



**HAL**  
open science

## Binding properties of the quaternary assembly protein SPAG1

Marie-Eve Chagot, Raphael dos Santos Morais, Sana Dermouche, Dorian Lefebvre, Xavier Manival, Christophe Chipot, François Dehez, Marc Quinternet

► **To cite this version:**

Marie-Eve Chagot, Raphael dos Santos Morais, Sana Dermouche, Dorian Lefebvre, Xavier Manival, et al.. Binding properties of the quaternary assembly protein SPAG1. *Biochemical Journal*, 2019, 476 (11), pp.1679-1694. 10.1042/BCJ20190198. hal-02140524

**HAL Id: hal-02140524**

**<https://hal.univ-lorraine.fr/hal-02140524v1>**

Submitted on 6 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Title : Binding properties of the quaternary assembly protein SPAG1

Marie-Eve Chagot<sup>1</sup>, Raphael Dos Santos Morais<sup>1,\*</sup>, Sana Dermouche<sup>1</sup>, Dorian Lefebvre<sup>2</sup>, Xavier Manival<sup>1</sup>, Christophe Chipot<sup>3,4</sup>, François Dehez<sup>3,4</sup> and Marc Quinternet<sup>2</sup>

<sup>1</sup>Université de Lorraine, CNRS, IMoPA, F-54000 Nancy, France; <sup>2</sup>Université de Lorraine, CNRS, INSERM, IBSLor, F-54000 Nancy, France; <sup>3</sup>LPCT, UMR 7019, Université de Lorraine, CNRS, Vandoeuvre-lès-Nancy, France; <sup>4</sup>Laboratoire International Associé CNRS and University of Illinois at Urbana-Champaign, Vandoeuvre-lès-Nancy, France

Correspondence: François Dehez (francois.dehez@univ-lorraine.fr) or Marc Quinternet (marc.quinternet@univ-lorraine.fr)

### Abstract

In cells, many constituents are able to assemble resulting in large macromolecular machineries possessing very specific biological and physiological functions, e.g. ribosome, spliceosome and proteasome. Assembly of such entities is commonly mediated by transient protein factors. SPAG1 is a multidomain protein, known to participate in the assembly of both the inner and outer dynein arms. These arms are required for the function of sensitive and motile cells. Together with RUVBL1, RUVBL2 and PIH1D2, SPAG1 is a key element of R2SP, a protein complex assisting the quaternary assembly of specific protein clients in a tissue-specific manner and associating with heat shock proteins (HSPs) and regulators. In this study, we have investigated the role of TPR domains of SPAG1 in the recruitment of HSP chaperones by combining biochemical assays, ITC, NMR spectroscopy and molecular dynamics (MD) simulations. First, we propose that only two, out of the three TPR domains, are able to recruit the protein chaperones HSP70 and HSP90. We then focused on one of these TPR domains and elucidated its 3D structure using NMR spectroscopy. Relying on an NMR-driven docking approach and MD simulations, we deciphered its binding interface with the C-terminal tails of both HSP70 and HSP90. Finally, we addressed the biological function of SPAG1 and specifically demonstrated that a SPAG1 sub-fragment, containing a putative P-loop motif, cannot efficiently bind and hydrolyze GTP *in vitro*. Our data challenge the interpretation of SPAG1 possessing GTPase activity. We propose instead that SPAG1 regulates nucleotide hydrolysis activity of the HSP and RUVBL1/2 partners.

## Introduction

Cilia are essential elements of motile and sensitive cells. In human, cilia are crucial for the function of specific tissues, including testis, kidney or lung and their dysfunction is responsible for various pathologies, chief among which are infertility, polycystic kidney disease and primary ciliary dyskinesia (PCD) [1,2]. Cilia are operated by means of a molecular motor composed of both outer and inner dynein arms combining several dozens of proteins. Assembly of these arms requires many protein factors called DNAAFs (dynein axonemal assembly factors) [3], which are not components of the mature arm. SPAG1 belongs to this family of proteins. Several mutations in the human SPAG1 gene were shown to induce PCD [4]. To be specific, these mutations can lead to frameshift in the gene, deletion in the gene, nonsense codons, missense codons, or loss of the start codon. Interestingly enough, the R2SP complex, which combines SPAG1, PIH1D2, RUVBL1 and RUVBL2 proteins, was shown to be involved in the assembly of different protein complexes [5]. In addition, RUVBL1, absent from mature dynein arms, was proven pivotal to maintain the integrity of cilia [6].

The R2SP complex is close to the R2TP complex, with RPAP3 and PIH1D1 adaptors being replaced by SPAG1 and PIH1D2, respectively [5,7]. Both R2SP and R2TP were proposed to work in a concerted fashion with protein chaperones HSP70 and HSP90, thereby facilitating quaternary structure organization of client proteins [8]. It is worth noting that R2TP, together with the RPAP3 interactor WDR92 (also known as MONAD), participates in dynein assembly, but at a stage distinct from SPAG1, thus, suggesting the involvement of HSP (heat shock protein) chaperones in this biological process [9,10]. Despite similarities, R2SP and R2TP target their own specific protein clients. In addition, expression of R2SP was shown to be tissue-specific with a peculiar enrichment in testis observed for SPAG1 [5].

Although the structure of the R2TP complex and its sub-complexes has been extensively studied by NMR (nuclear magnetic resonance), X-ray and cryo-EM, high-resolution three-dimensional data for the R2SP complex are still lacking [11–16]. SPAG1/RPAP3 and PIH1D1/PIH1D2 proteins have been shown to share common features. Whereas PIH1D1 and PIH1D2 are both composed

of one PIH and one CS domain, SPAG1 and RPAP3 both possess one conserved C-terminal domain, as well as several TPR domains [5]. TPR domains (for tetratricopeptide repeat) are repeats of at least two 34-residue motifs arranged in a helix-turn-helix fashion [17]. Interestingly enough, a multi-TPR protein, such as HOP, organizes the delivery of clients from HSP70 to HSP90 via simultaneous anchoring of the chaperones to dedicated TPR domains [18,19]. Although HSP70 and HSP90 are structurally very different, this anchoring is mainly driven by a common motif made of residues EEVD and located at the C-terminal end of the chaperones [20]. Similar functions have been deciphered for RPAP3 [14,21]. More precisely, we have demonstrated that RPAP3-TPR1 is likely to recruit HSP70, whereas RPAP3-TPR2 is responsible for anchoring HSP90 to a large complex involving PIH1D1. The CS domain of the latter protein strongly binds a 39-residue sequence specific to isoform 1 of RPAP3, and its PIH domain is recognized by RPAP3-TPR1. Thus, due to their organizational similarities, RPAP3:PIH1D1 and SPAG1: PIH1D2 could share common functions. However, SPAG1 has been proposed to possess a specific GTP hydrolysis activity, possibly providing the energy required for its co-chaperone function [22]. Moreover, SPAG1 seems also able to activate G-protein-mediated extracellular signal-regulated kinases (ERK1/2) by associating with protein G $\beta$ 1 [23], and other experimental evidence has highlighted its role in nucleotide-dependent pathways, such as the AMP-dependent kinase, mitogen-activated protein kinase and protein kinase C signaling pathways [24]. A P-loop motif (i.e. phosphate-binding loop, also known as Walker A motif), capable of binding specifically the phosphate moiety of nucleotides, and the consensus sequence of which is (G/A)-(X)<sub>4</sub>-GK-(T/S), was identified in the region between the last TPR domain and the C-terminal domain (from residue 781 to residue 788) [22]. It ought to be noted that the presence of this motif, not found in RPAP3, was also established in RUVBL proteins [25]. Indeed, RUVBLs can directly bind and hydrolyze nucleotides, inducing conformational changes that could putatively modify the specificity and/or the affinity of their protein clients [26].

In this work, we deciphered the binding features of SPAG1 toward the protein chaperones HSP70 and HSP90, combining biochemical, biophysical and in silico approaches. We specifically focused on the third TPR domain of SPAG1, exhibiting a high stability in vitro. Using NMR and MD

(molecular dynamics) simulations, we revealed at the molecular level the structure and dynamics of SPAG1 in complex with the tail of HSPs. We also investigated the ability of SPAG1 to bind and hydrolyze GTP. With this study, we bring to the fore new data, key for a better understanding of the involvement of HSP chaperones in dynein assembly.

