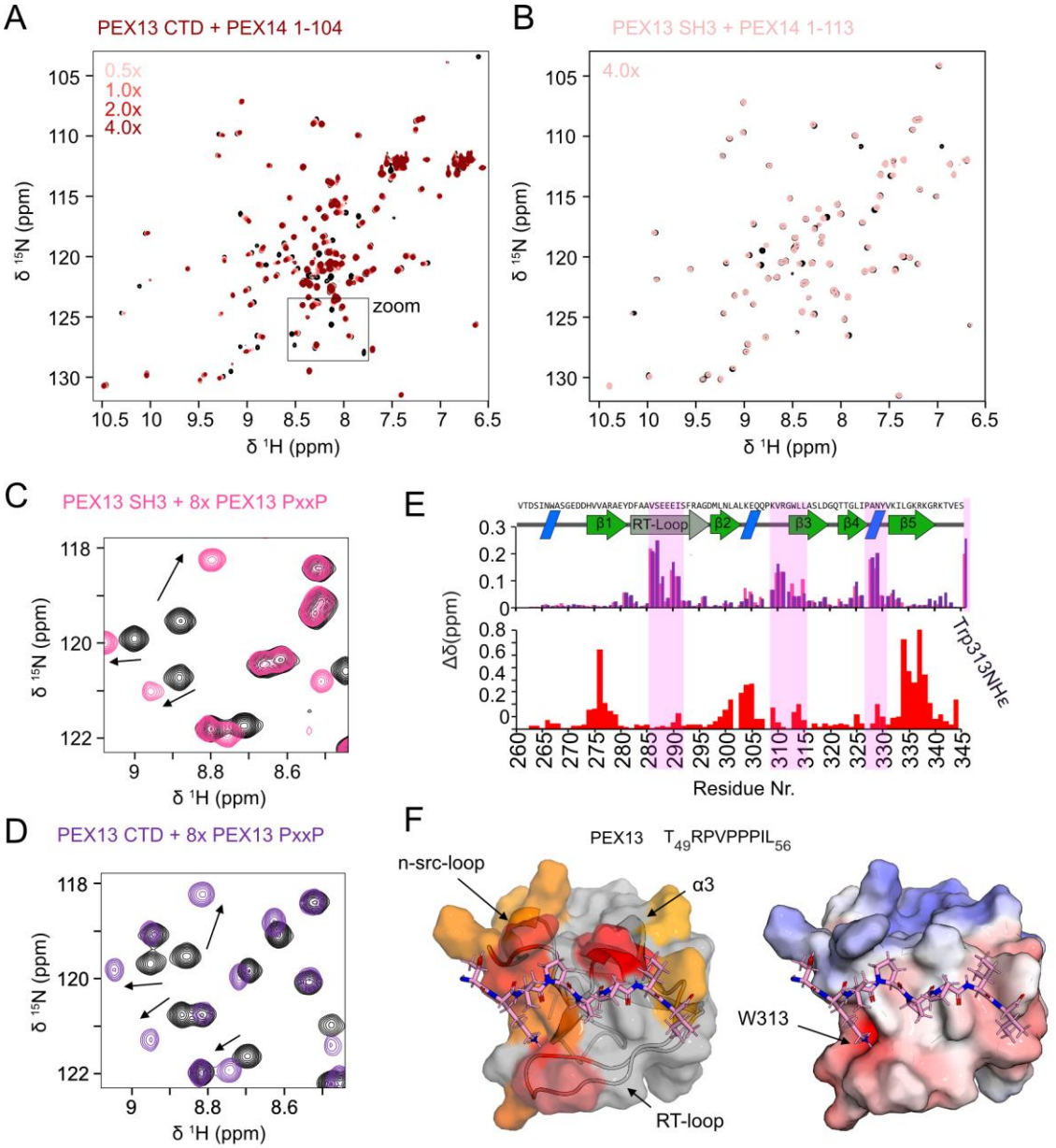

Figure 7.7. Interactions of PEX13 SH3 with PxxP peptides. (A) Overlaid 2D spectra of PEX13 CTD (black) titrated with increasing concentrations PEX14 NTD (1-104) (red scale). Zoom is shown in main figure. (B) 2D spectra of PEX13 SH3 with 4x excess of PEX14 NTD_{long} (1-113) (C) 2D spectra of PEX13 SH3 (black) with 8x excess of PEX13 PxxP peptide (49-61) (pink). (D) 2D spectra of PEX13 CTD (black) with 8x excess of PEX13 PxxP peptide (49-61) (dark purple) (E) Chemical shift perturbations from PxxP titration on SH3 (pink) and CTD (dark purple) (upper panel) and from FxxxF titration on SH3 (red/bottom panel) mapped on the PEX13 SH3 sequence demonstrating non- overlapping binding sites (F) Structure of PEX13 SH3 with modelled PEX13 PxxP (49-56) represented with transparent surface showing chemical shift perturbations (yellow to red gradient) from NMR titrations (left panel) and electrostatic surface (right panel). PxxP sequence shown above. Binding of PEX13 PxxP (49-61) induces typical CSPs in the regions of RT-loop, n-src-loop and α3 (α10) (Shi *et al*, 2012).

PEX5 WxxxF/Y motifs compete with the internal FxxxF motif on PEX13 SH3

We then characterized the molecular interactions between PEX13 and PEX5. The PEX5 NTD (1-315) but not PEX5 TPR (315-639) domain was found to bind to PEX13 CTD and PEX13 SH3 (**Figure 7.8**).

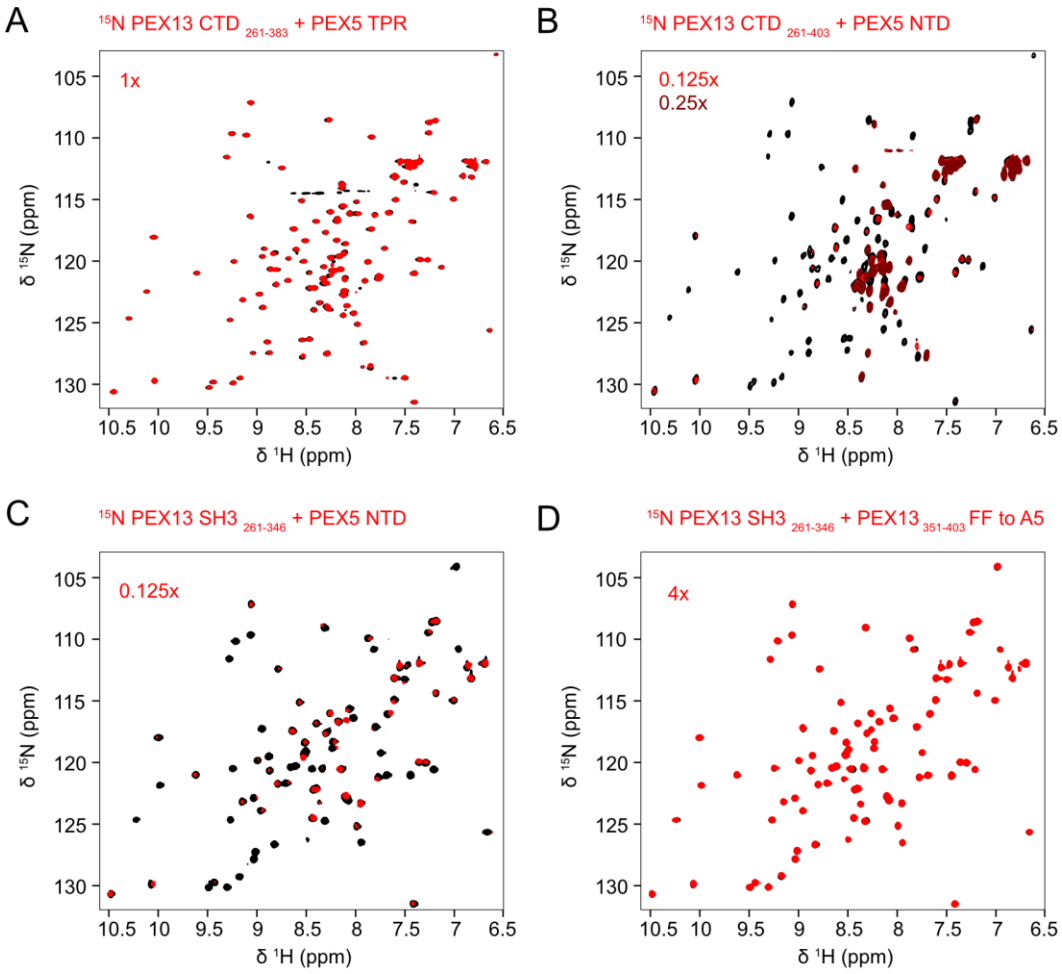

Figure 7.8. Titration of ^{15}N PEX13 CTD or PEX13 SH3 with PEX5 TPR, PEX5 NTD or PEX13 C-terminal peptide. (A) Overlaid 2D spectra of PEX13 CTD (261-383; missing the last 20 amino acids; black) titrated with equimolar concentration of PEX5 TPR domain (red). (B) Overlaid 2D spectra of PEX13 CTD (black) titrated with 0.125x (red) or 0.25x (dark red) excess of PEX5 NTD. Notably, resonances experience excessive line-broadening at 0.125 (1/8) ligand concentration. (C) Overlaid 2D spectra of PEX13 SH3 (black) titrated with 0.125x (red) excess of PEX5 NTD. The SH3 titration experiment shows very similar line-broadening effect as seen with PEX13 CTD indicating binding of PEX5 NTD to PEX13 SH3. (D) Overlaid 2D spectra of PEX13 SH3 (black) titrated with 4x (red) excess of PEX13 C-terminal peptide (351-403) with the FxxxF motif mutated to AAAAA.

Using NMR titrations we show that the eight (di)aromatic peptide motifs of PEX5, also known as W-motifs, bind to the PEX13 SH3 domain or the PEX13 CTD (**Figure 7.9A, Figure 7.10A, B**). The W4 motif was identified as the strongest binder followed by W2 and W3 (**Figure 7.9B**). These PEX5 motifs are able to compete with the PEX13 internal FxxxF motif, while other (di)aromatic peptide motifs only bind to PEX13 SH3 with low affinity and not to PEX13 CTD (**Figure 7.9C)**. PEX5 W3 induces small CSPs but significant line-broadening on PEX13 SH3 (**Figure 7.9B star**), which indicates strong binding as well. These observations are supported by ITC experiments, which show K_D's of 43, 88, and 102 μM for W4, W2 and W3, respectively. To relate these affinities, we evaluated the binding of the PEX13 internal FxxxF motif in *trans*, which shows a K_D of 27 μM **(Table 7.3)**, and thus stronger than any of the (di)aromatic peptide motifs of PEX5. Affinities of the other motifs were too weak to be measured accurately by ITC (**Figure 7.9. D, Figure 7.10 C, D, Table 7.3).** Binding of the same PEX5 motifs to PEX14 NTD isoverall much stronger, with a range of 21 nM to 3136 nM (**Figure 7.9D, E, Figure 7.11, Table 7.4**). These values are in agreement with previous studies in different buffer conditions (Gopalswamy et al., in preparation). Of note, amongst the eight (di)aromatic peptide motifs in PEX5, W4 shows the highest relative binding affinity for PEX13 SH3 and the weakest interaction with the PEX14 NTD (**Figure 7.6E, 5D, E**). To investigate the binding mode of PEX5 W4 to PEX13 SH3 in more detail, we crystallized a PEX13 SH3 GS W4 chimera and solved the structure at 2.3 Å by X-ray crystallography (**Figure 7.9F, Table 7.1**). In contrast to the internal FxxxF motif, the binding interface is limited to the core motif driven by hydrophobic interactions and few hydrogen bonds from R183, Y185 and Y188 to the backbone or K304 sidechain **(Figure 7.9G**). Nevertheless, polar backbone interactions involving G335 and K336 as well as coordination from K304 seem to be important since they are conserved from PEX13 FxxxF to PEX5 W4 (**Figure 7.9H, colored lines**). These results show that all (di)aromatic peptide motifs from PEX5 NTD are able to bind the

isolated PEX13 SH3 domain but only W4, W2 and W3 are able to compete with the internal FxxxF motif. The structure of the PEX13 SH3 - W4 complex shows a limited binding interface that lacks the electrostatic interactions that are seen with the PEX13 FxxxF motif (**Figure 7.4**).

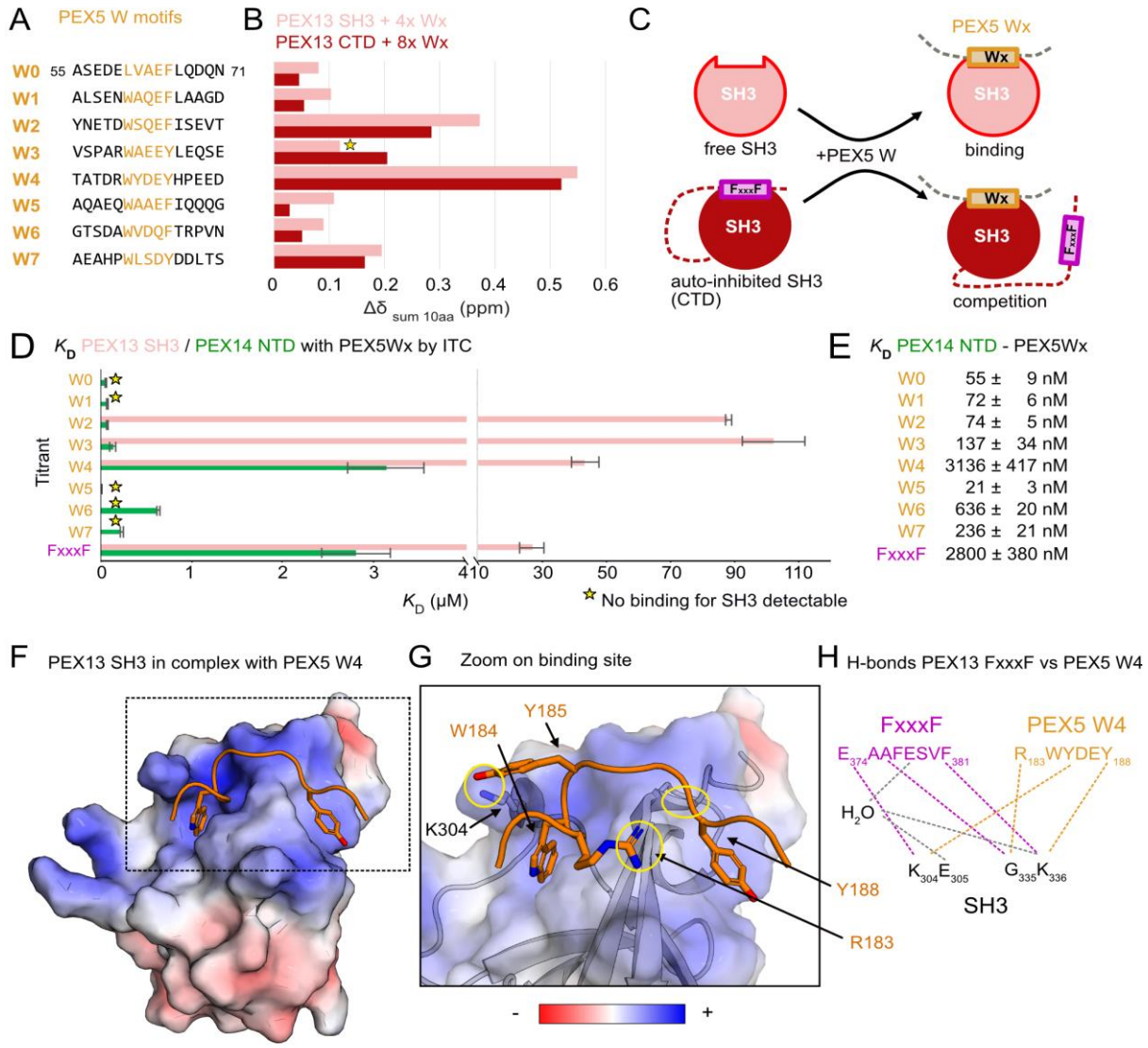

Figure 7.9. Interaction of PEX5 (di)aromatic peptide motifs with PEX13 SH3 or CTD. (A) Overview of PEX5 (di)aromatic peptide motifs. W0 was expressed as PEX5 1-76 while other W motifs were purchased as peptides as listed. (B) Induced Chemical shifts changes of PEX13 SH3 or CTD upon addition of 4x or 8x PEX5 (di)aromatic peptide motifs represented as the sum of 10 involved residues. The star indicates W3 which shows less chemical shift perturbation but extensive line-broadening. (C) Schematic representation of PEX13 SH3 / W peptide binding (top) or PEX13 CTD / W peptide competition. (D) Plot of triplicate K_D values from ITC experiments of PEX13 SH3 or PEX14 NTD with PEX5 W and FxxxF peptide motifs. (E) K_D values from ITC experiments in numbers. (F) Electrostatic surface representation of the PEX13 SH3 GS W4 structure at 2.3 Å. (G) Zoomed view of the W binding site showing polar interactions marked with yellow circles (H) Schematic representation of the conserved hydrogen-bond network of PEX13 SH3 / FxxxF (pink) and PEX5 W4 (orange) interaction. Additional contact sites are marked in gray.

