Mohanraj Gopalswamy, Chen Zheng, Stefan Gaussmann, Hamed Kooshapur, Eva Hambruch, Wolfgang Schliebs, Ralf Erdmann, Iris Antes and Michael Sattler*

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

https://doi.org/10.1515/hsz-2022-0177
Received May 8, 2022; accepted November 4, 2022;
published online November 28, 2022

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 **ExWΦxE(F/Y)Φ** for high affinity binding to the PEX14 NTD and discuss conservation of the (di)aromatic peptide recognition by PEX14 in other species.

Keywords: isothermal titration calorimetry; molecular dynamics; NMR; peroxisome biogenesis; protein-protein interactions; Wxxx(F/Y) motifs.

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 and Nagotu 2017). Peroxisomal proteins are nuclear-encoded and need to be imported into the organelle post-translationally (Emmanouilidis et al. 2016; Erdmann and Schliebs 2005; Giannopoulou et al. 2016; Lazarow and Fujiki 1985; 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. 1987; Gould et al. 1989). 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).

Mohanraj Gopalswamy, Stefan Gaussmann and Chen Zheng contributed equally to this work. **Iris Antes**: Deceased.

***Corresponding author: Michael Sattler**, Bavarian NMR Center, Department of Bioscience, School of Natural Sciences, Technical University of Munich, Lichtenbergstr. 4, D-85747 Garching, Germany; and Institute of Structural Biology, Molecular Targets and Therapeutics Center, Helmholtz Center Munich, Ingolstädter Landstr. 1, D-85764 Neuherberg, Germany,
E-mail: sattler@helmholtz-muenchen.de. https://orcid.org/0000-0002-1594-0527

Mohanraj Gopalswamy, Stefan Gaussmann and Hamed Kooshapur, Bavarian NMR Center, Department of Bioscience, School of Natural Sciences, Technical University of Munich, Lichtenbergstr. 4, D-85747 Garching, Germany; and Institute of Structural Biology, Molecular Targets and Therapeutics Center, Helmholtz Center Munich, Ingolstädter Landstr. 1, D-85764 Neuherberg, Germany

Chen Zheng and Iris Antes, TUM School of Life Sciences, Technical University of Munich, Emil-Erlenmeyer-Forum 8, D-85354 Freising, Germany; and TUM Center for Functional Protein Assemblies, Technical University of Munich, Ernst-Otto-Fischer-Straße 8, D-85748 Garching, Germany

Eva Hambruch, Wolfgang Schliebs and Ralf Erdmann, Institute of Biochemistry and Pathobiochemistry, Ruhr-Universität Bochum, Universitätsstr. 150, D-44780 Bochum, Germany. https://orcid.org/0000-0001-8380-0342 (R. Erdmann)

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). The PEX14 NTD 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 1) (Neuhaus et al. 2014; Saidowsky et al. 2001). These peptide motifs bind to the conserved PEX14 NTD (Supplementary Figure 1) (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. 2014).