## **Experimental**

### **Cloning**

DNA sequences of human HSP70, HSP90 and HSP90-MC as well as DNA sequences of the full-length human SPAG1 and of sub-fragments 206–327, 446–570, 440–573, 622–742 and 622–926 were cloned into pnEA-3CH, pnYK and pCARGHO vectors [27,28]. The pnEA and pCARGHO vectors introduce, respectively, a 6x-His-tag and a CRD-tag at the N-terminal position of the recombinant proteins. The pnYK vector produces native recombinant proteins.

### **Protein co-expression assays**

For each co-expression test, a Ca<sup>2+</sup>-competent *Escherichia coli* BL21(DE3) pRARE2 strain was co-transformed with plasmids encoding his-tagged ( pnEA) and native ( pnYK) proteins [27]. The positive co-transformations were selected on LB-agar medium supplemented with antibiotics. One single colony was used for IPTG-induced overexpression of recombinant proteins in a selective LB medium, overnight at 20°C. The cell pellet was then sonicated in lysis buffer [25 mM HEPES ( pH 7.5), 300 mM NaCl, 0.5 mM TCEP, 10 mM imidazole]. The cell lysate was placed in contact with TALON beads (Clontech) and left stirring for 1 h at 4°C.

Beads were washed three times with the lysis buffer and proteins retained on the beads were eluted using the lysis buffer supplemented with 300 mM imidazole. Samples were analyzed by SDS–polyacrylamide gels (10% SDS-PAGE). Control experiments were also performed with native proteins purified in the absence of his-tagged partners.

### **Purification of recombinant proteins**

Isolated recombinant proteins were overexpressed in *E. coli* BL21(DE3) pRARE2 grown in a

selective LB medium or in a selective M9 medium supplemented with 0.5 g/L  $15\text{NH}_4\text{Cl}$  and 2.0 g/L  $^{13}\text{C-d6-glucose}$  for NMR samples. Briefly, bacteria were transformed with the plasmid of interest, selected on an LB agar medium and placed in pre-culture in a selective LB medium at  $37^\circ\text{C}$ . One liter of LB or NMR growth medium was seeded with the pre-culture. When  $\text{OD}_{600\text{nm}}$  reached  $\sim 0.6$ , overexpression of the protein was induced with 0.2 mM IPTG and cell cultures were placed at  $20^\circ\text{C}$  for one night. Cells were harvested and sonicated in lysis buffer [25 mM HEPES (pH 7.5), 300 mM NaCl, 0.5 mM TCEP, 10 mM Imidazole]. The tagged proteins were retained on TALON (or lactose sepharose) beads that were washed three times with the lysis buffer. Elution from the beads was performed using the 3C-protease or with lysis buffer supplemented with 300 mM imidazole or D-lactose whether the 6x-His or CRD tags were kept or not. Finally, a size exclusion chromatography performed in 10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP on Superdex-75 or Superdex-200 columns permitted to isolate the protein of interest. The purification process was followed on SDS-PAGE (12.5%). When necessary, the final buffer was exchanged using several dilution/concentration cycles in Amicon Ultra Centrifugal Filters (Merck Millipore).

### **Dynamic light scattering**

DLS experiments were performed at  $20^\circ\text{C}$  on a Zetasizer Nano ZS (Malvern Panalytical) in low-volume quartz cell. Data were processed using the Zetasizer software (Malvern) with default parameters.

### **Size exclusion chromatography coupled to static light scattering**

Size exclusion chromatography coupled to static light scattering experiment (SEC-SLS) was performed at  $30^\circ\text{C}$  with an AKTA Explorer HPLC system (GE Healthcare) coupled with both RALS/LALS light scattering and refractive index detectors (Viscotek TDA302, Malvern Panalytical). The refractive index increment value ( $dn/dc$ ) used to determine the molecular weight was 0.185 mL/g. The SEC column (Shodex KW-804, 7  $\mu\text{m}$ , 8 mm ID  $\times$  300 mm, void

volume ~6 mL, total volume ~12.5 mL) was equilibrated with phosphate-buffered saline (pH 7.4, PBS, Gibco, Thermo Fischer Scientific). The flow rate was 0.5 mL/min. Data were processed using the OmniSEC software (v5.12, Malvern Panalytical).

### **Isothermal titration calorimetry**

The binding of synthetic peptides DASRMEEVD (C-terminal fragment 724–732 of human HSP90) and SGPTIEEVD (C-terminal fragment 638–646 of human HSP70) to the full-length SPAG1 or sub-fragments was measured at 288 K with an ITC-200 microcalorimeter (Malvern) in 10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP. Peptide and protein concentrations, around 1.5 mM and 0.15 mM, respectively, were calculated using absorbance at 205 and 280 nm [29]. Peptides and proteins were dispatched into the syringe and cell, respectively. Calorimetric data were treated with ORIGIN7 to derive the dissociation constant ( $K_d$ ), the stoichiometry ( $N$ ), and the variations in enthalpy ( $\Delta H$ ), entropy ( $\Delta S$ ) and free energy ( $\Delta G$ ).

### **Protein–protein cross-link assays**

His-tagged HSP70 and HSP90-MC proteins concentrated at 30  $\mu$ M were incubated for 1 h at room temperature with SPAG1-TPRs 1 and 3 concentrated at 120  $\mu$ M in 10 mM NaPi (pH 6.4), 150 mM NaCl supplemented with 20 mM EDC ((1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, Sigma Aldrich). The cross-linking reaction was stopped by addition of LDS buffer (Expedeon). Controls without EDC were performed. Samples were analyzed on precast SDS-PAGE Gradient 4–12% (Expedeon).

### **HSP binding monitored by NMR**

$^1\text{H}$ - $^{13}\text{C}$  METHYL-SOFAST-HMQC were recorded on  $^{13}\text{C}$ -labelled SPAG1-TPRs 1 and 3 concentrated at 30 mM in 10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP, 5% D<sub>2</sub>O in the absence or in the presence of a 3 molar excess of non-labeled his-tagged HSP70 or HSP90-MC.

Spectra were collected at 293 K on a 600 MHz Avance III spectrometer equipped with a TCI cryoprobe.

### **NMR analysis of SPAG1-TPR3**

An almost complete resonance assignment (>91%) of the free <sup>13</sup>C/<sup>15</sup>N-labelled SPAG1-TPR3 was obtained using a 1 mM sample in 10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP, 5% D<sub>2</sub>O. Classical 2D and 3D spectra were recorded at 293 K on a 600 MHz Avance III spectrometer equipped with a TCI cryoprobe. 3D structures were derived from <sup>1</sup>H-<sup>15</sup>N and <sup>1</sup>H-<sup>13</sup>C NOESYHSQC and <sup>1</sup>H-<sup>1</sup>H NOESY experiments recorded with a mixing time of 120 ms and from dihedral restraints coming from TALOS-N and PREDITOR [30,31]. CYANA3.0 was used for NOE assignment [32]. Finally, 100 structures were refined in water using RECOORD scripts [33]. The 20 structures with the lowest overall energies were selected as a representative ensemble of structures.

The NMR assignments of the HSP-bound forms of <sup>13</sup>C/<sup>15</sup>N-labelled SPAG1-TPR3 were obtained using classical heteronuclear 3D NMR spectra recorded on samples containing 1 mM protein and 3 mM peptides. Buffer and temperature conditions were the ones described above. 2D- and 3D-filtered NOESY experiments collected with a mixing time of 120 ms were recorded to select the <sup>1</sup>H signals coming from the peptide before transferring them either to the protein or to the peptide signals. Filtered TOCSY experiments collected with a mixing time of 20 and 70 ms were also recorded to obtain the NMR assignment of the bound peptide.

Chemical shift perturbations (CSPs) of backbone amide groups in SPAG1-TPR3 upon HSP peptides binding were determined by comparing <sup>1</sup>H-<sup>15</sup>N HSQC spectra of the free and bound forms of the <sup>15</sup>N-labelled protein.