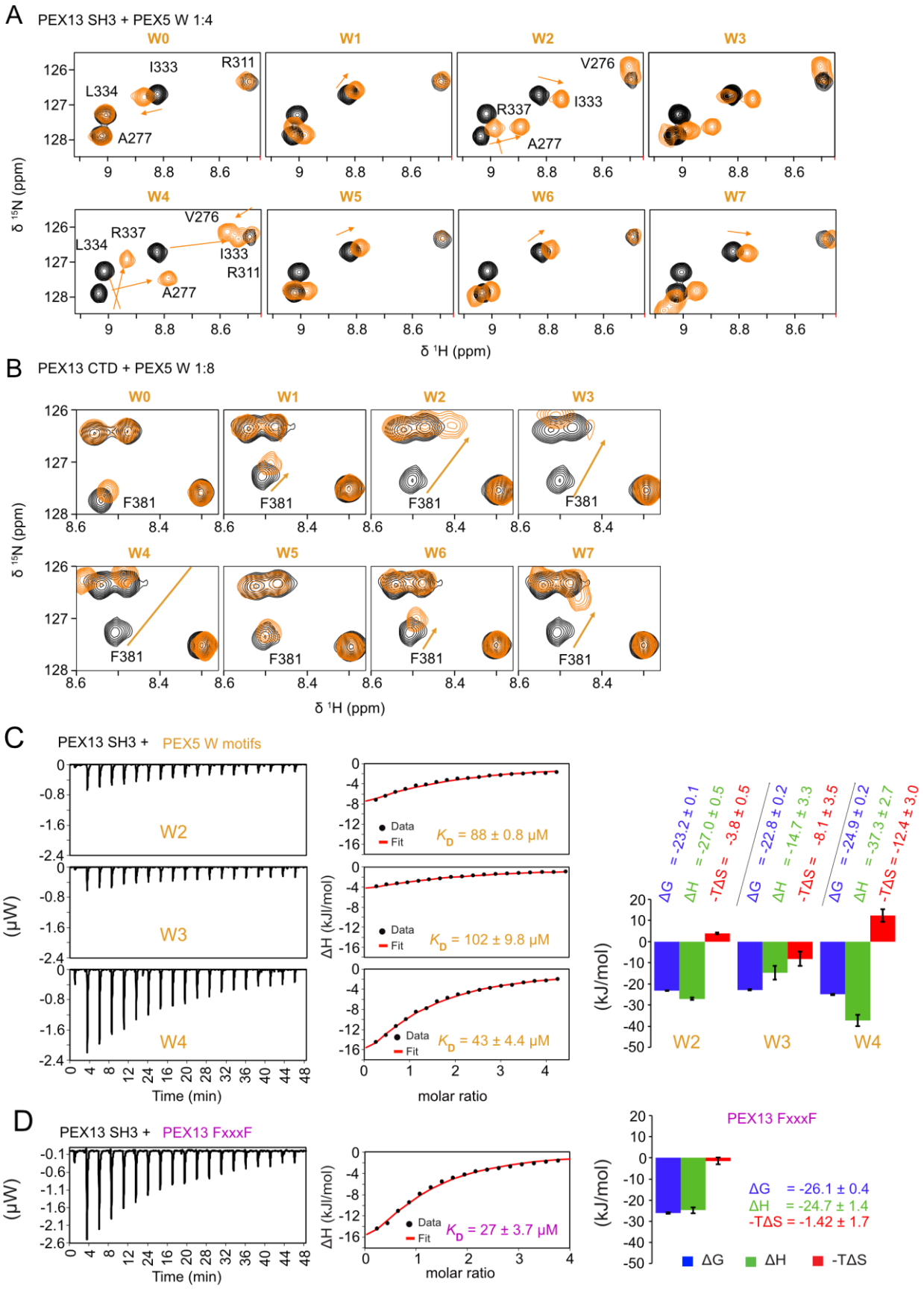
A
PEX13 SH3 + PEX5 W 1:4
W0
W1
W2
W3
W4
W5
W6
W7
I333
R311
L334
A277
R337
V276
δ ¹⁵N (ppm)
δ ¹H (ppm)
B
PEX13 CTD + PEX5 W 1:8
F381
C
PEX13 SH3 + PEX5 W motifs
(µW)
Time (min)
Data
Fit
K_D = 88 ± 0.8 µM
K_D = 102 ± 9.8 µM
K_D = 43 ± 4.4 µM
ΔH (kJ/mol)
molar ratio
ΔG = -23.2 ± 0.1
ΔH = -27.0 ± 0.5
-TΔS = -3.8 ± 0.5
ΔG = -22.8 ± 0.2
ΔH = -14.7 ± 3.3
-TΔS = -8.1 ± 3.5
ΔG = -24.9 ± 0.2
ΔH = -37.3 ± 2.7
-TΔS = -12.4 ± 3.0
(kJ/mol)
D
PEX13 SH3 + PEX13 FxxxF
K_D = 27 ± 3.7 µM
PEX13 FxxxF
ΔG = -26.1 ± 0.4
ΔH = -24.7 ± 1.4
-TΔS = -1.42 ± 1.7

Figure 7.10. (Competitive) binding of PEX13 SH3 and CTD with PEX5 W peptides (A) Overlaid 2D spectra of PEX13 SH3 (black) and with 4x excess of PEX5 W peptides (orange). (B) Overlaid 2D spectra of PEX13 CTD (black) and with 8x excess of PEX5 W peptides (orange). (C) ITC titration of PEX13 SH3 with PEX5 W2, W3 and W4 (D) ITC titration of PEX13 SH3 with FxxxF peptide (350-403)

Table 7.3 Isothermal titration calorimetry of PEX13 SH3 with PEX5 W2, W3, W4 or PEX13 C-terminal peptide (351-403).

Peptide	N	K_D (µM)	ΔG (kJ/mol)	ΔH (kJ/mol)	-TΔS (kJ/mol)
PEX5 W2	1	88.1 ± 0.1	-23.17 ± 0.04	-27.00 ± 0.47	3.81 ± 0.47
PEX5 W3	1	102.2 ± 9.9	-22.80 ± 0.20	-14.65 ± 3.25	-8.14 ± 3.46
PEX5 W4	1	43.4 ± 4.4	-24.93 ± 0.24	-37.30 ± 2.66	-12.38 ± 2.95
PEX13 FxxxF	1	26.9 ± 3.7	-26.13 ± 0.38	-24.73 ± 1.38	-1.42 ± 1.69

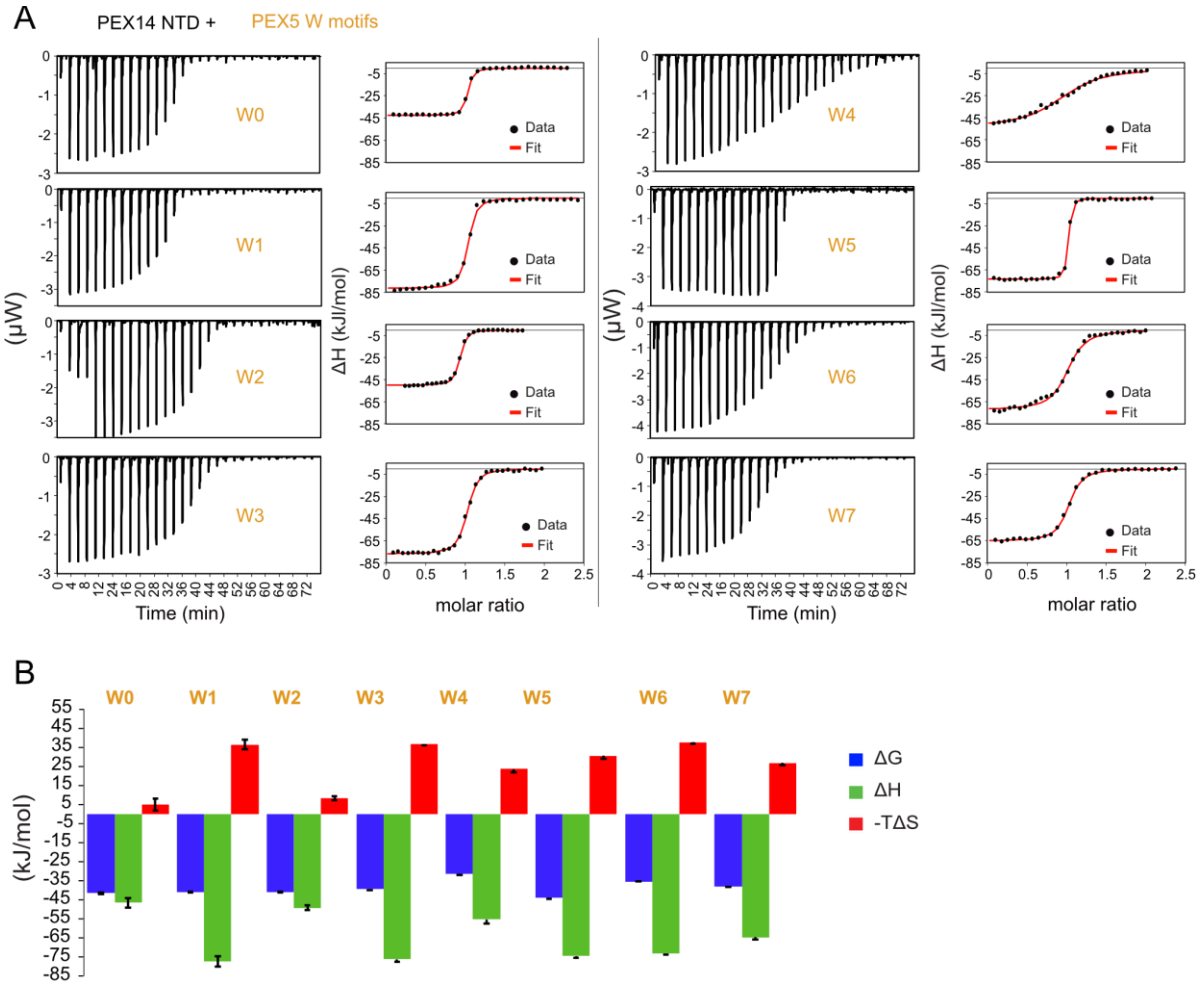

Figure 7.11. Isothermal titration calorimetry of PEX14 NTD/ PEX5 W interactions. (A) ITC titration of PEX14 NTD (1-104) with PEX5 W (-like) motifs. (B) Energetic contribution of ITC titration PEX14 NTD (1-104) with PEX5 W (-like) motifs.

Table 7.4 Isothermal titration calorimetry of PEX14 NTD with PEX5 W-peptide motifs.

Peptide	N	K_D (nM)	ΔG (kJ/mol)	ΔH (kJ/mol)	-TΔS (kJ/mol)
PEX5 W0	1	55.1 ± 0.1	-41.53 ± 0.44	-46.43 ± 2.56	4.91 ± 3.00
PEX5 W1	1	72.3 ± 6.4	-40.83 ± 0.24	-77.27 ± 2.89	36.43 ± 2.64
PEX5 W2	1	73.6 ± 5.9	-40.96 ± 0.18	-49.10 ± 1.13	8.33 ± 1.23
PEX5 W3	1	136.5 ± 34.3	-39.33 ± 0.62	-76.00 ± 1.27	36.63 ± 0.69
PEX5 W4	1	3136.7 ± 417.8	-31.50 ± 0.33	-55.17 ± 2.36	23.70 ± 2.00
PEX5 W5	1	20.5 ± 3.3	-43.96 ± 0.38	-74.47 ± 1.09	30.47 ± 1.49
PEX5 W6	1	636.3 ± 20.22	-35.40 ± 0.07	-73.13 ± 0.58	37.70 ± 0.60
PEX5 W7	1	236.0 ± 21.33	-37.90 ± 0.20	-64.63 ± 1.22	26.77 ± 1.31

The PEX13 FxxxF motif modulates PEX5 interactions in cells

The functional significance of the PEX13 SH3 autoinhibition on PEX5 interactions was then investigated in cellular assays. We expressed various constructs of PEX13 in T-REx PEX13 -/- (KO) cells generated by CRISPR/Cas9 genome engineering (**Figure 7.12A, B**). We tested PEX13-FL and a truncated form which ends after the SH3 domain to exclude any interaction of the C-terminal region (**Figure 7.12A**). Additionally, we generated a FxxxF to AAAAA mutant (FF/A5), which is expected to show none or only weak auto-inhibition. This is confirmed by NMR titrations with a C-terminal peptide harboring this FxxxF to A5 mutation does not induce significant chemical shift perturbations on the SH3 domain (**Figure 7.8D**). A control was generated by replacing the FxxxF motif by the PEX5 W4 motif (**Figure 7.12B**). Immunoprecipitation experiment with cellular PEX5 detected by a PEX5 antibody full-length (FL), revealed visible amounts of PEX13 (**Figure 7.12C, second lane IP**), which were further increased in lysates with the PEX13 1-346 truncation (**Figure 7.12C, third lane IP, D**) and the PEX13 FxxF to A5 mutation (**Figure 7.12D, fourth lane IP, D**). Lower amounts of PEX13 were detected for PEX13 harboring the FxxxF to W4 mutation compared to FL (**Figure 7.12C, fifth lane IP, D**). Noteworthy is the detection of PEX14 in all samples (**Figure 7.12C**). The controls with PEX5 -/- /PEX13 FL and PEX5 FL/PEX13 -/- show no endogenous level of PEX5 or PEX13 respectively. The load was controlled by detection of GAPDH (**Figure 7.12C, lower panel**). Repetitions of PD experiments are shown in **Figure 7.13.**

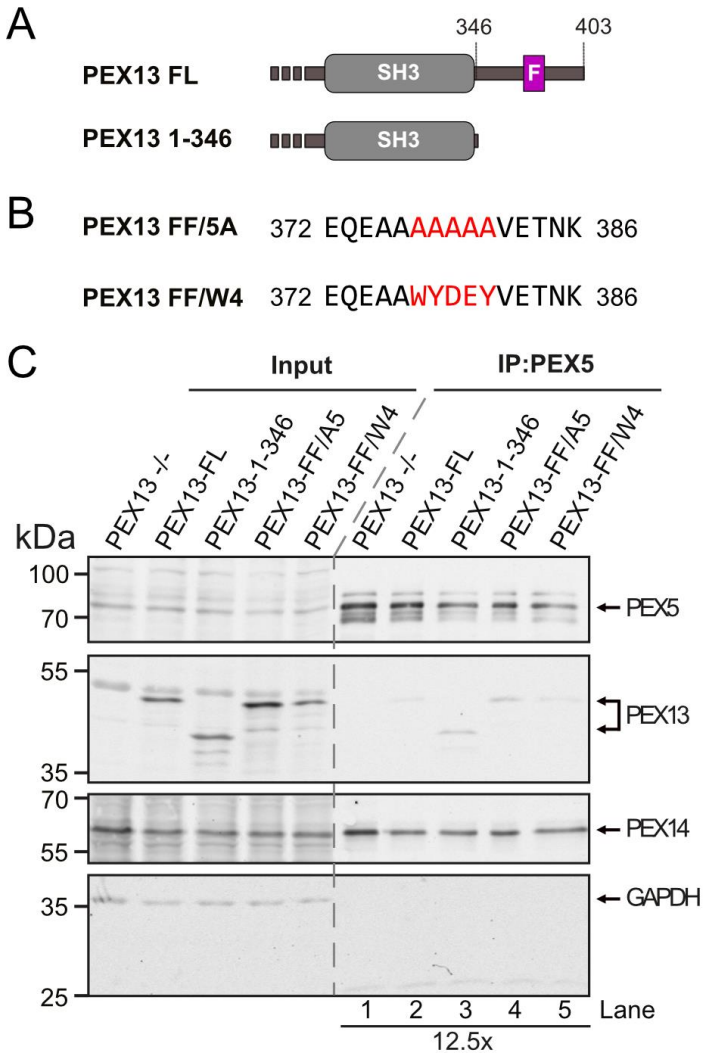

Figure 7.12. PEX13 FxxxF motif modulates PEX5 binding. PEX13 FL, PEX13 1-346 (A) and the two mutations FxxxF to A5 and FxxxF to W4 (B) were expressed in T-REx PEX13 -/- cells. The cell lysates were subjected pull down procedures and affinity purification with PEX5 antibody and analyzed via SDS PAGE and western blot (C).

