

Ligand Recognition by the TPR Domain of the Import Factor Toc64 from *Arabidopsis thaliana*

Rashmi Panigrahi¹, Abdussalam Adina-Zada¹, James Whelan^{2,3}, Alice Vrielink^{1*}

1 School of Chemistry and Biochemistry, University of Western Australia, Crawley, Western Australia, Australia, **2** ARC Centre of Excellence in Plant Energy Biology, University of Western Australia, Crawley, Western Australia, Australia, **3** Department of Botany, School of Life Science, La Trobe University, Bundoora, Victoria, Australia

Abstract

The specific targeting of protein to organelles is achieved by targeting signals being recognised by their cognate receptors. Cytosolic chaperones, bound to precursor proteins, are recognized by specific receptors of the import machinery enabling transport into the specific organelle. The aim of this study was to gain greater insight into the mode of recognition of the C-termini of Hsp70 and Hsp90 chaperones by the Tetratricopeptide Repeat (TPR) domain of the chloroplast import receptor Toc64 from *Arabidopsis thaliana* (*At*). The monomeric TPR domain binds with 1:1 stoichiometry in similar micromolar affinity to both Hsp70 and Hsp90 as determined by isothermal titration calorimetry (ITC). Mutations of the terminal EEVD motif caused a profound decrease in affinity. Additionally, this study considered the contributions of residues upstream as alanine scanning experiments of these residues showed reduced binding affinity. Molecular dynamics simulations of the TPR domain helices upon peptide binding predicted that two helices within the TPR domain move backwards, exposing the cradle surface for interaction with the peptide. Our findings from ITC and molecular dynamics studies suggest that *At*Toc64_TPR does not discriminate between C-termini peptides of Hsp70 and Hsp90.

Citation: Panigrahi R, Adina-Zada A, Whelan J, Vrielink A (2013) Ligand Recognition by the TPR Domain of the Import Factor Toc64 from *Arabidopsis thaliana*. PLoS ONE 8(12): e83461. doi:10.1371/journal.pone.0083461

Editor: Eugene A. Permyakov, Russian Academy of Sciences, Institute for Biological Instrumentation, Russian Federation

Received: August 1, 2013; **Accepted:** November 4, 2013; **Published:** December 31, 2013

Copyright: © 2013 Panigrahi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: alice.vrielink@uwa.edu.au

Introduction

Non-globular proteins, which contain repeated structural motifs arranged in tandem, are ubiquitous in nature. Many of these proteins have extended structures with exposed interaction surfaces. The common scaffold is known to interact with a variety of ligands [1] and the binding of the ligand does not induce any structural rearrangement [1,2]. Armadillo repeats (ARM), Ankyrins (Ank), Leucine-rich repeats (LRR), HEAT repeats and Tetratricopeptide repeats (TPR) are some examples of repeat protein [1,3,4]. Structural and binding studies of the Hsp70 and Hsp90 organizing protein (Hop) [5,6] have provided some knowledge of the interactions between the TPR domain and peptides, however, the role of conformational changes in these interactions has not been well characterized.

In this study, we focus on Tetratricopeptide Repeat (TPR) repeats and their versatility with respect to ligand recognition. These structural domains act as interaction scaffolds and mediators of multi-protein complexes and are found in all kingdoms of life [7]. TPR repeats form the 20 most common folds in the Pfam database [8]. They consist of multiple repeats of degenerate 34 amino acids, forming the canonical helix-turn-helix fold. Typically proteins with this motif contain 3–16 sequential TPR motifs, arranged in a tandem array. The anti-parallel packing of the helices forms a grooved surface with concave and convex faces, with ligand binding usually occurring within the concave surface [9]. TPR interaction with their ligand is usually specific. This specificity is provided by the unique geometry of the binding pocket and ionic, hydrogen bond and hydrophobic

interactions between amino acids residues of the TPR and the ligand. TPR domains span a range of oligomeric states from monomers [6,10] to higher oligomers [11–13]. Nuclear protein ssn6 [14], chromatin associated protein CDC23 [15] and mitotic chromosome disjunction protein nuc2 + [16] were among the first to be identified as TPR containing proteins. Furthermore, proteins with TPR domains play an essential role in the import of proteins into mitochondria [17], chloroplast [18] and peroxisomes [19].

Proteins destined to the chloroplast are nuclear encoded, synthesized in the cytosol, and transported as precursor proteins (or preproteins), usually with NH₂-terminal targeting sequence called transit peptides [20]. As both the mitochondria and chloroplast co-exist in plant cells, the sorting of protein between these organelles is a unique aspect of plant cell biology, and thus place a higher degree of stringency of protein targeting and sorting compared to non-plant cells. In chloroplasts, the process of recognition and translocation is initiated by subunits of the multimeric protein import complex called translocon at the outer envelope of chloroplast (Toc). This Toc core is comprised of the channel type, Toc75 [21], membrane anchored GTPases Toc159 [22], Toc34 [23] and an integral membrane protein Toc64 [24]. Topology analyses of Toc64 suggest that it consists of an amidase domain and a C-terminal TPR domain. The protein spans the membrane three times and positions the TPR domain facing the cytosol [20,25]. The TPR domain of Toc64 is composed of three TPR repeats followed by a solvation or capping helix (Figure S1A) [26]. Association of Toc64 with preproteins was found to be chaperone mediated [25,27]. Toc64 from *Pisum sativum* was found to interact with the C-terminal peptides of both Hsp70 and Hsp90

from wheat germ lysate, however, a preference was observed for interaction with the C-terminus of Hsp90 from human [25]. In the above interaction studies, Toc64, Toc64_TPR, and the C-terminal peptide of Hsp90 were matrix immobilized. It was also observed from *in vivo* experiments that the TPR domain of Toc64 exhibited a stronger interaction with Hsp90 whereas the transmembrane region acts as substrate for Hsp70 [25]. Thus, the TPR domain of Toc64 acts as a docking site, preferentially for Hsp90 bound preproteins, whereas Hsp70 was thought to be non-essential for interaction with preproteins [25,27]. The EEVD motif in the C-terminus of Hsp70 and Hsp90 families is highly conserved in all eukaryotes. Although this motif anchors to the TPR domain by a dicarboxylate clamp mechanism, residues N-terminal to this sequence have also been reported to contribute towards the specificity of interaction through hydrophobic and van der Waals contacts as in case of Hop [6]. The interaction analysis of the EEVD peptide with TPR domain showed a drastic reduction in binding affinity in Hop on removal of the above N terminal residues [5].

In the present study, we characterized the interaction of the TPR domain from *AtToc64* with C-terminal peptides from Hsp70 and Hsp90 by using biophysical approaches as a continuation of studies carried out by Qbadou S, *et al.* [28]. Using isothermal titration calorimetry studies and molecular dynamics simulations, we have investigated the mode of interaction of *AtToc64* with C-terminal octapeptides of Hsp70 and Hsp90. *AtToc64_TPR* interacts with the above octapeptides with similar affinity and 1:1 stoichiometry. Noticeably, point mutations of any of the last five residues of the peptides to alanine cause significant abolishment in peptide binding. Using molecular dynamics, we have delineated the contribution of residues involved in interactions and the dynamics mediating binding events. These studies showed that the Hsp70/90 octapeptide bound conformations are mediated by the outward movement of the terminal helices of the TPR, exposing the inner surface of the cradle for interaction with the peptide.

Materials and Methods

Expression clone of the TPR domain of *AtToc64*

An expression clone for the TPR domain of *AtToc64* was constructed using a cDNA encoding the full-length protein (accession number NP_188424, At3g17970). The region encompassing the TPR domain of *AtToc64* was composed of residues 474–589 (Figure 1). PCR amplification of this region used the forward primer (5'- TTACCATGGCCGAGATTGCCAAAGA-GAAGGGTAA -3', *NcoI* site underlined), and the reverse primer (5'- TTCACTCGAGCTGGAATTTTCTCAGTCTCTCTGC -3', *XhoI* site underlined). The PCR amplified product was digested with *NcoI* and *XhoI* and inserted into the pETM10 expression vector (Invitrogen), digested with the same restriction enzymes to generate the *AtToc64_TPR_pETM10* expression plasmid encoding the TPR domain containing a C-terminal hexa-histidine tag (*AtToc64_TPR-H6*).

Expression in *E. coli*

Overexpression of the TPR sequence was performed in *E. coli* BL21(DE3) (Novagen). 1 µg of plasmid DNA was used to transform 50 µL of competent cells. Single colonies were grown in Luria broth medium at 37°C supplemented with 40 µg/µL of kanamycin (Amresco). Protein expression was induced by the addition of 1 mM isopropyl- β-D-thiogalactopyranoside (IPTG). After 12 hours of induction at 20°C, cells were harvested by

centrifugation, and the pellet obtained was used for protein purification.

Protein purification

Recombinant expressed TPR domain was purified using nickel affinity chromatography (Ni-NTA). Briefly, 12 g of pellet was resuspended with 400 mL of lysis buffer (50 mM Hepes, 500 mM KCl, 10% glycerol, pH 7.3). The resuspended pellet was lysed using an Emulsiflex C5 high-pressure homogenizer (Avestin) at 13000 p.s.i. The lysate was clarified by centrifuging at 22,700 × g for 45 min followed by filtering the supernatant through a 0.22 µm membrane. The supernatant was loaded onto a 5 mL HisTrap FFcrude column using an ÄKTA purifier FPLC system (GE Healthcare). The Ni-NTA bound TPR was eluted using a concentration gradient of 40 mM to 1 M of imidazole (pH 7.3) and the concentration of the protein was quantified by measuring the absorbance at 280 nm using a calculated extinction coefficient of 1.09 M⁻¹ cm⁻¹. The protein purity was evaluated using standard SDS-PAGE analysis. The purified protein was concentrated using an Amicon centrifugal filter unit (30 kDa cutoff).

Relative molecular mass estimation by size exclusion chromatography

A superdex 75 prep grade size exclusion chromatography column (GE Healthcare) was equilibrated with superdex buffer (25 mM Hepes, 100 mM KCl, pH 7.3). The column was calibrated using gel filtration standards (Alcohol dehydrogenase, *M_r* 150,000; void volume; Bovine serum albumin, *M_r* 66,000 (1); Ovalbumin *M_r* 44,000 (2), Carbonic anhydrase *M_r* 29,000 (3); Lysozyme *M_r* 14,000 (4). 2 mL of concentrated protein at 8 mg/mL was applied to the equilibrated column and eluted using the above superdex buffer. The relative molecular masses of the peaks obtained were calculated using a logarithmic interpolation. The elution peak corresponding to non-aggregated *AtToc64_TPR-H6* was pooled and concentrated in the buffer systems required for various biophysical experiments.

Peptide synthesis

Synthetic C-terminal octapeptides to Hsp70 (GPTIEEVD) and Hsp90 (TSRMEEVD) were designed. In addition, shortened versions of these peptides were designed containing only the C-terminal pentapeptide sequences: Hsp70_C5 (IEEVD), Hsp90_C5 (MEEVD). Additionally, alanine mutations of each of the amino acids in the octapeptides were designed representing a library of 16 Ala-substituted peptides. Two random sequences of five residues (ASDTM) and eight residues (DMTSRGTQ) were also designed for negative control experiments. All the above peptides were synthesized and provided at >85% purity by Biomatik (Canada).