For this purpose, composite <sup>1</sup>H-<sup>15</sup>N chemical shifts were calculated as follows:  $\Delta\delta = \text{Square root}(0.1(\delta^{15}\text{N})^2 + (\delta^1\text{H})^2)$ .

### **HADDOCK modeling**



HADDOCK2.2 and CNS1.3 were used to determine the NMR-driven structure of the SPAG1-TPR3 bound to SGTPIEEVD and DASRMEEVD peptides [34]. The best NMR structure of SPAG1-TPR3 and the X-ray structures of the HSP70 and HSP90 peptides enclosed, respectively, in PDB entries 3Q49 and 4I2Z were used as starting structures [35,36]. Filtered NOE cross-peaks recorded on SPAG1-TPR3:peptide complexes were converted into distance restraints and injected in the docking procedure. Automated mode was chosen for the definition of the semi-flexible interfaces. 512 and 128 structures were calculated in the rigid-body and the semi-flexible simulated annealing steps, respectively. For the last 128 structures, a final flexible refinement in water was performed. Clustering of the docking solutions was performed using the FCC (Fraction of Common Contacts) algorithm with a cutoff of 0.75 [37]. The 3D structures were ranked according to their HADDOCK score.

### **Molecular dynamics simulation**

Molecular assays were built starting from the best Haddock predicted structures subsequently solvated in a cubic box (edge 80 Å) by ~15 300 water molecules in the presence of 150 mM NaCl. Ions, protein and the three peptides were described using the CHARMM36 force field [38] including CMAP corrections [39] whereas water was represented by the TIP3P model [40]. All MD trajectories were generated with the NAMD package (version 2.13) [41]. The equation of motions was integrated by means of a multiple time step algorithm [42] using a time step of 2 and 4 fs for short- and long-range interactions, respectively. Covalent bonds involving hydrogen atoms were constrained to their equilibrium values by means of the Rattle algorithm [43]. All trajectories were carried out in the isobaric–isothermal ensemble at 300 K under 1 atm using Langevin dynamics [44] (damping coefficient 1 ps<sup>-1</sup>) and the Langevin piston method [45]. Long-range electrostatic interactions were taken into account by means of the Particle Mesh Ewald (PME) algorithm [46]. Water and ions were first thermalized during 5 ns while restraining harmonically the heavy atoms of the complexes to their initial positions. Next, all positional restraints were removed. The distance between the protein residue I724 and the peptide residues

Met728/Ile642 was enforced to evolve at values equal or lower to that observed by NMR using the 'colvar' module of NAMD [47]. The molecular assays were further thermalized in this condition for 50 ns. Finally, MD trajectories of 550 ns were generated without any restraints and analyzed using VMD [48].

## **GTP binding and hydrolysis monitored by NMR**

<sup>1</sup>H-<sup>15</sup>N BEST-TROSY-HSQC and HSQC spectra were recorded on the <sup>15</sup>N-labelled fragment 622–926 of SPAG1 concentrated at 0.1 mM in the presence or absence of a 10, 20 and 30 molar excess of GTP using a 600 MHz Avance III spectrometer equipped with a TCI cryoprobe. Data were collected at 293 K in 10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP, 5% D<sub>2</sub>O supplemented or not with 5 mM MgCl<sub>2</sub>. Affinity value was derived from CSP using an in-house python script.

GTPase activity was followed using decoupled 1D <sup>31</sup>P NMR spectra recorded at 298 K in 10 mM HEPES (pH 7.5), 150 mM NaCl, 5% D<sub>2</sub>O, 5 mM MgCl<sub>2</sub>, 1 mM GTP supplemented or not with 0.1 mM <sup>15</sup>N-labelled SPAG1(622–926). Samples were incubated for 24 h at 298 K before data acquisition. Spectra were recorded on a 400 MHz Avance III spectrometer equipped with a BBFO probe.

## **Results**

### **SPAG1 and RPAP3 display similar structural organization**

SPAG1 and RPAP3 are homologous proteins, binding, respectively, PIH1D2 and PIH1D1 to form, with RUVBL1/2 proteins, on the one hand, the R2SP complex, and, on the other hand, the R2TP complex. When aligned, the amino acid sequences of SPAG1 and RPAP3 display similarity and identity scores of 34% and 23%, respectively (Figure 1A, Supplementary Figure S1). As previously described, SPAG1 and RPAP3 both possess TPR domains and one specific C-terminal domain, which represents the signature of R2TP-like quaternary chaperones [5]. Despite these similarities, SPAG1 is 30% longer than RPAP3. This difference in size is mainly

due to the presence of one supplemental TPR domain in SPAG1, compared to the two TPR domains found in RPAP3. More specifically, analysis of the amino acid alignment suggests that TPRs 1 and 2 in RPAP3 are homologous to TPRs 1 and 3 in SPAG1, leaving an orphan TPR2 in the middle of the SPAG1 sequence. It is noteworthy that five of the eight mutations described in the SPAG1 gene and involved in PCD, take place in TPRs 1 and 3, the three last mutations occurring within the N- and the C-terminal regions [4]. All of the three TPRs in SPAG1 hold five conserved residues defining a carboxylate-clamp motif, suggesting an ability to bind the EEVD sequence located at the C-terminal tail of the HSP family members [49].

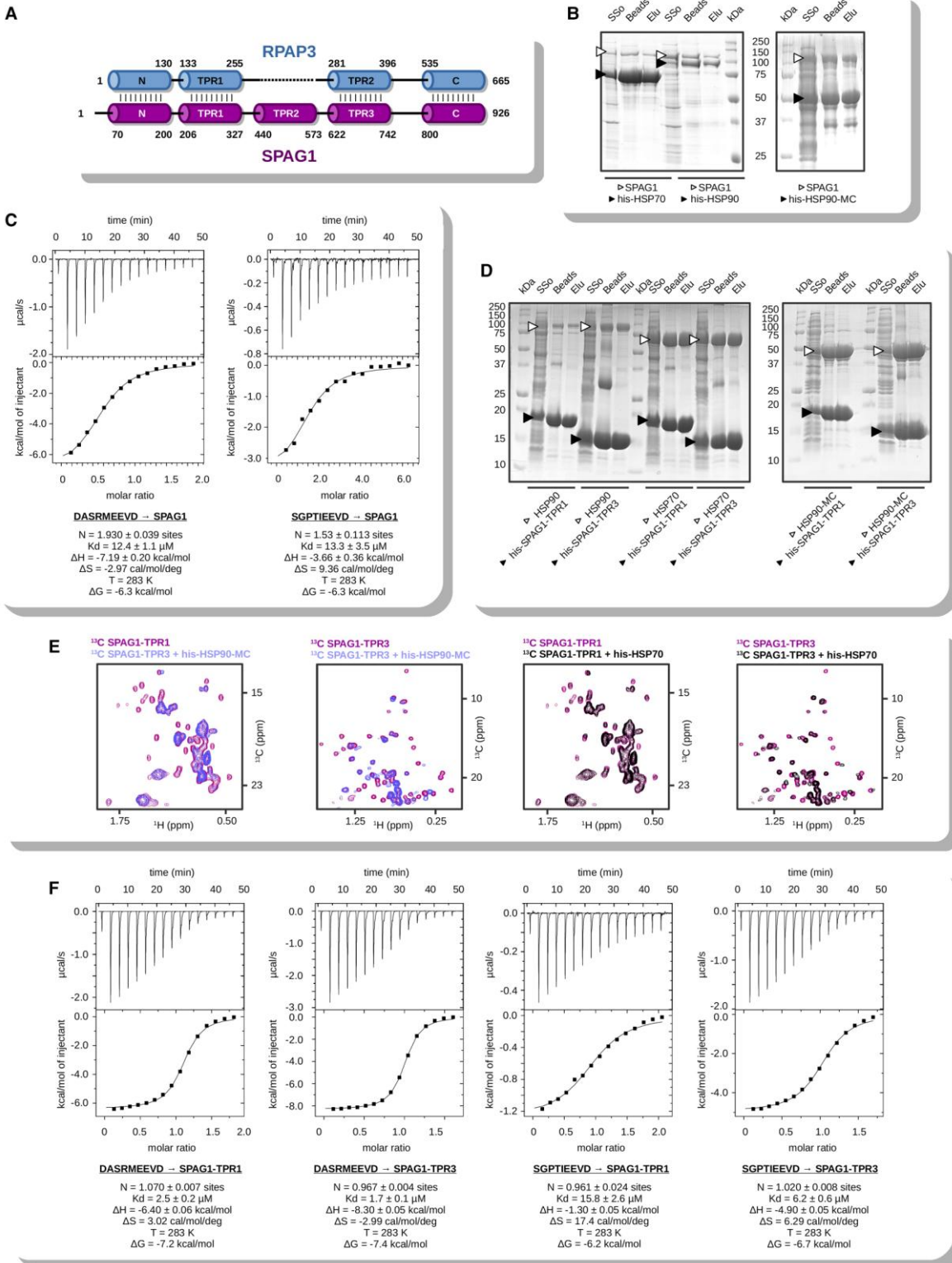
### **SPAG1 binds two EEVD motifs with micromolar affinities**