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

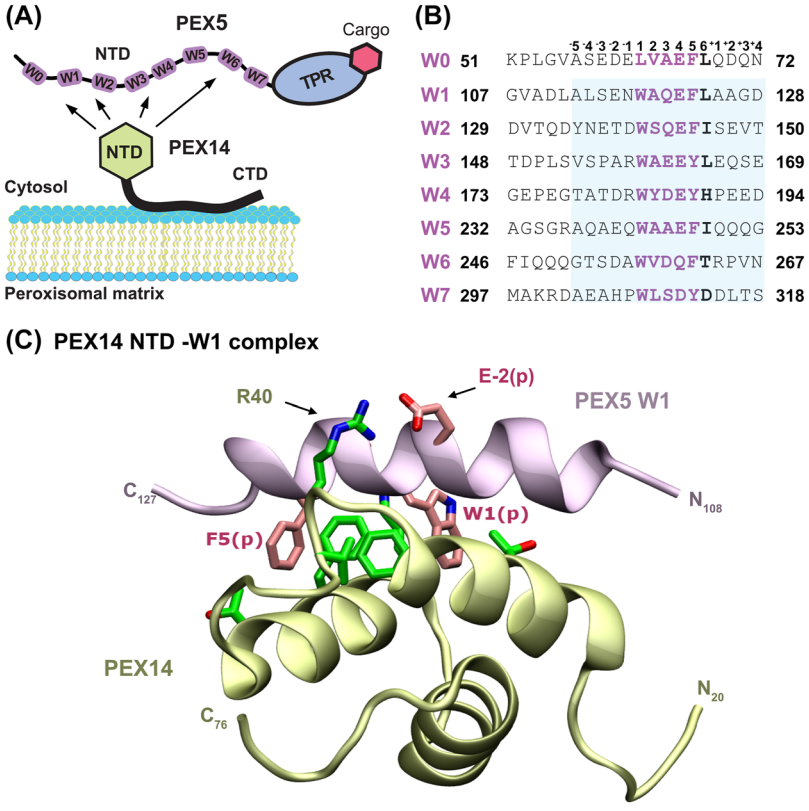

Figure 1: Interactions between PEX14 NTD and PEX (di)aromatic motifs. (A) The PEX14 NTD, which is located at the peroxisomal membrane, recognizes the W0–W7 motifs in the N-terminal region of PEX5. (B) 20-mer peptide comprising the W0–W7 motifs found in human PEX5 (UniProt ID: P50542). The 20-mer peptides 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.

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 and Jardim 2015; Kerssen et al. 2006; Watanabe et al. 2016).

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 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 four 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, **ExWΦxE(F/Y)Φ**, for high affinity 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 1B). The binding processes represent a single transition with exothermic binding enthalpies with dissociation constants ranging from 60 nM to 6 μM (Table 1, Supplementary Figure 2). 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 (Supplementary Figure 3), consistent with previous reports (Kerssen et al. 2006), suggesting further variations in the binding interface.

The energetics of binding are notably different for each motif (Figure 2A). Under all measurement conditions, binding between the peptides and PEX14 NTD is

Table 1: Isothermal titration calorimetry of PEX5 W0–W7 peptide binding to PEX14 NTD.

peptide	K_D (nM)	ΔH (kcal/mol)	Δ*G* (kcal/mol)	(−*T*Δ*S*) (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

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.