Circular Dichroism studies (CD)

Far-UV CD spectra corresponding to peptide bond absorption were recorded from 185 to 260 nm at 10°C using a Jasco-810 spectropolarimeter. Spectra were collected for 10 µM of protein in 10 mM of sodium phosphate buffer pH 7.3 in a Quartz SUPRASIL cuvette (Hellma) with a path length of 1 mm. Measurements were made with an increment step of 0.5 nm, an integration time of 4 sec per step and a bandwidth of 2 nm. The signal due to buffer alone was subtracted from that of the protein.

The proportions of secondary structures of the protein was estimated from the [θ] values between 190 and 240 nm using the DichroWeb server [29] (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>) and the CDSSTR algorithm [30–32]

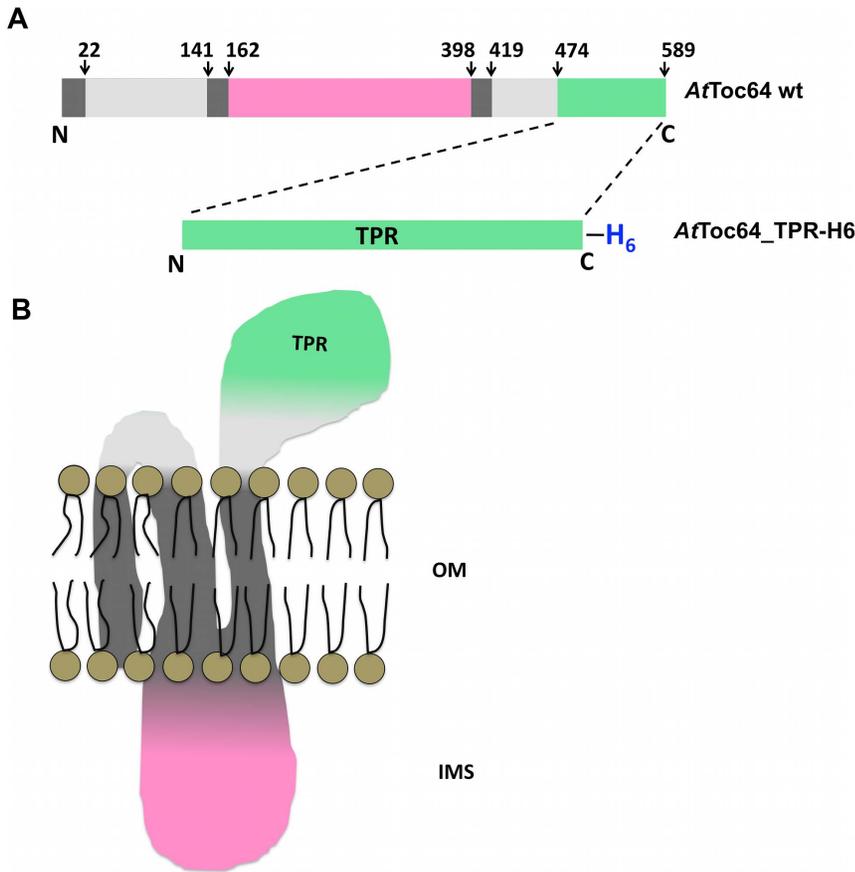


Figure 1. Topological model of Toc64 from *Arabidopsis thaliana*. **A.** Domain organization in AtToc64 from the uniprot database (http://www.uniprot.org/uniprot/Q9LVH5#section_name). Different regions shown are helical transmembrane regions (dark grey), cytoplasmic exposed regions shown (light grey), regions exposed to the intermembrane space (pink) and TPR domain (green). The TPR domain of the protein expressed for this study is shown in green with a hexahistidine tag at the C-terminus. **B.** Topological model of the domain organization. OM: outer membrane, IMS: Intermembrane space.

doi:10.1371/journal.pone.0083461.g001

The helical content of AtToc64_TPR-H6 was calculated from the molar ellipticity at 208 and 222 nm [33] using the following equation:

$$\% \alpha\text{-helix} = ([\Theta]_{222} + 3000) / (36000 + 3000) \times 100$$

The change in molar ellipticity was monitored at 222 nm while varying the temperature from 10–80°C at a rate of 1°C/min. The thermally denatured sample was cooled to 10°C at the same rate to observe the effect of temperature on the folding of recombinant expressed AtToc64_TPR-H6.

Dynamic Light Scattering studies (DLS)

DLS experiments were carried out with a Zetasizer Nano ZS (Malvern). 35 μM of protein in the superdex buffer was used for this study. Measurements were performed in triplicate consisting of 10 acquisitions per run.

Analytical Ultra-centrifugation (AUC)

These experiments were conducted on a Beckman XL-A analytical centrifuge, which was equipped with An60-Ti rotor and absorbance optics [34]. The experiments were carried out using 70 μM of protein samples in superdex buffer (390 μL) with superdex buffer (400 μL) as reference separately loaded into a

double sector centrepiece and built up in the rotor. Protein samples were monitored continuously by UV absorbance at 280 nm. Prior to the start of centrifugation, the rotor was equilibrated to 20°C and the vacuum was brought below 10 micron. The rotor was then set to spin at 60,000 rpm for 16 hrs. A total of 300 scans were collected. The acquired data were analysed using the SEDFIT program [35]. The c(s) method implemented in the program was used for the data analysis, where c(s) is the sedimentation coefficient distribution function of the macromolecule. The physical parameters of the sample solution used for the data analysis were partial specific volume (0.74 ml/g), buffer density (1.005 g/ml) and viscosity (0.01002 P). This provides excellent resolution and sensitivity for characterizing sample homogeneity. A confidence level of $p=0.95$ was used while solving the size distributions.

Isothermal Titration Calorimetry (ITC)

Peptides were dissolved in the superdex buffer at 6–14 mM concentration and 302 μL were titrated against 1.4 ml of protein at 0.1–0.3 mM in a VP-ITC Microcal Instrument (GE Healthcare). Titrations were carried out at 20°C using 30 injections of 10 μL each injected at interval of 200 seconds. Injections were continued beyond saturation levels to allow determination of the heats of ligand dilution. The non-linear least square curve-fitting algorithm (Microcal Origin) was used for data fitting. After

subtraction of the heat of dilution, three floating variables: stoichiometry (N), binding constant (K_d) and change in enthalpy of interaction (ΔH) were obtained. For subsequent alanine scan experiments the stoichiometry was fixed at $N = 1$.

Molecular Dynamics Simulation Studies

Initial Geometry. A reliable model of the TPR domain for *in silico* studies was obtained by submitting the *AtToc64_TPR* protein sequence for automated protein structure modeling using the I-TASSER pipeline [36,37]. A model was constructed using multiple threading alignments, which avoided bias towards a particular structural model as in homology modeling [38]. The molecular systems for protein-ligand complexes were built using the high-resolution crystal structure of 1ELW, the TPR domain of Hop in complex with the C-terminal octapeptide of Hsp70 [6] (obtained from the Protein Data Bank) as templates and a model obtained from the I-TASSER server. The TPR domains from Hop and *AtToc64* exhibited 50% sequence identity. The starting structures were refined in 200 independent FlexPepDock simulations, which consisted of a low-resolution pre-optimization step followed by a high-resolution refinement and high-resolution mode simulations using Rosetta FlexPepDock [39,40]. Thus two near native models of protein-peptide complexes were constructed and used for molecular dynamics simulations. In total, three molecular systems were prepared: *AtToc64_TPR* receptor (Apo), *AtToc64_TPR* receptor with C-term Hsp70 (octapeptide) bound (T_C70) and *AtToc64_TPR* receptor with C-term Hsp90 (octapeptide) bound (T_C90). Ramachandran plots (RAMPAGE, <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) [41], of the modeled structures are found in the supplementary materials.

Preparation for simulation. Graphics Processing Unit (GPU) accelerated Assisted Model Building with Energy Refinement (AMBER) suite version 12 [42] associated with the latest all-atom ff12SB force field was used for simulation studies [43]. The starting structures were neutralized using Na^+ and Cl^- ions and hydrogen atoms positioned using the *tleap* module from AMBER-TOOLS12. The protein was centered in a solvent truncated octahedron box, which was made of TIP3P (3-point charged) triangulated water molecules with a 12 Å cut off in all directions [44]. The total number of atoms including water molecules was approximately 20,000 across various systems. The systems were minimized using a two-phase energy minimization approach which included 2500 cycles of steepest descent and 2500 cycles of conjugate gradient with solute atoms restrained by harmonic potentials with force constants of $50 \text{ kcal mol}^{-1} \text{ \AA}^2$. This was followed by 5000 steps of unrestrained whole system minimization. 50 ps of density equilibration with weak harmonic restraints of $2 \text{ kcal mol}^{-1} \text{ \AA}^2$ on the solute molecule was performed followed by unrestrained equilibration for 500 ps under constant pressure and temperature conditions. All simulations were performed using the SHAKE algorithm [45] with constraints on hydrogen-linked bonds (allowing a tolerance of 0.0001). To evaluate long-range electrostatic interactions, the Particle Mesh Ewald (PME) method [46] was used with a cut-off of 9 Å. An integration time step of 2 fs was used to numerically solve Newton's equations of motion. Langevin dynamics was used to maintain a constant temperature of 300 K throughout the simulations. All the simulations were performed using the PMEMD module in AMBER12. Production runs were carried out for 50 ns in an explicit solvent environment and with an isothermal-isobaric (NPT) ensemble.

Analysis of trajectory. After the production runs were completed, each of the trajectories were analyzed based on variation in kinetic and potential energies using the *ptraj* program. Root-mean-square deviation (rmsd) and atomic positional fluctuation per

residue (rmsf) were analyzed to understand the overall conformational changes throughout the trajectory. The number of hydrogen and other non-bonded bond interactions were calculated using HBPLUS [47]. Interactions of arginine or lysine side chains with oxygen atoms of aspartate or glutamate residues which were from 3–4 Å in distance and exhibited angles (donor – H – acceptor) less than 90° were not classified as hydrogen bonds, but considered as electrostatic interactions. In addition, the solvent accessible surface area and binding energies for the protein-peptide interactions were calculated using NACCESS [48] and MM/PB(GB)SA tools [49,50] respectively. Computational Alanine scanning was performed with the peptide, in order to understand the contribution of individual amino acid residues towards binding. When calculating the difference in free energies ($\Delta\Delta G$) between the wild type and the mutants, the results of using Generalized Born calculations (GB) were taken into consideration, as they are well suited for protein-ligand and protein-protein interaction calculations.

These values were calculated according to the following equation:

$$\Delta\Delta G = \Delta G^{wild} - \Delta G^{mutant}$$

For all of the above analyses, the last 5 ns of the trajectory was analyzed. The APBS software was used to compute the electrostatic potential surface [51]. Ligplot was used to map the hydrogen and hydrophobic bonding patterns between the peptide residues and the residues of TPR involved in interaction [52]. Visual Molecular Dynamics [53] was used for visualizing the trajectories of the simulations of the three systems. The dictionary of protein secondary structure (DSSP) [54] was used to assign the secondary structure per amino acid along different trajectories.