In RPAP3, TPRs 1 and 2 display significantly different affinities for HSP90 C-terminal peptides ( $K_d \sim 100$  and  $\sim 5$  mM, respectively) and similar affinities for HSP70 peptides ( $K_d \sim 50$  mM) [14]. As previously performed for RPAP3, we determined the HSP-binding properties of SPAG1 to get insight into the specific role of its three TPR domains in the putative recruitment of HSP chaperones.

First, we performed protein co-expression assays (Figure 1B, Supplementary Figure S2A) and showed that the full-length SPAG1 co-elutes with his-tagged HSP70, HSP90 and HSP90-MC (the latter protein corresponding to HSP90 truncated of its N-terminal domain). Our results suggest that the C-terminal tails of HSP proteins are sufficient to recruit SPAG1. We further carried out ITC (isothermal titration calorimetry) analyses with the full-length SPAG1 and long peptides mimicking the HSP tails (HSP90: DASRMEEVD and HSP70: SGPTIEEVD) (Figure 1C). Using static light scattering coupled to exclusion size chromatography, we first checked that the full-length SPAG1 was monomeric (Supplementary Figure S2B). Similar  $K_d$  values of  $\sim 13$  mM were then measured for the two peptides. More strikingly, a stoichiometry value of  $\sim 2$  was obtained for the two peptides, meaning that two binding sites in SPAG1, with similar affinities toward peptides, are able to provide ITC signal. These results strongly suggest that only two out of the three TPR domains are available in SPAG1 for the binding of HSP tails.

To assess this property, we tried to overexpress isolated TPR domains of SPAG1 in *E. coli*. Whereas TPRs 1 and 3 were found soluble (i.e. fragments 206–327 and 622–742) and suitable for biophysical studies, we failed to solubilize TPR2 with the desired efficiency, despite several attempts toward this end (Supplementary Figure S3A–D). This shortcoming suggests that TPR2 in SPAG1 displays hydrophobic features that cause it to dive mainly into the bacteria pellet after cell lysis and centrifugation. Insofar as recombinant TPRs 1 and 3 are concerned, their purification was successfully performed. TPR3 was found highly soluble and stable, while TPR1 exhibited a marked propensity toward aggregation and instability upon concentration (i.e. for concentration >150 mM, Supplementary Figure S3E). Next, using protein co-expression assays, we showed that isolated his-tagged TPRs 1 and 3 co-eluted with full-length HSP70, HSP90 and HSP90-MC (Figure 1D, Supplementary Figure S2A). These interaction data were strengthened, first, using NMR experiments, wherein drastic changes in the SOFAST-METHYL-HMQC spectra of TPRs 1 and 3 occurred upon addition of HSP70 and HSP90-MC (Figure 1E, Supplementary Figure S4A), and, second, using protein–protein cross-link assays (Supplementary Figure S4B). To acquire additional information about the interaction, we then conducted ITC experiments with the HSP peptides (Figure 1F). The stoichiometry values of  $\sim 1$  demonstrate that each TPR domain possesses one HSP-binding site. With the HSP90 peptide, the  $K_d$  values were similar for TPR1 ( $K_d = 2.5 \pm 0.2$  mM) and for TPR3 ( $K_d = 1.7 \pm 0.1$  mM). With the HSP70 peptide, the  $K_d$  values were slightly higher than the ones measured with the HSP90 peptide. Specifically, a 3- and a 6-fold increase were obtained, respectively, for TPR3 ( $K_d = 6.2 \pm 0.6$  mM) and for TPR1 ( $K_d = 15.8 \pm 0.2$  mM).

Altogether, these data suggest that only TPRs 1 and 3 within the full-length SPAG1 could efficiently recruit the EEVD motif of the HSP family members. Indeed, considering the unexpected insolubility of the TPR2 domain, we can assume that its carboxylate-clamp motif is not accessible within the full-length protein. Last, working with isolated TPR domains slightly enhances the affinity toward HSP90 peptides but should reflect the behavior of these modules in the context of the full-length SPAG1. We, thus, tried to establish the molecular basis underlying the recognition between SPAG1 and members of the HSP family using NMR spectroscopy.



**Figure 1. SPAG1 interacts with HSP70 and HSP90.**

(A) Schematic comparison between the human RPAP3 and SPAG1 proteins. Limitations of the sub-domains are indicated. (B) Protein co-expression assays between SPAG1 and his-HSP70, his-HSP90 and his-HSP90-MC. 'SSo', 'Beads' and 'Elu' relate respectively to the supernatant sonicate, talon beads and elution from the beads with imidazole. The co-purified proteins are indicated with arrows. (C) ITC experiments recorded between SPAG1 and peptides DASRMEEVD (HSP90) and SGPTIEEVD (HSP70). Stoichiometry (N), dissociation constant ( $K_d$ ), temperature (T), variations in enthalpy ( $\Delta H$ ), in entropy ( $\Delta S$ ) and in free energy ( $\Delta G$ ) are reported. (D) Protein co-expression assays between his-SPAG1-TPRs 1 and 3 and HSP90, HSP70 and HSP90-MC (E) SOFAST-METHYL-HMQC spectra of isolated TPR1 and TPR3 of SPAG1 in the absence (purple) and in the presence of his-tagged HSP90-MC (blue) and HSP70 (black). In all cases, methyl resonances are significantly shifted or highly broaden upon addition of HSP proteins. (F) ITC experiments recorded between isolated TPR1 and TPR3 of SPAG1 and peptides DASRMEEVD (HSP90) and SGPTIEEVD (HSP70). Same parameters as those described in C are reported

### **NMR analysis of the free and HSP-bound forms of SPAG1-TPR3**

We set out to perform NMR experiments on TPR1. In our hands, this domain was somewhat unstable, thus, precluding a clear NMR assignment. We then focused on TPR3, fully stable and soluble when purified from a  $^{13}\text{C}/^{15}\text{N}$ -enriched M9 medium.

First, most of the  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  resonances in SPAG1-TPR3 were assigned and derived into a well-resolved ensemble of three-dimensional structures (Table 1). SPAG1-TPR3 is composed of seven  $\alpha$ -helices, the first six organized into three TPR motifs (Figure 2A). Helix  $\alpha 7$  plays the role of a capping helix, which may enhance the solubility by shielding several hydrophobic residues located on helix  $\alpha 6$  (Figure 2B). The residues of the carboxylate-clamp, namely K627, N631, N662, K692 and R696, are located on the concave face of the domain and are ideally positioned in a cradle ready to bind the EEVD tail of HSP proteins (Figure 2C).

Next, building on our ITC experiments, we monitored by NMR the binding of HSP peptides to  $^{15}\text{N}$ -labelled SPAG1-TPR3, using  $^1\text{H}$ - $^{15}\text{N}$  CSPs. This method reveals the footprint left by the peptide on the protein domain. Strikingly, the two peptides provided very similar CSP profiles (Figure 2D). Helices  $\alpha 1$ ,  $\alpha 3$ ,  $\alpha 5$  (which hold the five carboxylate-clamp residues) and  $\alpha 7$  were similarly perturbed upon peptide binding. These data suggest a common binding mode for HSP90 and HSP70, but also, and at first sight, a minor contribution to the binding mode from residues lying upstream the EEVD sequence. Some slight differences could, nonetheless, be observed at the root of helix  $\alpha 7$ , when HSP70 and HSP90 footprints are compared, which could be correlated to the difference in amino acid sequence of the two peptides and to the slight difference in affinity values measured by ITC (Figure 1F). Thus, in order to understand the molecular mechanisms that drive these binding modes, we modeled SPAG1-TPR3:HSP-peptide complexes and performed MD simulations.