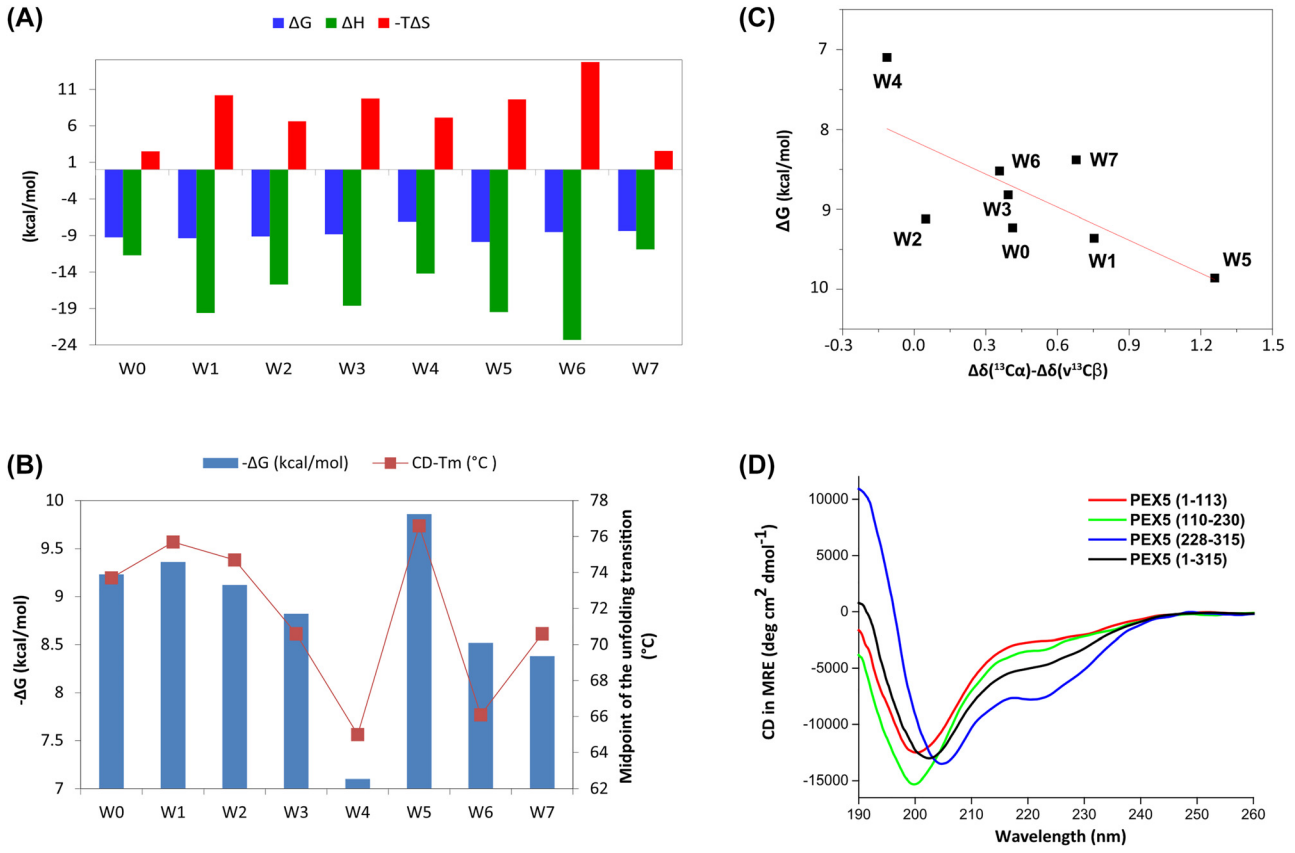

Figure 2: Thermodynamic and conformational features of the PEX14 NTD peptide interaction. (A) Bar Graph comparing thermodynamic parameters for the PEX14 NTD interaction with different (di)aromatic peptide ligands. Color codes are Δ*G* in blue, Δ*H* in green and $-T\Delta S$ in red. (B) Correlation between free binding energy ($-\Delta G$), 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 ($-\Delta G$) and average ^{13}C NMR secondary chemical shifts for the peptide motifs. The extent 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 and 222 nm, consistent with the presence of partial helical folding.

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 2B). 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 and 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 2B, Supplementary Figure 4). 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

Table 2: Comparison of experimental binding enthalpies and calculated interaction energies.

Peptide	Sequence[a]	ΔH (kcal/mol)	Interaction Energy (kcal/mol)	$\Delta\delta$ 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

[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 six residues in the core motif.

temperatures and thermodynamic stabilities of the peptide complexes show a clear correlation with the free energies calculated from the ITC data (Figure 2B).

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 2D). These results are in agreement with our previous reported NMR analysis, where we identified defined α-helical propensities within the PEX5 NTD by analysis of ^{13}C secondary chemical shift (Supplementary Figure 5) (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 (Supplementary Figure 5G). Notably, in the context of the full PEX5 NTD W5 motif exhibits the largest extent of α-helical conformation (Table 2). There is clear correlation of the free energy of association ($-\Delta G$) determined by ITC and helical propensity of the (di) aromatic motifs from average ^{13}C secondary chemical shift values (Figure 2C). 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.

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; Hartmann et al. 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 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 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 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 W and F residues in the W1 peptide (WAQEFL) correspond to W1(p) and F5(p), respectively (Figure 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 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.

From a comprehensive analysis of the PEX14-peptide interaction (Tables 3, 4; Supplementary Table 1) three key features can be identified that are important for binding to the PEX14 NTD (Figure 3, top left): (i) a hydrophobic core interface, (ii) electrostatic interactions with K56 and (iii) electrostatic interactions with R40 and N38, as illustrated for the PEX14-W1 complex in Figure 3.

Table 4: Population of hydrogen bonds or salt bridges during the simulation.

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

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 one corresponds to 0–100%.

Table 3: Binding site features.

	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

[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.

(i) 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. Note, that the K56 and R40 contribute additional hydrophobic interactions via their bulky aliphatic side chains contacting W1(p) and L6(p)/F5(p), respectively. Thus, the hydrophobic area in PEX14 NTD stretches over the whole binding site and forms complementary hydrophobic pockets for the corresponding peptide residues (Figure 3i). If this stable hydrophobic interaction pattern exists, a strong intra-peptide backbone hydrogen bond between the flanking residue –2(p) and residue 3(p) of the core motif can be observed (Figure 3iii), which stabilizes the helical conformation of the peptide.

(ii) The hydrophobic interface is flanked on either side by the two charged PEX14 residues R40 and K56, respectively. These residues contribute not only hydrophobic interactions but mediate strong electrostatic contacts involving their positively charged side chains with negatively charged side chains in the PEX14 NTD, where intra-ligand interactions can further stabilize these interactions. The hydrophobic/hydrophilic interaction pattern of K56 is conserved in all peptides containing the W1(p) and E4(p) residues and thus presents a second key feature (Figure 3, ii).

