J Mol Model (2009) 15:971–982
DOI 10.1007/s00894-008-0449-y

ORIGINAL PAPER

Evolutionarily evolved discriminators in the 3-TPR domain of the Toc64 family involved in protein translocation at the outer membrane of chloroplasts and mitochondria

Oliver Mirus · Tihana Bionda · Arndt von Haeseler · Enrico Schleiff

Received: 23 October 2008 / Accepted: 16 December 2008 / Published online: 6 February 2009

Abstract Transport of polypeptides across membranes is a general and essential cellular process utilised by molecular machines. At least one component of these complexes contains a domain composed of three tetratricopeptide repeat (3-TPR) motifs. We have focussed on the receptor Toc64 to elucidate the evolved functional specifications of its 3-TPR domain. Toc64 is a component of the Toc core complex and functionally replaces Tom70 at the outer membrane of mitochondria in plants. Its 3-TPR domain recognises the conserved C-terminus of precursor-bound chaperones. We built homology models of the 3-TPR domain of chloroplastic Toc64 from different species and of the mitochondrial isoform from *Arabidopsis*. Guided by modelling, we identified residues essential for functional discrimination of the differently located isoforms to be located almost exclusively on the convex surface of the 3-TPR domain. The only exception is *at*568Ser/*ps*557Met, which is positioned in the ligand-binding groove. The functional implications of the homology models are discussed.

Keywords Molecular dynamics simulation · Homology modeling · Toc64 · Tetratricopeptide repeat · Mitochondria · Chloroplasts · Chaperones

Introduction

Mitochondria [1] and plastids [2] were derived by an endosymbiotic process, necessitating the development of protein transport and translocation systems to ensure the exchange of proteins between cell compartments [3]. The evolution of translocation systems involved the invention of novel systems such as translocon of matrix-targetted proteins at the inner mitochondrial membrane [4] and "recycling" of prokaryotic systems [5]. The functional capacity of prokaryotic translocation machineries was subsequently fine-tuned by inserting additional components [6]. During this process of functional "optimisation" of protein translocation, several well known events of protein evolution took place, for instance domain stealing [7], enzyme recycling by modulation of enzymatic activities, and gene duplication [8]. A prominent example of the latter is Toc64.

Toc64 has an amidase domain silenced by a single point mutation in the catalytic centre [9], which still recognises a putative substrate, but does not process it [10]. Fused to the amidase, Toc64 contains a cytosolically exposed tetratricopeptide repeat (TPR) domain. TPR motifs were first identified in the nuclear protein ssn6 [11]; the chromatin-associated replication protein CDC23 [12]; and nuc2 +, a protein required for mitotic chromosome disjunction [13]. In recent years, the list of proteins carrying TPR motifs has grown steadily [14–18]. Receptors containing a 3-TPR domain that are involved in protein translocation include the mitochondrial receptors Tom20 [19], Tom70 [19, 20] and Tom34 [21–24], the peroxisomal receptor Pex5 [25, 26],

Electronic supplementary material The online version of this article (doi:10.1007/s00894-008-0449-y) contains supplementary material, which is available to authorized users.

O. Mirus · T. Bionda · E. Schleiff (✉)
Cluster of Excellence and Center of Membrane Proteomics,
Institute for Molecular Biosciences,
Johann Wolfgang Goethe University,
Max-von-Laue-Str. 9, N200, 60438 Frankfurt am Main, Germany
e-mail: schleiff@bio.uni-frankfurt.de

A. von Haeseler
Center for Integrative Bioinformatics Vienna, Max F. Perutz Laboratories, University of Vienna,
Vienna, Austria

and the endoplasmic reticulum translocon component Sec72 [27, 28].

The 34 amino acids of the TPR motif form two helices—'a' and 'b' (Fig. 1a)—connected by a short intrarepeat loop [14–16]. Three tandemly arranged TPR motifs, which form a right-handed superhelix [29] capped by a so-called solvation helix at the C-terminal end, are defined as a 3-TPR domain [30]. Within this 3-TPR domain, the motifs are again connected by short interrepeat loops, and the individual motifs are termed motifs I, II and III, respectively. The short interrepeat loops show a slightly higher mobility than the helices as revealed by NMR measurements [31]. Additionally, the amino terminal end and the solvation helix are found to be more flexible in terms of positioning and folding [31]. Finally, 3-TPR domains involved in Hsp70/Hsp90 (heat shock protein) recognition form a conserved dicarboxylate clamp, which specifically binds the two carboxylate groups of the, also conserved, C-terminal aspartate of Hsp70/Hsp90 [29]. Recently, the 3-TPR domain of Toc64 was described as being of the dicarboxylate clamp type [32].

Interestingly, in *Arabidopsis thaliana*, three isoforms of Toc64 (I, III, V, reflecting their chromosomal localisation) were initially described [33, 34]. Whereas the first isoform (*at*Toc64-I) behaves as a typical amidase [35, 36], the second functions as an import receptor of chloroplasts (*at*Toc64-III [8, 32]), and the third is located in the outer membrane of mitochondria, functionally replacing the mitochondrial Tom70 found in yeast and mammals (*at*Toc64-V [8, 37–38]). The chloroplast-localised receptor recognises Hsp90 loaded with precursor protein destined for chloroplast import via its cytosolically exposed 3-TPR domain [10, 32]. Nevertheless, Toc64 represents a non-essential receptor family [39, 40], which parallels findings for TPR motif-containing Tom components [19, 20, 34].

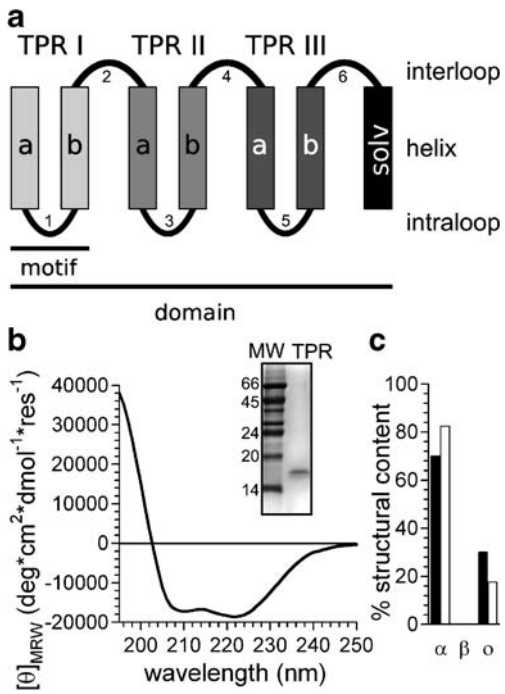


Fig. 1a–c Architecture of the three-tetratricopeptide repeat (3-TPR) domain of *ps*Toc64TPR. **a** Schematic representation of the 3-TPR domain showing the nomenclature used in this paper. **b** Expression and purification of recombinant 3-TPR domain of Toc64; purity was confirmed by SDS-PAGE analysis and Coomassie Blue staining (*inset*). The CD-spectrum (*line*) was recorded as described in Materials and methods. **c** Secondary structure content determined from the CD-spectrum (*black bars*) and according to the generated homology model (*white bars*, see Fig. 2); α α-helical content, β β-content, *o* other structural elements

The presence of a similar receptor in two different membranes/organelles, and their interaction with chaperones rather than with incoming precursor proteins raises the question of how the receptor—in particular its Hsp-recognising 3-TPR domain—provides the specificity required to discriminate between Hsps loaded with preprotein destined for their respective organelles? Using homology modelling, we investigated the structural properties of the 3-TPR domain of Toc64 from *Pisum sativum* (*ps*Toc64TPR), localised in the outer envelope of plastids [9], and of Toc64 from *Arabidopsis thaliana*, which is encoded on chromosome V (*at*Toc64-V TPR) and localised in the outer membrane of mitochondria. [8] At present, the most ancestral Toc64 investigated is that of the moss *Physcomitrella patens* [39]; we therefore included this Toc64-1 3-TPR domain (*pp*Toc64-1 TPR) in this study. Our analysis allowed us to derive the structural positioning of conserved amino acids, putative discriminative positions of mitochondrial and chloroplasts receptors, and the internal motions of the 3-TPR domain. The results can be applied to guide future biochemical experiments.