## **HADDOCK modeling and molecular dynamics simulation of the SPAG1-TPR3: HSP-peptide complexes**

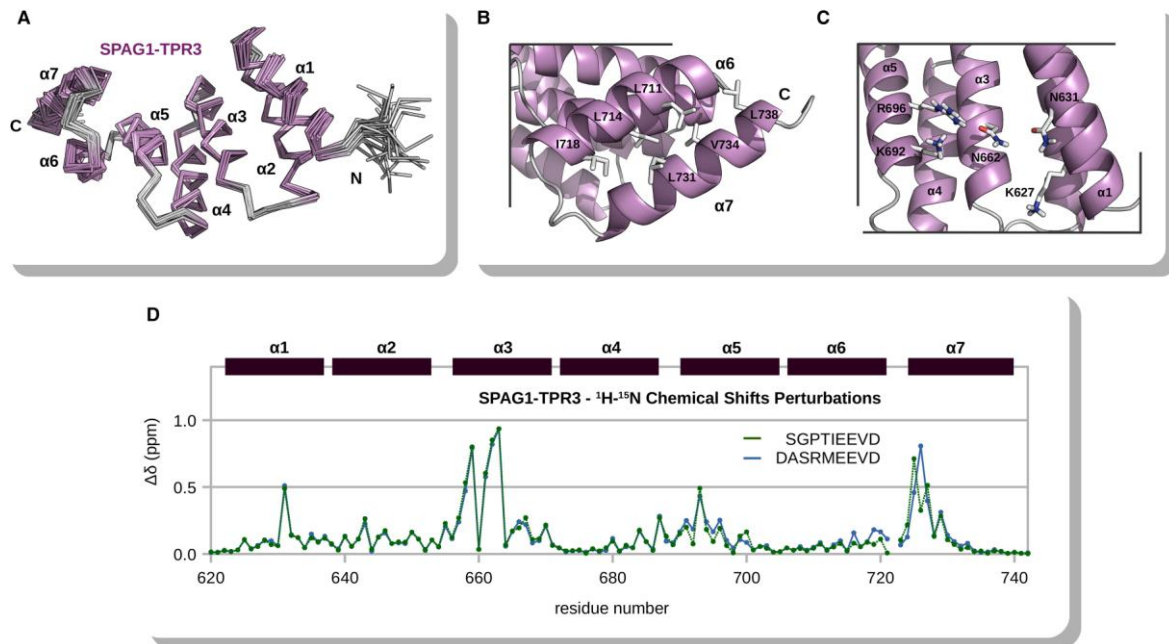
We started our structural study by re-assigning the NMR frequencies of  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  nuclei in the HSP-bound forms of SPAG1-TPR3. We then recorded intermolecular NMR signals using filtered NOESY experiments (Figure 3). HSP90 and HSP70 peptides provided common intermolecular NOEs, in agreement with similar CSP profiles imprinted on SPAG1-TPR3. We aimed to use these unambiguous NMR data to drive a docking procedure using the HADDOCK protocol [34,50].

Toward this end, we screened the Protein Data Bank in search of three-dimensional structures of HSP peptides that could globally satisfy both intermolecular data and the CSP profiles. For the HSP90 peptide, we found entry 4I2Z, which depicts the crystal structure of the HSP90-bound form of myosin chaperone UNC-45 from *Caenorhabditis elegans* [35]. We extracted the ASRMEEVD segment from this entry and fulfilled the N-terminal position with an Asp residue. For the HSP70 peptide, we found entry 3Q49, which depicts the crystal structure of the TPR domain of CHIP complexed to peptide PTIEEVD [36]. We extracted the latter segment and fulfilled the N-terminal position with a Ser and a Gly residue. Each peptide was docked onto the best NMR structure of SPAG1-TPR3, using 49 and 40 intermolecular distance restraints for HSP90 and HSP70, respectively (Supplementary Figure S5).

**Table 1 NMR-derived restraints and structural statistics for the 20 best structures of SPAG1-TPR3 calculated with RECOORD scripts**

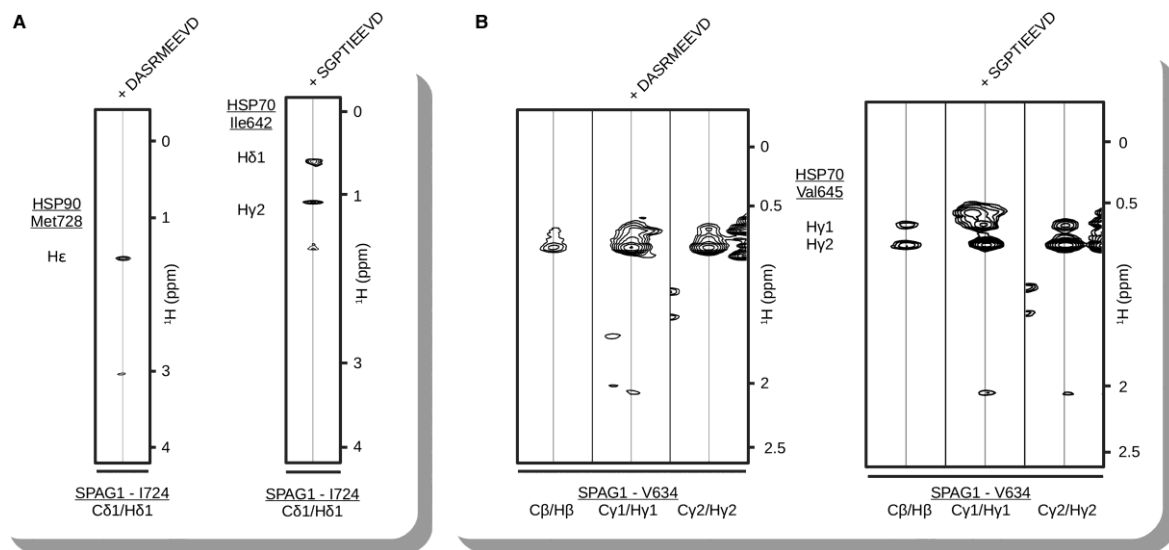
NMR distances and dihedral constraints		
SPAG1-TPR3	Total NOEs	3488
	Short range ( $ i-j  \leq 1$ )	1605
	Medium range ( $ i-j  < 5$ )	943
	Long range ( $ i-j  \geq 5$ )	940
	Total dihedral angle restraints	239
<b>f</b>		107
$\psi$		105
$\chi^1$		27
Structure statistics		
	Violation occurrences	0
	Distance constraints ( $>0.5 \text{ \AA}$ )	0
	Dihedral angle constraints ( $>5^\circ$ )	0
Deviations from idealized geometry		
	Bond lengths ( $\times 10^{-3} \text{ \AA}$ )	$2.743 \pm 0.169$
	Bond angles ( $^\circ$ )	$0.433 \pm 0.013$
	Impropers ( $^\circ$ )	$1.083 \pm 0.055$
RMSD to best structure ( $\text{\AA}$ )		
	All backbone atoms	$1.17 \pm 0.28$
	All heavy atoms	$1.40 \pm 0.22$
	Backbone atoms in secondary structures	0.
	$45 \pm 0.09$ Heavy atoms in secondary structures	0.
	$93 \pm 0.07$	
Ramachandran statistics		
	Residues in most favored regions (%)	92.1
	Residues in additional allowed regions (%)	7.9
	Residues in generously allowed regions	0
	Residues in disallowed regions (%)	0