(iii) The third important binding feature involves R40, which establishes hydrophilic interactions with E–2(p) and N38 through its terminal guanidino group (Figure 3iii), featuring a similar but not that as highly conserved interaction pattern as seen for K56. Overall, the two flanking hydrophilic interactions additionally stabilize the central hydrophobic region such that the peptide is effectively locked in its bound position, which may explain the large enthalpy contributions.

Detailed analyses of the binding modes of the individual PEX14-peptide complexes are given in Tables 3, 4 and Supplementary Table 1. In addition, representative structures from the last 10 ns of simulation (see Materials and Methods) are shown in Figure 4 and Supplementary Figure 6. We find that the strong binding peptides, i.e. W1, W3, W5 and W6 exhibit all three interaction features (Table 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 4) and thus a very stable helical conformation of the bound peptide. For the peptides W3 and W6 lacking the glutamate at position –2(p), R40 stabilizes the binding to the PEX14 NTD by interacting with residues located more C-terminal (Supplementary Figure 6).

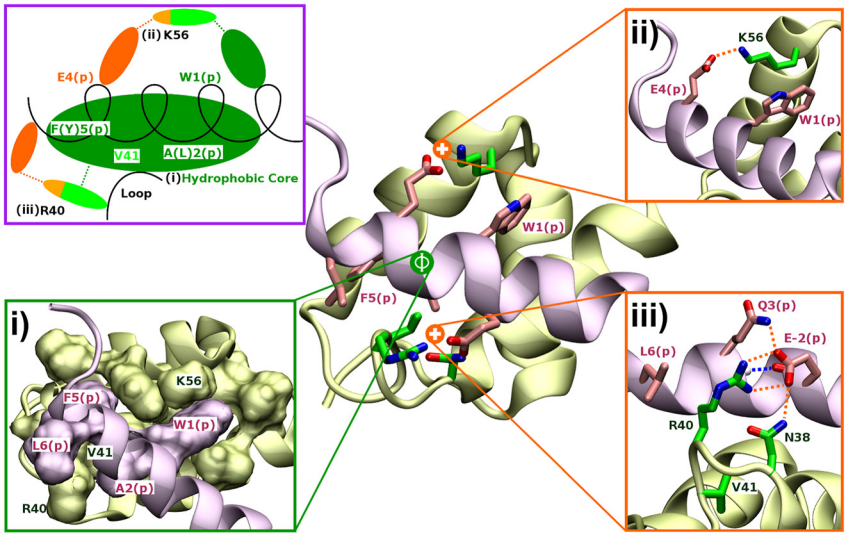

Figure 3: Key interactions in the PEX14 NTD/PEX5 W1-peptide complex. The PEX14 NTD is shown in yellow/green, the PEX5 peptide in pink. Important residues are given in licorice representation. The central picture shows the most prominent structure of W1 during the last 10ns of MD simulation. Salt bridges and hydrogen bonds are indicated by colored dashed lines (orange: Sidechain contacts, blue: backbone contacts). A schematic sketch of the binding pattern of the PEX14 NTD/PEX5 complex is shown on the top left, illustrating the three binding features (i), (ii) and (iii). (i) Surface representation of the residues contributing to the hydrophobic core. (ii) Hydrophilic interactions of K56. (iii) Hydrophilic interactions of R40.