Materials and methods

Circular dichroism analysis

Circular dichroism (CD) analysis was performed as previously described [54, 55] using the heterologously expressed 3-TPR domain of *ps*Toc64 purified as published [32]. Briefly, the CD-spectrum was recorded in 10 mM Hepes/KOH, pH 7.6, and 100 mM NaCl on a Jobin Yvon CD6 spectrometer (Jobin YVON-SPEX, Edison, NJ) in a cuvette with 1 mm path length for far-UV at 22°C with 1 nm steps, 1 s integration time, and a slit width of 2 nm.

Homology modelling

The Phyre server [56] (http://www.sbg.bio.ic.ac.uk/~phyre/) was used to find suitable templates for modelling. The crystal structures of the TPR1 domain (1ELW) and the

TPR2a domain (1ELR) of Hop [29], and of protein phosphatase 5 [57] (1A17) were used as templates to model the 3-TPR domain of the different Toc64 proteins [32] (Table 1). Alignments of target and template sequences are available as supplementary material (Fig. S1). The homology model of the 3-TPR domain of *ps*Toc64 based on 1ELW was generated with WHAT IF [43]. We constructed a model of the 3-TPR domain from *ps*Toc64 based on 1ELR with Modeller v8.2 [42, 58]. Swissmodel [59] was applied to the 3-TPR domain of *at*Toc64-V and *pp*Toc64-1. All homology models were refined with Yasara's Yamber2 forcefield [45] (http://www.yasara.org). Ramachandran plots (RAMPAGE, http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) [60] of the homology models as well as the results of the structure validation server Verify3D (http://nihserver.mbi.ucla.edu/Verify_3D/) [61, 62] are available as supplementary material (Figs. S11, S12).

Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed with GROMACS [63–65] v3.2.1 or v3.3.1 and the GROMOS96 force field [66] (43a1) on AMD Opteron and Intel Woodcrest clusters. In total, we simulated 370 ns. The protonation states of histidine, lysine, aspartate, glutamate and cysteine were chosen to correspond to pH 7. As we are modelling C-terminally located protein domains, artificial N-termini were neutralised. Each protein was placed together with simple point charge (SPC) water [67] in a rhombic dodecahedron. The proteins were at least 1.4 nm away from the periodic boundaries of the box defined for a simulation. The systems were neutralised by addition of sodium or chloride ions, and subjected to a steepest descents energy minimisation. Next, a position restrained MD run of 100 ps with the heavy atoms of the protein immobilised was performed to relax the protein's hydrogen atoms and the water molecules.

Initial velocities of atoms in the system were taken randomly from a Maxwell-Boltzmann distribution at 300 K. We employed the Berendsen weak coupling scheme [68] to maintain the temperature at 300 K with a coupling time of 0.1 ps, the pressure at 1 atm with a coupling time of 0.5 ps, and an isothermal compressibility of 4.5×10^{-5} bar^{-1}. Long-range electrostatic interactions were calculated by the Particle Mesh Ewald (PME) method [69] with a cut off of 1.0 nm for the direct space sum and a spacing of 0.1 nm for the FFT (fast Fourier transform) grid. The short-range neighbour list was cut off at 1.0 nm and updated every fifth step. Van der Waals forces were cut off at 1.4 nm. Bond distances and bond angles were constrained by the SETTLE [70] and LINCS [71] algorithm. The integration time step was 2 fs. Snapshots of the simulated systems were recorded every 5 ps and protein coordinates were saved every 1 ps.

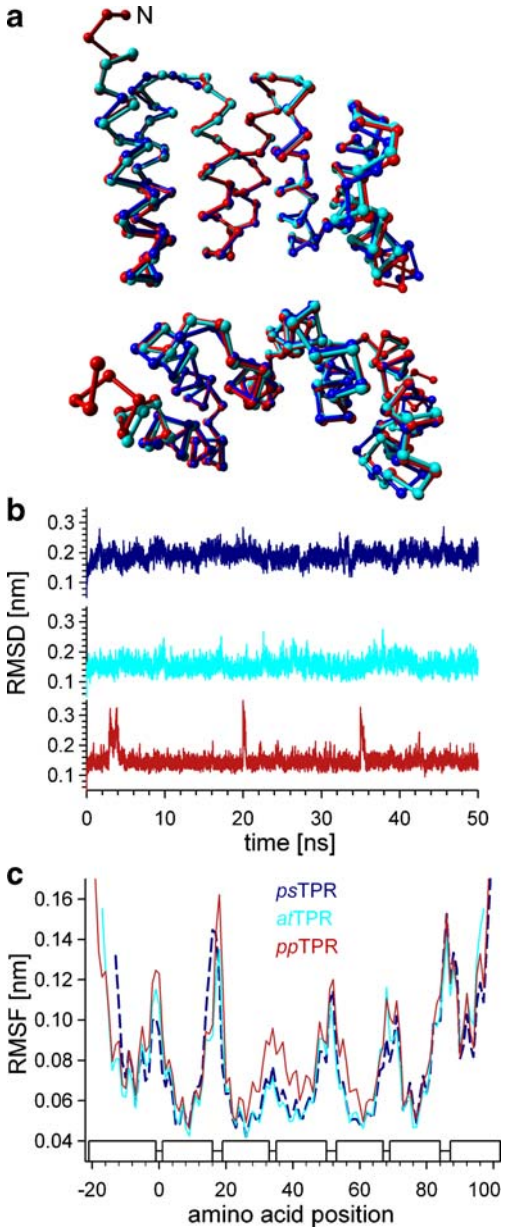
Each of these sets of structural coordinates represents one frame. As every protein was simulated for at least 50 ns, we accumulated at least 50,001 frames in each trajectory. The global rotational and translational degrees of freedom of a protein were removed by least squares fitting the frames on the starting structure of the respective trajectory by keeping the relative positions of the atoms within each structure. Each trajectory was split in up to 2,048 subintervals with steps of a factor of two.