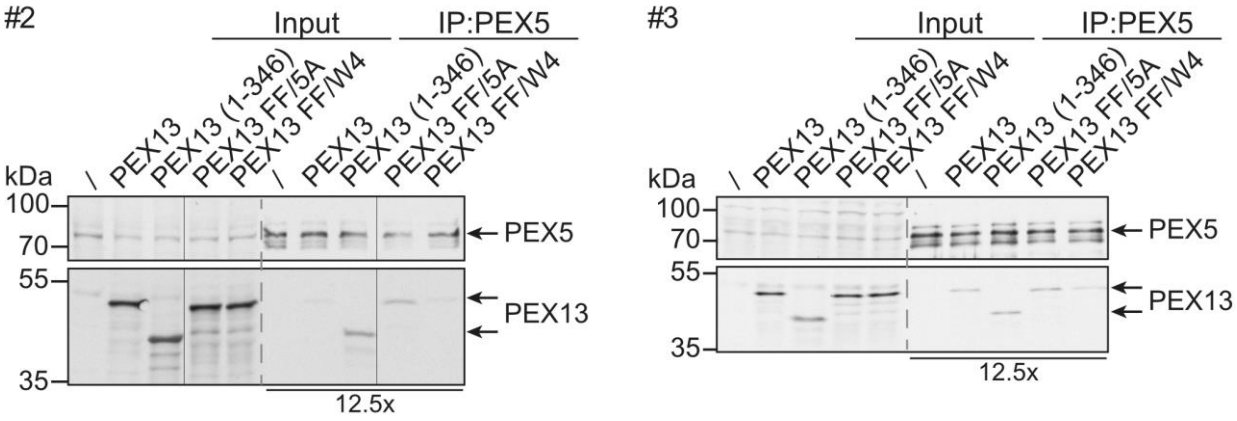

Figure 7.13. Repetitions two (#2) and three (#3) of the PEX13 pull-down experiments. PEX13 FL, PEX13 1-346 and the two mutations FF/A5, FF/W4 (B) were expressed in T-Rex PEX13 KO cells. The cell lysates were subjected pull down procedures and affinity purification with PEX5 antibody and analyzed via SDS PAGE and western blot.

Mutation or truncation of the PEX13 FxxxF motif down regulates PTS1 import

The functional significance of the autoinhibition of the SH3 domain by the proximal FxxxF motif was addressed using a cellular import complementation assay. Here the complementing activity of PEX13 variants was analyzed in a cell-based model using T-REx PEX13 -/- cells with non-functional PTS1 import. The PEX13-deficient cells were transfected with a bicistronic vector encoding for i) eGFP-SKL as marker for peroxisomal PTS1 import and ii) full-length PEX13 or PEX13 truncation and mutation variants.

Cells solely expressing eGFP-SKL showed no import activity indicated by the diffuse distribution of eGFP in the cytosol (**Figure 7.14A**). In contrast, transfection of PEX13 FL results in functional import represented by the congruent punctate pattern of eGFP and co-localization with the peroxisomal marker protein PMP70 (**Figure 7.14B)**. Besides functional and non-functional import, cells with partial import with diffuse and punctate eGFP pattern were observed. These cells were counted as import-competent, as they were capable of importing at least some of the eGFP-PTS1 (**Figure 7.14E**). We first assessed the importance of the full C-terminal region. The re-introduction of a PEX13 truncation (1-264) lacking the SH3 domain and C-terminal region reduced the import capability drastically to 24% **(Figure 7.14C)** compared to wild-type **(Figure 7.14B)**. We further investigated the role of the intrinsic FxxxF motif by truncation or mutation. Interestingly, transfection of PEX13 -/- cells with the PEX13 1-346 truncation lacking the FxxxF motif showed a restored import of 88% (**Figure 7.14D)**. Substitution of the FxxxF motif by polyA (FF/A5), which supposed to have non or reduced inhibition ability, led to an import efficiency of 65% (**Figure 7.14E)**. Complementation experiments with the PEX13 FF to W4 mutant, which is expected to exhibit enhanced autoinhibition, showed a total import efficiency of 61 % **(Figure 7.15**).

These results highlight the importance of the PEX13 C-terminal region including the SH3 domain and the FxxxF motif as regulatory element in PTS1 import. Our cellular assays demonstrate a drastic effect on import upon deletion of the C-terminal region and a modulation of import by the SH3-proximal FxxxF motif.

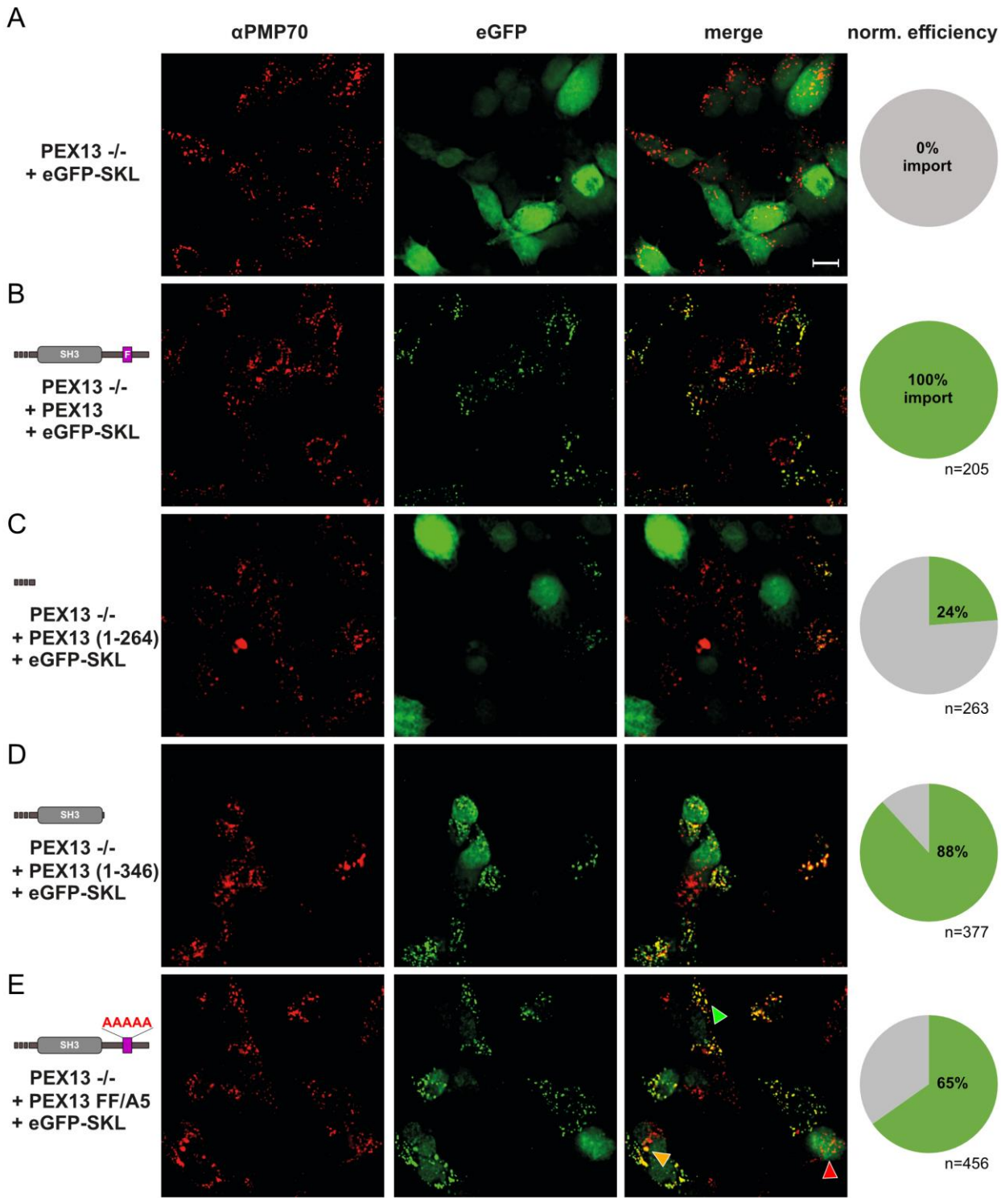

Figure 7.14. PEX13 SH3-proximal FxxxF motif modulates PTS1 import. PEX13-deficient T-REx cells were transfected with different PEX13 truncation and mutation variants as indicated on the left side in (A-E) to monitor the rescue of PTS1 import defect via fluorescence microscopy. Detection of peroxisomal membranes was achieved via PMP70 specific antibody indicated as red congruent punctate patterns (left panels). Scale bar: 10 µm (A) Expression of solely eGFP-SKL did not rescue PTS1 import indicated by a diffuse cytosolic green eGFP signal (middle and right

panels) and served as negative control. (B) Transfection of the cells with PEX13 FL located on a bicistronic vector together with eGFP-SKL led to a congruent punctate pattern colocalizing with PMP70 indicating rescued import. Import efficiencies of other PEX13 variants were normalized to PEX13 FL. (C) Truncation of the full CTD (1-264) had a large effect on import which was determined only for 24% of the cell. (D) Expression of the truncated PEX13 1-346 showed an import efficiency of 88% whereas expression of PEX13 FF to A5 (E) reduced the rescue efficiency to 65%. Besides the phenotype of import (E, green arrow) and non-import (E, red arrow), a third phenotype with partial import (E, orange arrow) indicated by diffuse and punctuate eGFP signal was observed.

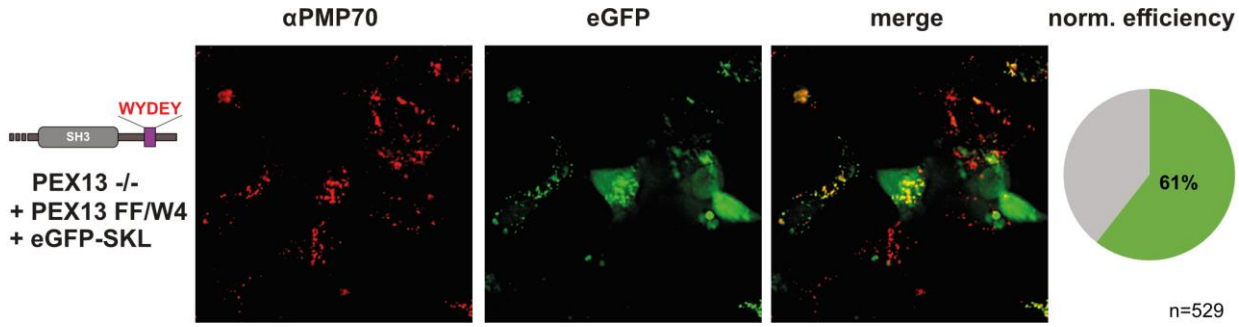

Figure 7.15. Complementation assay with PEX13 FF/W4 mutant. The W4 mutation of the FxxxF motif reduces import efficiency to 61%.

Discussion

Here, we present a comprehensive structural and biophysical analysis of the C-terminal domain of human PEX13 and its role in peroxisome biogenesis and function. We discovered the presence of an internal FxxxF peptide motif that mediates an unexpected autoinhibition with the flanking PEX13 SH3 domain. We further investigated binding sites between the core components of the human peroxisomal docking and translocation machinery PEX13, PEX14 and PEX5. Our results identify a novel interaction of the human PEX13 CTD with the PEX14 NTD, that is mediated by the FxxxF motif. Importantly, we demonstrate that the SH3 domain of human PEX13 SH3 shows binding to (di)aromatic peptide motifs in the PEX5 NTD, similar to yeast PEX13, demonstrating evolutionary conservation of these interactions. Moreover, these interactions are further modulated by the SH3-proximal FxxxF motif in human PEX13. The internal FxxxF motif shows significant effects on PTS1 import efficiency, and shows reduced import in a cell-based assay upon deletion or mutation of the FxxxF motif. This suggests that the modulation of the PEX5, PEX13 and PEX14 interactions by the FxxxF motif play an important role in fine-tuning peroxisome biogenesis.

SH3 domains are commonly known to bind PxxP peptides to mediate complex formation as adaptor proteins (Birge *et al*, 1996) acting in signal transduction (Schlessinger, 1994). The ability of FxxxF

motif binding has not been reported so far, except for the yeast homologue Pex13 SH3 which binds a di-aromatic peptide motif of Pex5 (WxxxF/Y) (Barnett *et al.*, 2000; Bottger *et al*, 2000; Douangamath *et al.*, 2002). . It seems that the binding of (di)aromatic peptide motifs is a unique and conserved feature of PEX13 SH3 domains since there is no other case reported yet. However, binding of yeast the Pex5 WxxxF/Y (similar to W6 in human) motif has been mapped to the residues L308, F310, L333, M334, A335, V352, T354 and N356 which are mostly located in β1 and β2 of the yeast Pex13 SH3 domain (Douangamath *et al.*, 2002) whereas binding of (di)aromatic peptide motifs to human PEX13 SH3 was mapped to β1 and β5 **(Figure 7.4)**. Moreover, residues of the binding surface in yeast are poorly conserved to human **(Figure 7.2D**). This demonstrates that of important functions can be conserved over sequence evolution. In both organims, the Pex13 SH3 domain acts as regulator of the import mechanism, rather than in a cellular signal transduction. This is somewhat reminiscent of the SNX9 SH3 domain which regulates nuclear import of pSMAD3 by binding of the HIV TAT peptide, which is neither proline rich nor (di)aromatic (YGRKKRRQRRR) (Kang *et al*, 2017). It will be interesting to see if further SH3 domains exhibit affinity for unrecognized peptide motifs and a potential role in protein import.

Several interactions involving the PEX13 SH3, PEX14 NTD and PEX5 (di)aromatic peptide motifs have been reported for the yeast system (Barnett *et al.*, 2000; Douangamath *et al.*, 2002; Erdmann & Blobel, 1995). These interactions stabilize a ternary complex of Pex13, Pex14 and Pex5, which involves distinct interactions of Pex5 WxxxF/Y motifs with the Pex13 SH3 and the Pex14 NTD, respectively (Barnett *et al.*, 2000; Bottger *et al.*, 2000; Douangamath *et al.*, 2002). This complex is stabilized by an additional interaction of a Pex14 PxxP motif binding to the SH3 domain. Our results show that the specific details of these interactions are distinct in the human systems, where the PEX14/PEX13 interaction is mediated by an internal FxxxF motif of PEX13. Moreover, the interaction of the PEX13 SH3 domain with PxxP poly-proline peptide motifs is distinct and shows non-canonical contacts, and thus no binding with the PxxP motif in the N-terminal region of PEX14 is observed in the human system (**Figure 7.7B)**, in contrast to yeast.

Not unexpected, the PEX13 FxxxF motif binds to the same binding interface on the PEX14 NTD and is structurally similar to the recognition of the (di)aromatic PEX5 peptide motifs and an FxxxF motif in the peroxisomal membrane protein transport factor PEX19 (Neufeld *et al.*, 2009). The recognition of PEX5 (di)aromatic peptide motifs by the PEX13 SH3 on a surface that is opposite and non-overlapping with the PxxP binding site is conserved from yeast to human. However, a

notable difference is that the interaction in human is modulated by the autoinhibition mechanism of PEX13. However, there is no clear evidence if the motif up or down regulates the import efficiency since truncation of the motif as well as mutation to a weaker or stronger binding reduces import capability in our cell assay which hints to a fine tuning mechanism. The PEX13 SH3 domain itself on the contrary seems to be essential for successful peroxisomal import, since truncation of the full C-terminal region leads to a drastic reduction in import capabilities. Already Gould *et al.* (1996a) described the importance of the PEX13 SH3 domain and proposed an interaction toward PEX5 NTD. However, in this study a truncated and probably unstable construct of the SH3 domain (277-333) was used, which might rationalize why PEX5 binding was not detected in pulldown experiments. Another possible explanation is the low affinity of the interaction, as our data show that only PEX5 W4, W2 and W3 motifs show significant binding to the PEX13 SH3 domain (**Fig 5D; Figure 7.10C**). A later study mapped the PEX13/PEX5 interaction to the N-terminal region of PEX13 and the WxxxF/Y motifs W2, W3 and W4 from PEX5 by GST pulldown experiments (Otera *et al.*, 2002). So far, we could not verify the proposed interaction of the PEX5 N-terminal region as the PEX13 NTD is prone to aggregate in our hands. Interestingly, Krause et al. showed that the PEX13 N-terminal region mediates homo-oligomerization (Krause *et al.*, 2013). Thus future studies should aim to characterize this additional potential PEX5/PEX13 interaction, possibly in a membrane-mimicking environment, considering that PEX13 is an integral membrane protein (Barros-Barbosa *et al.*, 2019).