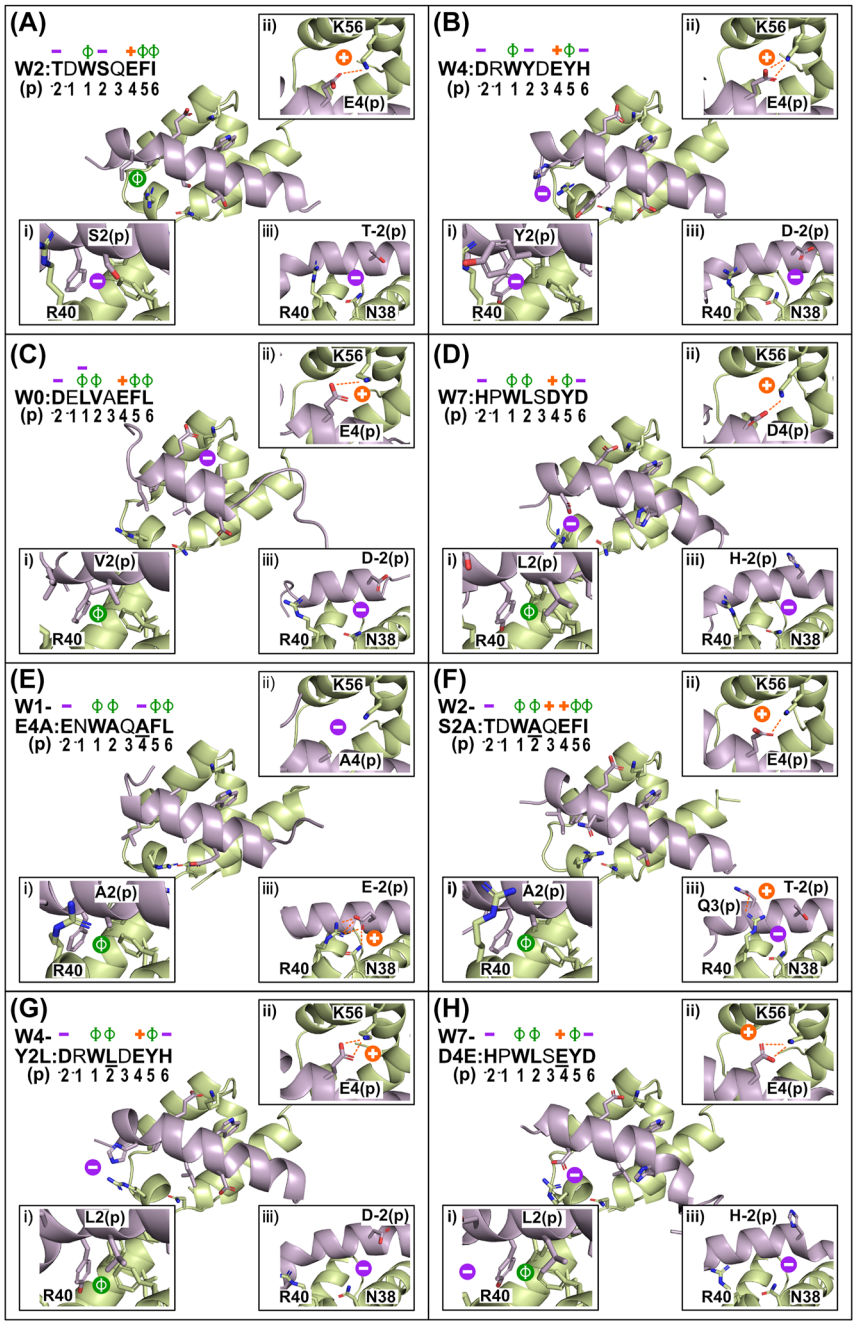

Figure 4: Molecular dynamics simulations. Most prominent structures observed during the last 10 ns of the MD simulations of PEX14 NTD bound to wildtype or mutant PEX5 peptides are shown. The PEX14 NTD and PEX5 peptides are colored in green and pink, respectively. The three key features hydrophobic core, electrostatic interactions with K56 and interactions with R40/*N*38 are shown in boxes (i), (ii) and (iii) respectively. Hydrophobic interactions are indicated by green ϕ, charged and polar interactions with an orange "+" sign. Interaction features that are missing are indicated by a purple "–" sign. Salt bridges and hydrogen bonds are indicated by dashed lines. Symbols above the respective peptide sequences indicate existing or missing features of the individual residues. Residues mutated compared to wildtype motifs are underlined (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.

Analysis of the complexes with intermediate and weak binding peptides (Table 3 and Figure 4) 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 3, 4; Figure 4A, B) and thus crucial for

optimal packing. Residues in the flanking regions can also affect the binding affinity. The threonine residue (T) at –2(p) in W2 abolishes the favored R40 and N38 interactions seen with W1 (Figure 3D) while the H mutation at position 6(p) in W4 causes a charge clash with R40 which destabilizes the binding (Figure 4B).

The effects of these amino acid substitutions on the overall stability of the complexes are summarized in Table 4. 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 further perturb the already weakened central hydrophobic core (Figure 4A). 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 4B). 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 1). The less favorable ΔH likely reflects that the replacement of W by L reduces the hydrophobic contact surface (Figure 4C). 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 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 4). These results are comparable with W1 showing similar affinities (Table 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 (Figure 4C, iii), suggesting conformational entropy compensation. This is likely also reflected in the fast off-rate (k_{off}) observed in SPR experiments (Neuhaus et al. 2014).

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 4D). 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 4).