Principal components analysis

Roughly the first nanosecond was excluded from the principal components analysis (PCA) to remove equilibration effects. N- and C-terminal residues of the proteins were excluded from the PCA, and only the C_α atoms of the remaining residues were chosen for further processing. We applied the PCA as described by Hess [44, 72]. We calculated the covariance matrix of the coordinates of the C_α atoms for all subintervals (prepared as described in the previous paragraph) and the full trajectory. The covariance matrix is a symmetrical matrix with a dimension corresponding to three times the number of C_α atoms analysed. The covariance matrix was diagonalised by an orthonormal transformation matrix. This diagonal matrix contains the eigenvalues; the columns of the transformation matrix are the corresponding eigenvectors. The columns of

Table 1 Homology modelling targets and templates. The name of the protein family, the name of the protein, the accession number, the amino acids of the sequence modelled, the PDB ID (http://www.pdb.org) of the template, the amino acids of the template used for modelling and the percentage of sequence identity between query and template are given

Protein	Target			Template		Identity
	Name	Accession no.	Amino acids	PDB ID	Amino acids	
Toc64	*ps*Toc64 (I)	AAF62870	474–593	1ELR	222–347	31%
	*ps*Toc64 (II)	AAF62870	477–593	1ELW	2–118	31%
	*at*Toc64-V	NP_196504	485–603	1A17	25–143	34%
	*pp*Toc64	AAS47584	466–592	1A17	19–145	34%



◀ **Fig. 2a–c** Homology models of the 3-TPR domains. **a** Homology models of the 3-TPR domains were built as described in Materials and methods. The structures shown are average structures derived from molecular dynamics (MD) simulations. *Cyan at*Toc64-V TPR from *Arabidopsis thaliana*, *blue ps*Toc64TPR (Model II) from *Pisum sativum*, *red pp*Toc64-1 TPR from *Physcomitrella patens*. Molecular graphics created with YASARA (www.yasara.org) and PovRay (www.povray.org). **b** The root mean square deviation (RMSD) for the individual MD simulations are shown (*black ps*Toc64 model II, *dark grey at*Toc64-V, *grey pp*Toc64-1, colour code the same in **a** and **b**). The trajectory of the MD simulation of *pp*Toc64-1 TPR was constructed by concatenating trajectories of three MD simulations. The 2nd and 3rd MD simulation (20–35 ns and 35–50 ns) were started from picoseconds 2,757 and 3,856, respectively. These two frames correspond to the sharp peaks during the first 5 ns. **c** The root mean squared fluctuations (RMSF) of the models were calculated from the trajectories of the performed MD simulations. Above the *x*-axis, the secondary structure elements of the average structures of the simulated models are given as *boxes* (α-helices) and *lines* (loop regions) connecting the boxes. Residues *at*502Lys, *ps*492Lys and *pp*489Lys are aligned to position 0 on the abscissa

the eigenvector matrix were sorted by size of the corresponding eigenvalues, so that the eigenvector with the highest eigenvalue constitutes the first and the eigenvector with the lowest eigenvalue, the last eigenvector.

The average overlap of the sampling of each subinterval with the sampling of the full-length trajectory was calculated for the smallest set of number of eigenvectors describing at least 80% of the internal motions of the simulated protein. The average cosine content of the 1st principal component and the average root mean squared fluctuations (RMSF) of the subintervals were calculated. To normalise the results, the autocorrelation (*a*) of the smallest set of number of principal components describing at least 80% was determined.

In the following, we performed a pair-wise comparison of the MD simulations. By calculating pair-wise inner-products of the first few eigenvectors of MD simulation pairs, we constructed matrices that reflect the similarity of the internal motions of the simulated proteins. Coordinates of the initial models and average structures of the MD simulations are available upon request.

MD analysis and visualisation

A visual inspection of the MD trajectories was performed with VMD [73]. The root mean square deviation (RMSD) traces and the RMSF with respect to the starting structure of each simulation were calculated for the full length of the trajectories.

We visualised the principal modes of motion of the first few eigenvectors as porcupine plots [74]. These porcupine plots can also be automatically generated by the DynaTraj server developed by Barret and co-workers [75]. Each element of an eigenvector holds the direction and strength of motion of the corresponding C_α atom. These features are visualised as cones originating from each C_α atom of the average structure obtained from the MD simulation. The cone length represents the strength of the motion. In order to obtain visible cones, the eigenvectors were scaled by the inverse cubed square root of the number of residues (see DynaTraj server online documentation). Such a porcupine plot is identical to a plot in which the direction of all cones has been exactly reversed. The porcupine plots were visualised with VMD and rendered with PovRay (http://www.povray.org).

H-bonds and water-inserted H-bonds were detected with the g_hbond tool of Gromacs 3.1.4. The donor-hydrogen-

acceptor cutoff angle was 60° and the hydrogen-acceptor distance cutoff 0.25 nm. A donor atom donates a polar hydrogen and an acceptor atom binds this hydrogen via a free electron pair. A Python script was developed to merge H-bond data for donor/acceptor atoms of the same residue.

The interaction energies of receptor residues with the ligand were calculated by rerunning the MD simulations with the recorded trajectories. For the reruns, we employed a simple twin-range cutoff for the coulomb (2.25 nm) and van der Waals (1.4 nm) forces to avoid time-consuming multiple reruns of each simulation to extract the long-range interaction energies from the PME energy of reciprocal space. We are aware that this approach reduces the precision of the results, but for our purposes these errors are acceptable. The electrostatic energies were not multiplied by the common scaling factor of 0.5, which would take into account a distribution of the energy over both interacting partners.

A Yasara Python plugin was developed to label and colour those residues for which interaction energies with the ligand have been calculated according to the strength of this interaction. Yasara's colour circle goes from 0 to 360 and is composed of the following colours: blue (0/360), cyan (300), green (240), yellow (180), red (120), and magenta (60). Green was defined as zero energy, blue and red as maximal absolute energy. The interaction energies for the receptor residues were scaled by the following scheme: colourID=240+x, for x>0 and colourID=240−x for x<0 with $x=120*\log_2(\mathrm{abs}([\mathrm{energy}]) + 1)/\log_2(\mathrm{abs}([\mathrm{energy}_{max}])+1)$.

Results

Homology modelling and MD simulations of the 3-TPR domain of Toc64

We applied a homology modeling approach to explore structural features of the Toc64 3-TPR domain as no crystal or NMR structure is available yet. This task was supported by secondary structure analysis of the Toc64 3-TPR domain from *Pisum sativum* (*ps*Toc64TPR). Circular dichroism measurements of this domain (Fig. 1b) revealed an α-helical content of about 70% and no indications of β-strands (Fig. 1c). This finding parallels the structural composition of typical 3-TPR domains [29]. The specificity of our homology modeling approach was determined by building two models of the 3-TPR domain of *ps*Toc64 utilising either PDB:1ELR (model I) or PDB:1ELW (model II) as a template, which were ranked 1st and 2nd by the PHYRE server, respectively (Table 1; Fig. 2a). For both, the amino acid identity between *ps*Toc64TPR and the template lies within the homology modelling zone [41]. The homology model of *ps*Toc64TPR based on the template 1ELR was built using Modeller [42], because two loops had to be reconstructed. In contrast, the alignment of *ps*Toc64TPR and 1ELW has no gaps (Fig. S1), and the model was built utilising WHAT IF [43]. With both homology models, MD simulations of 60 and 50 ns, respectively, were performed (Table S1). The structural changes are measured as the RMSD of the C_α atoms with respect to the starting structure [Fig. 2b, blue (model II), see Fig. S2a for model I]. A comparison of the average structures (derived from MD simulations) with the templates 1ELR and 1ELW revealed similar RMSDs (1.7 Å or 1.6 Å for model I, 1.6 Å or 1.6 Å for model II; 1.9 Å for 1ELR vs 1ELW). The RMSD between the two average structures equals 0.5 Å, which suggests that very similar structures were adopted independent of the templates and the homology modelling strategy.

Next, we built homology models of the 3-TPR domain of the *A. thaliana* protein with putative mitochondrial localisation (*at*Toc64-V), and of the *Physcomitrella patens* protein (*pp*Toc64-1) with plastid localisation (Table 1, Fig. 2a). Like the models of *ps*Toc64TPR, the homology models of the 3-TPR domain of *at*Toc64-V and *pp*Toc64-1 were stable during the MD simulations as reflected by their respective RMSD traces (Fig. 2b, cyan and red, respectively) with an average RMSD of ~1.5 Å for both *at*Toc64-V and *pp*Toc64-1. For all models, the highest fluctuations of the C_α-atoms were observed in loop regions, the solvation helix and in the amino terminal helix (Fig. 2c, Fig. S2b), which parallels findings by NMR-spectroscopy [31]. Furthermore, from the PCA (see Supplementary material) [44] of the models we conclude that we have sufficiently sampled conformations associated with the free energy minima that have been visited during the MD simulations (Table S2). Summarising, the homology models passed all of our tests, enabling further analyses.