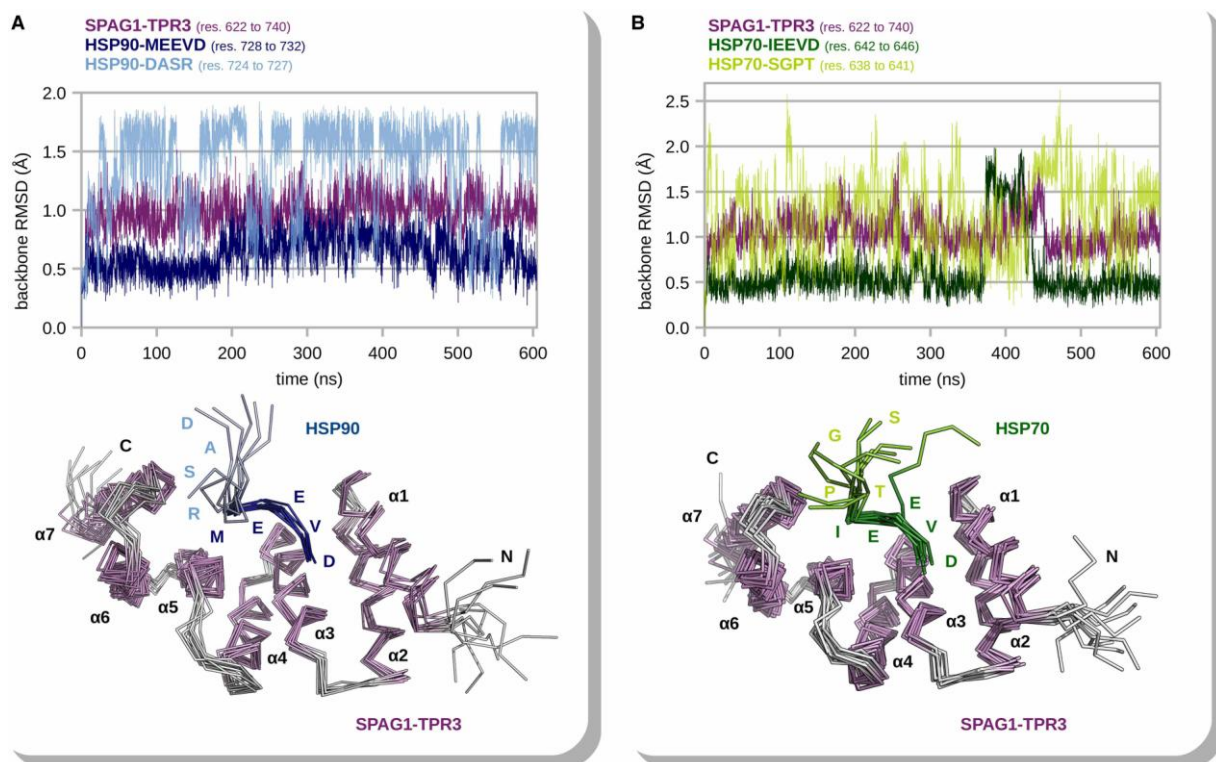
**Figure 2. NMR analysis of the free and HSP-bound SPAG1-TPR3.**

(A) Backbone representation of the 20 best NMR structures of SPAG1-TPR3. (B) Zoom on the capping helix  $\alpha 7$ . (C) Zoom on the conserved residues K627, N631, N662, K692 and R696 of the carboxylate-clamp. In (B) and (C), the best NMR structure was used and only polar hydrogens were represented. (D) Chemical shift perturbations ( $\Delta\delta$  in ppm) of backbone amide groups in  $^{15}\text{N}$ -labelled SPAG1-TPR3 in the presence of peptides DASRMEEVD (blue line) and SGPTIEEVD (green line) were plotted against the protein sequence number. The limitations of SPAG1-TPR3  $\alpha$ -helices are indicated on the top of the diagrams.



**Figure 3. SPAG1-TPR3 shares a common pattern of intermolecular NOEs with the two HSP peptides.**

(A) Intermolecular NOEs detected between Ile724 (A) and Val634 (B) in  $^{13}\text{C}$ -labelled SPAG1-TPR3 and HSP peptides using 3D-filtered NOESY experiments.



**Figure 4. Evolution of backbone RMSDs in SPAG1-TPR3:HSP complexes.**

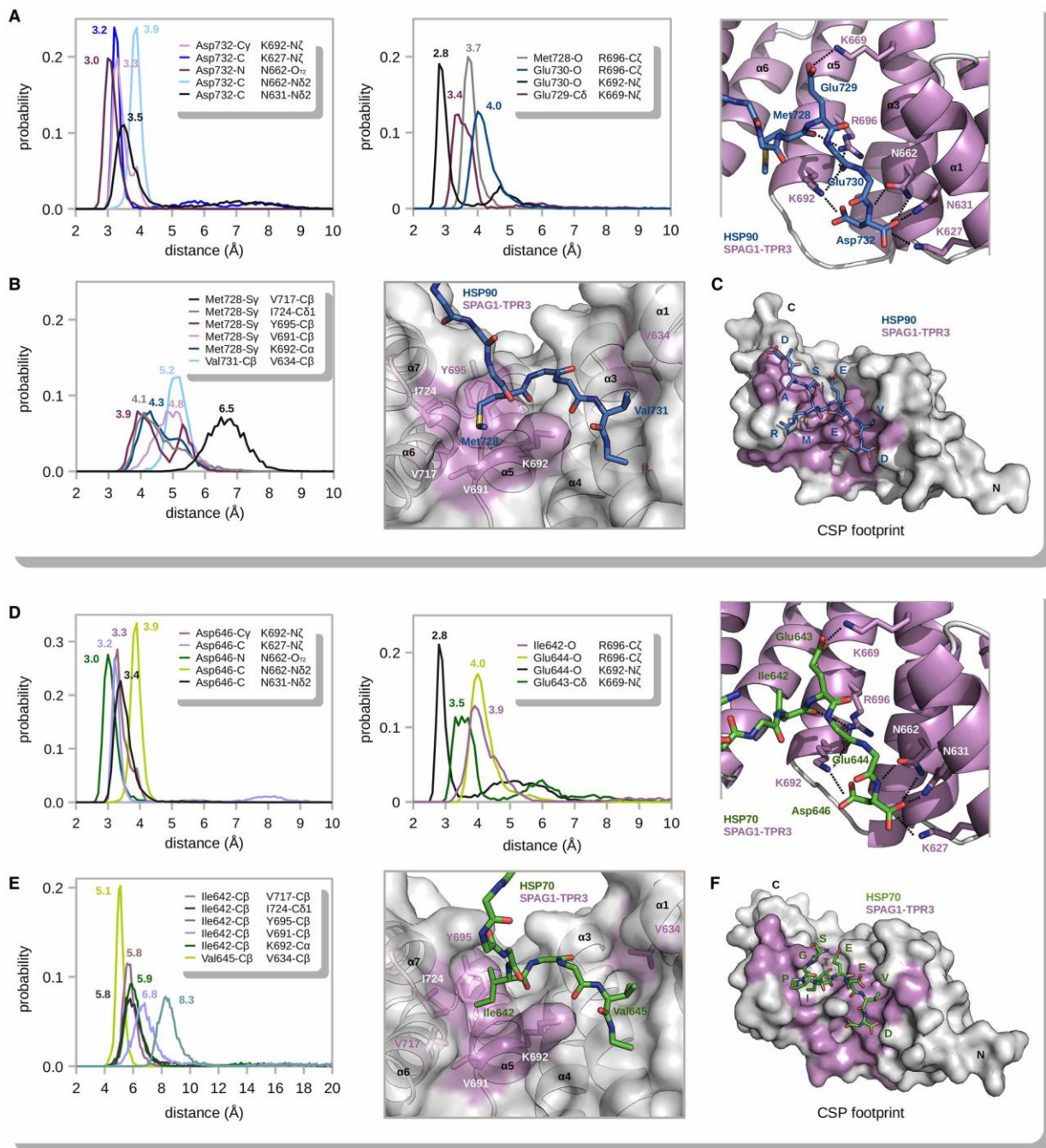
(A) Analysis of the SPAG1-TPR3:DASRMEEVD complex. The backbone RMSD values of SPAG1-TPR3 (residues 622 to 740, in purple) and of the DASR and MEEVD segments in the peptide (respectively, in light and dark blue) are plotted against the simulation time. (B) Analysis of the SPAG1-TPR3:SGPTIEEVD complex. The backbone RMSD values of SPAG1-TPR3 (residues 622–740, in purple) and of the SGPT and IEEVD segments in the peptide (respectively, in light and dark green) are plotted against the simulation time. In (A) and (B), the first frame of the trajectory was the reference structure. In (A) and (B), 3D structures were extracted from the trajectories at 60-ns intervals, superimposed and represented as ribbons.