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 2) and a partial opening of the binding pocket (compare Figure 4E and

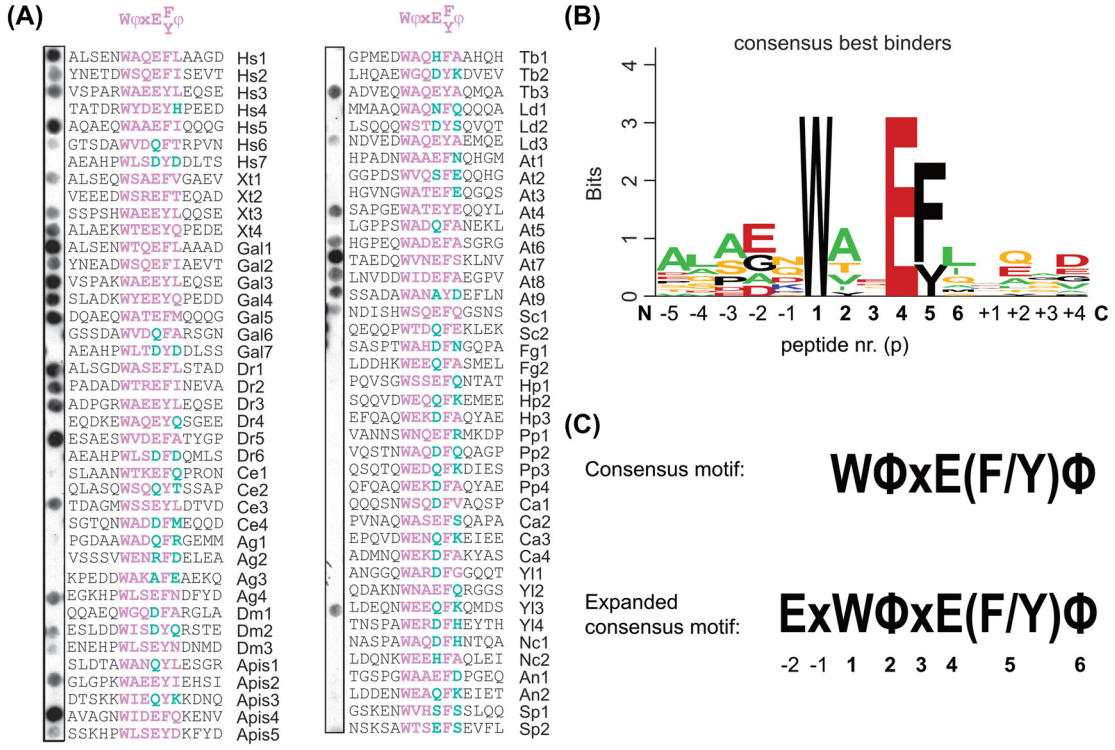

Figure 5: Peptide overlay binding assay with PEX14 NTD for various diaromatic peptide motifs and sequence conservation analysis of all PEX14 NTD binding motifs. (A) Peptide spot overlay assay with His-tagged PEX14 NTD and immobilized peptides representing Wxxx(F/Y) motifs of PEX5 of different species. 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 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.* (C) The refined core and expanded consensus motif for high affinity PEX14 binding.

Supplementary Figure 6D with Figure 3). 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 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 4). 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 (Figures 3iii, 4F, iii; Supplementary Figure 6B, iii). In the variant **W2_S2(p)L**, the aliphatic side chain of R40 contributes to the hydrophobic core (Supplementary Figure 6E, i).

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 4G, i; Supplementary Figure 6A, i, 6C, i). 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 4), and of hydrophobic interactions with R40 (Table 3). The R40/6(p) and K56-4(p) interactions for the D4(p)E variant are shown in (Figure 4H). 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 4D).

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 5). 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 5A). In accordance with results obtained by ITC (Table 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 (Supplementary Table 2), from which 12 showed a strong interaction with human PEX14 NTD (Figure 5B, Supplementary Table 2). 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 (Supplementary Figure 3) and other previous reports on Pex14-Pex5 interactions (Cyr et al. 2008; Hojjat and Jardim 2015; Watanabe et al. 2016).

Analysis of the peptide binding (Figure 5A, B) 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. With these data we propose a refined core consensus motif **WΦxE(F/Y)Φ** further expanded to **ExWΦxE(F/Y)Φ** for high affinity binding peptides (Figure 5C).

There are a few 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 4; Supplementary Figure 6). 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 5B). 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 core motif as **WΦxE(F/Y)Φ** (Figure 5C), 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 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 2A). 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. The observation that the E–2(p) position is strongly conserved in high affinity binding peptide motifs, and that experimental structures and our computational analysis indicate that it mediates important interactions with the conserved N38 and R40 side chains, defines an expanded consensus motif for high affinity interactions, **ExWΦxE(F/Y)Φ**, which includes the E–2(p) position (Figure 5C).