Analysis of 3-TPR domain models complexed with the C-terminus of Hsp90

To explore the functional implications of our model, we analysed the flexibility of the 3-TPR domain in the presence of a bound octapeptide. For this purpose, the last eight amino acids of the Hsp90 C-terminus were modelled based on homology to the crystal structure of the Hop TPR1/Hsp70 C-terminal 8-mer complex (PDB: 1ELW). The homology models of *at*Toc64-V TPR and *ps*Toc64 (model II) were superimposed onto this crystal structure. The sequence of the octapeptide was mutated with the program Yasara to correspond to Hsp90 [32]. Afterwards, we performed energy minimisations of the protein–peptide complexes with the Yamber2 force field [45] followed by MD simulations of 80 ns each. The average RMSD of the

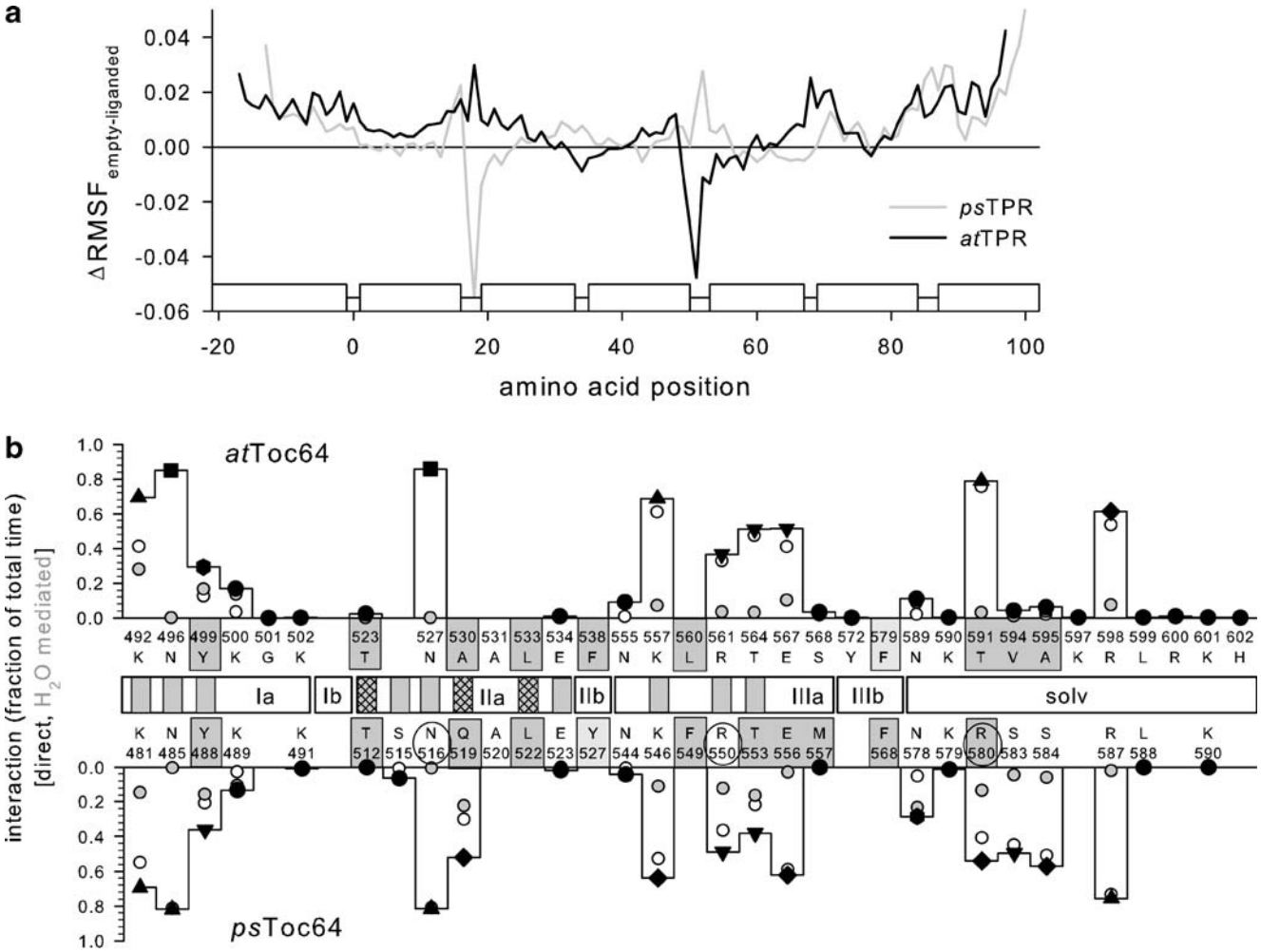


Fig. 3a,b Modelled peptide binding by the 3-TPR domain of *at*Toc64-V and *ps*Toc64. **a** Time-dependent RMSF of C_α atoms as a function of the length of the subintervals calculated for *at*Toc64-V (*black*) and *ps*Toc64 (*grey*) complexed with the peptide. **b** Map of residues of the 3-TPR domains forming either H-bonds or putative van der Waals contacts. The graph shows all residues of the liganded 3-TPR domain of *at*Toc64-V (*top*) or *ps*Toc64 (*bottom*) involved in H-bond formation (*dots* in graph) or van der Waals contacts (residues in *grey boxes*) with the ligands. For H-bonds, the relative number of frames where a direct (*black circle*) or water-mediated H-bond (*grey circle*) exists during the MD simulation (excluding the initial equilibration phase) is shown (please note that one amino acid can form direct or mediated H-bonds at the same time). A detailed H-bond plot is available as Supplementary material. The relative positions of the helices are indicated in the central part of the figure. *Grey regions* in the helices indicate electrostatic interactions and *hatched regions* hydrophobic interactions with the corresponding positions in the crystal structure of Hop [29]. *Circled positions* (bottom) indicate residues previously mutated within the 3-TPR domain of *ps*Toc64 [32]

liganded 3-TPR domain of *at*Toc64-V (Fig. S3a) or *ps*Toc64 (not shown) was lower than observed during the MD simulation of the 3-TPR domain alone (Table S2). Further, the RMSD of the bound octapeptide was quite low (Fig. S3b, black). The analyses of the RMSD distribution of both the *Arabidopsis* and the pea system show that the internal motions of the 3-TPR domain are—as expected [46]—more restricted when the peptide is bound in the binding pocket (Fig. S3c), although to a slightly different extent in the two systems (see Supplementary material). Performing a PCA of the complexed 3-TPR systems, we obtained a quality of sampling similar to that of the free 3-TPR domain (Fig. S4, Table S1).