Principal Component Analysis (PCA) was also performed in order to substantiate the above results [55]. GROMOS 4.5 [56] was used on the protein backbone atoms ($\text{C}\alpha$) to obtain the collective coordinates for protein motions from covariance matrices in the form of sets of eigen values (amount of motion) and eigen vectors (direction of motion). In this way all the linearly correlated motions were analyzed.

Correlated atomic motion in the apo and ligand bound forms were obtained by analyzing the dynamical cross correlation map (DCCM) of $\text{C}\alpha$ atoms using Bio3D [57,58]. This provided a means to understand the correlation of motions of neighbouring and/or distant residues [59]

All graphical representations were generated using PyMOL [60].

Results

Purification and characterization of TPR domain of *AtToc64*

Recombinant *AtToc64_TPR*-H6 (Figure 1) was cloned, over-expressed and purified using Ni-NTA affinity chromatography. The purity was accessed by SDS-PAGE where a single band, corresponding to 14.3 kDa, was observed (Figure S1B). The observed molecular mass agreed with that predicted from the sequence of his-tagged *AtToc64_TPR*. Further size exclusion chromatography was performed which showed that the protein exists predominantly as a dimer calculated from gel filtration protein standards (Figure 2A). The dimer fractions were pooled, concentrated and used for further studies. Based on these results, we hypothesize that the TPR domain is either globular and dimeric or elongated and monomeric.

The TPR repeats are known to adopt an alpha-helical solenoid structure; this is supported by analysis of the protein sequence by

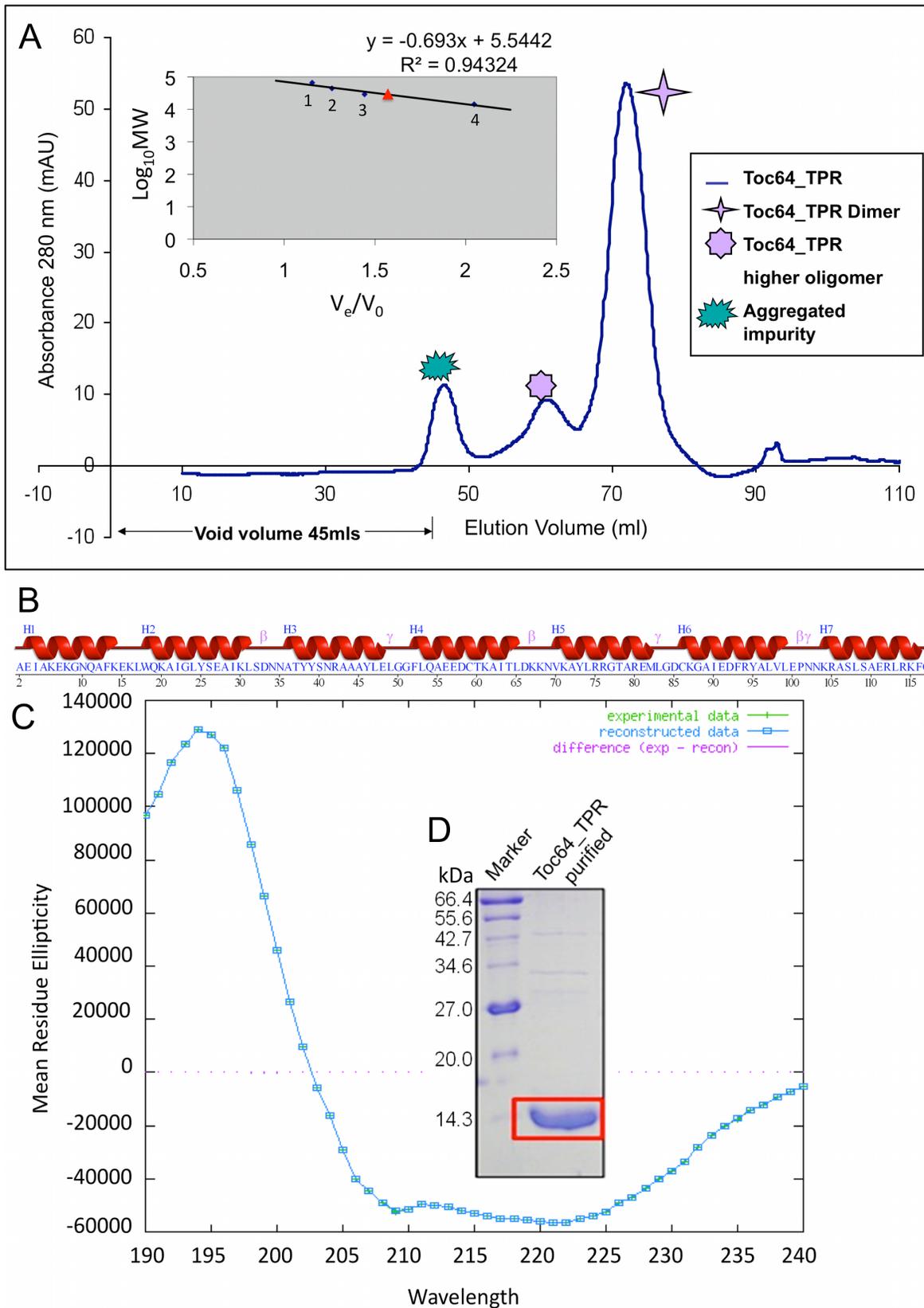


Figure 2. Biophysical characterization of *AtToc64_TPR-H6* using size exclusion chromatography and circular dichroism. **A.** Elution profile of *AtToc64_TPR-H6* from size exclusion chromatography using a Superdex 75 prep grade column. The different shaped stars represent different forms of the protein as described in the figure. The V_e/V_0 versus Log_{10}MW plot for superdex standards is shown as an inset. Numbers 1–4 represent different standard proteins used as described in the methods. Red triangle represents *AtToc64_TPR-H6*. **B.** Schematic representation of secondary structure of *AtToc64_TPR*. **C.** CD spectrum profile of the protein obtained after analysis with DICHROWEB. doi:10.1371/journal.pone.0083461.g002

PSIPRED [61,62] (Figure 2B). Circular dichroism spectra can be used to estimate the regular secondary structural elements in a protein. The spectral signature revealed that protein adopted a predominantly (80%) helical structure (Figure 2C), in agreement with that observed in the PSIPRED prediction. Thermal denaturation experiments showed that the protein had a T_m of $\sim 35^\circ\text{C}$ (Figure S1C).

Estimation of polydispersity by DLS and AUC

In order to understand the oligomeric state of the protein, DLS studies were performed. This analysis on a $35\ \mu\text{M}$ solution of freshly purified recombinant *AtToc64_TPR-H6* showed that the protein was predominantly monodispersed with a polydispersity value of 13%. Storage of the protein at 4°C for 24 hours caused an increase in the extent of polydispersity (52%) as compared to the freshly purified sample. AUC experiments were carried out to independently establish the homogeneity of protein molecules in solution and determine their molecular mass [63]. The protein existed as a mixture of monomeric and dimeric forms, where the former was predominantly present (Table 1). These results also explained the observed sample polydispersity by DLS; this polydispersity was irreversible. Additionally, AUC experiments were performed in combination with the two peptides (octapeptides). Interestingly, it was observed that in all cases the monomeric form was predominant (Table 1). The result of the AUC studies suggests that the behaviour of *AtToc64_TPR-H6* by size exclusion chromatography is likely due to the non-globular, ellipsoid shape and not due to the presence of dimers in solution. The higher oligomers observed in the AUC is likely due to nonspecific self-association of protein molecules with time.

ITC characterization of peptide interaction with *AtToc64_TPR-H6*

In order to understand the energetics of binding between the C-termini of Hsp70/Hsp90 and the TPR domain of *AtToc64*, binding studies using ITC were performed [64,65]. The curves obtained were fit using a 1:1 binding model. The binding stoichiometries (n) were found to be 1.2 and 0.94, when protein at $300\ \mu\text{M}$ in the superdex buffer was titrated with $10\ \text{mM}$ Hsp70 octapeptide or $14\ \text{mM}$ Hsp90 octapeptide solutions respectively at 20°C . Similar experiments were carried with pentapeptide versions of Hsp70 and Hsp90, wherein n values were 1.2 and 1.04 respectively. Hence in all of the above cases, the thermodynamics of binding events exhibited exothermic behaviour. The heat of reaction per injection was calculated from the area under the peak, which gradually decreased with complex formation and reached the heat of dilution of the respective peptide when the protein was saturated. After ~ 20 and ~ 11 injections in case of Hsp70 and Hsp90 respectively, no unbound TPR was present. The thermodynamic parameters and the final titration curve

(Figure 3A and 3B) were computed, resulting in similar binding affinities for the two octapeptides to TPR. Reducing the length of the peptide to the C-terminal five residues containing the EEVD motif decreased the binding affinity two and half fold in the case of Hsp70 and four fold in the case of Hsp90 (Table 2). The Gibb's free energy (ΔG) of the above binding events were similar in magnitude to each other, ranging from -4.90 to $-4.94\ \text{kcal/mol}$. In the case of the octapeptide of Hsp90, a favourable enthalpic contribution to binding was observed ($\Delta H = -6.18\ \text{kcal/mol}$) with a decrease in entropy ($\Delta S = -4.33\ \text{cal/mol}\cdot\text{K}$). This is characteristic of an enthalpy driven process for binding of the peptide to the protein, where possible changes in conformation of one or both components in the binding event is expected. In the case of the octapeptide of Hsp70, the enthalpic contribution to the binding event was quite low ($\Delta H = -1.0\ \text{kcal/mol}$) accompanied by an increase in entropy ($\Delta S = 13.2\ \text{cal/mol}\cdot\text{K}$). An increase in entropy was also seen in the binding of the pentapeptide of both Hsp70 ($\Delta S = 9.29\ \text{cal/mol}\cdot\text{K}$) and Hsp90 ($\Delta S = 5.03\ \text{cal/mol}\cdot\text{K}$), suggesting an entropy driven event which may be due to a combination of the release of water molecules from the binding site upon ligand binding and a classical hydrophobic effect. In summary, a favourable enthalpic binding process, in spite of a decrease in entropy, have aided the binding of the protein with the Hsp90 octapeptide; whereas a favourable entropic process, despite a decrease in enthalpy have aided the binding of protein with Hsp70 octapeptide.