For the two HADDOCK calculations, the FCC clustering approach (cutoff = 0.75), subsequent to the final water refinement stage, placed all the 128 calculated structures of each complex in a unique ensemble [37]. No distance restraint violations greater than 0.5 Å were detected. C $\alpha$  RMSD values of  $0.76 \pm 0.11$  Å and  $0.77 \pm 0.23$  Å were measured on the top-10 structures of SPAG1-TPR3 complexed with HSP90 and HSP70, respectively (Supplementary Figure S6). No significant change was observed between the bound and the unbound conformations of SPAG1-TPR3 (C $\alpha$  RMSDs  $\sim 0.5$  Å).

Next, to assess the structure and the dynamics of the binding interface, a 605-ns MD simulation was performed on both SPAG1-TPR3:DASRMEEVD and SPAG1-TPR3:SGPTIEEVD immersed in a water box, using the best HADDOCK solutions as a starting point. In the first 5 ns of the simulations, distance restraints were enforced to limit the motions of backbone atoms, while permitting reorganization of the side chains. A single distance restraint based on intermolecular

NOEs was then used to immobilize the peptide over the 50 following ns. Finally, all geometric restraints were removed for the last 550 ns of the simulations. With respective mean backbone RMSD values of  $0.99 \pm 0.15 \text{ \AA}$  and  $1.05 \pm 0.19 \text{ \AA}$ , SPAG1-TPR3 bound to HSP90 and HSP70 peptides showed only minor deviations from the initial position along the trajectories (Figure 4, Supplementary Movies MS1 and MS2). Similarly, the MEEVD and IEEVD parts, in the HSP90 and HSP70 peptides, respectively, were shown to be stable throughout the trajectories, as depicted in the mean backbone RMSD values of  $0.63 \pm 0.15 \text{ \AA}$  and  $0.61 \pm 0.32 \text{ \AA}$  measured for these segments (Figure 4, Supplementary Movies MS1 and MS2). We could, nevertheless, note a short and partial release of the IEEVD segment between 370 and 430 ns.

Stability of the interface in the two complexes was ensured by a common pattern of interactions. For clarity in the following description, the one-letter code and three-letter code were, respectively, used for the denomination of residues in SPAG1-TPR3 and in HSP peptides. First, an electrostatic network was woven between the five residues K627, N631, N662, K692 and R696 of the carboxylate-clamp motif from SPAG1-TPR3 and the residues Met728 to Asp732 in HSP90 and the corresponding amino acids Ile642 to Asp646 in HSP70 (Figure 5). Specifically, the carboxylate groups of the main and side chains of Asp732/Asp646 in the HSP90/70 peptides were highly stabilized by the side chains of K627, N631, N662 and K692 in SPAG1. R696 in SPAG1 and the backbones of Met728/Ile642 and Glu730/Glu644 in the peptides contributed efficiently to the interface as well as the side chains of K669 and Glu729/Glu643. Second, hydrophobic contacts were formed between the side chains of V634 in SPAG1 and Val731/Val645 in the peptides (Figure 5). This binding mode matches well with that commonly observed in the other three-dimensional structures of HSP-bound TPR domains. Third, supplemental stabilization forces were provided by a hydrophobic pocket made of side chains from V691, K692, Y695, V717 and I724 in SPAG1 harboring the methyl groups of Met728 and Ile642 in the HSP peptides (Figure 5). Similar pockets were also observed in other complexes, such as RPAP3, CHIP and FKBP38 bound to similar peptides [14,36,51].



**Figure 5. MD analysis of SPAG1-TPR3 in complex with the DASMEEVD and SGTPIEEVD peptides.**

(A,D) Distribution of intermolecular distances involving carboxylate-clamp residues in SPAG1-TPR3 and atoms in the peptides. For each distribution, the value of the most populated class of distance is indicated. A view of the polar network of interaction is shown on a representative 3D structure extracted from the trajectory. (B,E) Distribution of intermolecular distances involving residues forming hydrophobic pockets in SPAG1-TPR3 and hosted residues in the peptides. For each distribution, the value of the most populated class of distance is indicated. A view of the pockets is shown on a representative 3D structure extracted from the trajectory. (C,F) Residues in SPAG1-TPR3 displaying  $^1\text{H}$ - $^{15}\text{N}$  CSP values in the 75<sup>th</sup> percentile upon peptide binding are colored in purple on the molecular surface of a representative 3D structure of the complexes (see also Figure 2). For convenience, hydrogens were not represented.

However, in sharp contrast with the latter MEEVD and IEEVD segments, the DASR and the SGPT extremities in HSP90 and HSP70 exhibited more flexibility along the trajectories (Figure 4, Supplementary Movies MS1 and MS2). The backbone RMSD values of  $1.39 \pm 0.38 \text{ \AA}$  and  $1.22 \pm 0.41 \text{ \AA}$  measured in the DASR and SGPT parts of the HSP90 and HSP70 peptides reflected their higher mobility. In both complexes, the N-terminal part of the peptides (i.e. the first four residues) appeared more solvent-exposed than the C-terminal part (i.e. the five last residues). However, we were able to highlight in the trajectories significant transient salt bridges established between Arg727 in HSP90 and E721 and E726 in SPAG1-TPR3 (Supplementary Figure S7). Similar interactions could not be observed with the HSP70 peptide, the first residues of which lack strong electron attracting groups. This difference could explain the slight lower affinity measured for the HSP70 peptide in complex with SPAG1-TPR3 (Figure 1F). However, the dynamic behavior observed within the N-terminal segment of the peptides suggests that the majority of the binding energy arises from the stable interface involving both polar and hydrophobic anchoring of the (M/I)EEVD segment into the cradle of SPAG1-TPR3.

Last, the evolution of the intermolecular distances agrees nicely with the experimental NOE and CSP data (Figure 5, Supplementary Figure S6), suggesting that our MD simulations rendered a realistic representation of the three-dimensional structure of the complexes at hand.

### **SPAG1 does not bind and hydrolyze GTP efficiently**

Having deciphered the binding mode of HSPs onto TPR domains at the atomic level, we focused next on another putative substrate of SPAG1 with the primary idea of understanding, at the same atomic scale, the nucleotidase function of the protein. Indeed, GTPase activity of SPAG1 was demonstrated by measuring the release of inorganic phosphate from GTP in the presence of *E. coli* extracts overexpressing or not a recombinant fragment of SPAG1, going from residues 577 to 821 [22]. Authors showed an increase in absorbance at 700 nm in the presence of SPAG1, suggesting a direct contribution of the overexpressed protein in GTP hydrolysis. Using GTP and ATP blot assays, SPAG1 was proposed to bind GTP, while protein sequence analysis highlighted