Detailed analysis (see Supplementary material) revealed that the majority of peptide conformations of both 3-TPR domains' ligands was assigned to a single cluster, but a portion of conformations of the ligand of *ps*Toc64TPR was also found to be significantly different from the *Arabidopsis* system (Fig. S5a,b). This might be a first hint that the mode of association of these two 3-TPR domains with their chaperone might differ. In line with this notion, we observed that the fluctuation restriction of residues of the 3-TPR domain by peptide binding differs between *at*Toc64-V and *ps*Toc64 (Fig. 3a). In *ps*Toc64TPR, loop 4 appears to be restricted in its motion by the bound peptide, whereas loop 2 shows a higher mobility; in *at*Toc64-V TPR, the

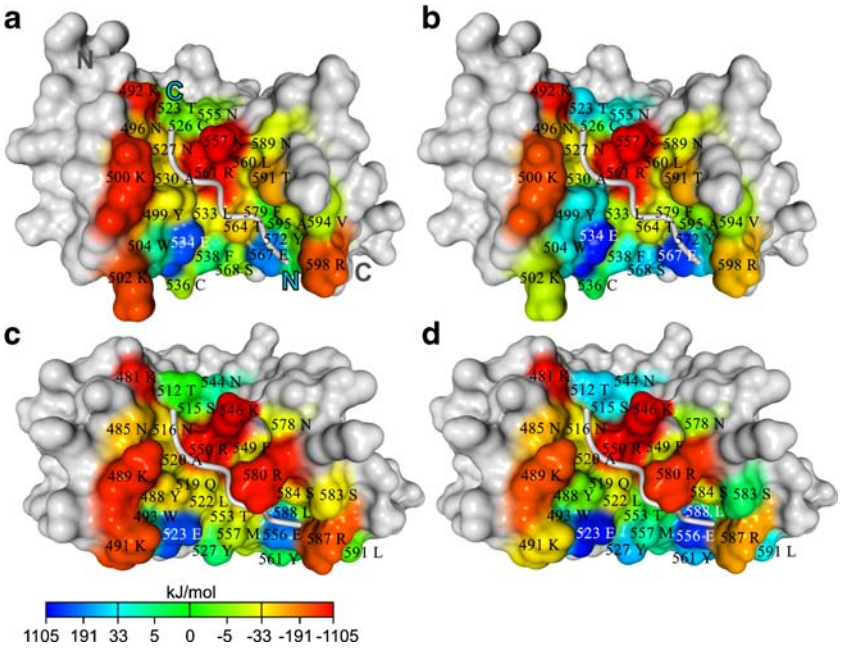


Fig. 4a–d Electrostatic interaction energies. Non-bonded Coulomb and Lennard-Jones (LJ) interaction energies were calculated for several residues of the receptor domain with its ligand as described in Materials and methods. The homology models of the receptor–ligand complexes were taken from late time points of their respective trajectory from MD simulations. The 3-TPR domain is shown in *surface representation* and the ligand as a *ribbon model*. Residues are named by their position in the amino acid sequence of the respective protein. **a**, **b** *at*Toc64-V TPR; **c**, **d** *ps*Toc64TPR. In **a**, N and C termini of the 3-TPR domain and the ligand are indicated in *grey* and *cyan* letters, respectively. In **a** and **c**, the receptor residues interacting with the ligand are coloured by the sum of their non-bonded Coulomb and LJ interaction energies: *bluish colour* disfavoured interaction, *yellowish* or *reddish colours* favoured interaction. The darker the blue/red colour, the more disfavoured/favoured the interaction. In **b** and **d**, differences in the interaction energies of receptor residues with the ligand and water are mapped onto the surface of the 3-TPR domain. A *bluish colour* shows that the strength of interaction of a receptor residue with water is stronger than with the ligand; the opposite situation is indicated by *yellowish* and *reddish colours*. The darker the blue/red colour, the stronger/weaker the interaction of the residue with water. Molecular graphics created with YASARA (www.yasara.org) and PovRay (www.povray.org)

converse is true. For both, the strongest restriction of movement is found for the solvation helix and helix Ia, which is as expected as both are terminal helices also involved in ligand binding.

Looking at intramolecular H-bonds, and selecting those that are present in at least 20% of the recorded frames in the corresponding trajectory (excluding the first 5 ns as equilibration time), we observed changes in the H-bond network when comparing the free 3-TPR domain with its ligand-bound state (Tables S3–S6). On the convex side of the 3-TPR domain we found that the conserved residues *at*528/*ps*517Arg and *at*562/*ps*551Arg act as a sort of middleman, bridging the distance between helices Ib-IIb and IIb-IIIb, respectively, by forming H-bonds with *at*512/*ps*501Thr (Ib), *at*544/*ps*533Asp (IIb) and *at*546/*ps*535Thr (IIb), *at*578/*ps*567Asp (IIIb), respectively. In *at*Toc64-V, *at*566Arg also bridges helices IIb-IIIb, contacting *at*542Glu and *at*578Asp. However, the side-chains of *at*566Arg as well as the corresponding amino acid in *ps*Toc64TPR, *ps*555Arg, show stronger motion than *at*562Arg and *ps*551Arg (data not shown). *At*566Arg and *ps*555Arg form H-bonds with several residues, but the frequencies of these interactions stay below our threshold for *ps*555Arg (data not shown). On the concave side—the binding groove—the H-bond network switches certain H-bonds upon ligand binding. In the free 3-TPR domain of *ps*Toc64, this network reaches from residue *ps*485Asn over *ps*500Tyr, *ps*516Asn, *ps*519Gln, *ps*488Tyr to *ps*523Glu. Also, *ps*515Ser forms an H-bond with *ps*550Arg. In the ligand-bound state, H-bonds *ps*500Tyr–*ps*516Asn and *ps*488Tyr–*ps*523Glu are diminished, but *ps*485Asn–*ps*516Asn and *ps*519Gln–*ps*550Arg increase their frequency of interaction. In contrast to *ps*Toc64, the 3-TPR domain of *at*Toc64-V does not possess such a distinctive H-bond network. In *at*Toc64-V TPR, the position equivalent to *ps*515Ser is a cysteine (*at*526Cys), and that of *ps*519Gln an alanine (*at*530Ala), both of which interrupt the network. In *at*Toc64-V we also observe a network connecting helices Ia, Ib and IIa via several residues in the order *at*518Gln, *at*488Ser, *at*524Tyr, *at*489Glu, *at*492Lys.

We next analysed MD simulations with respect to the dynamics of the H-bond network between the 3-TPR

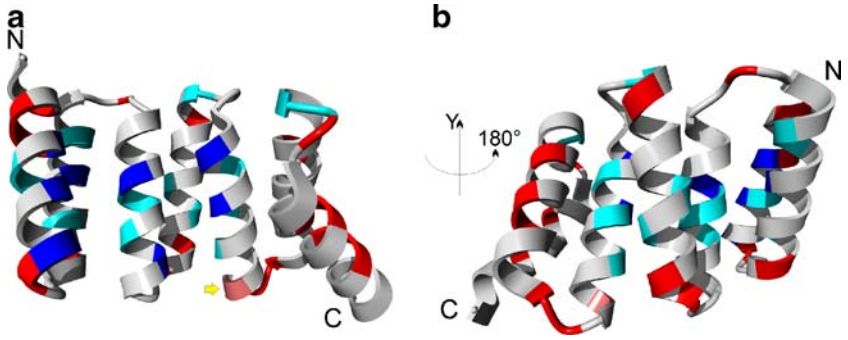


Fig. 5a,b Conserved amino acids and amino acids involved in peptide interactions. The positioning of conserved amino acids within the 3-TPR domain of mitochondrial and plastidic Toc64 proteins was analysed (Fig. S10). **a** The model I of *ps*Toc64TPR shown represents the view into the chaperone binding pocket. **b** 180° rotation around the *y*-axis. *Blue positions* are conserved in the 3-TPR domain of both mitochondrial and plastid-localised Toc64 proteins and are involved in substrate recognition. *Cyan positions* are conserved in the 3-TPR domain of both mitochondrial and plastid-localised Toc64 proteins, but these residues are not involved in substrate recognition. Positions of differing amino acid compositions between the 3-TPR domains of plastidic and mitochondrial localised Toc64 proteins are coloured in *red*. The *yellow arrow* points to a position that differs between the two sub-trees [51] but is involved in chaperone recognition. Molecular graphics created with YASARA (www.yasara.org) and PovRay (www.povray.org)