Experimental alanine scanning

The EEVD motifs in the peptides have been shown previously to act as an anchor to the protein [5]. As both the octapeptides bound to the protein with similar affinity, it was of interest to investigate the contribution of each residue towards the interaction. For this work alanine scanning mutagenesis was carried out on each of the octapeptides used for the study and the binding was characterized by ITC studies (Figure 4). For this description, the N-terminal Gly residue is termed Gly¹¹⁷, and the following residues are arranged in ascending order, e.g. Pro¹¹⁸, Thr¹¹⁹, Ile¹²⁰, Glu¹²¹, Glu¹²², Val¹²³, Asp¹²⁴, in case of Hsp70 and a similar nomenclature for Hsp90 is used. Alanine substitution of Asp¹²⁴ abolished binding in Hsp70 and caused a significant decrease in case of Hsp90, suggesting the occurrence of electrostatic interactions. A strong contribution of hydrophobic interactions to the binding was hypothesized because of a significant increase in K_d or the complete loss of binding in the case of Ile¹²⁰/Met¹²⁰ and Val¹²³ respectively. As expected, mutation of Glu¹²¹ and Glu¹²² to Alanine abolished or significantly reduced the binding affinity for Hsp70 and Hsp90 respectively. Mutation of the first four N-terminal residues in both the peptides to Alanine caused a decrease in binding affinity with the exception of the Arg¹¹⁹ in case of Hsp90 (Figure 4, Table S1).

Table 1. Analytical ultracentrifugation.

Components	Sedimentation Coefficient	Molecular mass (kDa)	Monomer %
<i>AtToc64_TPR</i>	1.60 ± 0.17	15 ± 2	67
<i>AtToc64_TPR</i> with C-terminal Hsp70 octapeptide	1.9 ± 0.3	**	83
<i>AtToc64_TPR</i> with C-terminal Hsp90 octapeptide	1.9 ± 0.2	14.9 ± 2	71

Analytical ultracentrifugation experiments were performed at 20°C in superdex buffer. As the frictional coefficient was high ($f/f_0 = 1.54$) in the Hsp70 bound form hence adequate determination of the molecular mass was not possible. However the sedimentation coefficient obtained suggests that it is a monomer.

doi:10.1371/journal.pone.0083461.t001

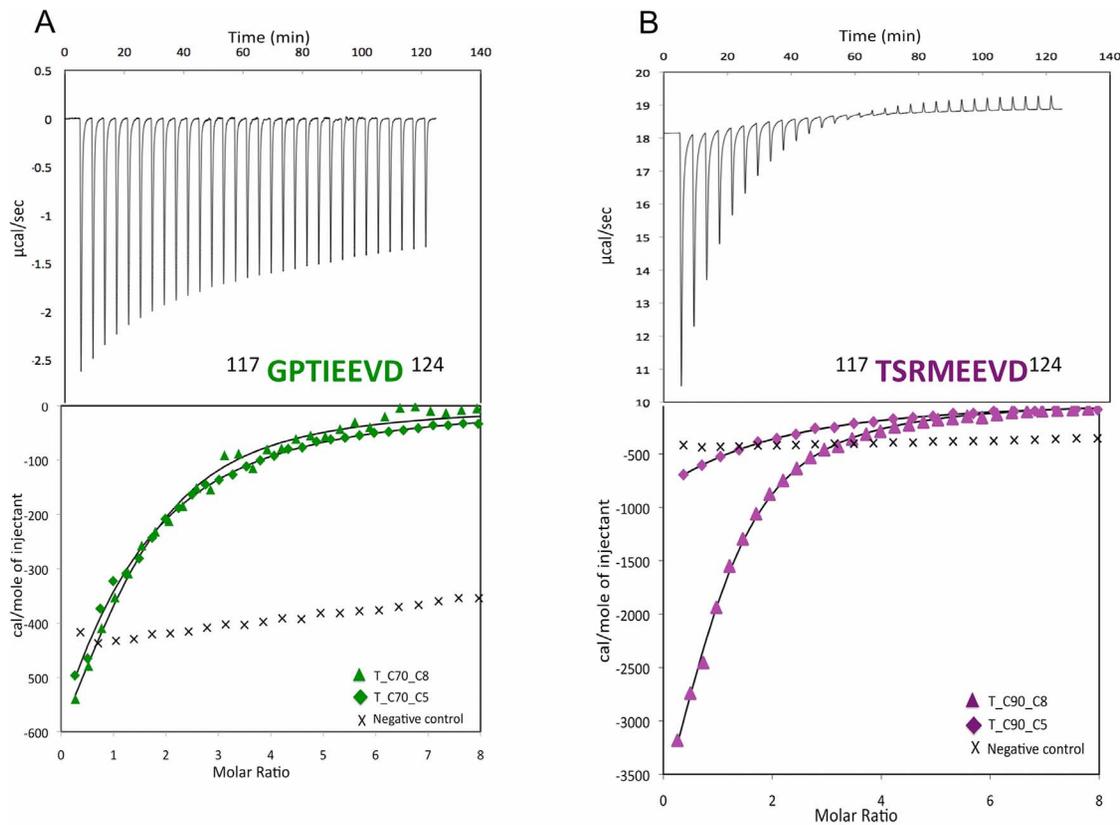


Figure 3. Binding isotherms for the interaction of *AtToc64_TPR-H6* with Hsp70 and Hsp90. The ITC isotherms obtained for the Hsp70 octapeptide binding to *AtToc64_TPR-H6* (in green) and the Hsp90 octapeptide binding to *AtToc64_TPR-H6* (in magenta). The bottom panels show the curves obtained for titration of the octapeptides of Hsp70/90 into TPR (triangles); the C-terminal pentapeptides of Hsp70/90 into TPR (diamonds) and random peptides into TPR (crosses).
doi:10.1371/journal.pone.0083461.g003

Simulation studies

The quality of the models generated were checked using Ramachandran plot where $\sim 94\%$ of the residues were present in the favoured region (Figure S3, S4 and S5), Further the RMSD between the *AtToc64_TPR* modelled using ITASSER and the TPR domain of Hop (1ELW) was found to be 1.4 \AA (Figure S2). Three systems were simulated based on the isothermal and isobaric (NPT) ensemble for 50 ns each: apo *AtToc64_TPR* (apo), *AtToc64_TPR* with the octapeptide of Hsp70 bound (T_C70) and *AtToc64_TPR* with the octapeptide of Hsp90 bound (T_C90). The RMS fluctuation for each system at the end of the

equilibration was minimal. The overall structural stability of the systems throughout the simulation was assessed by calculating the RMSD of the $C\alpha$ atoms from the appropriate starting structures for each simulation. The average $C\alpha$ -RMSD were found to be 2.15 \AA , 1.38 \AA , and 1.39 \AA for the apo, T_C70 and T_C90 systems respectively (Figure 5A). This suggested that the complexed structures were stable and retained their overall structure during the simulation. Further analysis of the three independent simulations for each of the systems indicated that, in all cases, the cradle topology of the TPR was maintained. An analysis of $C\alpha$ atomic positional fluctuation (RMSF) for each of the systems showed that the N and C termini exhibited higher temperature factors (B factor). Additionally, higher B factors were also observed for TPR residues 29–33 in the T_C70 system, corresponding to a loop connecting the H2 and H3 helices (Figure 5B). Finally, the RMSF of the terminal residues were decreased relatively upon ligand binding.

Table 2. Thermodynamic parameters obtained for TPR-Hsp interaction using isothermal titration calorimetry.

Systems	N	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol)	K_d (μM)
T_C70_C5	0.93 \pm 0.13	-4.37	-1.63	2.72	568
T_C70_C8	1.2 \pm 0.12	-4.9	-1.0	3.86	230
T_C90_C5	1.04 \pm 0.13	-4.17	-2.68	1.47	800
T_C90_C8	0.94 \pm 0.04	-4.94	-6.18	-1.26	218

N is the stoichiometry; ΔG is the calculated change in Gibb's free energy; ΔH is the change in enthalpy; ΔS is the change in entropy and K_d is the binding affinity.

doi:10.1371/journal.pone.0083461.t002

Analysis of protein-peptide interactions

The concave surface of the TPR cradle was crucial for peptide binding as observed in the average structures (Figure 6A and 6B). This is exemplified through intermolecular hydrogen bonding and nonbonding interactions with the residues that line the inner surface of the cradle. Approximately 15 interactions of each type are found in both of the peptide bound forms (Figure 6C). Notably, two types of hydrogen bonding interactions existed on the protein-peptide interface: sequence specific interactions, which

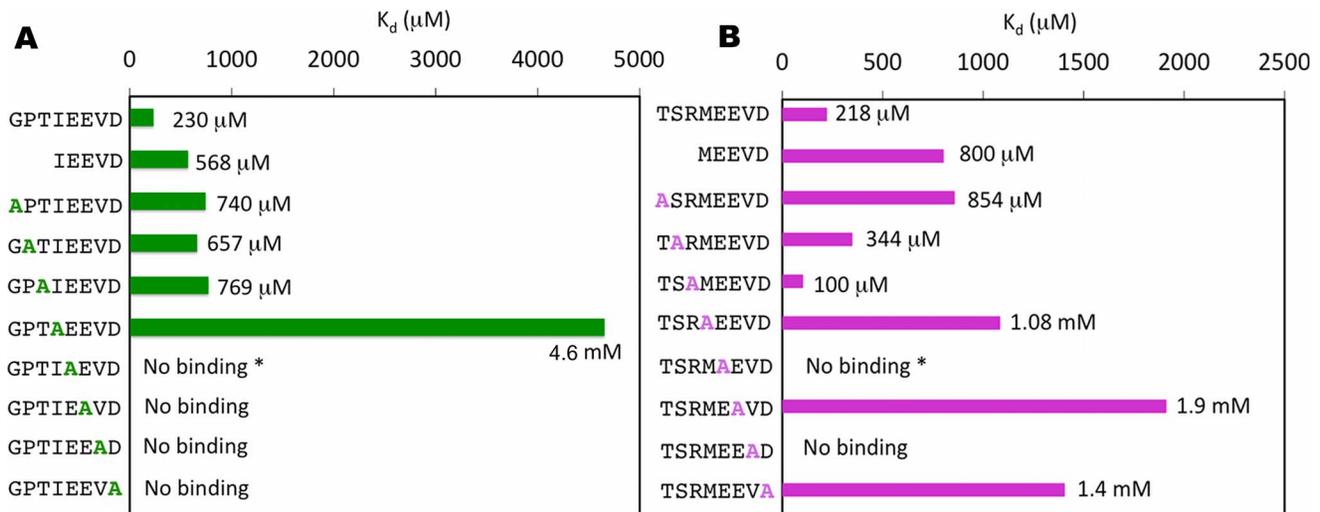


Figure 4. The ITC binding data from alanine scanning mutagenesis of the peptide interaction with *AfToc64_TPR-H6*. **A.** Data using the C-terminal octapeptide from Hsp70. **B.** Data using the C-terminal octapeptide of Hsp90. Asterisk (*) suggests that the heat change during binding event is quite low and the signal to noise ratio is high. Hence the K_d is considered as no binding. doi:10.1371/journal.pone.0083461.g004

involved the side chains of the peptide residues and sequence independent interactions, which involved the peptide main chain. Intra-residue interactions are also observed and are likely to be necessary to maintain the proper conformation of the peptides necessary for interactions with the protein.