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 (Supplementary Figure 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 5), 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 (Supplementary Figure 1) showing that *N*38 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 threonine in position 2(p) a glutamate in position 6(p) where an aliphatic residue is preferred, and a proline in the –1(p) position, which disfavors a helical conformation. We found that a reverse motif with the sequence S**DFQEVW**DS in ScPex5 mediates binding to ScPEX14 NTD (Supplementary Figure 3). This motif is conserved among yeast and partly matches the consensus sequence when inverted (i.e. SD**WVEQFD**). The relatively weak, micromolar affinity for this motif may reflect a non-optimal sequence lacking E4(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. The low affinity binding is further consistent with the importance of N38 and R40 in Pex14 NTDs as D(E)38 and S(T)40 mutations as seen in yeast are indeed not compatible with high affinity binding.

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 5A) and the conservation of the PEX14 NTD (Supplementary Figure 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) 4(p) and 6(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.

Materials and methods

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 His6 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 His_6-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/L $^{15}NH_4Cl$ (Cambridge Isotope Laboratories), 2 g/L [U-^{13}C]-glucose-d12 (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 14–16 h 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 μ L volumes at an interval of 150 s 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% D_2O 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 $^{13}C\alpha$, $^{13}C\beta$ chemical shifts (Kjaergaard and 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 1 h 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, $Y = A2 + (A1-A2)/(1+\exp[(x-x_0)/dx)]$. Where $A1$ and $A2$ are the folding and unfolding intercept respectively. x is the midpoint of the curve and dx 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 His_6-tagged PEX14 NTD in TBS. Bound PEX14 NTD was immunodetected by monoclonal anti-His_6 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 1B bound to PEX14. As initial structures the complex of PEX14 NTD with PEX5 (PDB-ID: 2W84, peptide sequence W1, Table 2) was used. The sequence of PEX5 was mutated to the corresponding sequence for all other motifs using the IRECS (Hartmann et al. 2009; Hartmann et al. 2007) 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, 10,100 steps of restraint minimization (100 steps with the steepest descent algorithm and 10,000 steps with the conjugate gradient method) were done with the protein atoms restrained using a 50 kcal mol^{-1} $Å^{-2}$ force constant. Second, 100,100 steps of energy minimization (100 steps with the steepest descent and 100,000 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 45×10^{-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 and 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).

Acknowledgements: The authors dedicate this work to the memory of Prof. Iris Antes, who sadly passed away on August 4, 2021.

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.

Research funding: This work was supported by the Deutsche Forschungsgemeinschaft, FOR1905 (R.E. & M.S.), SFB1035 (M.S. & I.A.) and the Cluster of Excellence CIPSM (DFG EXC114 to I.A. & M.S.).

Conflict of interest statement: The authors declare that they have no conflicts of interest with the contents of this article.

References

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

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81: 3684–3690.

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

Brocard, C., Lametschwandtner, G., Koudelka, R., and Hartig, A. (1997). Pex14p is a member of the protein linkage map of Pex5p. EMBO J. 16: 5491–5500.

Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, T.E., Darden, T.A., Duke, R.E., and Gohlke, H. (2014). *Amber 14*. University of California, San Francisco, CA.

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

Choe, J., Moyersoen, J., Roach, C., Carter, T.L., Fan, E., Michels, P.A., and Hol, W.G. (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.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. 14: 1188–1190.

Cyr, N., Madrid, K.P., Strasser, R., Aurousseau, M., Finn, R., Ausio, J., and Jardim, A. (2008). *Leishmania donovani* peroxin 14 undergoes a marked conformational change following association with peroxin 5. J. Biol. Chem. 283: 31488–31499.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: AnN-log(N) method for Ewald sums in large systems. J. Chem. Phys. 98: 10089–10092.

Dawidowski, M., Emmanouilidis, L., Kalel, V.C., 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, V.C., 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.

Deb, R. and Nagotu, S. (2017). Versatility of peroxisomes: an evolving concept. Tissue Cell 49: 209–226.

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

Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., et al. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J. Comput. Chem. 24: 1999–2012.

Effelsberg, D., Cruz-Zaragoza, L.D., Tonillo, J., Schliebs, W., and Erdmann, R. (2015). Role of Pex21p for piggyback import of Gpd1p and Pnc1p into peroxisomes of. Saccharomyces cerevisiae J Biol Chem. 290: 25333–25342.