domain and the octapeptide (Fig. S6a). For clarity, we have linked the H-bond data for every receptor residue (Fig. 3b). We identified 27 residues in the 3-TPR domain of *at*Toc64-V that form H-bonds with the octapeptide, either directly or mediated by water molecules. However, only 10 residues were found to have H-bonds that are present in more than 20% of the stored frames of the MD simulation (Fig. 3b). In addition, we analysed the distances between residues of the 3-TPR domain and residues of the octapeptide. We found six residues of the 3-TPR domain positioned such that they can form van der Waals interactions with residues of the octapeptide (Fig. S6b,c; Fig. 3b, amino acids in grey boxes). This analysis shows that most H-bonds formed between the 3-TPR domain and the ligand are due to residues in helices Ia, IIIa and the solvation helix. Three residues of helix IIa (residues 523, 530 and 533) are putatively involved in van der Waals interactions with the ligand. In the pea 3-TPR domain peptide complex, we identified 24 residues forming H-bonds with the ligand (Fig. 3b, bottom), of which 14 are present in more than 20% of the frames. The van der Waals interactions are quite similar for *ps*Toc64TPR, although a few differences should be noted. The side-chain of *ps*Gln519 forms H-bonds with the ligand and provides a larger hydrophobic surface than *at*Ala530 at the corresponding position in *at*Toc64-V TPR. Also, *ps*Met557 is capable of stronger van der Waals interactions than *at*Ser568. Even in the solvation helix, the capabilities for hydrophobic interactions are increased. Although *at*Val594 is replaced by the small *ps*Ser583, the presence of *ps*Arg580 instead of *at*Thr591 more than compensates this change. To conclude, the overall distribution of interacting amino acids is quite similar for the 3-TPR domain of *at*Toc64 and *ps*Toc64. However, the pea 3-TPR binds its ligand with more stable H-bonds formed by residues of the N-terminal part of the solvation helix (Fig. 3b).

We calculated matrices of the mean smallest pairwise residue distances of the 3-TPR domain of *at*Toc64-V and *ps*Toc64 in the free and ligand-bound state from the trajectories of the MD simulations (Fig. S8a,b). The results are as expected for contact maps of the 3-TPR domains [47]. The only peculiarity is the short distance (slightly below the cutoff of 0.5 nm) between the residue pair *at*499Tyr and *at*561Arg in the ligand-bound state of *at*Toc64-V. Both residues are involved in binding the Hsp90 C-terminal peptide (Fig. 3b). When Hsp90 is bound in the binding groove, the frequency of the presence of the H-bond *at*499Tyr–*at*534Glu decreases (data not shown), as also observed for *ps*Toc64TPR, and the interaction with the ligand brings the sidechains of *at*499Tyr and *at*561Arg closer together. But why are the corresponding residues in *ps*Toc64TPR further apart? The answer is again the highly conserved alanine in *at*Toc64-V TPR at position 530, which in *ps*Toc64 is occupied by a glutamine (*ps*519Gln). So, in *at*Toc64-V, the sidechains of *at*499Tyr and *at*561Arg can come closer together, because there is no bulky glutamine sidechain in the way. We calculated a matrix containing the difference of the amino acid distance in the free and ligand-bound state for both *at*Toc64-V TPR and *ps*Toc64TPR (Fig. S8c). This matrix reflects the observed differences in the helical cross-angles of the 3-TPR domains in the free and the ligand-bound state (Fig. S7), and confirms that the distances of the amino acids are predominantly enlarged.

We also took a look at the electrostatic interactions of receptor residues in contact with the ligand (Fig. 4, Tables S7, S8). As expected, the residues that form the dicarboxylate clamp show high interaction energies. Also, *ps*Arg580, which was shown to be crucial for Hsp-binding [32], exhibits interaction energies similar to Arg and Lys residues (*ps*Lys481, *ps*Lys489, *ps*Lys546, *ps*Arg550) of the conserved clamp-type motif. Quite surprisingly, the two

glutamates (*at*Glu534/*ps*Glu523, *at*Glu567/*ps*Glu556) show stronger binding to H_2O than to the peptide.

Discussion

Flexibility of the 3-TPR domain of Toc64

Cliff et al. [48] proposed a coupled folding and binding mechanism for ligand binding by 3-TPR domains. They observed a partially unfolded protein phosphatase 5 (PP5) 3-TPR domain in solution, which is fully folded upon ligand binding. This result was challenged by Cortajarena and Regan [49], who showed that the 3-TPR domain of PP5 is fully folded in solution and does not undergo major structural changes upon ligand binding; instead, only minor changes were observed. The latter parallels our findings, because the 3-TPR domain of *ps*Toc64 is folded in solution (Fig. 1). Our MD simulations of the *at*Toc64-V TPR domain with and without ligand further support the previously proposed model [49] as we also observed only small changes of the helix cross-angles within the 3-TPR domain upon ligand binding (Fig. S7). In addition, a designed 3-TPR domain, which specifically binds to the C-terminal peptide of Hsp90, was shown to unfold at temperatures >70°C [50]. This further supports the notion that 3-TPR domains might be folded at room temperature in general. In agreement with the NMR analysis of Cheng et al. [31], for all homology models we obtained the highest flexibility enforced by the peptide in the amino terminal helix and by the solvation helix (Fig. S9).

Analysis of the MD simulations of the liganded 3-TPR domains revealed the differential behaviour of the mitochondrial *at*Toc64-V and the plastidic *ps*Toc64 systems (Fig. 3a). Introducing the chaperone C-terminus into the binding groove of the 3-TPR domain of *at*Toc64-V shifts the emphasis within the first few dominating eigenvectors from motions along helix IIb, or orthogonal to the superhelical axis of the 3-TPR domain, towards motions perpendicular to the longitudinal axis of the ligand. Thus, the dominating motions of an unliganded 3-TPR domain open up the binding site to ease recognition of the chaperone C-terminus. In contrast, the motions of the pea system are not as strongly altered, as visualised by the inner-product of the first few eigenvectors (Fig. S9a). As expected, a reduction in the flexibility of the liganded 3-TPR domain was obtained for both systems (Fig. S3).

3-TPR domains of Toc64 proteins with different localisation

The distinct localisations of the mitochondrial and plastidic isoforms of Toc64, and therefore the interaction with different partners, leads to selective pressure on these sequences, which has resulted in the divergence of the mitochondrial and plastidic Toc64 TPR cluster [51]. Examination of the amino acid composition revealed that the mitochondrial-localised Toc64 isoforms contain overall less negative (aspartate: −19%; glutamate: −13%) and more positive charges (arginine: +21%) in their 3-TPR domains, and they also prefer leucine (+36%) over isoleucine (−31%). Mapping conserved residues (Fig. S10) onto the modelled structure, we found that they are involved mostly in stabilising intramolecular interactions within the 3-TPR domain or in the recognition of the Hsp C-terminus (Fig. 5a,b; blue). This agrees with the assumption that Toc64 isoforms at both the chloroplastic and the mitochondrial outer membrane facilitate the recognition of Hsp90 delivering precursor proteins [8, 20, 32]. The positions at which significant difference in amino acid frequency occurs between the mitochondrial and plastidic isoforms are located almost exclusively at the back side with respect to the binding groove of the 3-TPR domain (Fig. 5b). Hence, these residues might be involved in the association of the 3-TPR domain with the Toc [32] or Tom complex as found for Tom70 [52]. Thus, the evolutionary development leading to localisation-specific features was driven by alteration of positions putatively interacting with other complex members, rather than by alteration of amino acids in the chaperone binding pocket.