Carboxylate clamp. Interaction analysis of the crystal structure of TPR domains from Hop liganded with the C-terminal octapeptide and pentapeptide of Hsp70 and Hsp90 respectively, have shown that the carboxylate moieties of the highly conserved terminal Asp residue of the peptides form electrostatic interactions with the protein residue [6]. This has been referred to as the two-carboxylate clamp. The conserved Asp¹²⁴ in both the complexes are clamped to the TPR by the terminal carboxylate moieties and held in place by a myriad of interactions to the protein (Figure 7A, 7B). In case of the T_C70 system, the two carboxylates are held in place by hydrogen bonding interaction with Lys⁵ and Asn⁹ of helix H1. Additionally, the main chain carboxylate is stabilized by interaction with Asn⁴⁰ of helix H3 whereas Thr³⁶ of helix H3 and Lys⁷⁰ of helix H5 help in stabilizing the side chain carboxylate (Figure 8A). In the T_C90 system, the terminal carboxylate of Asp¹²⁴ are clamped by hydrogen bonding interactions with Lys⁵, Asn⁹ of helix H1 and Asn⁴⁰ of helix H3 (Figure 8B). Thus the carboxylate clamp, which is highly conserved in both the systems, anchors the peptides to the TPR.

Other key hydrogen bonding interactions. The hydrogen bonding interactions from the TPR domain to the peptide in each complex are directed predominantly towards side chains, hence exploit sequence-specific features. Residues N terminal to Glu¹²¹ of the C-term Hsp70 peptide do not display any hydrogen bonding interaction with the TPR. In T_C90 system, Ser¹¹⁸ forms hydrogen-bonding interaction with Ser¹⁰⁸ and Arg¹¹¹ and Met¹²⁰ displays hydrogen bonding interaction with Arg¹⁰⁴. Side chain hydrogen bonding interactions of Glu¹²¹ exist with side chains of Arg⁷⁴/Arg⁷⁴, Thr⁷⁷/Asn¹⁰² and Arg¹⁰⁴/Arg¹⁰⁴ in the T_C70/T_C90 systems respectively (Figure 7A, 7B). The carboxylate oxygen atoms of Glu¹²² form hydrogen bond acceptors for Lys⁷⁰, Arg⁷⁴ and Arg¹⁰⁴ nitrogen containing groups in both the systems and additionally with Arg¹¹⁹ in the T_C90 system. Val¹²³ acts as hydrogen bond acceptor for Lys⁷⁰, exclusively in T_C90 system.

A groove for Valine. Among the charged residues of the EEVD motif in the peptides, valine is the only hydrophobic residue. The Val¹²³ is held in place by a number of van der Waals and hydrophobic interaction with different residues of the TPR domain, which form a groove. The interacting residues forming van der Waals contacts are Asn⁹ and Asn⁴⁰ in case of T_C70 system and Asn⁴⁰, Lys⁷⁰ and Arg⁷⁴ in case of T_C90 system. Similarly, Phe¹² and Ala⁴³ in case of T_C70 and Phe¹² in case of T_C90 form hydrophobic interactions with Val¹²³.

Intra-peptide interactions. Intra-peptide interactions play a key role in giving a particular conformation to the peptide. This helps in the proper presentation of the peptide residues, which aid in interaction to the protein (Figure 9A, 9B). In the T_C70 system, the key intra-peptide hydrogen bonding interactions are as below: the main chain nitrogen of Ile¹²⁰ with main chain oxygen of Pro¹¹⁸; the main chain nitrogen of Val¹²³ with main chain oxygen of Glu¹²¹ and the main chain nitrogen of Glu¹²² with side chain oxygen of Glu¹²². Similarly in the T_C90 system, the key intra-peptide hydrogen bonding interactions are: the amino group of Arg¹¹⁹ with the side chain oxygen of Glu¹²² and the main chain nitrogen of Glu¹²² with its own side chain oxygen.

Computational alanine scanning

Though absolute free energies from simulations are not achievable [66], the underlying trend caused due to mutation has been reliably captured. Computational Alanine scanning was performed on the trajectories of the two complex systems. Six of the eight residues in the each of the peptides were mutated individually to Alanine. The terminal amino acids could not be mutated by MMPB/GBSA due to limitation of the program. The calculated binding free energies for the Hsp70 and Hsp90 liganded systems were found to be -12.6 and -13.8 kcal/mol respectively. Mutations of Val¹²³, Glu¹²², Glu¹²¹, Met¹²⁰ and Ile¹²⁰ individually to Ala caused a decrease in binding, which was in accordance with our experimental ITC findings (Table 3). Alanine mutation of Arg¹¹⁹ slightly improved binding as observed in our experiments. In contrast, mutation of Pro¹¹⁸ and Ser¹¹⁸ to Alanine had only a minimal effect on binding, a result that differed from the ITC results. The slight change in conformation of the peptides caused

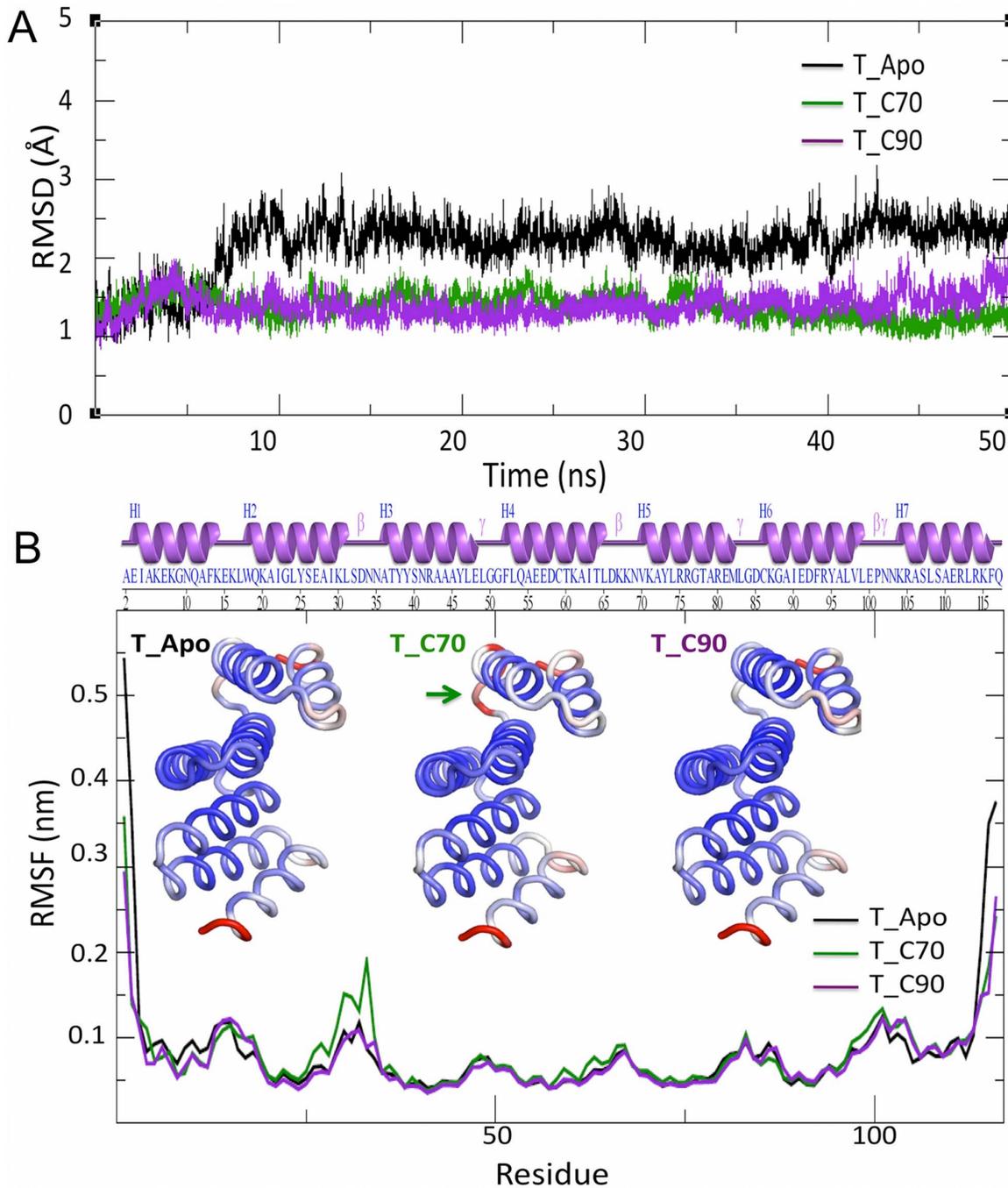


Figure 5. Analysis of the molecular dynamics trajectory obtained after 50 ns of simulation. **A.** Root mean square deviation (RMSD) plot for the apo TPR receptor (black trace), the Hsp70 C-terminal octapeptide bound form of the receptor (T_C70: green trace) and the Hsp90 C-terminal octapeptide bound form of the receptor (T_C90: magenta trace). **B.** The atomic positional fluctuation (RMSF) plot obtained for each of the above systems. The loop with high B factors in T_C70 is shown with a green arrow.
doi:10.1371/journal.pone.0083461.g005

by the mutations may have resulted in changes to the binding affinity; these have not been taken into consideration by the computational alanine scanning studies. However, the differences found in the later case was quite small, and hence the trend of change in binding free energies could be considered acceptable. In all computational studies the secondary structure of the protein is conserved all throughout the simulation and there is very little change between alpha, pi and 3_{10} helices (Figure S5).

Dynamic cross-correlation map (DCCM)

Dynamic cross-correlation maps suggest that distinct correlation exist between the helices (Figure S7). In the ligand unbound form, the majority of residues show a positive correlation between H1 and H2; H3 and H4; H4 and H5 (Figure S7A). Fewer residues show positive correlation between H5 and H6. Negative correlated motion is observed in H1 with respect to H4 and H6. Also the segment, residues between 40–60, which comprises of a part of