The network of interactions identified in our study show that PEX13, PEX5 and PEX14 interactions are modulated by binding of (di)aromatic peptide motifs with overlapping and thus competing binding sites. This implies that the formation of a simple ternary complex as observed in yeast is not possible and shows the evolutionary increased complexity, where functional activity is not regulated by distinct binding motifs but by relative affinity and avidity effects. This may be reflected by the increase of (di)aromatic peptide motifs from three to eight in PEX5 (Kerssen *et al.*, 2006; Otera *et al.*, 2002; Saidowsky *et al.*, 2001). The different affinities between PEX13/PEX14 **(Figure 7.6)**, PEX13/PEX5 and PEX14/PEX5 (**Figure 7.9**) hint a potential sequential binding model. The strong affinity of the PEX5 NTD towards PEX14 NTD is required for the initial docking event in the cytosol, which may be later replaced by PEX13 SH3/PEX5 NTD and PEX13 FxxxF/PEX14 NTD at the membrane. This may be depending on the presence of a high molecular excess of PEX13 due to the lower binding affinity towards PEX5 and PEX14. In line with this, a

high molecular mass complex containing only PEX13 has been reported (Reguenga *et al.*, 2001). Handing PEX5 from PEX14 over to PEX13 could happen in two different steps of peroxisomal import; i) as part of the docking event before cargo translocation into the peroxisomal matrix, or ii) to enhance cargo release inside the lumen. In this case, PEX13 could play a similar role as Pex8 in yeast, considering that a human homologue of Pex8 has not been identified. Involvement in docking would require the location of the SH3 domain faced to the cytosol which was first proposed by Gould *et al.* (1996a). However, more recent studies from rat liver locate the SH3 domain inside the peroxisome (Barros-Barbosa *et al.*, 2019; Reguenga *et al.*, 2001), which is consistent with the cargo release hypothesis. Taking observations from Otera *et al.* (2002) of a PEX13/PEX5 interaction via PEX13 N-terminal region in account, PEX13 could even act in docking via its N-terminal domain and in cargo release via its C-terminal domain.

In conclusion, our results highlight the importance of the PEX13 SH3 domain for peroxisomal import and demonstrate a regulatory function of the newly identified FxxxF motif considering autoinhibitory interactions with the SH3 domain. However, the interaction network and competitive interactions (di)aromatic peptide motifs of varying affinities in solution is likely further modulated by differential localization and concentration of PEX5, PEX13 and PEX14 at the membrane and the presence of cargo bound to the PEX5 TPR domain. Thus, future studies should analyze the PEX interaction network and complexes at the membrane.

Experimental procedures

Molecular cloning

For recombinant expression in bacteria, the full length genes of human PEX13 (UniProtKB no. Q92968), human PEX14 (UniProtKB no. O75381) and human PEX5 (UniProtKB no. P50542) were optimized according to the codon usage of E. coli and synthesized by IDT (IDT Europe GmbH, Germany). These sequences were used as templates to generate PEX13 fragments PEX13 SH3 (261-346), FxxxF (261-383), CTD (261-403), 2GSc-FxxxF (chimera), GSc-W4 (chimera) as well as the PEX14 fragments PEX14 NTD (1-104), NTD_{long} (1-113) and PEX5 W0 (1-76) in a His_6-SUMO-tag modified pETM13 vector (pETM13S). The cloning was done using site directed ligase independent mutagenesis (SLIM) (Chiu, 2004) in an extended version to implement inserts using the same fashion of short and tail primers. The vector backbone and inserts were amplified

by polymerase chain reaction (PCR) amplification using the according short and tail primers (**Table 7.5**) to generate overlaps with sticky ends. The backbone amplificates and inserts were mixed with a 5-fold molar excess of insert and annealed during the SLIM cycle (Chiu, 2004). The annealed vector was directly transformed into DH10b cells for DNA amplification.

The same protocol was used to create the PEX13 constructs FL (1-403), trunc1 (1-264), trunc2 (1-346), trunc3, FxxxF to A5 and FxxxF to W4 substitution in the bi-cistronic mammalian expression vector pIRES2-EGFP. Primer are listed in **Table 7.6**.

Table 7.5. Primer list for cloning into pETM13S vector.

Construct	Bound-aries	Forward short	Forward tail	Reverse short	Reverse tail
pETM13S	backbone	GGAAGC TGAGTTG GCTGCTG CCAC	TAACAAAGCCC GAAAGGAAGCT GAGTTG	GCGGTGA GCCTCAAT AATATCGT TATCC	GCCACCAATCTGTT CGCGGTG
PEX13 SH3	261-346	GTTACTG ATTCCAT CAACTG GGCCTC	GAACAGATTGG TGGCGTTACTG ATTCCATCAACT GGGCCTC	TGAGCTTT CTACTGTT TTACGTCC TTTGC	TTTCGGGCTTTGTT ATGAGCTTTCTACT GTTTTACGTCCTTT GC
PEX13 FxxxF	261-383	GTTACTG ATTCCAT CAACTG GGCCTC	GAACAGATTGG TGGCGTTACTG ATTCCATCAACT GGGCCTC	CTCAACG AAGACGG ATTCGAA GGC	TTTCGGGCTTTGTT ACTCAACGAAGAC GGATTCGAAGGC
PEX13 CTD	261-403	GTTACTG ATTCCAT CAACTG GGCCTC	GAACAGATTGG TGGCGTTACTG ATTCCATCAACT GGGCCTC	TTTGGTAC CTCACAA ATCCTGTT TTTCACCG	TTTCGGGCTTTGTT ATTTGGTACCTCAC AAATCCTGTTTTTC ACCG
PEX13 2GSc FxxxF	chimera	TAACAA AGCCCG AAAGGA AGCTGA GTTG	GGCGGTGGAGG CAGCGGAGGTG GAGGAAGCGAC GAGCAAGAAGC CGCCTTCGAAT CCGTCTTCGTTG AGTAACAAAGC CCGAAAGGAAG CTGAGTTG	TGAGCTTT CTACTGTT TTACGTCC TTTGC	CTCAACGAAGACG GATTCGAAGGCGG CTTCTTGCTCGTCG CTTCCTCCACCTCC GCTGCCTCCACCGC CTGAGCTTTCTACT GTTTTACGTCCTTT GC
PEX13 Gsc PEX5 W4	chimera	TAACAA AGCCCG AAAGGA AGCTGA GTTG	ACCGCGACCGA TCGCTGGTATG ATGAATATCAT CCGGAAGAAGA TTAACAAAGCC CGAAAGGAAGC TGAGTTG	TGAGCTTT CTACTGTT TTACGTCC TTTGC	ATCTTCTTCCGGAT GATATTCATCATAC CAGCGATCGGTCG CGGTGCTGCCTCCA CCGCCTGAGC

PEX14 NTD	1-104	ATGGCTAGCAGCGAACAGGCC	GAACAGATTGGTGGCAATGGCTAGCAGCGAACAGGCC	ACTACCCGCCGGAGAATACGG	TTTCGGGCTTTGTTAACTACCCGCCGGAGAATACGG
PEX14 NTD long	1-113	ATGGCTAGCAGCGAACAGGCC	GAACAGATTGGTGGCAATGGCTAGCAGCGAACAGGCC	CGCCAGTGCGCCAATCG	TTTCGGGCTTTGTTACGCCAGTGCGCCATAATCG
PEX5 W0	1-76	ATGGCAATGCGGGAGCTG	GAACAGATTGGTGGCATGGCAATGCGGGAGCTG	GGACACAAGGGGTGCATTCTG	TTTCGGGCTTTGTTAGGACACAAGGGGTGCATTCTG

Table 7.6. Primer list for cloning into bi-cistronic vector pIRES2 - GFP.

Time	Bound-aries	Forward short	Forward tail	Reverse short	Reverse tail
pEB100 pIres 2	backbone	GCGGGCCCGGGATCC	GCAGTCGACGGTACCGCGGGCCCGGGATCC	CTGAGTCCGGTAGCGCTAGC	GCTTGAGCTCGAGATCTGAGTCCGGTAGCGCTAGC
PEX13 FL	1-403	ATGGCGTCCCAGCCG	ATCTCGAGCTCAAGCATGGCGTCCCAGCCG	TTAAAGATCTTGCTTTTCTCCATCTTTCCC	GGTACCGTCGACTGCTTAAAGATCTTGCTTTTCTCCATCTTTCCC
PEX13 trunc1	1-264	GCAGTCGACGGTACCGC	GAAGTAACAGACAGCTAAGCAGTCGACGGTACCGCGG	ATCACTGTGAGTAGACAATAGTTTCC	TTAGCTGTCTGTTACTTCATCACTGTGAG
PEX13 trunc2	1-346	GCAGTCGACGGTACCGC	GAAGTAACAGACAGCTAAGCAGTCGACGGTACCGCGG	ACTTGATTCCACCGTTTTCCTAC	GGTACCGTCGACTGCACTTGATTCCACCGTTTTCCTAC
PEX3 FxxxF / 5A	1-403 +mutation	GTTGAAACTAATAAGGTTCCAGTTGCAC	GCAGCCGCTGCAGCTGTTGAAACTAATAAGGTTCCAGTTGCAC	GGCAGCTTCCTGTTCATCCAAAG	AGCTGCAGCGGCTGCGGCAGCTTCCTGTTCATCCAAAG
PEX3 FxxxF / W4	1-403 +mutation	GTTGAAACTAATAAGGTTCCAGTTGCAC	TGGTACGACGAATACGTTGAAACTAATAAGGTTCCAGTTGCAC	GGCAGCTTCCTGTTCATCCAAAG	GTATTCGTCGTACCAGGCAGCTTCCTGTTCATCCAAAG

Protein sample preparation

PEX constructs were transformed into *Escherichia coli* BL21 (DE3) cells and expressed in LB or isotope-enriched M9 minimal medium. Uniformly ^{15}N or ^{15}N, ^{13}C labeled proteins were expressed in H_2O M9 minimal medium supplemented with 50 µg/ml kanamycin, 1g/liter [U-^{15}N] ammoniumchloride and 2g/liter hydrated [U-^{13}C] glucose as the sole sources of nitrogen and carbon, respectively. After transformation, single colonies were picked randomly and cultured in the medium of choice overnight at 37°C. On the next day, cultures were diluted to an optical density of 600 nm (OD_{600}) of 0.1, and grown up to a OD_{600} of 0.4-0.6. Protein expression was induced with 0.5 mM IPTG and was carried out for 4h at 37°C.

The cells were harvested by centrifugation at 6000 g for 20 min at 4°C. For protein purification the cell pellets were resuspended in lysis buffer (50mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole) substituted with lysozyme (from chicken), DNAse and protease inhibitor mix (Serva, Heidelberg, Germany) and lysed by pulsed sonication (10 min, 40% power, large probe, Fisher Scientific model 550) followed by centrifugation at 38000 g for 45 min. All proteins were purified using gravity flow Ni-NTA (Qiagen, Monheim, Germany) affinity chromatography. The supernatant of the lysis was incubated with Ni-NTA beads (2ml/1l culture) for 20 min at 4 °C while rotating. Subsequently to incubation, the protein-laden beads were washed with 7 CV high salt buffer (50mM Tris pH 7.5, 750 mM NaCl, 20 mM imidazole) and 10 CV wash buffer (50 mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole). The elution was performed with 3-5 CV elution buffer (50 mM Tris pH 7.5, 300 mM NaCl, 500 mM imidazole). Dialysis and SUMO cleavage was executed over night at 4 °C in 20 mM Tris pH 7.5, 150 mM NaCl. Further purification was done with a reverse Ni-NTA column where the flow through containing the cleaved protein of interest was collected and concentrated for size exclusion chromatography using a Superdex S75, 16/600 (GE Healthcare, Rosenberg, Sweden). The size exclusion chromatography as last step of the purification was performed directly in NMR buffer.

Peptide preparation

All 15-mer peptides used in this study were purchased from PSL (Peptide Specialty Laboratories GmbH, Heidelberg, Germany). Peptides delivered in TFA salt were dissolved in H2O and the pH

adjusted to 7.5 using 1 M NaOH. If used for ITC, the peptides were dialyzed against ITC buffer (20 mM Tris pH 7.5, 50 mM NaCl).

PEX5 W1	ALSENWAQEFLAAGD
PEX5 W2	YNETDWSQEFISEVT
PEX5 W3	VSPARWAEEYLEQSE
PEX5 W4	TATDRWYDEYHPEED
PEX5 W5	AQAEQWAAEFIQQQG
PEX5 W6	GTSDAWVDQFTRPVN
PEX5 W7	AEAHPWLSDYDDLTS

NMR spectroscopy

NMR data were collected on Bruker Avance III spectrometers operating at 500, 600, 800, 900 or 950 MHz, equipped with cryogenic probes. The sequential assignment of backbone resonances for PEX13 CTD was performed based on heteronuclear experiments such as ^{1}H-^{15}N-HSQC, HNCA, HN(CO)CA, CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, HN(CA)NNH and H(NCA)NN (Sattler M *et al.*, 1999; Weisemann *et al.*, 1993). {^{1}H}-^{15}N heteronuclear NOE (hetNOE) experiments (Farrow *et al.*, 1994) were performed using the pulse sequence hsqcnoef3gpsi (Bruker, Avance version 12.01.11) with a 4.5 s interscan delay. NOE values are given simply by the ratio of the peak heights in the experiment with and without proton saturation (hetNOE = I_{sat}/I_0) (Renner *et al.*, 2002). ^{15}N HSQC-based T_1 and T_2 experiments used sequences developed from (Farrow *et al.*, 1994) with water-control during the relaxation period in the T_1 sequence using a cosine-modulated IBURP-2 pulse (Gairí *et al.*, 2015) and modifications in the T_2 sequences based on (Lakomek *et al.*, 2012). For both T_1 and T_2 experiments 8 time points with delays of 80, 160, 240, 320, 400, 64, 800, 1000ms (T_1) and 14.4, 28.8, 43.2, 57.6, 72.0, 86.4, 100.8, 115.2ms (T_2) were measured respectively. NMR-Spectra were processed using Topspin (Bruker Biospin, Rheinstetten, Germany) or NMRPipe (Delaglio *et al.*, 1995) and analyzed using CcpNMR Analysis 2.4.2 (Vranken *et al.*, 2005).

All NMR experiments were performed at 298°K. PEX13 and PEX14 spectra were recorded in 20 mM Tris pH 7.5, 50 mM NaCl and 50 mM NaP pH 6.5 and 100 mM NaCl respectively.

For all titration experiments a reference protein concentration 100 µM was used. For protein/protein titrations, every titration point was prepared as individual sample to avoid dilution effects. Protein/peptide titrations such as titration of PEX5 (di)aromatic peptide motifs with high concentrated peptides (10 -15mM) were performed in a single NMR tube. Ligands were added

with increasing concentrations up to an excess of 8 fold. The chemical shift perturbation ($\Delta\delta_{avg}$) was calculated by using formula $\Delta\delta_{avg}=[(\Delta\delta_H)^2+(\Delta\delta_N*0.159)^2]^{0.5}$. Chemical shift perturbations in **Figure 7.9B** are illustrated with the sum of 10 residues which are affected upon (di)aromatic peptide motif binding. NMR chemical shift assignments are available at the BMRB, accession code: 51336.

Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) measurements were performed as triplicates at 25°C using a MicroCal PEAQ-ITC (Malvern Instruments Ltd. U.K) calorimeter. Buffer conditions were 20 mM Tris pH 7.5, 50 mM NaCl. For all titrations a titrant dilution control experiment was performed and subtracted before the data were fitted to a one-site binding model using the Malvern Analysis software.

PEX 13 SH3 at a concentration of 35-48 µM was titrated with PEX5 W peptides or PEX5 1-76 (W0) at concentration of 0.7-0.6 mM and PEX13 350-403 (FxxxF) at a concentration of 0.9-1 mM. PEX14 (1-104) at a concentration of 20-60 µM was titrated with PEX5 W peptides or PEX5 W0 at concentration of 0.4 to 0.6 mM, PEX13 FxxxF at a concentration of 0.46mM and with PEX13 CTD at a concentration of 1 mM. The concentration of PEX14 was corrected with the fit, since it cannot be accurate measured at 280 nm owing to the extinction coefficient of only 1490.

ITC data are summarized in Tables 1, 2 and 3.