Intramolecular H-bonds in the 3-TPR domain of Toc64 proteins

Our MD simulations identified a distinct network of H-bonds in the peptide-binding groove as well as on the convex side of the 3-TPR domain (Tables S3–S6). Upon ligand binding, certain H-bonds in this network are diminished or amplified. This “switching” of H-bonds is more pronounced on the convex than on the concave side of the 3-TPR domain. Interesting to note is the dramatic effect of two residues, which differ between *at*Toc64-V and *ps*Toc64, on the H-bond network within the binding groove: the presence of *ps*519Gln instead of the usual *at*530Ala (Fig. S10) and *ps*515Ser instead of *at*526Cys lead to the gain of five and three H-bonds in the free and ligand-bound state of *ps*Toc64TPR, respectively. However, the consequence of this enhanced H-bonding in *ps*Toc64TPR remains to be investigated.

Chaperone binding site of the 3-TPR domain of Toc64 proteins

Statistical analysis of a databank of 3-TPR domains identified positions in TPR motifs that are conserved in

Hsp90-binding 3-TPR domains [50]. In our model of the 3-TPR domain of *ps*Toc64, these positions include residues 516 and 550, which show a pronounced interaction with the octapeptide via H-bonds (Fig. 3). Mutations of these particular amino acids were previously found to inhibit the interaction with Hsp90 [32]. Comparing the profile of the interacting amino acids obtained for the Hsp90-recognising motif of Hop [29] with our interaction map, we find a significant consensus, but also distinct differences. For instance, position *at*Cys526/*ps*Ser515 (aligned to position 263 of Hop) forms a water-inserted H-bond in the crystal structure of Hop. In our MD simulations, however, other than sporadic H-bonds of *ps*Ser515 distributed over the whole trajectory (Fig. S6a), we do not observe formation of H-bonds with the octapeptide at this position of the 3-TPR domain. In addition, in the "chloroplastic" sequence, a glutamine at position *ps*519 forms H-bonds that are not present in *at*Toc64-V or Hop [29], but this is rather an exception, because this position in Toc64 is occupied mostly by an alanine (position 46 in Fig. S10). Furthermore, in the crystal structure, no contacts to residues in helices Ib and IIb were reported [29], although the methionine of the Hsp90 peptide clearly contacts residue Phe240 in the TPR2a domain of Hop (PDB: 1ELR). In Toc64, this position is occupied mostly by a bulky tryptophan. In our simulations, we observed that the peptide's methionine is bound in a largely hydrophobic pocket formed by *at*533Leu/*ps*522Leu, *at*534Glu/*ps*523Glu, *at*538Phe/*ps*527Tyr, *at*564Thr/*ps*553Thr, and *at*568Ser/*ps*557Met.

In this context it is interesting to note that at position *at*568/*ps*557 (position 84 in Fig. S10), the putatively plastid-localised proteins contain a methionine in the majority of cases (13/18). In the mitochondrial localised proteins this position is quite variable, and in 8 out of 14 cases the above mentioned binding pocket is opened up by the presence of a rather small amino acid (serine, alanine). In Hop, this residue is involved in chaperone binding [29], whereas in PP5 residues on the side of its 3-TPR domain exposing the intrarepeat loops are responsible for intramolecular binding to and autoinhibition of PP5 [53]. Our MD simulations show that the interaction energies of *at*568Ser/*ps*557Met with the chaperone C-terminus are smaller than with water (Fig. 4b,d). Hence, this position might not be so important for chaperone recognition after all. Therefore, it is tempting to assume a role for this position in intramolecular interaction analogous to that of PP5. Indeed, at least most of the first ~150 residues of *ps*Toc64 are exposed to the cytosol [10], presenting a possible intramolecular interaction partner for the 3-TPR domain. Another alternative explanation might be that this residue serves as a specificity factor for interaction with other members of the Toc or Tom complex. Hence, our analysis pinpoints a functionally relevant position, which should be approached experimentally in future.

Acknowledgements We would like to thank Lutz Voigt for technical help, and Thomas Schlegel and Nicole Scherer for their bioinformatic support. Special thanks to Thomas Becker, Joanna Tripp and Jason Young for helpful discussions regarding the project. The work was supported by grants from the Deutsche Forschungsgemeinschaft (SFB-TR01) and from the Volkswagenstiftung to E.S. and Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF) to A.v.H.

References

1. Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481
2. McFadden GI (1999) Endosymbiosis and evolution of the plant cell. Curr Opin Plant Biol 2:513–519
3. Wickner W, Schekman R (2005) Protein translocation across biological membranes. Science 310:1452–1456
4. Dolezal P, Likic V, Tachezy J et al (2006) Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318
5. Löffelhardt W, von Haeseler A, Schleiff E (2007) The β-barrel shaped polypeptide transporter, an old concept for precursor protein transfer across membranes. Symbiosis 44:33–42
6. Kalanon M, McFadden GI (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. Genetics 179:95–112
7. Hulett JM, Lueder F, Chan NC et al (2008) The transmembrane segment of Tom20 is recognized by Mim1 for docking to the mitochondrial TOM complex. J Mol Biol 376:694–704
8. Chew O, Lister R, Qbadou S et al (2004) A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. FEBS Lett 557:109–114
9. Sohrt K, Soll J (2000) Toc64, a new component of the protein translocon of chloroplasts. J Cell Biol 148:1213–1221
10. Qbadou S, Becker T, Bionda T (2007) Toc64—a preprotein receptor at the outer membrane with bipartide function. J Mol Biol 367:1330–1346
11. Schultz J, Marshall-Carlson L, Carlson M (1990) The N-terminal TPR region is the functional domain of SSN6, a nuclear phosphoprotein of *Saccharomyces cerevisiae*. Mol Cell Biol 10:4744–4756
12. Sikorski RS, Boguski MS, Goebl M et al (1990) A repeating amino acid motif in CDC23 defines a family of proteins and a new relationship among genes required for mitosis and RNA synthesis. Cell 60:307–317
13. Hirano T, Kinoshita N, Morikawa K et al (1990) Snap helix with knob and hole: essential repeats in *S. pombe* nuclear protein nuc2+. Cell 60:319–328
14. Goebl M, Yanagida M (1991) The TPR snap helix: a novel protein repeat motif from mitosis to transcription. Trends Biochem Sci 16:173–177
15. Lamb JR, Tugendreich S, Hieter P (1995) Tetratrico peptide repeat interactions: to TPR or not to TPR? Trends Biochem Sci 20:257–259
16. D'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. Trends Biochem Sci 28:655–662
17. Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P (2004) SMART 4.0: towards genomic data integration. Nucleic Acids Res 32:D142–D144