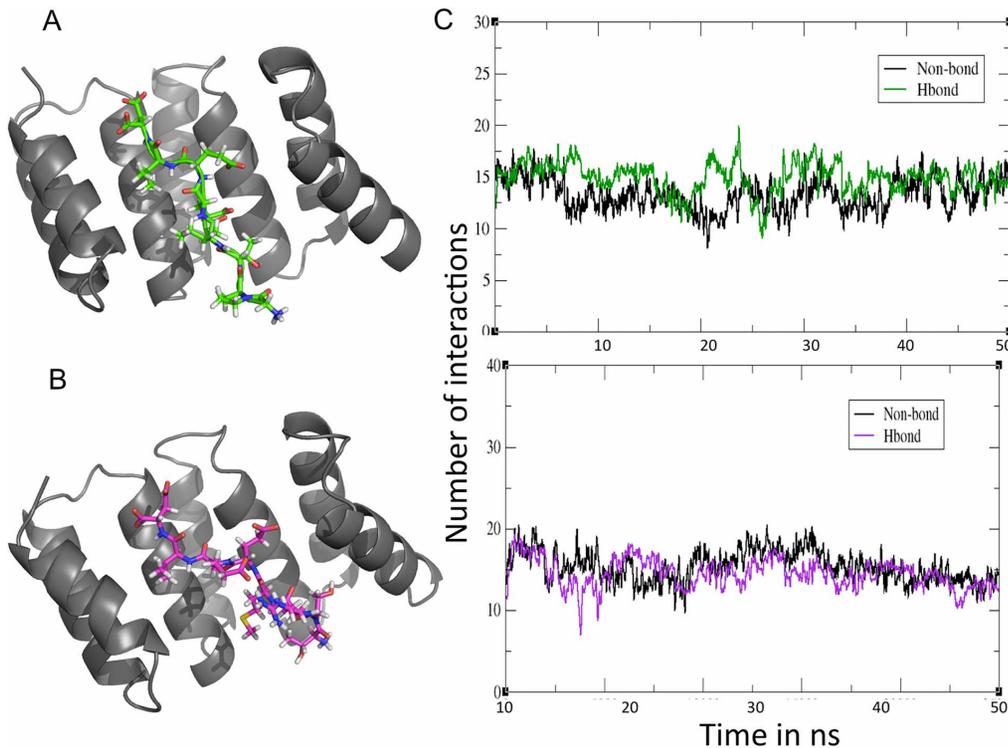


Figure 6. Characterization of the average structures obtained after simulation of the complexes. **A.** The TPR domain with bound octapeptide from Hsp70 (green) **B.** The TPR domain with bound octapeptide from Hsp90 (magenta) **C.** Number of interactions in the T_C70 system (green) and in the T_C90 system (magenta) during the simulations. doi:10.1371/journal.pone.0083461.g006

H3, a loop and a part of H4 show negative correlation with H7. Similarly, H5 shows negative correlation with H7. Ligand binding reduces all the observed inter helical positive correlations, further enhancing negative correlated motions (Figure S7B, S7C). In T_C70, additional negative correlation is seen between H6 and the region consisting of residues 60–70 (Figure S7B).

Principal component analysis (Essential dynamics)

Although DCCM, is a good way to analyze motions between the pair of atoms, the complexity of collective principal atomic motion could not be visualized. To understand this, essential dynamics was performed. The overall protein motions were decomposed into a set of eight eigenvectors. The first two eigenvectors (most significant principal components, PC1 and PC2) accounted for ~65% of the total fluctuation. The 3D plot of these two components projected, with the potential energy on the z-axis, is represented as a 2D plot (Figure 10A–C). The plots depict that there is clearly more conformational space sampled in the case of Apo form. The dominant motions in the apo form are confined to the terminal residues of the H1 and H7 helices (Figure 10D), which correlate with the rmsd and positional fluctuation plots. The distributions of both the complexes are smaller than that observed in the apo form. In the T_C90 system, the conformational spread is more restricted (Figure 10C) than observed for the T_C70 system (Figure 10B). An analysis of the individual principal components in the complexes suggests that protein in the peptide bound forms do not show loss of alpha helicity and suggests that peptide binding may act to stabilize the TPR. The most prominent motions are observed in the H1 and H7 helices with very little motion in concave inner surface of the TPR cradle (Figure 10D–F).

Discussion

In the present study, we purified recombinant *A/Toc64*_TPR-H6 to homogeneity. Analytical ultracentrifugation experiments estimate that the TPR domain exists predominantly as a monomer. Isothermal calorimetry studies indicate that the domain interacts with the C-terminal regions of Hsp70 and Hsp90 with similar micromolar affinity and 1:1 stoichiometry. Molecular dynamics studies have been performed to provide atomic level descriptions for the protein-peptide interactions. The terminal aspartate in both Hsp70 and Hsp90, is anchored to the TPR by a dicarboxylate clamp, supporting previous findings [6]. Electrostatic potential surface representations of the ligand bound form of *A/Toc*-TPR shows that the peptides are bound to a predominantly positively charged surface within the cradle of the TPR domain. In the case of Hsp70, the N-terminal residues are exposed and not interacting with the protein (Figure S8). Alanine scanning mutations of residues 120 to 124 in both peptides significantly perturb binding to the TPR domain. Intramolecular hydrogen bonding interactions between different residues of the peptides are observed, which might be necessary for providing suitable conformations to the peptides needed for binding to the receptor. Though hydrogen bonds, electrostatic and hydrophobic contacts exist, van der Waals interactions play a major role in positioning the EEVD motif. Alanine scanning mutation of either Gly¹¹⁷ or Pro¹¹⁸ using ITC reduces the binding affinity of the peptide, however the MD studies do not show that these residues interact specifically with the TPR domain. A possible explanation for this observation may be that residues contribute to the ideal conformation of the peptide necessary for interaction with the protein or they can prevent sampling of unwanted conformational

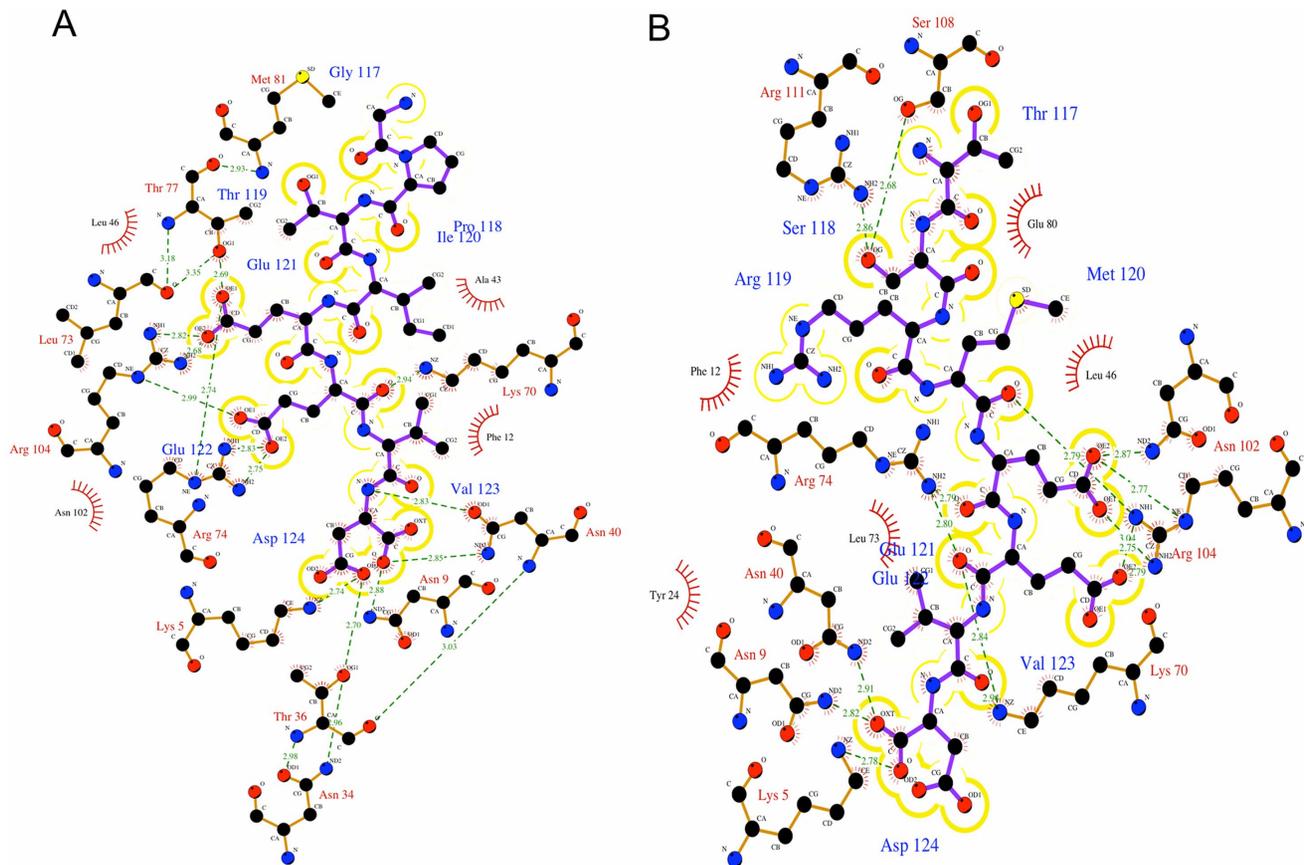


Figure 7. Interactions occurring in the protein-peptide interface generated by Ligplot. A. Interactions between the TPR domain and the C-terminal octapeptide of Hsp70; **B.** Interactions between the TPR domain and C-terminal octapeptide of Hsp90. The peptide is shown in purple bonds and the protein in brown bonds. doi:10.1371/journal.pone.0083461.g007

space by the peptide. Proline commonly adopts a cis conformation. Glycine, as a small and achiral amino acid, can occupy a larger volume of conformational space without unfavourable steric interactions with other amino acids. Because of these unique features, glycine and proline residues often occur in turns and loops [67,68].

Using MMPB/GBSA for computational alanine scanning, the change in binding energy due to mutation of Gly117 and Pro118 to alanine could not be captured. The plausible reason for the above observation is that in MD experiments, the peptide is initially docked to the protein. The binding energy change upon mutation can be reliably captured for residues of the peptide, which directly interact with the protein. Although residues such as

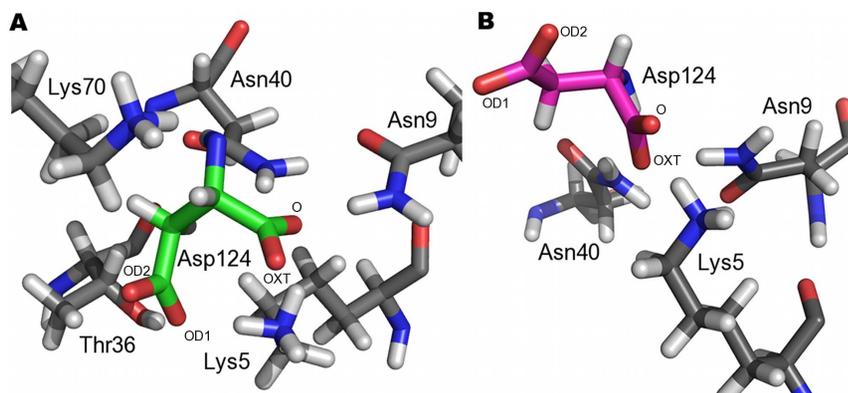


Figure 8. The carboxylate clamp. A. Key interacting residues of the TPR domain with the carboxylate of Asp¹²⁴ of Hsp70 (shown in green bonds); **B.** Key interacting residues of the TPR domain with the carboxylate of Asp¹²⁴ of Hsp90 (shown in magenta bonds). The oxygen atoms of the carboxylate moieties are labeled. The TPR domain is shown with grey bonds. doi:10.1371/journal.pone.0083461.g008