X-ray crystallography

All crystals were grown using the vapor diffusion sitting drop method in 96 well plates. Therefore, the proteins were purified in 5 mM Tris pH 7.5, 50 mM NaCl and later 1:1 diluted (200 nl) with crystallization buffer which were supplied from NeXtal. All proteins were crystallized at a concentration of 20mg/ml. Apo PEX 13 SH3 domain was crystallized in 0.01 M Zinc chloride, 0.1 M sodium acetate pH 5 and 20 % (w/v) PEG6000 (PACT suite, A12). PEX13 SH3 2GSc FxxxF chimera was crystallized in 0.2 M sodium chloride, 0.1 M Bis-Tris pH 6.5 and 25% PEG3350 (ClassicsII suite, F11). PEX13 SH3 GSc PEX5 W4 chimera was crystallized in 0.2 M sodium sulfate, 0.1 M Bis Tris propane pH 6.5 and 20% (w/v) PEG3350. Cryoprotectant for all proteins was 25% (v/v) ethylene glycol (PACT suite, F8). Data of apo PEX13 SH3 and PEX13 SH3-2GCs-FxxxF chimera were collected on SLS beamline X06DA (Paul-Scherer-Institute, Villingen, Ch)

while data of PEX13 SH3-GSc-W4 chimera was collected on beamline P11 at PETRA III (EMBL, Hamburg, DE).

Collected data were processed using CCP4i2 suite (Potterton *et al.*, 2018; Winn *et al.*, 2011) using XDS package (Kabsch, 2010) for data reduction, aimless for data scaling, MOLREP (Vagin & Teplyakov, 2010) or PHASER (McCoy *et al.*, 2007) for molecular replacement, COOT (Emsley & Cowtan, 2004) for model building and REFMAC5 (Murshudov *et al.*, 2011) for refinement. Refined structures were uploaded to wwPDBdeposition using pdb_extract (Yang *et al.*, 2004).

Size exclusion chromatography - static light scattering (SEC-SLS)

SLS on PEX13-CTD was done using an OmniSEC Resolve and Reveal device (Malvern Panalytics. Malvern, Uk) equipped with a Superdex 75 increase 10/300 GL column (Cytiva). First 70 µl 2 mg/ml BSA standard (column calibration) and then 70 µl of 2 mg/ml PEX13-CTD in NMR/ITC buffer (20 mM Tris pH 7.5, 50 mM NaCl) was passed though the column with a constant flow of 0.3 ml/min. The concentration was monitored via absorption at 280nm and the refractive index. The molecular weight was calculated with the RALS signal using Omnisec software (version 11.01, Malvern Panalytics, Malvern, Uk).

Computational Modelling

The structure of PI3K SH3 domain in complex with a PxxP ligand (PDB ID: 3I5R) was identified as similar to PEX13 SH3 domain by sequence search. Both domains were aligned and the PxxP ligand was copied to the PEX13 SH3 structure and subsequently mutated to PEX13 PxxP (TRVPPPIL) using Maestro (Schrodinger suite). The energy of the complex was then minimized using OPLS2005 force field and letting all residues in a radius of 4 Å freely rotate

Multiple sequence alignments

Multiple sequence alignment of PEX13 SH3 with other human SH3 domains. First, a RCSB PDB databank search for structures of human SH3 domains was performed. Eight human SH3 domains from different proteins as well as yeast Pex13p SH3 domain were selected (**Supplementary Table 4**). Then the according sequences including the ± 10 flanking amino acids were selected and

with the sequence from the human PEX13 SH3 aligned using clustalΩ (https://toolkit.tuebingen.mpg.de/tools/clustalo) and visualized using Jalview (version 2.11.2.0)

Multiple sequence alignment of PEX13 with mammalian sequences. Mammalian sequences which are similar to PEX13 were found using PSI-BLAST from ncbi blastp (https://blast.ncbi.nlm.nih.gov/) selecting the non-redundant protein sequences database with the organism restricted to mammals. The maximal target sequences were adjusted to 250 and the run started with preselected standards. From the hits, 186 sequences with a query coverage of minimum 90% were selected and aligned using clustalΩ (https://toolkit.tuebingen.mpg.de/tools/clustalo) before being visualized using ConSurf web server (Ashkenazy *et al.*, 2016; Ashkenazy *et al.*, 2010; Celniker *et al.*, 2013) (https://consurf.tau.ac.il/).

Cell culture

T-REx^{TM}293 (Invitrogen, USA) and T-REx293 PEX13KO cells were grown in Dulbecco's Modified Eagle's Medium high glucose (DMEM) supplemented with 10 % fetal calf serum, 4 mM L-glutamine, 100,000 U/l penicillin and 100 mg/l streptomycin at 37°C and 8.5% CO_2. For cell transfections X-tremeGENE HP DNA Transfection Reagent (Roche, Germany) was used according to the manufacturer's instructions.

Fluorescence microscopy

To perform immunofluorescence microscopy cells were seeded on coverslips to appropriate density and fixed with 3% formaldehyde/D'PBS for 20 min. After membrane permeabilization with 1% Triton X-100/D'PBS for 5 min cells were incubated for 30 min in the primary antibody aPMP70 (rabbit, 1:500, Invitrogen) in D'PBS supplemented with 1% BSA. The incubation with the secondary antibody goat arabbit IgG (H+L) Alexa Fluor 594 (Invitrogen) was done for 10 min under light protection. Cells were mounted on glass slides with Mowiol 4-88 (Calbiochem, USA) supplemented with DAPI. Imaging was performed using the Axioplan 2 (Zeiss).

The quantification of import-competent cells was performed manually by cell counting in randomly taken images.

Quantification of rescued import was done over three biological replicates in total numbers. However, the import efficiency of PEX13 FL with 81% was normalized to 100% and import efficiencies from PEX13 variants were normalized to FL.

Co-immunoprecipitation

To study interactions between PEX5 and PEX13 dynabeads were coupled with an aPEX5 antibody (mouse, Cizmowski et al., 2011) using the dynabeads™ Antibody Coupling Kit (Invitrogen, USA) according to the manufacturer's instructions.

Cells were seeded on 10 cm dishes and transfected with different PEX13 truncations or mutations. 48 h after transfection cells were incubated in IP lysis buffer (25 mM Tris/HCl pH 7.4, 150 mM NaCl, 1 % NP-40, 5 % glycerol, cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche), 25 µg/ml DNase) for 15 min on ice. After removal of the cell debris by centrifugation (13,000 rpm, 5 min, 4°C) equal amounts of lysates were incubated with dynabeads on a rotating disk for 1 h at 4°C. Next the dynabeads were washed three times with lysis buffer and bound proteins were eluted with 0.1 M Glycin pH 2.8. Samples were collected and analyzed by SDS-PAGE and immunoblotting using the following antibodies: rabbit aPEX5 (1:5000, Fodor et al., 2015), rabbit aPEX13 (1:1000, Proteintech, Germany), chicken aPEX14 (1:1000, Ruhr-University Bochum), mouse aGAPDH (1:7500, Proteintech, Germany). Band intensities on immunoblots were quantified using densitometry (ImageJ, NIH).

Data deposition and availability

NMR chemical shift assignments are available at the BMRB, accession code: 51336. Structural coordinates and restraint files for the human PEX13 SH3 domain, alone, and chimeric constructs with the FxxxF and W4 peptide motifs are available at the PDB, accession codes 7Z0I, 7Z0J and 7Z0K.

Author Contributions

MS, SG and RE designed the study. SG performed protein expression, biophysical and NMR experiments. SG and KZ performed X-ray crystallography, SG, KZ and GP analyzed crystallographic data. JO performed pulldown and complementation assays. MS and SG wrote the manuscript, all authors commented and approved the manuscript.

Chapter 8:

Conclusions and Outlook

Conclusions and Outlook

This thesis reports on the molecular mechanisms involved in peroxisomal matrix import and biogenesis described by structural, biochemical and biophysical studies. The peroxisomal matrix import can be described as sequential events of cargo recognition, docking, translocation and release. The focus of this work was to characterize structural features and binding interfaces of the peroxins PEX5, PEX14 and PEX13. To this end, solution state NMR spectroscopy was applied in combination with X-ray crystallography, biophysical and biochemical methods.

NMR analysis of the intrinsically disordered domain (NTD) of the soluble receptor PEX5 identified a number of amphipathic α-helical regions, which partly overlap with (di)aromatic peptide motifs (WxxxF/Y) and show weak interactions with membrane mimics. Noteworthy, not only PEX5 NTD but also PEX14-NTD is found to weakly interact with bicelles with a surface that partially overlaps with the WxxxF/Y binding site. However, the binding affinity of the PEX5 NTD and PEX14 NTD is not altered in the presence of a constant bicelle concentration. This hints to a regulatory mechanism to prevent unspecific binding from low affinity peptide such as PEX19 FxxxF motif. The Mechanistic details of the (di)aromatic peptide motif – PEX14 NTD interaction were studied with biophysical (ITC, NMR, CD), biochemical and computational methods. The analysis of affinities and thermodynamics identified key features of the ligand peptides, which suggest a refined consensus binding motif **WΦxE(F/Y)Φ** for PEX14-NTD binding. This result will allow us to specifically predict good binders for PEX14 NTD based on sequence. Future investigations should focus on a sequence based search for potential binders and cargoes that possibly translocate into the peroxisome without the support from PEX5 describing a new import route. Vice versa, a structural and sequence based investigation should focus on the identification of a PEX14 NTD fold in other proteins, which might be involved in similar translocation processes.

A structural analysis of the PEX13 C-terminal domain (CTD) identified an autoinhibition mechanism of the PEX13-SH3 domain by a flanking FxxxF motif, which was shown to modulate PTS1 import in a cell based assay. Investigation of the interaction sites between PEX13 CTD and the core components of the peroxisomal docking and translocation machinery PEX5 and PEX14 demonstrate an interaction of the PEX13 CTD with the PEX14 NTD via FxxxF motif and of the PEX13-SH3 domain with PEX5-NTD via WxxxF/Y motifs. Both interactions are modulated by the autoinhibition of PEX13.

Results of this work will have major impact in the peroxisome field as it first identified interactions between PEX13, PEX5 and PEX14, which are crucial for peroxisomal matrix import. Mapping of the characterized interaction sites between the peroxins PEX5, PEX13 and PEX14 revealed a complex network of (di)aromatic peptide motifs composed of high affinity (PEX5-PEX14), mid affinity (PEX13-PEX14) and mid to low affinity binding (PEX5-PEX13) events (**Figure 8.1**). Moreover, the interaction between PEX14 and PEX13 involving the PEX 13 FxxxF motif is not conserved from yeast to human. In consequence, a "docking complex" as known from yeast cannot exists in human. The interactions between PEX5 (di)aromatic peptide motifs and the PEX13 SH3 domain, on the other hand, are similar to the one found in yeast. However, these interactions are modulated by the autoinhibition of the PEX13 SH3 by the intrinsic FxxxF increasing the complexity of the human system. Only excess of PEX5 or the presence of PEX14 NTD is able to release that autoinhibited state.

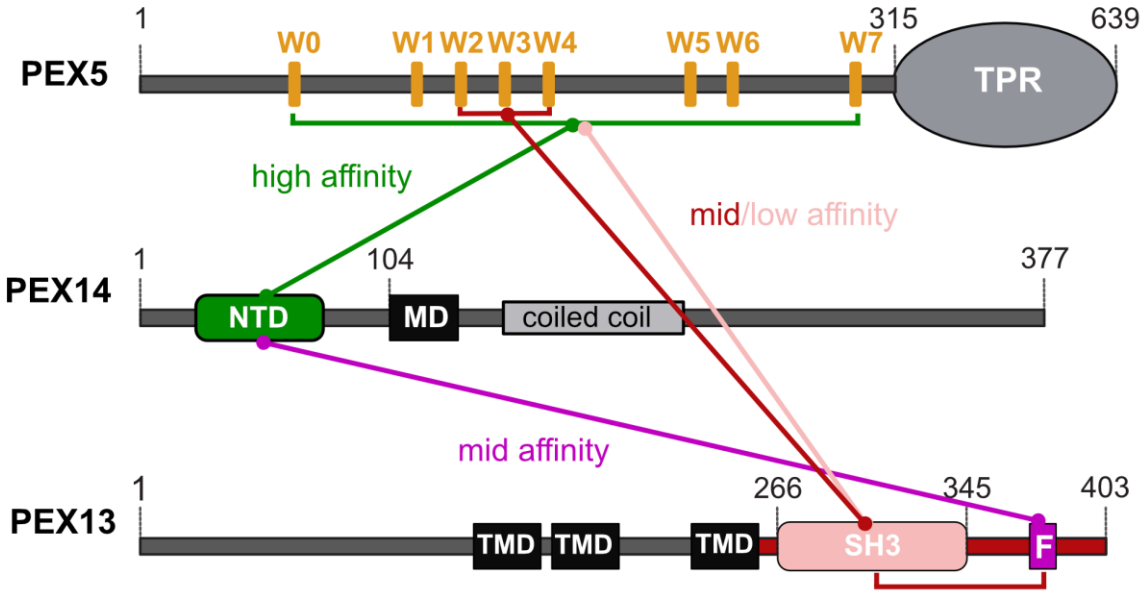

Figure 8.1: Interaction network of PEX5, PEX13 and PEX14. (Di)aromatic peptide motifs of PEX5 NTD bind to PEX14 NTD with high affinity and to PEX13 SH3 with mid/low affinity. PEX13 SH3 is autoinhibited by the flanking FxxxF motif, which can be released by an excess of PEX14 or a yet higher excess of PEX5 W2, W3 and W4. The opening of the PEX13 SH3 domain by PEX14 is driven by binding of the PEX13 FxxxF motif to the PEX14 NTD at the same binding site where PEX5 binds.

With the results of this study one can propose a model for the sequential steps in docking, translocation and cargo release. The affinity of PEX14 NTD towards membranes hint to a control mechanism to prevent unrelated, low affinity binding in the absence of PEX5. The high affinity binding of the first W-motifs (especially W0) towards PEX14 NTD tethers and stabilizes PEX5 -

membrane interaction to promote pore formation **(Figure 8.2**) (Gaussmann *et al.*, 2021; Neuhaus *et al.*, 2014). The obtained data are unfortunately not sufficient to develop a model for the pore formation that possibly involves PEX14 and or PEX13 oligomers. However, my model based on the obtained data suggests that PEX14 NTD can co-translocate into the peroxisomal lumen where molecular excess of PEX13 replaces PEX5 on PEX14 via the FxxxF motif and simultaneously allows PEX5 – PEX13 SH3 binding, which triggers the pore disorganization and cargo release **(Figure 7.2)**. This model is supported by the inconsistency of the reported PEX14 topology and finding of a high molecular mass complex consistent of PEX13 only (Barros-Barbosa *et al.*, 2019; Bharti *et al.*, 2011; Neufeld *et al.*, 2009; Neuhaus *et al.*, 2014; Reguenga *et al.*, 2001; Shimizu *et al.*, 1999; Will *et al.*, 1999).