18. Finn RD, Tate J, Mistry J, Coggill PC, Sammut JS, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A (2008) The Pfam protein families database. Nucleic Acids Res 36: D281–D288
19. Haucke V, Horst M, Schatz G et al (1996) The Mas20p and Mas70p subunits of the protein import receptor of yeast mitochondria interact via the tetratricopeptide repeat motif in Mas20p: evidence for a single hetero-oligomeric receptor. EMBO J 15:1231–1237
20. Young JC, Hoogenraad NJ, Hartl FU (2003) Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell 112:41–50
21. Nuttall SD, Hanson BJ, Mori M et al (1997) hTom34: a novel translocase for the import of proteins into human mitochondria. DNA Cell Biol 16:1067–1074
22. Young JC, Obermann WM, Hartl FU (1998) Specific binding of tetratricopeptide repeat proteins to the C-terminal 12-kDa domain of hsp90. J Biol Chem 273:18007–18010
23. Chewawiwat N, Yano M, Terada K et al (1999) Characterization of the novel mitochondrial protein import component, Tom34, in mammalian cells. J Biochem (Tokyo) 125:721–727
24. Yang CS, Weiner H (2002) Yeast two-hybrid screening identifies binding partners of human Tom34 that have ATPase activity and form a complex with Tom34 in the cytosol. Arch Biochem Biophys 400:105–110
25. Gatto GJ Jr, Geisbrecht BV, Gould SJ et al (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. Nat Struct Biol 7:1091–1095
26. Kumar A, Roach C, Hirsh IS et al (2001) An unexpected extended conformation for the third TPR motif of the peroxin PEX5 from *Trypanosoma brucei*. J Mol Biol 307:271–282
27. Feldheim D, Schekman R (1994) Sec72p contributes to the selective recognition of signal peptides by the secretory polypeptide translocation complex. J Cell Biol 126:935–943
28. Ponting CP (2000) Proteins of the endoplasmic-reticulum-associated degradation pathway: domain detection and function prediction. Biochem J 351:527–535
29. Scheufler C, Brinker A, Bourenkov G et al (2000) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell 101: 199–210
30. Jabet C, Sprague ER, VanDemark AP et al (2000) Characterization of the N-terminal domain of the yeast transcriptional repressor Tup1. Proposal for an association model of the repressor complex Tup1 x Ssn6. J Biol Chem 275:9011–9018
31. Cheng CY, Jarymowycz VA, Cortajarena AL et al (2006) Repeat motions and backbone flexibility in designed proteins with different numbers of identical consensus tetratricopeptide repeats. Biochem 45:12175–12183
32. Qbadou S, Becker T, Mirus O et al (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. EMBO J 25:1836–1847
33. Jackson-Constan D, Keegstra K (2001) Arabidopsis genes encoding components of the chloroplastic protein import apparatus. Plant Physiol 125:1667–1676
34. Oreb M, Reger K, Schleiff E (2006) Chloroplast protein import: reverse genetic approaches. Curr Genom 7:235–244
35. Pollmann S, Neu D, Weiler EW (2003) Molecular cloning and characterization of an amidase from *Arabidopsis thaliana* capable of converting indole-3-acetamide into the plant growth hormone, indole-3-acetic acid. Phytochemistry 62:293–300
36. Pollmann S, Neu D, Lehmann T et al (2006) Subcellular localization and tissue specific expression of amidase 1 from *Arabidopsis thaliana*. Planta 224:1241–1253
37. Chan NC, Likic VA, Waller RF et al (2006) The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. J Mol Biol 358:1010–1022
38. Lister R, Carrie C, Duncan O et al (2007) Functional definition of outer membrane proteins involved in preprotein import into mitochondria. Plant Cell 19:3739–3759
39. Rosenbaum Hofmann N, Theg SM (2005) Toc64 is not required for import of proteins into chloroplasts in the moss *Physcomitrella patens*. Plant J 43:675–687
40. Aronsson H, Boij P, Patel R et al (2007) Toc64/OEP64 is not essential for the efficient import of proteins into chloroplasts in *Arabidopsis thaliana*. Plant J 52:53–68
41. Rost B (1999) Twilight zone of protein sequence alignments. Protein Eng 12:85–94
42. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815
43. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8:52–56
44. Hess B (2002) Convergence of sampling in protein simulations. Phys Rev E 65:031910
45. Krieger E, Darden T, Nabuurs S et al (2004) Making optimal use of empirical energy functions: force field parameterization in crystal space. Proteins 57:678–683
46. Otto S (2006) Reinforced molecular recognition as an alternative to rigid receptors. Dalton Trans 23:2861–2864
47. Main ER, Xiong Y, Cocco MJ, D'Andrea L, Regan L (2003) Design of stable alpha-helical arrays from an idealized TPR motif. Structure 11:497–50848
48. Cliff MJ, Harris R, Barford D et al (2006) Conformational diversity in the TPR domain-mediated interaction of protein phosphatase 5 with Hsp90. Structure 14:415–426
49. Cortajarena AL, Regan L (2006) Ligand binding by TPR domains. Protein Sci 15:1193–1198
50. Cortajarena AL, Kajander T, Pan W, Cocco MJ, Regan L (2004) Protein design to understand peptide ligand recognition by tetratricopeptide repeat proteins. Protein Eng Des Sel 17: 399–409
51. Schlegel T, Mirus O, von Haeseler A et al (2007) The tetratricopeptide repeats of receptors involved in protein translocation across membranes. Mol Biol Evol 24:2763–2774
52. Fan AC, Bhangoo MK, Young JC (2006) Hsp90 functions in the targeting and outer membrane translocation steps of Tom70-mediated mitochondrial import. J Biol Chem 281:33313–33324
53. Yang J, Roe SM, Cliff MJ et al (2005) Molecular basis for TPR domain-mediated regulation of protein phosphatase 5. EMBO J 24:1–10
54. Schleiff E, Turnbull JL (1998) Functional and structural properties of the mitochondrial outer membrane receptor Tom20. Biochemistry 37:13043–13051
55. Schleiff E, Soll J, Sveshnikova N et al (2002) Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. Biochemistry 41:1934–1946
56. Bennett-Lovsey RM, Herbert AD, Sternberg MJE et al (2008) Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. Proteins 70:3
57. Das AK, Cohen PW, Barford D (1998) The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. EMBO J 17:1192–1199
58. Fiser A, Do RK, Sali A (2000) Modeling of loops in protein structures. Protein Sci 9:1753–1773
59. Schwede T, Kopp J, Guex N et al (2003) SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 31:3381–3385
60. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure

validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 50:437–450
61. Bowie JU, Lüthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253:164–170
62. Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356:83–85
63. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comp Phys Commun 91:43–56
64. Lindahl E, Hess B, van der Spoel D (2001) Gromacs 3.0: a package for molecular simulation and trajectory analysis. J Mol Model 7:306–317
65. van der Spoel D, Lindahl E, Hess B et al (2005) GROMACS: fast, flexible, and free. J Comput Chem 26:1701–1718
66. van Gunsteren WF, Billeter SR, Eising AA et al (1996) Biomolecular asimulation: the GROMOS96 manual and user guide. ETH, Zürich
67. Berendsen HJC, Postma JPM, van Gunsteren WF et al (1981) Interaction models for water in relation to protein hydration. In: Pullman B (ed) Intermolecular forces. Reidel, Dordrecht, pp 331–342
68. Berendsen HJC, Postma JPM, van Gunsteren WF et al (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
69. Essmann U, Perera L, Berkowitz ML et al (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593
70. Miyamoto S, Kollman P (1992) Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comp Chem 13:952–962
71. Hess B, Bekker H, Berendsen HJC et al (1997) LINCS: a linear constraint solver for molecular simulations. J Comp Chem 18:1463–1472
72. Hess B (2000) Similarities between principal components of protein dynamics and random diffusion. Phys Rev E 62:8438–8448
73. Humphrey W, Dalke A, Schulten K (1996) VMD - Visual Molecular Dynamics. J Mol Graph 14:33–38
74. Tai K, Shen T, Börjesson U et al (2001) Analysis of a 10-ns molecular dynamics simulation of mouse acetylcholinesterase. Biophys J 81:715–724
75. Barrett CP, Hall BA, Noble ME (2004) Dynamite: a simple way to gain insight into protein motions. Acta Crystallogr D Biol Crystallogr 60:2280–2287