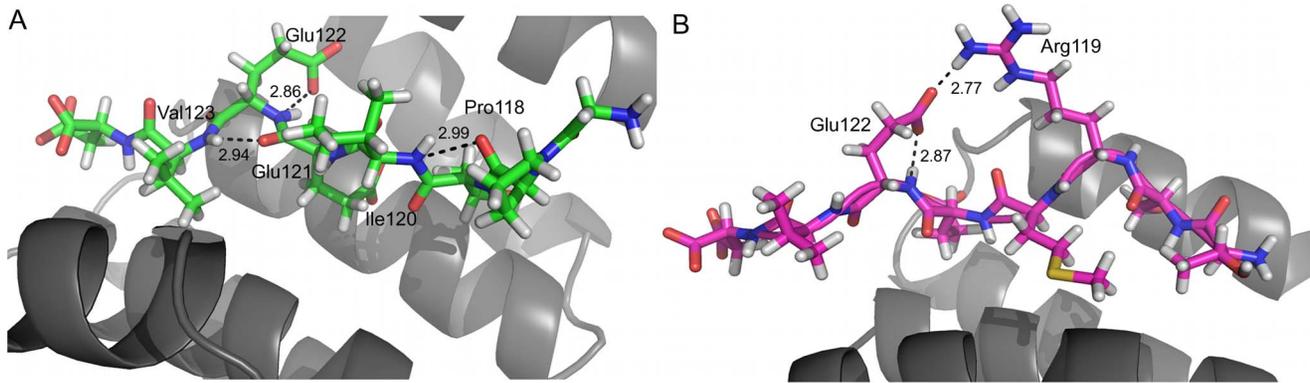


Figure 9. Key intrapeptide interactions. **A.** Intra-peptide interaction in the T_C70 system (coloured green) **B.** Intra-peptide interaction in the T_C90 system (coloured magenta). The TPR domain is shown as a ribbon diagram in grey.
doi:10.1371/journal.pone.0083461.g009

glycine and proline in the Hsp70 C-terminal peptide, do not interact directly with the TPR domain, they are likely to play an important role in providing the required conformation to the peptide for recognition by the TPR domain. Thus the binding energy changes ($\Delta\Delta G$) cannot be reliably computed for alanine mutations of residues, which are not directly interacting with the protein.

Unfavorable entropic contributions upon ligand binding are often compensated for by increased dynamic motion in distant regions of the protein. Essential dynamics shows motion in the terminal helices upon binding with either peptide. The H1 and H7 helices move backwards in order to expose the inner surface of the cradle for peptide binding. It is known that TPRs have a rigid conformation [26]. In contrast, this study provides evidence that the change in curvature by the concerted movement of secondary structural elements may be necessary for ligand binding. This observation supports a recent study on a TPR containing protein, MamA, where 3 Å radial movement by two N-terminal TPR motifs was observed in different crystal structures upon ligand

imitator binding [12]. Unlike the case of Hop, where there are two TPR domains with specificity for Hsp binding, proteins such as CHIP and CNS1 [28,69] recognize both Hsp70 and Hsp90 with single TPR domain. In CHIP, I¹²⁰ (Hsp70)/M¹²⁰ (Hsp90) are accommodated in a hydrophobic pocket, which is absent in Hop [6]. In our case, Ile¹²⁰ and Val¹²³ sandwich the Phe¹² of AtToc_TPR, thus enhancing the interaction through van der Waals contacts.

In the current work, the two peptides used for this study exhibit a conserved EEVD motif, however residues N-terminal to this conserved sequence differ between Hsp70 (sequence: G¹¹⁷P¹¹⁸T¹¹⁹I¹²⁰) and Hsp90 (sequence: T¹¹⁷S¹¹⁸R¹¹⁹M¹²⁰). Despite these differences, the two peptides bind to the AtToc64_TPR with similar affinity. The differences in the observed binding mechanism (entropy *vs* enthalpy driven reactions) is likely to be a consequence of the contributions of the different residues located at the N-terminus, where peptides with different sequences bind with similar affinity to the same protein. Interestingly, the pentapeptide versions of Hsp70 (sequence: I¹²⁰E¹²¹E¹²²V¹²³D¹²⁴) and Hsp90 (sequence: M¹²⁰E¹²¹E¹²²V¹²³D¹²⁴) demonstrate entropically driven interactions as observed by ITC studies, although with weak micromolar affinity than the octapeptides. Additionally the residues glycine and proline do not interact with the protein in the case of the TPR-Hsp70 system. It is tempting to conclude from the above observed results that the interactions involved with the N-terminal residues of the Hsp90 octapeptide, namely threonine, serine and arginine, contribute towards an enthalpically driven interaction. Additionally, one might hypothesize that the interaction of hydrophobic residues such as I¹²⁰ and M¹²⁰ might assist in the release of water molecules from the interaction interface.

Furthermore, computational studies using essential dynamics suggests that a higher conformational spread is observed in the TPR-Hsp70 system than in the TPR-Hsp90 system resulting in an entropy driven process in the former system. The observed flexibility in the terminal helices, H1 and H7, upon ligand binding may further contribute to the entropy of the system. However, in the TPR-Hsp90 system, the interaction with the N-terminal residues (T¹¹⁷S¹¹⁸R¹¹⁹) appears to dominate over the entropy due to the increased flexibility thus resulting in an enthalpy driven process.

This is the first study reported till date, where we throw light on the AtToc64_TPR and C-term Hsp70/Hsp90 interaction at a molecular level using *in vitro* studies with ITC and *in silico* studies with molecular dynamics. Our *in vitro* and *in silico* studies suggest

Table 3. Computational Alanine Scanning for both ligand bound systems.

Hsp70 mutants	$\Delta\Delta G$
GATIEEVD	-0.39
GPAIEEVD	-0.04
GPTAEEVD	-2.01
GPTIAEVD	-17.6
GPTIEAVD	-15.9
GPTIEEAD	-3.66
Hsp90 Mutants	
TARMEEVD	-0.14
TSAMEEVD	-1.21
TSRAEEVD	-3.08
TSRMAEVD	-12.0
TSRMEAVD	-3.37
TSRMEEAD	-3.97

Point mutations to alanine were performed computationally for each residue (in bold) of the Hsp70 and Hsp90 octapeptides. Changes in binding free energy ($\Delta\Delta G$) in kcal/mol is calculated using MMGBSA.

doi:10.1371/journal.pone.0083461.t003

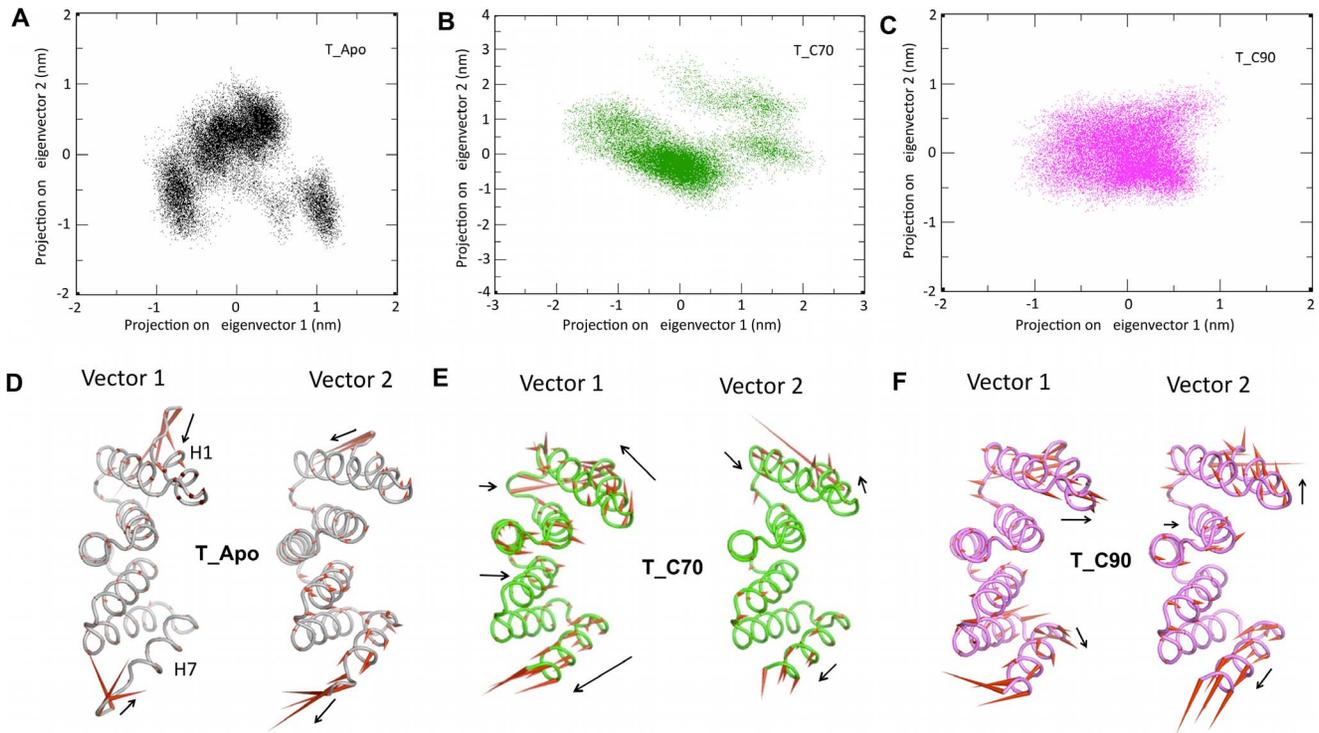


Figure 10. Principal component analysis for the three systems. Essential dynamics -2D projection of individual trajectories with their first two eigen vectors: vector 1 and vector 2 (A–C) and their corresponding porcupine plots (D–F) for T_Apo (black), T_C70 (green) and T_C90 (magenta). Porcupine plots of the three systems displayed with a cone model. The length and orientation of the cone (red) is positively correlated with the magnitude and direction of motion.

doi:10.1371/journal.pone.0083461.g010

that *At*Toc64_TPR can recognize the C-terminal octapeptide of both Hsp70 and Hsp90 with similar affinity and mutation of residues within this peptide to alanine destabilizes the interactions.