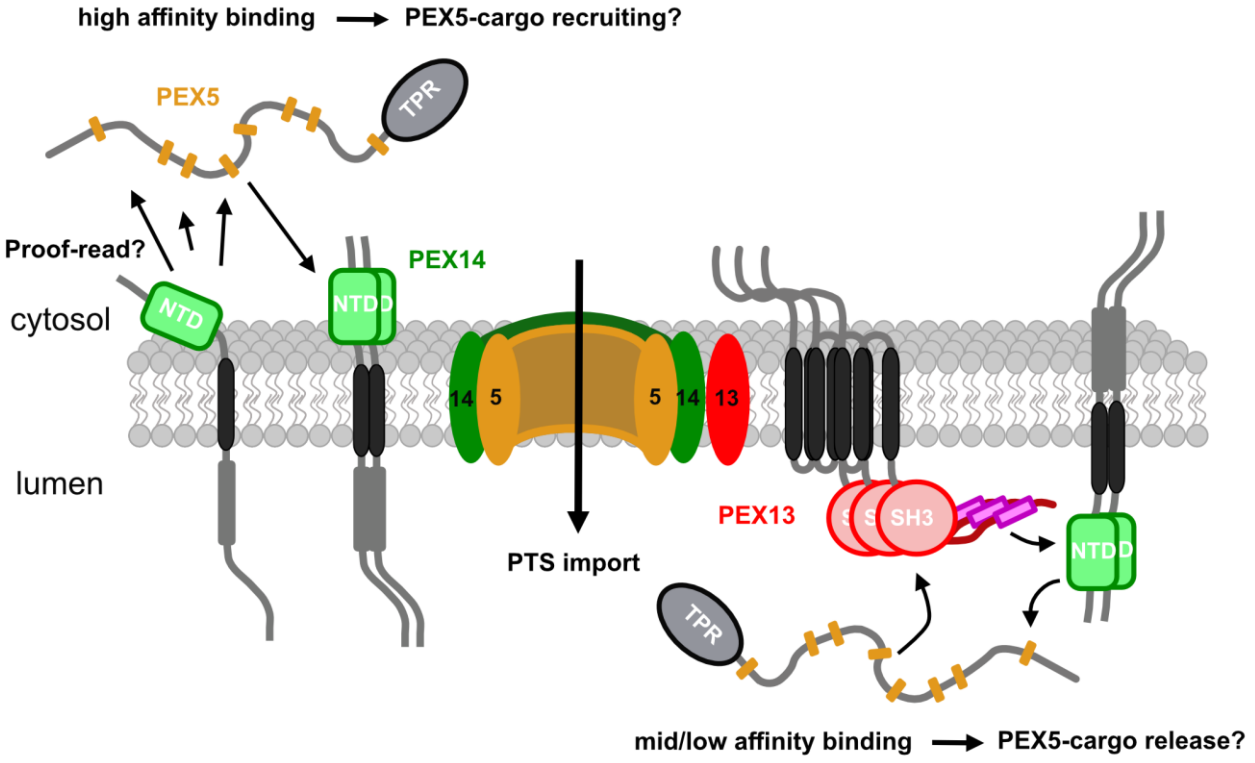

Figure 8.2: Model for docking and cargo release. The strong affinity from PEX5 NTD towards PEX14 NTD is needed for the initial docking event in the cytosol to tether and stabilize PEX5 – membrane interactions. Upon pore formation and cargo translocation, PEX14 NTD co-translocalizes into the lumen. The PEX14 – PEX5 interaction is then replaced by PEX13 FxxxF and simultaneous PEX13 SH3 -PEX5 NTD binding, which triggers pore disassembly and cargo release.

Future studies should focus on full length proteins and assemblies of complexes in native membranes or membrane-like environments. Such experiments can be carried out in membrane-mimicking environments, such as nanodiscs using integrative structural biology, focusing on NMR spectroscopy and cryo-EM to capture the presence of dynamic conformational states. PEX13 and PEX14 can be assembled into nanodiscs and analyzed via NMR spectroscopy and cryo-EM. Such an experiment will yield information about spontaneous complex formation in a membrane like environment which gives hints on the domain topology of the complex. Further insights into domain topology can be obtained by crosslinking experiments and native mass spectrometry.

Data reported in this thesis will be useful to design mutants of PEX13, which could stall peroxisomal import by disrupting cargo release. In that case, an *in vitro* assembled pore could be trapped in a certain state and analyzed by cryo-EM. These potential mutations can be screened by cellular assays by inhibiting proteasomal degradation either by inhibiting the proteasome directly or inhibiting poly-ubiquitination of PEX5. To link these experiments with a more functional context they should be combined with cell based visualization and functional assays, for examples, utilizing high resolution fluorescence microscopy like the recently introduced minimal photon fluxes concept (MINFLUX) or complementation assays (Schmidt *et al*, 2021). MINFLUX microscopy can be able to detect domain topologies in import active cells if the size of the fluorophore is small enough not to falsify the detection.
Once the minimal requirements for a stable pore formation are found, the size and permeability, in respect to cargo translocation, can be studied by atomic force microscopy and electrophysiological experiments.

Abbreviations

2D	Two dimensional
3D	Three dimensional
ACOX1	Acyl-CoA oxidase 1
ADHAPS	Alkyl-dihydroxyaceton-phosphate synthase
AGT	Alanine glyoxylate aminotransferase
ALDP	X-linked adrenoleukodystrophy
AMACR	2-methylacyl-CoA racemase
CAT	Catalase
CD	Circular dichroism
Cryo-EM	Cryo-electron microscopy
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbations
CTD	C-terminal domain
D7PC	1,2-Diheptanoyl-sn-Glycero-3-Phosphocholine
DBP	D-bi-functional protein
DD	Dipole-dipole
DHAPAT	dihydroxyacetone-phosphate acyltransferase
DHCA	Di-Hydroxycholestanioc Acid
DLP1	Dynamin-1-like protein
DMPC	1,2-Dimyristoyl-sn-Glycero-3-Phosphocholine
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DOPE	1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine
DPC	dodecylphosphocholine
ER	Endoplasmatic reticulum
FID	Free induction decay
Fis1	Mitochondrial Fission Protein 1
GTP	Guanosine-5'-triphosphate
GTP	Guanosintriphosphate
H_2O_2	Hydrogen Peroxide
HS	Heimler Syndrome
HSQC	Heteronuclear single quantum coherence
IDP	Intrinsically disordered protein
INEPT	Insensitive nuclei enhanced by polarization transfer
IPTG	Isopropyl-β-d-thiogalactopyranosid
IRD	Infantile refsum disease
ITC	Isothermal titration calorimetry
JEP	Juxtaposed elongated peroxisomes
LB	Lysogeny broth

MD	Molecular dynamics
Mff	Mitochondrial fission factor
mPTS	Membrane peroxisomal targeting signal
NALD	Neonatal adrenoleukodystrophy
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NTA	Nitrilotriacetic acid
NTD	N-terminal domain
PBD	Peroxisome biogenesis disorder
PC	phosphatidylcholine
PE	phosphatidylethanolamine
PED	Peroxisomal enzyme deficiency
PEG	Polyethylene glycol
P-ER	Peroxisomales endoplasmatic reticulum
PEX	Peroxin
PEX5L	PEX5 long isoform
PEX5S	PEX5 short isoform
PI	phosphatidylinositol
PMP	Peroxisomal membrane protein
PP	Pre-peroxisomes
ppm	parts per million
PS	phosphatidylserine
PTS1	Peroxisomal targeting signal 1
PTS2	Peroxisomal targeting signal 2
RCDP	Rhizomelic chondrodysplasia punctate
SCPx	sterol carrier protein X
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SLIM	Site-directed, Ligase-Independent Mutagenesis
SLS	Static light scattering
SPM	Sphingomyelin
THCA	Tri-hydroxycholestanioc acid
TMP	Transmembrane protein
TMS	Tetramythysilane
TPA	Tubular peroxisomal accumulation
TPR	Tetratricopeptide
TROSY	Transverse relaxation-optimized spectroscopy
VLCFA	Very-long chain fatty acid
ZS	Zellweger syndrome
ZSS	Zellweger syndrome spectrum

References

Agrawal G, Subramani S (2016) De novo peroxisome biogenesis: Evolving concepts and conundrums. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1863: 892-901

Amery L, Brees C, Baes M, Setoyama C, Miura R, Mannaerts GP, Veldhoven PPV (1998) C-terminal tripeptide Ser-Asn-Leu (SNL) of human D-aspartate oxidase is a functional peroxisome-targeting signal. *Biochemical Journal* 336: 367-371

Antes I (2010) DynaDock: A new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility. *Proteins* 78: 1084-1104

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* 44: W344-W350

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research* 38: W529-W533

Azevedo JE, Schliebs W (2006) Pex14p, more than just a docking protein. *Biochimica et biophysica acta* 1763: 1574-1584

Bakker BM, Mensonides FI, Teusink B, van Hoek P, Michels PA, Westerhoff HV (2000) Compartmentation protects trypanosomes from the dangerous design of glycolysis. *Proc Natl Acad Sci U S A* 97: 2087-2092

Barnett P, Bottger G, Klein AT, Tabak HF, Distel B (2000) The peroxisomal membrane protein Pex13p shows a novel mode of SH3 interaction. *EMBO J* 19: 6382-6391

Barøy T, Koster J, Strømme P, Ebberink MS, Misceo D, Ferdinandusse S, Holmgren A, Hughes T, Merckoll E, Westvik J *et al* (2015) A novel type of rhizomelic chondrodysplasia punctata, RCDP5, is caused by loss of the PEX5 long isoform. *Human Molecular Genetics* 24: 5845-5854

Barros-Barbosa A, Ferreira MJ, Rodrigues TA, Pedrosa AG, Grou CP, Pinto MP, Fransen M, Francisco T, Azevedo JE (2019) Membrane topologies of PEX13 and PEX14 provide new insights on the mechanism of protein import into peroxisomes. *FEBS J* 286: 205-222

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 81: 3684-3690

Bharti P, Schliebs W, Schievelbusch T, Neuhaus A, David C, Kock K, Herrmann C, Meyer HE, Wiese S, Warscheid B *et al* (2011) PEX14 is required for microtubule-based peroxisome motility in human cells. *J Cell Sci* 124: 1759-1768

Birge RB, Knudsen BS, Besser D, Hanafusa H (1996) SH2 and SH3-containing adaptor proteins: redundant or independent mediators of intracellular signal transduction. *Genes to Cells* 1: 595-613

Bloom M, Reeves LW, Wells EJ (1965) Spin Echoes and Chemical Exchange. *The Journal of Chemical Physics* 42: 1615-1624

Bolik-Coulon N, Bouvignies G, Carlier L, Ferrage F (2019) Experimental characterization of the dynamics of IDPs and IDRs by NMR. In: pp. 65-92. Elsevier:

Bottger G, Barnett P, Klein AT, Kragt A, Tabak HF, Distel B (2000) Saccharomyces cerevisiae PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, non-PXXP-related manner. *Mol Biol Cell* 11: 3963-3976

Bowen P, Lee CS, Zellweger H, Lindenberg R (1964) A Familial Syndrome of Multiple Congenital Defects. *Bull Johns Hopkins Hosp* 114: 402-414

Braverman N, Dodt G, Gould SJ, Valle D (1998) An isoform of pex5p, the human PTS1 receptor, is required for the import of PTS2 proteins into peroxisomes. *Hum Mol Genet* 7: 1195-1205

Brocard C, Hartig A (2006) Peroxisome targeting signal 1: Is it really a simple tripeptide? *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1763: 1565-1573

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

Brunelle JL, Green R (2014) Coomassie blue staining. *Methods Enzymol* 541: 161-167

Case D, Babin V, Berryman J, Betz R, Cai Q, Cerutti D, Cheatham III T, Darden T, Duke R, Gohlke H (2014) Amber 14.

Castro JA, de Mecca MM, Bartel LC (2006) Toxic side effects of drugs used to treat Chagas' disease (American trypanosomiasis). *Hum Exp Toxicol* 25: 471-479

Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, Pupko T, Ben-Tal N (2013) ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Israel Journal of Chemistry* 53: 199-206

Chen X, Zaro JL, Shen WC (2013) Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* 65: 1357-1369

Chiu J (2004) Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Research* 32: e174-e174

Choe J, Moyersoen J, Roach C, Carter TL, Fan E, Michels PA, Hol WG (2003) Analysis of the sequence motifs responsible for the interactions of peroxins 14 and 5, which are involved in glycosome biogenesis in Trypanosoma brucei. *Biochemistry* 42: 10915-10922

Cornell RB, Taneva SG (2006) Amphipathic helices as mediators of the membrane interaction of amphitropic proteins, and as modulators of bilayer physical properties. *Curr Protein Pept Sci* 7: 539-552

Coura JR, Viñas PA (2010) Chagas disease: a new worldwide challenge. *Nature* 465: S6-S7

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188-1190

Cyr N, Madrid KP, Strasser R, Aurousseau M, Finn R, Ausio J, Jardim A (2008) Leishmania donovani peroxin 14 undergoes a marked conformational change following association with peroxin 5. *The Journal of biological chemistry* 283: 31488-31499

Dammai V, Subramani S (2001) The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell* 105: 187-196

Daragan VA, Mayo KH (1997) Motional model analyses of protein and peptide dynamics using ^{13}C and ^{15}N NMR relaxation. *Prog NMR Spectrosc* 31: 63-105

Darden T, York D, Pedersen L (1993) Particle mesh Ewald: AnN·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* 98: 10089-10092

Dawidowski M, Emmanouilidis L, Kalel VC, Tripsianes K, Schorpp K, Hadian K, Kaiser M, Maser P, Kolonko M, Tanghe S *et al* (2017) Inhibitors of PEX14 disrupt protein import into glycosomes and kill Trypanosoma parasites. *Science* 355: 1416-1420

Dawidowski M, Kalel VC, Napolitano V, Fino R, Schorpp K, Emmanouilidis L, Lenhart D, Ostertag M, Kaiser M, Kolonko M *et al* (2020) Structure-Activity Relationship in Pyrazolo[4,3-c]pyridines, First Inhibitors of PEX14-PEX5 Protein-Protein Interaction with Trypanocidal Activity. *J Med Chem* 63: 847-879

De Duve C, Baudhuin P (1966) Peroxisomes (microbodies and related particles). *Physiol Rev* 46: 323-357

de Jesus AJ, Allen TW (2013) The role of tryptophan side chains in membrane protein anchoring and hydrophobic mismatch. *Biochimica et biophysica acta* 1828: 864-876

Deb R, Nagotu S (2017) Versatility of peroxisomes: An evolving concept. *Tissue Cell* 49: 209-226

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* 6: 277-293

Delille HK, Agricola B, Guimaraes SC, Borta H, Lüers GH, Fransen M, Schrader M (2010) Pex11pβ-mediated growth and division of mammalian peroxisomes follows a maturation pathway. *Journal of Cell Science* 123: 2750-2762

Demangeat JL (2013) Nanosized solvent superstructures in ultramolecular aqueous dilutions: twenty years' research using water proton NMR relaxation. *Homeopathy* 102: 87-105

Desjeux P (2004) Leishmaniasis: current situation and new perspectives. *Comp Immunol Microbiol Infect Dis* 27: 305-318

Distel B, Erdmann R, Gould SJ, Blobel G, Crane DI, Cregg JM, Dodt G, Fujiki Y, Goodman JM, Just WW *et al* (1996) A unified nomenclature for peroxisome biogenesis factors. *J Cell Biol* 135: 1-3

Dodt G, Braverman N, Wong C, Moser A, Moser HW, Watkins P, Valle D, Gould SJ (1995) Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the peroxisome biogenesis disorders. *Nat Genet* 9: 115-125

Dodt G, Gould SJ (1996) Multiple *PEX* genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: Evidence that PTS1 protein import is mediated by a cycling receptor. *J Cell Biol* 135: 1763-1774

Douangamath A, Filipp FV, Klein AT, Barnett P, Zou P, Voorn-Brouwer T, Vega MC, Mayans OM, Sattler M, Distel B *et al* (2002) Topography for independent binding of alpha-helical and PPII-helical ligands to a peroxisomal SH3 domain. *Mol Cell* 10: 1007-1017

Duve CD (1969) The peroxisome: a new cytoplasmic organelle. *Proceedings of the Royal Society of London Series B Biological Sciences* 173: 71-83

Effelsberg D, Cruz-Zaragoza LD, Tonillo J, Schliebs W, Erdmann R (2015) Role of Pex21p for Piggyback Import of Gpd1p and Pnc1p into Peroxisomes of Saccharomyces cerevisiae. *The Journal of biological chemistry* 290: 25333-25342

El Magraoui F, Brinkmeier R, Schrötter A, Girzalsky W, Müller T, Marcus K, Meyer HE, Erdmann R, Platta HW (2013) Distinct Ubiquitination Cascades Act on the Peroxisomal Targeting Signal Type 2 Co-receptor Pex18p. *Traffic* 14: 1290-1301

Emmanouilidis L, Gopalswamy M, Passon DM, Wilmanns M, Sattler M (2016) Structural biology of the import pathways of peroxisomal matrix proteins. *Biochimica et biophysica acta* 1863: 804-813

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* 60: 2126-2132

Erdmann R, Blobel G (1995) Giant peroxisomes in oleic acid-induced Saccharomyces cerevisiae lacking the peroxisomal membrane protein Pmp27p. *Journal of Cell Biology* 128: 509-523

Erdmann R, Schliebs W (2005) Peroxisomal matrix protein import: the transient pore model. *Nat Rev Mol Cell Biol* 6: 738-742

Erdmann R, Veenhuis M, Kunau W-H (1997a) Peroxisomes: organelles at the crossroads. *Trends Cell Biol* 7: 400-407

Erdmann R, Veenhuis M, Kunau WH (1997b) Peroxisomes: Organelles at the crossroads. *Trends Cell Biol* 7: 400-407

Fang Y, Morrell JC, Jones JM, Gould SJ (2004) PEX3 functions as a PEX19 docking factor in the import of class I peroxisomal membrane proteins. *J Cell Biol* 164: 863-875

Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* 33: 5984-6003

Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K (2002) NMR analysis of a 900K GroEL GroES complex. *Nature* 418: 207-211

Fransen M, Terlecky SR, Subramani S (1998) Identification of a human PTS1 receptor docking protein directly required for peroxisomal protein import. *Proc Natl Acad Sci U S A* 95: 8087-8092

Freitas MO, Francisco T, Rodrigues TA, Alencastre IS, Pinto MP, Grou CP, Carvalho AF, Fransen M, Sa-Miranda C, Azevedo JE (2011) PEX5 protein binds monomeric catalase blocking its tetramerization and releases it upon binding the N-terminal domain of PEX14. *The Journal of biological chemistry* 286: 40509-40519

Fujiki Y, Lazarow PB (1985) Post-translational import of fatty acyl-CoA oxidase and catalase into peroxisomes of rat liver in vitro. *The Journal of biological chemistry* 260: 5603-5609

Fujiki Y, Matsuzono Y, Matsuzaki T, Fransen M (2006) Import of peroxisomal membrane proteins: the interplay of Pex3p- and Pex19p-mediated interactions. *Biochimica et biophysica acta* 1763: 1639-1646

Fujiki Y, Rachubinski RA, Lazarow PB (1984) Synthesis of a major integral membrane polypeptide of rat liver peroxisomes on free polysomes. *Proceedings of the National Academy of Sciences* 81: 7127-7131

Furuya T, Kessler P, Jardim A, Schnaufer A, Crudder C, Parsons M (2002) Glucose is toxic to glycosome-deficient trypanosomes. *Proc Natl Acad Sci U S A* 99: 14177-14182

Gairí M, Dyachenko A, González MT, Feliz M, Pons M, Giralt E (2015) An optimized method for 15N R1 relaxation rate measurements in non-deuterated proteins. *Journal of biomolecular NMR* 62: 209-220

Gáspári Z, Perczel A, 2010. Protein Dynamics as Reported by NMR, Annual Reports on NMR Spectroscopy. Elsevier, pp. 35-75.