Emmanouilidis, L., Gopalswamy, M., Passon, D.M., Wilmanns, M., and Sattler, M. (2016). Structural biology of the import pathways of peroxisomal matrix proteins. Biochim. Biophys. Acta. 1863: 804–813.

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

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

Giannopoulou, E.A., Emmanouilidis, L., Sattler, M., Dodt, G., and Wilmanns, M. (2016). Towards the molecular mechanism of the integration of peroxisomal membrane proteins. Biochim. Biophys. Acta. 1863: 863–869.

Gould, S.J., Keller, G.A., Hosken, N., Wilkinson, J., and Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. J. Cell Biol. 108: 1657–1664.

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

Greenfield, N.J. (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.

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

Hartmann, C., Antes, I., and 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.

Hojjat, H. and Jardim, A. (2015). The Leishmania donovani peroxin 14 binding domain accommodates a high degeneracy in the pentapeptide motifs present on peroxin 5. Biochim. Biophys. Acta 1850: 2203–2212.

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

Jansen, R.L.M., Santana-Molina, C., van den Noort, M., Devos, D.P., and van der Klei, I.J. (2021). Comparative genomics of peroxisome biogenesis proteins: making sense of the PEX proteins. Front. Cell Dev. Biol. 9: 654163.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79: 926–935.

Kerssen, D., Hambruch, E., Klaas, W., Platta, H.W., de Kruijff, B., Erdmann, R., Kunau, W.H., and Schliebs, W. (2006). Membrane association of the cycling peroxisome import receptor Pex5p. J. Biol. Chem. 281: 27003–27015.

Kjaergaard, M. and Poulsen, F.M. (2011). Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. J. Biomol. NMR 50: 157–165.

Lazarow, P.B. and Fujiki, Y. (1985). Biogenesis of peroxisomes. Annu. Rev. Cell Biol. 1: 489–530.

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

Neufeld, C., Filipp, F.V., 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., Kooshapur, H., Wolf, J., Meyer, N.H., 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. J. Biol. Chem. 289: 437–448.

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

Otera, H., Setoguchi, K., Hamasaki, M., Kumashiro, T., Shimizu, N., and 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.

Rees, D.C. and Robertson, A.D. (2001). Some thermodynamic implications for the thermostability of proteins. Protein Sci. 10: 1187–1194.

Reuter, M., Kooshapur, H., Suda, J.G., 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. J. Mol. Biol. 433: 166765.

Roe, D.R. and Cheatham, T.E., 3rd (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J. Chem. Theor. Comput. 9: 3084–3095.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23: 327–341.

Saidowsky, J., Dodt, G., Kirchberg, K., Wegner, A., Nastainczyk, W., Kunau, W.H., and Schliebs, W. (2001). The di-aromatic pentapeptide repeats of the human peroxisome import receptor PEX5 are separate high affinity binding sites for the peroxisomal membrane protein PEX14. J. Biol. Chem. 276: 34524–34529.

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

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

Schwarzinger, S., Kroon, G.J., Foss, T.R., Chung, J., Wright, P.E., and Dyson, H.J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. J. Am. Chem. Soc. 123: 2970–2978.

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

Shiozawa, K., Konarev, P.V., Neufeld, C., Wilmanns, M., and Svergun, D.I. (2009). Solution structure of human Pex5.Pex14.PTS1 protein complexes obtained by small angle X-ray scattering. J. Biol. Chem. 284: 25334–25342.

Su, J.R., Takeda, K., Tamura, S., Fujiki, Y., and 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.

Tol, M.B., Deluz, C., Hassaine, G., Graff, A., Stahlberg, H., and Vogel, H. (2013). Thermal unfolding of a mammalian pentameric ligand-gated ion channel proceeds at consecutive, distinct steps. J. Biol. Chem. 288: 5756–5769.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59: 687–696.

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

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

Will, G.K., Soukupova, M., Hong, X., Erdmann, K.S., Kiel, J.A., Dodt, G., Kunau, W.H., and Erdmann, R. (1999). Identification and characterization of the human orthologue of yeast Pex14p. Mol. Cell Biol. 19: 2265–2277.

Yang, X., Purdue, P.E., and Lazarow, P.B. (2001). Eci1p uses a PTS1 to enter peroxisomes: either its own or that of a partner, Dci1p. Eur. J. Cell Biol. 80: 126–138.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/hsz-2022-0177).