Supporting Information

Table S1 Alanine scanning for both ligands (Hsp70/90) using ITC.
(DOCX)

Figure S1 Characterization of the *At*Toc64_TPR-H6. A. Cradle structure of the 3-TPR domain. Helices are numbered as H1 to H7. Each TPR forms a helix turn helix structure which repeats itself three times followed by a solvation or capping helix. **B.** 16% denaturing SDS PAGE gel showing the purification of *At*Toc64_TPR-H6 after Ni-NTA chromatography. The band observed in the red box indicates the presence of *At*Toc64_TPR-H6. **C.** Thermal denaturation curve of *At*Toc64_TPR-H6. T_m was calculated to be $\sim 35^\circ\text{C}$.
(TIF)

Figure S2 Overlay of the ITASSER modelled structure on the Hop crystal structure. Tube depiction of the superposition of the TPR model obtained from ITASSER on the high-resolution crystal structure of the TPR domain from Hop complexes with the Hsp70 octapeptide (1ELW). The modelled structure of the TPR domain is shown in blue color and that of the crystal structure is shown in yellow. The RMSD was found to be 1.36 \AA , suggesting that the modeled structure was quite reliable for use in MD studies.
(TIF)

Figure S3 Ramachandran plot of modelled *At*-Toc64_TPR (Apo).
(TIF)

Figure S4 Ramachandran plot of C-Hsp70 (octapeptide) bound form (T_C70).
(TIF)

Figure S5 Ramachandran plot of C-Hsp90 (octapeptide) bound form (T_C90).
(TIF)

Figure S6 Secondary structure map of all the trajectories created by DSSP. Color codes: green represents alpha helices, black represents turns, grey represents 3_{10} helices and white represents coils.
(TIF)

Figure S7 Dynamic cross-correlation maps. A. Apo, **B.** C-Hsp70 bound form of TPR and **C.** C-Hsp90 bound form of TPR using C α atoms.
(TIF)

Figure S8 Molecular surface representations. A. C-Hsp70 octapeptide is bound to *At*Toc64_TPR; **B.** C-Hsp90 octapeptide is bound to *At*Toc64_TPR. The averaged pdb of the final 5ns of the simulation from each trajectory was used to create these maps. Electronegative and electropositive charges are colored in red and blue respectively.
(TIF)

Acknowledgments

The authors gratefully acknowledge iVEC for allocating computing time on supercomputers under the Merit Allocation Scheme at the iVEC's Pawsey Centre in Perth. We also acknowledge Dr. Senthil Kumar DK (Indian Institute of Science) for useful discussions on the molecular dynamics aspects of the work.

References

- Grove TZ, Cortajarena AL, Regan L (2008) Ligand binding by repeat proteins: natural and designed. *Curr Opin Struct Biol* 18: 507–515.
- Morlot C, Thielens NM, Ravelli RBG, Hemrika W, Romijn RA, et al. (2007) Structural insights into the Slit-Robo complex. *Proceedings of the National Academy of Sciences* 104: 14923–14928.
- Boersma YL, Plückhun A (2011) DARPinS and other repeat protein scaffolds: advances in engineering and applications. *Current Opinion in Biotechnology* 22: 849–857.
- Zeytuni N, Zarivach R (2012) Structural and Functional Discussion of the Tetra-Trico-Peptide Repeat, a Protein Interaction Module. *Structure (London, England : 1993)* 20: 397–405.
- Brinker A, Scheufler C, Von Der Mulbe F, Fleckenstein B, Herrmann C, et al. (2002) Ligand discrimination by TPR domains. Relevance and selectivity of EEVD-recognition in Hsp70×Hop×Hsp90 complexes. *J Biol Chem* 277: 19265–19275.
- Scheufler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, et al. (2000) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* 101: 199–210.
- Allan RK, Ratajczak T (2011) Versatile TPR domains accommodate different modes of target protein recognition and function. *Cell Stress Chaperones* 16: 353–367.
- Kajander T, Cortajarena AL, Main ER, Mochrie SG, Regan L (2005) A new folding paradigm for repeat proteins. *J Am Chem Soc* 127: 10188–10190.
- D'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. *Trends in Biochemical Sciences* 28: 655–662.
- Sampathkumar P, Roach C, Michels PA, Hol WG (2008) Structural insights into the recognition of peroxisomal targeting signal 1 by *Trypanosoma brucei* peroxin 5. *J Mol Biol* 381: 867–880.
- Lunelli M, Lokareddy RK, Zychlinsky A, Kolbe M (2009) IpaB-IpgC interaction defines binding motif for type III secretion translocator. *Proc Natl Acad Sci U S A* 106: 9661–9666.
- Zeytuni N, Ozyamak E, Ben-Harush K, Davidov G, Levin M, et al. (2011) Self-recognition mechanism of MamA, a magnetosome-associated TPR-containing protein, promotes complex assembly. *Proc Natl Acad Sci U S A* 108: E480–487.
- Zhang Z, Kulkarni K, Hanrahan SJ, Thompson AJ, Barford D (2010) The APC/C subunit Cdc16/Cut9 is a contiguous tetratricopeptide repeat superhelix with a homo-dimer interface similar to Cdc27. *EMBO J* 29: 3733–3744.
- 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.
- Sikorski RS, Boguski MS, Goebel M, Hieter P (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.
- Hirano T, Kinoshita N, Morikawa K, Yanagida M (1990) Snap helix with knob and hole: essential repeats in *S. pombe* nuclear protein nuc2+. *Cell* 60: 319–328.
- Baker MJ, Frazier AE, Gulbis JM, Ryan MT (2007) Mitochondrial protein-import machinery: correlating structure with function. *Trends in Cell Biology* 17: 456–464.
- Mirus O, Bionda T, Haesler A, Schlieff E (2009) Evolutionarily evolved discriminators in the 3-TPR domain of the Toc64 family involved in protein translocation at the outer membrane of chloroplasts and mitochondria. *J Mol Model* 15: 971–982.
- Brocard C, Hartig A (2006) Peroxisome targeting signal 1: Is it really a simple tripeptide? *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1763: 1565–1573.
- Li HM, Chiu CC (2010) Protein transport into chloroplasts. *Annu Rev Plant Biol* 61: 157–180.
- Hinnah SC, Wagner R, Sveshnikova N, Harrer R, Soll J (2002) The Chloroplast Protein Import Channel Toc75: Pore Properties and Interaction with Transit Peptides. *Biophysical Journal* 83: 899–911.
- Smith MD, Rounds CM, Wang F, Chen K, Afithile M, et al. (2004) atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *J Cell Biol* 165: 323–334.
- Koenig P, Oreb M, Höfle A, Kaltofen S, Rippe K, et al. (2008) The GTPase Cycle of the Chloroplast Import Receptors Toc33/Toc34: Implications from Monomeric and Dimeric Structures. *Structure* 16: 585–596.
- Sohr K, Soll J (2000) Toc64, a new component of the protein translocon of chloroplasts. *J Cell Biol* 148: 1213–1221.
- Qbadou S, Becker T, Bionda T, Reger K, Ruprecht M, et al. (2007) Toc64 - A Preprotein-receptor at the Outer Membrane with Bipartite Function. *J Mol Biol* 367: 1330–1346.
- Cheng CY, Jarymowycz VA, Cortajarena AL, Regan L, Stone MJ (2006) Repeat Motions and Backbone Flexibility in Designed Proteins with Different

Author Contributions

Conceived and designed the experiments: RP AV JW. Performed the experiments: RP AA. Analyzed the data: RP AA. Contributed reagents/materials/analysis tools: AV JW. Wrote the paper: RP AV JW.

- Numbers of Identical Consensus Tetratricopeptide Repeats†. *Biochemistry* 45: 12175–12183.
- Qbadou S, Becker T, Mirus O, Tews I, Soll J, et al. (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO J* 25: 1836–1847.
- Qbadou S, Becker T, Mirus O, Tews I, Soll J, et al. (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO J* 25: 1836–1847.
- Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89: 392–400.
- Compton LA, Johnson WC Jr (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem* 155: 155–167.
- Manavalan P, Johnson WC Jr (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem* 167: 76–85.
- Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287: 252–260.
- Greenfield N, Fasman GD (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8: 4108–4116.
- Ralston G Introduction to analytical ultracentrifugation. 1–99.
- Laue TM, Shah BD, Ridgeway TM, Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding SE, Rowe AJ, Horton JC, editors. *Analytical ultracentrifugation in biochemistry and polymer science*. Cambridge, UK: Royal Society of Chemistry. pp. 90–125.
- Roy A, Xu D, Poisson J, Zhang Y (2011) A protocol for computer-based protein structure and function prediction. *J Vis Exp*: e3259.
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5: 725–738.
- Venclovas C, Margelevicius M (2005) Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment. *Proteins* 61 Suppl 7: 99–105.
- London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O (2011) Rosetta FlexPepDock web server—high resolution modeling of peptide-protein interactions. *Nucleic Acids Res* 39: W249–253.
- Raveh B, London N, Schueler-Furman O (2010) Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* 78: 2029–2040.
- Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, et al. (2003) Structure validation by Alpha geometry: phi, psi and Cbeta deviation. *Proteins* 50: 437–450.
- Goetz AW, Williamson MJ, Xu D, Polle D, Grand SL, et al. (2012) Routine microsecond molecular dynamics simulations with AMBER - PartI: Generalized Born. *Journal of Chemical Theory and Computation* 8: 1542–1555.
- Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, et al. (2012) AMBER12. University of California, San Francisco.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW (1983) Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* 79: 1407–1413.
- Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* 23: 327–341.
- Salomon-Ferrer R, Goetz AW, Poole D, Grand LS, Walker RC (2012) Routine microsecond molecular dynamics simulations with AMBER- PartII: Particle Mesh Ewald. *Journal of Chemical Theory and Computation* (in Prep).
- McDonald IK, Thornton JM (1994) Satisfying hydrogen bonding potential in proteins. *Journal of Molecular Biology* 238: 777–793.
- Hubbard SJ, Thornton JM (1993) NACCESS. London: Department of Biochemistry and Molecular Biology University College.
- Kollman PA, Massova I, Reyers C, Kuhn B, Huo S, et al. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research* 33: 889–897.
- Srinivasan J, Miller J, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of RNA hairpin loops and helices. *J Biomol Struct Dyn* 16: 671–682.
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98: 10037–10041.

52. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8: 127–134.
53. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14: 33–38, 27–38.
54. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577–2637.
55. Amadei A, Linssen ABM, Berendsen HJC (1993) Essential dynamics of proteins. *Proteins* 17: 412–425.
56. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, et al. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29: 845–854.
57. Grant BJ, Rodrigues AP, ElSawy KM, McCammon JA, Caves LS (2006) Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 22: 2695–2696.
58. Ichiye T, Karplus M (1991) Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins* 11: 205–217.
59. Harte WE Jr, Swaminathan S, Mansuri MM, Martin JC, Rosenberg IE, et al. (1990) Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proc Natl Acad Sci U S A* 87: 8864–8868.
60. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA, USA.
61. Buchan DW, Ward SM, Lobley AE, Nugent TC, Bryson K, et al. (2010) Protein annotation and modelling servers at University College London. *Nucleic Acids Res* 38: W563–568.
62. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* 292: 195–202.
63. Lebowitz J, Lewis MS, Schuck P (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* 11: 2067–2079.
64. Campoy AV, Freire E (2006) Isothermal titration calorimetry to determine association constants for high affinity ligands. *Nature Protocols* 1: 186–191.
65. Freyer MW, Lewis EA (2008) Isothermal titration calorimetry: Experimental design, data analysis, probing macromolecular/ligand binding and kinetic interactions. *Methods in cell biology* 84: 79–113.
66. Deng Y, Roux B (2009) Computations of standard binding free energies with molecular dynamics simulations. *J Phys Chem B* 113: 2234–2246.
67. Krieger F, Moglich A, Kiefhaber T (2005) Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J Am Chem Soc* 127: 3346–3352.
68. Trevino SR, Schaefer S, Scholtz JM, Pace CN (2007) Increasing protein conformational stability by optimizing beta-turn sequence. *J Mol Biol* 373: 211–218.
69. Hainzl O, Wegele H, Richter K, Buchner J (2004) Cns1 is an activator of the Ssa1 ATPase activity. *J Biol Chem* 279: 23267–23273.