Gatto GJ, Jr., Geisbrecht BV, Gould SJ, Berg JM (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nature structural biology* 7: 1091-1095

Gaussmann S, Gopalswamy M, Eberhardt M, Reuter M, Zou P, Schliebs W, Erdmann R, Sattler M (2021) Membrane Interactions of the Peroxisomal Proteins PEX5 and PEX14. *Front Cell Dev Biol* 9: 651449

Geerlof A, Brown J, Coutard B, Egloff MP, Enguita FJ, Fogg MJ, Gilbert RJC, Groves MR, Haouz A, Nettleship JE *et al* (2006) The impact of protein characterization in structural proteomics. *Acta Crystallographica Section D Biological Crystallography* 62: 1125-1136

Ghosh D, Berg JM (2010) A proteome-wide perspective on peroxisome targeting signal 1(PTS1)-Pex5p affinities. *Journal of the American Chemical Society* 132: 3973-3979

Giannopoulou EA, Emmanouilidis L, Sattler M, Dodt G, Wilmanns M (2016) Towards the molecular mechanism of the integration of peroxisomal membrane proteins. *Biochimica et biophysica acta* 1863: 863-869

Gimenez-Andres M, Copic A, Antonny B (2018) The Many Faces of Amphipathic Helices. *Biomolecules* 8: 45

Glover JR, Andrews DW, Rachubinski RA (1994) Saccharomyces cerevisiae peroxisomal thiolase is imported as a dimer. *Proc Natl Acad Sci U S A* 91: 10541-10545

Göbl C, Madl T, Simon B, Sattler M (2014) NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Progress in nuclear magnetic resonance spectroscopy* 80: 26-63

Goldfischer S, Moore CL, Johnson AB, Spiro AJ, Valsamis MP, Wisniewski HK, Ritch RH, Norton WT, Rapin I, Gartner LM (1973) Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome. *Science* 182: 62-64

Gootjes J, Schmohl F, Mooijer PAW, Dekker C, Mandel H, Topcu M, Huemer M, Von Schütz M, Marquardt T, Smeitink JA *et al* (2004) Identification of the molecular defect in patients with peroxisomal mosaicism using a novel method involving culturing of cells at 40°C: Implications for other inborn errors of metabolism. *Human Mutation* 24: 130-139

Gould SJ, Kalish JE, Morrell JC, Bjorkman J, Urquhart AJ, Crane DI (1996a) Pex13p is an SH3 protein of the peroxisome membrane and a docking factor for the predominantly cytoplasmic PTs1 receptor. *J Cell Biol* 135: 85-95

Gould SJ, Kalish JE, Morrell JE, Bjorkman J, Urquhart AJ, Crane DI (1996b) The SH3 domain of the Saccharomyces cerevisiae peroxisomal membrane protein Pex13p functions as a docking site for Pex5p, a mobile receptor for the import of PTS1-containing proteins. *J Cell Biol* 135: 85-95

Gould SJ, Keller GA, Hosken N, Wilkinson J, Subramani S (1989) A conserved tripeptide sorts proteins to peroxisomes. *J Cell Biol* 108: 1657-1664

Gould SJ, Keller GA, Schneider M, Howell SH, Garrard LJ, Goodman JM, Distel B, Tabak H, Subramani S (1990) Peroxisomal protein import is conserved between yeast, plants, insects and mammals. *EMBO J* 9: 85-90

Gould SJ, Keller GA, Subramani S (1987) Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *J Cell Biol* 105: 2923-2931

Gouveia AM, Reguenga C, Oliveira ME, Sa-Miranda C, Azevedo JE (2000) Characterization of peroxisomal Pex5p from rat liver. Pex5p in the Pex5p-Pex14p membrane complex is a transmembrane protein. *The Journal of biological chemistry* 275: 32444-32451

Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* 1: 2527-2535

Grou CP, Carvalho AF, Pinto MP, Wiese S, Piechura H, Meyer HE, Warscheid B, Sá-Miranda C, Azevedo JE (2008) Members of the E2D (UbcH5) Family Mediate the Ubiquitination of the Conserved Cysteine of Pex5p, the Peroxisomal Import Receptor. *Journal of Biological Chemistry* 283: 14190-14197

Haanstra JR, Gonzalez-Marcano EB, Gualdron-Lopez M, Michels PA (2016) Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites. *Biochimica et biophysica acta* 1863: 1038-1048

Hagen S, Drepper F, Fischer S, Fodor K, Passon D, Platta HW, Zenn M, Schliebs W, Girzalsky W, Wilmanns M *et al* (2015) Structural Insights into Cargo Recognition by the Yeast PTS1 Receptor. *Journal of Biological Chemistry* 290: 26610-26626

Hagn F, Nasr ML, Wagner G (2018) Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR. *Nat Protoc* 13: 79-98

Hardeman D, Zomer HW, Schutgens RB, Tager JM, van den Bosch H (1990) Effect of peroxisome proliferation on ether phospholipid biosynthesizing enzymes in rat liver. *Int J Biochem* 22: 1413-1418

Hartmann C, Antes I, Lengauer T (2007) IRECS: a new algorithm for the selection of most probable ensembles of side-chain conformations in protein models. *Protein Sci* 16: 1294-1307

Hartmann C, Antes I, Lengauer T (2009) Docking and scoring with alternative side-chain conformations. *Proteins* 74: 712-726

Hashimoto T, Fujita T, Usuda N, Cook W, Qi C, Peters JM, Gonzalez FJ, Yeldandi AV, Rao MS, Reddy JK (1999) Peroxisomal and Mitochondrial Fatty Acid β-Oxidation in Mice Nullizygous for Both Peroxisome Proliferator-activated Receptor α and Peroxisomal Fatty Acyl-CoA Oxidase. *Journal of Biological Chemistry* 274: 19228-19236

Hettema EH, Girzalsky W, van Den Berg M, Erdmann R, Distel B (2000) Saccharomyces cerevisiae pex3p and pex19p are required for proper localization and stability of peroxisomal membrane proteins. *EMBO J* 19: 223-233

Hildebrand PW, Preissner R, Frommel C (2004) Structural features of transmembrane helices. *FEBS letters* 559: 145-151

Hoepfner D, Schildknegt D, Braakman I, Philippsen P, Tabak HF (2005) Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* 122: 85-95

Hojjat H, Jardim A (2015) The Leishmania donovani peroxin 14 binding domain accommodates a high degeneracy in the pentapeptide motifs present on peroxin 5. *Biochimica et biophysica acta* 1850: 2203-2212

Hu J, Baker A, Bartel B, Linka N, Mullen RT, Reumann S, Zolman BK (2012) Plant Peroxisomes: Biogenesis and Function. *The Plant Cell* 24: 2279-2303

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14: 33-38, 27-38

Islinger M, Li KW, Seitz J, Völkl A, Lüers GH (2009) Hitchhiking of Cu/Zn Superoxide Dismutase to Peroxisomes - Evidence for a Natural Piggyback Import Mechanism in Mammals. *Traffic* 10: 1711-1721

Jansen RLM, Santana-Molina C, van den Noort M, Devos DP, van der Klei IJ (2021) Comparative Genomics of Peroxisome Biogenesis Proteins: Making Sense of the PEX Proteins. *Front Cell Dev Biol* 9: 654163

Jedd G, Chua N-H (2000) A new self-assembled peroxisomal vesicle required for efficient resealing of the plasma membrane. *Nature Cell Biology* 2: 226-231

Jonassen JA, Kohjimoto Y, Scheid CR, Schmidt M (2005) Oxalate toxicity in renal cells. *Urological Research* 33: 329-339

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) COMPARISON OF SIMPLE POTENTIAL FUNCTIONS FOR SIMULATING LIQUID WATER. *Journal of Chemical Physics* 79: 926-935

Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66: 125-132

Kalel VC, Emmanouilidis L, Dawidowski M, Schliebs W, Sattler M, Popowicz GM, Erdmann R (2017) Inhibitors of glycosomal protein import provide new leads against trypanosomiasis. *Microb Cell* 4: 229-232

Kalel VC, Li M, Gaussmann S, Delhommel F, Schäfer A-B, Tippler B, Jung M, Maier R, Oeljeklaus S, Schliebs W *et al* (2019) Evolutionary divergent PEX3 is essential for glycosome biogenesis and survival of trypanosomatid parasites. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1866: 118520

Kang J-H, Jung M-Y, Yin X, Andrianifahanana M, Hernandez DM, Leof EB (2017) Cell-penetrating peptides selectively targeting SMAD3 inhibit profibrotic TGF-β signaling. *Journal of Clinical Investigation* 127: 2541-2554

Karplus M (1963) Vicinal Proton Coupling in Nuclear Magnetic Resonance. *Journal of the American Chemical Society* 85: 2870-2871

Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28: 8972-8979

Kelley RI, Moser HW (1984) Hyperpipecolic acidemia in neonatal adrenoleukodystrophy. *American Journal of Medical Genetics* 19: 791-795

Kelly SM, Jess TJ, Price NC (2005) How to study proteins by circular dichroism. *Biochimica et biophysica acta* 1751: 119-139

Kerssen D, Hambruch E, Klaas W, Platta HW, de Kruijff B, Erdmann R, Kunau WH, Schliebs W (2006) Membrane association of the cycling peroxisome import receptor Pex5p. *The Journal of biological chemistry* 281: 27003-27015

Kim PK, Mullen RT, Schumann U, Lippincott-Schwartz J (2006) The origin and maintenance of mammalian peroxisomes involves a de novo PEX16-dependent pathway from the ER. *Journal of Cell Biology* 173: 521-532

Kindl H, Kruse C (1983) [55] Biosynthesis of glyoxysomal proteins. In: pp. 700-715. Elsevier:

Kjaergaard M, Poulsen FM (2011) Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of biomolecular NMR* 50: 157-165

Kleckner IR, Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et biophysica acta* 1814: 942-968

Klöpfer K, Hagn F (2019) Beyond detergent micelles: The advantages and applications of non-micellar and lipid-based membrane mimetics for solution-state NMR. *Progress in nuclear magnetic resonance spectroscopy* 114-115: 271-283

Klouwer FCC, Berendse K, Ferdinandusse S, Wanders RJA, Engelen M, Poll-The BT (2015) Zellweger spectrum disorders: clinical overview and management approach. *Orphanet Journal of Rare Diseases* 10

Knecht S, Ricklin D, Eberle AN, Ernst B (2009) Oligohis-tags: mechanisms of binding to Ni2+-NTA surfaces. *Journal of Molecular Recognition* 22: 270-279

Kneller JM, Lu M, Bracken C (2002) An effective method for the discrimination of motional anisotropy and chemical exchange. *Journal of the American Chemical Society* 124: 1852-1853

Koch J, Pranjic K, Huber A, Ellinger A, Hartig A, Kragler F, Brocard C (2010) PEX11 family members are membrane elongation factors that coordinate peroxisome proliferation and maintenance. *Journal of Cell Science* 123: 3389-3400

Korepanova A, Matayoshi ED (2012) HPLC-SEC Characterization of Membrane Protein-Detergent Complexes. *Current Protocols in Protein Science* 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 pex3Δ Strain. *Journal of Biological Chemistry* 280: 34350-34357

Krause C, Rosewich H, Thanos M, Gärtner J (2006) Identification of novel mutations inPEX2,PEX6,PEX10,PEX12, andPEX13in Zellweger spectrum patients. *Human Mutation* 27: 1157-1157

Krause C, Rosewich H, Woehler A, Gärtner J (2013) Functional analysis of PEX13 mutation in a Zellweger syndrome spectrum patient reveals novel homooligomerization of PEX13 and its role in human peroxisome biogenesis. *Human Molecular Genetics* 22: 3844-3857

Kunze M (2020) The type-2 peroxisomal targeting signal. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1867: 118609

Lakomek N-A, Ying J, Bax A (2012) Measurement of 15N relaxation rates in perdeuterated proteins by TROSY-based methods. *Journal of biomolecular NMR* 53: 209-221

Lazarow PB (2006) The import receptor Pex7p and the PTS2 targeting sequence. *Biochimica et biophysica acta* 1763: 1599-1604

Lazarow PB, Fujiki Y (1985) Biogenesis of peroxisomes. *Annual review of cell biology* 1: 489-530

Lee D, Hilty C, Wider G, Wüthrich K (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *Journal of Magnetic Resonance* 178: 72-76

LeMaster DM (1989) Deuteration in protein proton magnetic resonance. *Methods Enzymol* 177: 23-43

Li M, Gaussmann S, Tippler B, Ott J, Popowicz GM, Schliebs W, Sattler M, Erdmann R, Kalel VC (2021) Novel Trypanocidal Inhibitors that Block Glycosome Biogenesis by Targeting PEX3-PEX19 Interaction. *Front Cell Dev Biol* 9: 737159

Liu Y, Bjorkman J, Urquhart A, Wanders RJ, Crane DI, Gould SJ (1999) PEX13 is mutated in complementation group 13 of the peroxisome-biogenesis disorders. *Am J Hum Genet* 65: 621-634

Loria JP, Rance M, Palmer AG, 3rd (1999) Transverse-relaxation-optimized (TROSY) gradient-enhanced triple-resonance NMR spectroscopy. *J Magn Reson* 141: 180-184

Ma C, Agrawal G, Subramani S (2011) Peroxisome assembly: matrix and membrane protein biogenesis. *J Cell Biol* 193: 7-16

Mannaerts GP, Van Veldhoven PP (1993) Metabolic pathways in mammalian peroxisomes. *Biochimie* 75: 147-158

Marcotte I, Auger Ml (2005) Bicelles as model membranes for solid- and solution-state NMR studies of membrane peptides and proteins. *Concepts in Magnetic Resonance Part A* 24A: 17-37

Marshall PA, Dyer JM, Quick ME, Goodman JM (1996) Redox-sensitive homodimerization of Pex11p: a proposed mechanism to regulate peroxisomal division. *Journal of Cell Biology* 135: 123-137

Matsumoto N, Tamura S, Fujiki Y (2003) The pathogenic peroxin Pex26p recruits the Pex1p–Pex6p AAA ATPase complexes to peroxisomes. *Nature Cell Biology* 5: 454-460

Matsuzono Y, Kinoshita N, Tamura S, Shimozawa N, Hamasaki M, Ghaedi K, Wanders RJ, Suzuki Y, Kondo N, Fujiki Y (1999) Human PEX19: cDNA cloning by functional complementation, mutation analysis in a patient with Zellweger syndrome, and potential role in peroxisomal membrane assembly. *Proc Natl Acad Sci U S A* 96: 2116-2121

Matthews BW (1968) Solvent content of protein crystals. *Journal of molecular biology* 33: 491-497

Maxwell M, Bjorkman J, Nguyen T, Sharp P, Finnie J, Paterson C, Tonks I, Paton BC, Kay GF, Crane DI (2003) Pex13 inactivation in the mouse disrupts peroxisome biogenesis and leads to a Zellweger syndrome phenotype. *Mol Cell Biol* 23: 5947-5957

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phasercrystallographic software. *Journal of Applied Crystallography* 40: 658-674

McPherson A, Gavira JA (2014) Introduction to protein crystallization. *Acta Crystallographica Section F Structural Biology Communications* 70: 2-20

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

Metzler WJ, Constantine KL, Friedrichs MS, Bell AJ, Ernst EG, Lavoie TB, Mueller L (1993) Characterization of the three-dimensional solution structure of human profilin: proton, carbon-13, and nitrogen-15 NMR assignments and global folding pattern. *Biochemistry* 32: 13818-13829

Montilla-Martinez M, Beck S, Klumper J, Meinecke M, Schliebs W, Wagner R, Erdmann R (2015) Distinct Pores for Peroxisomal Import of PTS1 and PTS2 Proteins. *Cell Rep* 13: 2126-2134

Morris GA, Freeman R (1979) Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* 101: 760-762

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography* 67: 355-367

Nair DM, Purdue PE, Lazarow PB (2004) Pex7p translocates in and out of peroxisomes in Saccharomyces cerevisiae. *J Cell Biol* 167: 599-604

Neufeld C, Filipp FV, Simon B, Neuhaus A, Schuller N, David C, Kooshapur H, Madl T, Erdmann R, Schliebs W *et al* (2009) Structural basis for competitive interactions of Pex14 with the import receptors Pex5 and Pex19. *EMBO J* 28: 745-754

Neuhaus A, Eggeling C, Erdmann R, Schliebs W (2016) Why do peroxisomes associate with the cytoskeleton? *Biochimica et biophysica acta* 1863: 1019-1026

Neuhaus A, Kooshapur H, Wolf J, Meyer NH, Madl T, Saidowsky J, Hambruch E, Lazam A, Jung M, Sattler M *et al* (2014) A novel Pex14 protein-interacting site of human Pex5 is critical for matrix protein import into peroxisomes. *The Journal of biological chemistry* 289: 437-448

Nuttall JM, Motley A, Hettema EH (2011) Peroxisome biogenesis: recent advances. *Current opinion in cell biology* 23: 421-426

Okamura H, Garcia-Rodriguez C, Martinson H, Qin J, Virshup DM, Rao A (2004) A conserved docking motif for CK1 binding controls the nuclear localization of NFAT1. *Mol Cell Biol* 24: 4184-4195

Okumoto K (1998) Mutations in PEX10 is the cause of Zellweger peroxisome deficiency syndrome of complementation group B. *Human Molecular Genetics* 7: 1399-1405

Okumoto K, Misono S, Miyata N, Matsumoto Y, Mukai S, Fujiki Y (2011) Cysteine Ubiquitination of PTS1 Receptor Pex5p Regulates Pex5p Recycling. *Traffic* 12: 1067-1083

Okumoto K, Shimozawa N, Kawai A, Tamura S, Tsukamoto T, Osumi T, Moser H, Wanders RJA, Suzuki Y, Kondo N *et al* (1998) PEX12 , the Pathogenic Gene of Group III Zellweger Syndrome: cDNA Cloning by Functional Complementation on a CHO Cell Mutant, Patient Analysis, and Characterization of Pex12p. *Molecular and Cellular Biology* 18: 4324-4336

Opperdoes FR, Borst P (1977) Localization of nine glycolytic enzymes in a microbody-like organelle in Trypanosoma brucei: the glycosome. *FEBS letters* 80: 360-364

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. *The Journal of biological chemistry* 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

Passarge E, McAdams AJ (1967) Cerebro-hepato-renal syndrome. *The Journal of Pediatrics* 71: 691-702

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* 94: 12366-12371

Petriv OI, Tang L, Titorenko VI, Rachubinski RA (2004) A new definition for the consensus sequence of the peroxisome targeting signal type 2. *Journal of molecular biology* 341: 119-134

Pick U (1981) Liposomes with a large trapping capacity prepared by freezing and thawing of sonicated phospholipid mixtures. *Archives of biochemistry and biophysics* 212: 186-194

Platta HW, Hagen S, Erdmann R (2013) The exportomer: the peroxisomal receptor export machinery. *Cellular and molecular life sciences : CMLS* 70: 1393-1411

Poltera AA, Owor R, Cox JN (1977) Pathological aspects of human African trypanosomiasis (HAT) in Uganda. *Virchows Archiv A Pathological Anatomy and Histology* 373: 249-265

Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, Jenkins HT, Keegan R, Krissinel E, Stevenson K *et al* (2018) CCP4i2: the new graphical user interface to theCCP4 program suite. *Acta Crystallographica Section D Structural Biology* 74: 68-84

Poulos A, Sharp P, Whiting M (1984) Infantile Refsum's disease (phytanic acid storage disease): a variant of Zellweger's syndrome? *Clin Genet* 26: 579-586

Purdue PE, Lazarow PB (2001) Peroxisome biogenesis. *Annu Rev Cell Dev Biol* 17: 701-752

Pusey PN, Van Megen W (1986) Phase behaviour of concentrated suspensions of nearly hard colloidal spheres. *Nature* 320: 340-342

Rassi A, Jr., Rassi A, Marin-Neto JA (2010) Chagas disease. *Lancet* 375: 1388-1402

Ratbi I, Kim, Sommen M, Al-Sheqaih N, Guaoua S, Vandeweyer G, Jill, Kate, Simon, Neil *et al* (2015) Heimler Syndrome Is Caused by Hypomorphic Mutations in the Peroxisome-Biogenesis Genes PEX1 and PEX6. *The American Journal of Human Genetics* 97: 535-545

Reddy JK, Mannaerts GP (1994) Peroxisomal Lipid Metabolism. *Annual Review of Nutrition* 14: 343-370

Rees DC, Robertson AD (2001) Some thermodynamic implications for the thermostability of proteins. *Protein Sci* 10: 1187-1194

Reguenga C, Oliveira ME, Gouveia AM, Sa-Miranda C, Azevedo JE (2001) Characterization of the mammalian peroxisomal import machinery: Pex2p, Pex5p, Pex12p, and Pex14p are subunits of the same protein assembly. *The Journal of biological chemistry* 276: 29935-29942

Rehling P, Skaletz-Rorowski A, Girzalsky W, Voorn-Brouwer T, Franse MM, Distel B, Veenhuis M, Kunau WH, Erdmann R (2000) Pex8p, an intraperoxisomal peroxin of Saccharomyces cerevisiae required for protein transport into peroxisomes binds the PTS1 receptor pex5p. *The Journal of biological chemistry* 275: 3593-3602

Renner C, Schleicher M, Moroder L, Holak TA (2002) Practical aspects of the 2D 15N-[1h]-NOE experiment. *Journal of biomolecular NMR* 23: 23-33

Reumann S (2004) Specification of the Peroxisome Targeting Signals Type 1 and Type 2 of Plant Peroxisomes by Bioinformatics Analyses. *Plant physiology* 135: 783-800

Reuter M, Kooshapur H, Suda JG, Gaussmann S, Neuhaus A, Bruhl L, Bharti P, Jung M, Schliebs W, Sattler M *et al* (2021) Competitive Microtubule Binding of PEX14 Coordinates Peroxisomal Protein Import and Motility. *Journal of molecular biology* 433: 166765

Rhodin JAG, 1954. Correlation of ultrastructural organization : and function in normal and experimentally changed proximal convoluted tubule cells of the mouse kidney: an electron microscopic study. Dept. of Anatomy, Karolinska Institutet, Stockholm.

Roe DR, Cheatham TE, 3rd (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* 9: 3084-3095

Rucktaschel R, Girzalsky W, Erdmann R (2011) Protein import machineries of peroxisomes. *Biochimica et biophysica acta* 1808: 892-900

Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* 23: 327-341

Saidowsky J, Dodt G, Kirchberg K, Wegner A, Nastainczyk W, Kunau WH, 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. *The Journal of biological chemistry* 276: 34524-34529

Saksela K, Permi P (2012) SH3 domain ligand binding: What's the consensus and where's the specificity? *FEBS letters* 586: 2609-2614

Salzmann M, Pervushin K, Wider G, Senn H, Wuthrich K (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci U S A* 95: 13585-13590

Saryi NAA, Hutchinson JD, Al-Hejjaj MY, Sedelnikova S, Baker P, Hettema EH (2017) Pnc1 piggy-back import into peroxisomes relies on Gpd1 homodimerisation. *Scientific Reports* 7: 42579

Sattler M, J S, C G (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog NMR Spectrosc* 34: 93-158

Sattler M, Fesik SW (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure* 4: 1245-1249

Schafmeister C, Ross W, Romanovski V (1995) LEaP. *University of California, San Francisco*

Schlessinger J (1994) SH2/SH3 signaling proteins. *Current Opinion in Genetics & Development* 4: 25-30

Schliebs W, Saidowsky J, Agianian B, Dodt G, Herberg FW, Kunau WH (1999) Recombinant human peroxisomal targeting signal receptor PEX5. Structural basis for interaction of PEX5 with PEX14. *The Journal of biological chemistry* 274: 5666-5673

Schmidt R, Weihs T, Wurm CA, Jansen I, Rehman J, Sahl SJ, Hell SW (2021) MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope. *Nature Communications* 12

Schrader M, Bonekamp NA, Islinger M (2012) Fission and proliferation of peroxisomes. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1822: 1343-1357

Schrader M, Fahimi HD (2006) Growth and division of peroxisomes. *Int Rev Cytol* 255: 237-290

Schwarzinger S, Kroon GJ, Foss TR, Chung J, Wright PE, Dyson HJ (2001) Sequence-dependent correction of random coil NMR chemical shifts. *Journal of the American Chemical Society* 123: 2970-2978

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498-2504

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* 44: 213-223

Shi X, Betzi S, Lugari A, Opi S, Restouin A, Parrot I, Martinez J, Zimmermann P, Lecine P, Huang M *et al* (2012) Structural recognition mechanisms between human Src homology domain 3 (SH3) and ALG-2-interacting protein X (Alix). *FEBS letters* 586: 1759-1764

Shimizu N, Itoh R, Hirono Y, Otera H, Ghaedi K, Tateishi K, Tamura S, Okumoto K, Harano T, Mukai S *et al* (1999) The peroxin Pex14p. cDNA cloning by functional complementation on a Chinese hamster ovary cell mutant, characterization, and functional analysis. *The Journal of biological chemistry* 274: 12593-12604

Shimozawa N (2000) Identification of PEX3 as the gene mutated in a Zellweger syndrome patient lacking peroxisomal remnant structures. *Human Molecular Genetics* 9: 1995-1999

Shimozawa N, Tsukamoto T, Nagase T, Takemoto Y, Koyama N, Suzuki Y, Komori M, Osumi T, Jeannette G, Wanders RJ *et al* (2004) Identification of a new complementation group of the peroxisome biogenesis disorders and PEX14 as the mutated gene. *Hum Mutat* 23: 552-558

Shimozawa N, Zhang Z, Suzuki Y, Imamura A, Tsukamoto T, Osumi T, Fujiki Y, Orii T, Barth PG, Wanders RJ *et al* (1999) Functional heterogeneity of C-terminal peroxisome targeting signal 1 in PEX5-defective patients. *Biochemical and biophysical research communications* 262: 504-508

Shiozawa K, Konarev PV, Neufeld C, Wilmanns M, Svergun DI (2009) Solution structure of human Pex5.Pex14.PTS1 protein complexes obtained by small angle X-ray scattering. *The Journal of biological chemistry* 284: 25334-25342

Smith DW, Opitz JM, Inhorn SL (1965) A syndrome of multiple developmental defects including polycystic kidneys and intrahepatic biliary dysgenesis in 2 siblings. *The Journal of Pediatrics* 67: 617-624

Some D, Amartely H, Tsadok A, Lebendiker M (2019) Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *Journal of Visualized Experiments*

Sommer JM, Cheng QL, Keller GA, Wang CC (1992) In vivo import of firefly luciferase into the glycosomes of Trypanosoma brucei and mutational analysis of the C-terminal targeting signal. *Mol Biol Cell* 3: 749-759

Sommer LA, Meier MA, Dames SA (2012) A fast and simple method for probing the interaction of peptides and proteins with lipids and membrane-mimetics using GB1 fusion proteins and NMR spectroscopy. *Protein Sci* 21: 1566-1570

Spera S, Bax A (1991) Empirical correlation between protein backbone conformation and C.alpha. and C.beta. 13C nuclear magnetic resonance chemical shifts. *Journal of the American Chemical Society* 113: 5490-5492

Stanley WA, Filipp FV, Kursula P, Schuller 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

Steinberg SJ, Dodt G, Raymond GV, Braverman NE, Moser AB, Moser HW (2006) Peroxisome biogenesis disorders. *Biochimica et biophysica acta* 1763: 1733-1748

Stone MJ, Fairbrother WJ, Palmer AG, Reizer J, Saier MH, Wright PE (1992) Backbone dynamics of the Bacillus subtilis glucose permease IIA domain determined from nitrogen-15 NMR relaxation measurements. *Biochemistry* 31: 4394-4406

Su JR, 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 U S A* 106: 417-421

Szilágyi L (1995) Chemical shifts in proteins come of age. *Progress in nuclear magnetic resonance spectroscopy* 27: 325-442

Tam YY, Fagarasanu A, Fagarasanu M, Rachubinski RA (2005) Pex3p initiates the formation of a preperoxisomal compartment from a subdomain of the endoplasmic reticulum in Saccharomyces cerevisiae. *The Journal of biological chemistry* 280: 34933-34939

Tol MB, Deluz C, Hassaine G, Graff A, Stahlberg H, Vogel H (2013) Thermal unfolding of a mammalian pentameric ligand-gated ion channel proceeds at consecutive, distinct steps. *The Journal of biological chemistry* 288: 5756-5769

Tolbert NE, Essner E (1981) Microbodies: peroxisomes and glyoxysomes. *Journal of Cell Biology* 91: 271s-283s

Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66: 22-25

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59: 687-696

Walton PA, Hill PE, Subramani S (1995) Import of stably folded proteins into peroxisomes. *Mol Biol Cell* 6: 675-683

Wanders RJ (2004) Metabolic and molecular basis of peroxisomal disorders: a review. *Am J Med Genet A* 126A: 355-375

Wanders RJ, Ferdinandusse S, Brites P, Kemp S (2010) Peroxisomes, lipid metabolism and lipotoxicity. *Biochimica et biophysica acta* 1801: 272-280

Wanders RJ, Vreken P, Ferdinandusse S, Jansen GA, Waterham HR, van Roermund CW, Van Grunsven EG (2001) Peroxisomal fatty acid alpha- and beta-oxidation in humans: enzymology, peroxisomal metabolite transporters and peroxisomal diseases. *Biochem Soc Trans* 29: 250-267

Wanders RJ, Waterham HR (2006) Biochemistry of mammalian peroxisomes revisited. *Annual review of biochemistry* 75: 295-332

Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand E, Marcotte I (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochimica et biophysica acta* 1808: 1957-1974

Watanabe Y, Kawaguchi K, Okuyama N, Sugawara Y, Obita T, Mizuguchi M, Morita M, Imanaka T (2016) Characterization of the interaction between Trypanosoma brucei Pex5p and its receptor Pex14p. *FEBS letters* 590: 242-250

Waterham HR, Ebberink MS (2012) Genetics and molecular basis of human peroxisome biogenesis disorders. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1822: 1430-1441

Waterham HR, Ferdinandusse S, Wanders RJA (2016) Human disorders of peroxisome metabolism and biogenesis. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1863: 922-933

Weisemann R, Ruterjans H, Bermel W (1993) 3D triple-resonance NMR techniques for the sequential assignment of NH and 15N resonances in 15N- and 13C-labelled proteins. *Journal of biomolecular NMR* 3: 113-120

Will GK, Soukupova M, Hong X, Erdmann KS, Kiel JA, Dodt G, Kunau WH, Erdmann R (1999) Identification and characterization of the human orthologue of yeast Pex14p. *Mol Cell Biol* 19: 2265-2277

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al* (2011) Overview of theCCP4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography* 67: 235-242

Wishart DS, Sykes BD, Richards FM (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of molecular biology* 222: 311-333

Yang H, Guranovic V, Dutta S, Feng Z, Berman HM, Westbrook JD (2004) Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallographica Section D Biological Crystallography* 60: 1833-1839

Yang J, Pieuchot L, Jedd G (2018) Artificial import substrates reveal an omnivorous peroxisomal importomer. *Traffic* 19: 786-797

Yang X, Purdue PE, Lazarow PB (2001) Eci1p uses a PTS1 to enter peroxisomes: either its own or that of a partner, Dci1p. *Eur J Cell Biol* 80: 126-138