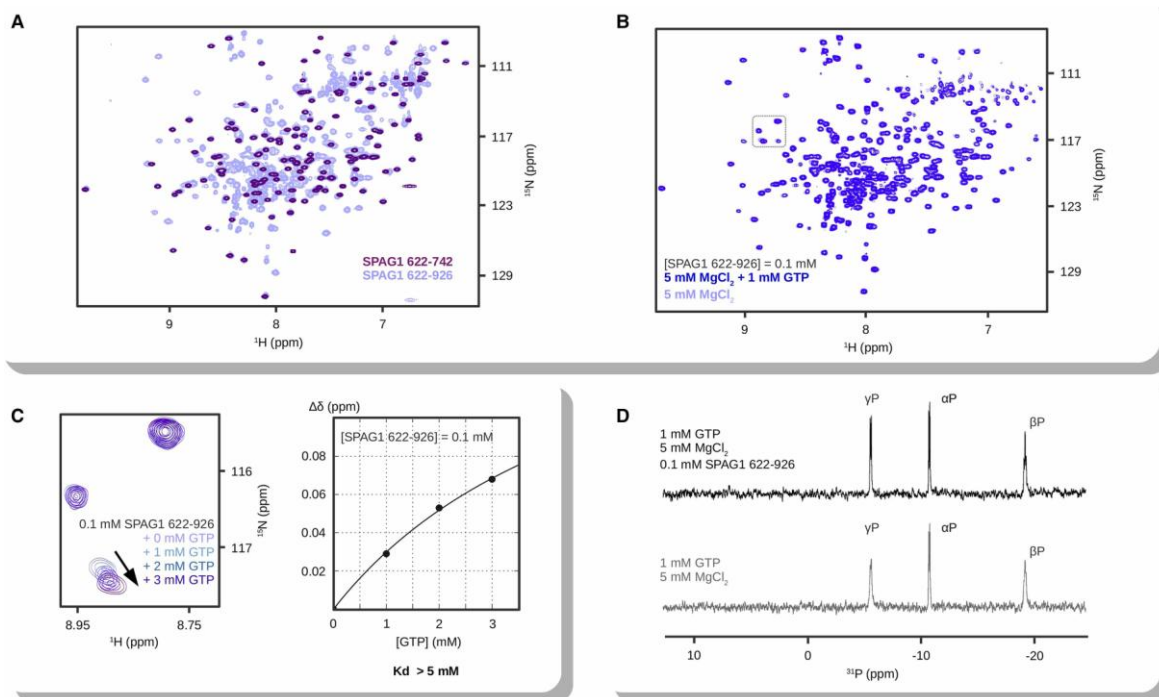
a P-loop motif between residues 781 and 788 (ASEKGGKS). To gain additional information about this function, we aimed to monitor the strength of the binding of GTP to the stable C-terminal fragment 622–926 of SPAG1, using NMR spectroscopy.

First, the HSQC <sup>1</sup>H-<sup>15</sup>N spectrum that we recorded on SPAG1(622–926) overlapped well with the spectrum of SPAG1-TPR3 that corresponds to region 622–742 (Figure 6A). Two sets of supplemental peaks could also be observed in this spectrum. The first one, with NMR signals of moderate intensities spreading over the spectrum widths, displays the features of well-folded protein regions. The second one, with intense <sup>1</sup>H and <sup>15</sup>N resonances centered, respectively, around 8 ppm and 120 ppm, is characteristic of disordered regions in proteins. These spectral features led us to reason that SPAG1(622–926) could behave as a protein made of two independent domains devoid of strong contacts between each other. Due to the multi-modular organization of SPAG1 (Figure 1A, Supplementary Figure S1), we could reasonably assign these peaks to the C-terminal domain of the protein and to the linker upstream that makes the connection with TPR3.

Surprisingly, we did not observe any drastic change in the NMR spectrum of SPAG1(622–926) in the presence of a 10-molar excess of GTP, even when Mg<sup>2+</sup> cations were added (Figure 6B). An efficient binding of the nucleotide should have significantly modified the chemical shift values and/or the intensities of those residues involved in the binding of the nucleotide. Cautious inspection of the SPAG1(622–926) spectra recorded with a huge excess of nucleotide showed, nonetheless, that few isolated NH resonances underwent observable chemical shift modifications upon addition of GTP (Figure 6B,C). These CSPs were derived into a K<sub>d</sub> greater than 5 mM, a value for which the question of biological relevance could be raised. This result strongly suggests that the 622–926 fragment of SPAG1 was not able to bind GTP efficiently in vitro (at least with a sub-millimolar affinity).

As very low binding could not totally be related to low hydrolysis efficiency, we monitored the release of inorganic phosphate by SPAG1(622–926) from GTP in the presence of MgCl<sub>2</sub>, using <sup>31</sup>P NMR. Even after 24 h of incubation, no <sup>31</sup>P peak between 0 and 5 ppm appeared, thus, indicating that hydrolysis of GTP could not be observed (Figure 6D). These data call into question

the efficiency of GTPase activity in SPAG1.



**Figure 6. Insights into the GTPase activity of SPAG1.**

(A) Superimposition of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the  $^{15}\text{N}$ -labelled fragments 622–742 (in purple) and 622–926 (in light blue) of SPAG1 recorded in the NMR buffer (10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP) at 293 K. (B) BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of SPAG1(622–926) recorded in the NMR buffer supplemented with 5 mM  $\text{MgCl}_2$  in the presence (marine blue) or absence (light blue) of 1 mM GTP. (C) Chemical shift perturbations upon GTP addition on SPAG1(622–926) monitored in experimental conditions mentioned in (B) and curve fitting of the data. The spectral region that was studied is highlighted in the gray box drawn in (B). (D) 1D  $^{31}\text{P}$  NMR spectra of 1 mM GTP in the presence (black) or absence (gray) of 0.1 mM  $^{15}\text{N}$ -labelled SPAG1(622–926) recorded at 298 K in 10 mM HEPES (pH 7.5), 150 mM NaCl, 5 mM  $\text{MgCl}_2$  and 10%  $\text{D}_2\text{O}$ . Reference (gray) and test (black) samples were prepared 24 h prior to data acquisition. Assignment of  $^{31}\text{P}$  nuclei in GTP is indicated.

## Discussion

### SPAG1 is a platform for HSP proteins

SPAG1 has been proposed to be involved in the assembly of dynein arms, and thus, in the function of cilia. Quaternary assembly of this kind of macromolecular complex is supervised by proteins with chaperoning activities. We have demonstrated here the ability of SPAG1 to recruit directly HSP70 and HSP90. We have shown that SPAG1 binds the tails of HSP70 and HSP90 with comparable affinities, either for the full-length protein or for the isolated TPR domains. The affinities measured between isolated SPAG1-TPRs 1 and 3 and the DASRMEEVD peptide (HSP90) are slightly higher than the ones measured with the SGPTIEEVD peptide (HSP70), thereby indicating that none of the TPR domains seems dedicated to one specific chaperone.

This idea is supported by our synergistic NMR and MD study of SPAG1-TPR3 and HSP peptides, illuminating that DASRMEEVD and SGPTIEEVD peptides can accommodate their protein receptor in a similar way. Specifically, on one side, the EEVD residues of the peptide are consensually bound to the carboxylate-clamp motif of the TPR domain and, on the other side, the Met728 and Ile642 residues (which were proposed as key residues for HSP specificity) are stably hosted in the same pocket in SPAG1-TPR3. This observation is in contrast with multi-TPR HSP co-chaperone proteins such as RPAP3 or HOP, the TPR domains of which have been proposed devoted to HSP70 or HSP90 in order to facilitate protein–client exchange inside a chaperone/ co-chaperone complex [19]. For example, we have previously documented two different binding modes for HSP70 and HSP90 peptides on RPAP3-TPR2 and related these results to an enhanced affinity of RPAP3-TPR2 for HSP90. Here, our results indicate that SPAG1 could be potentially saturated by one or the other HSP proteins. It is noteworthy that previous IP-SILAC experiments using PIH1D2 as bait have revealed that HSP70 was strongly associated with SPAG1, RUVBL1, RUVBL2 and PIH1D2 proteins [5], suggesting that HSP70 could be the privileged partner of SPAG1 in vivo. This result also raises the question of the putative regulating effect of SPAG1 on HSP chaperones. Indeed, one might wonder whether SPAG1 is a simple carrier for HSP proteins or is able to enhance or decrease the chaperoning abilities of the latter. To address this question, it would be of interest to test the effect of the full-length SPAG1 and its sub-fragments on the ATPase activity of HSPs. We propose in the present work that SPAG1-TPR2 could not participate in the binding of the HSP C-terminal tails due to its biochemical features. Still, one might question whether TPR2 only plays an organizational role in SPAG1, facilitating the proper folding of the protein, or act as a regulatory module that modifies the ATPase activity of the HSP chaperones through contacts that do not involve its carboxylate-clamp motif. Such regulatory roles were highlighted for the TPR domain of Sec72 that is able to interact with the ATPase domain of Ssb1, a yeast HSP70 homolog lacking the EEVD motif [52] and for Sti1, whose TPR2B domain interacts with the M domain of HSP90 [19]. In the same way, the formation of HSP70 dimers, and consequently the chaperoning activity of the protein, is also differentially regulated by TPR-containing proteins, Chip and Tomm34 [53].



## **Is SPAG1 a GTPase?**

Using the sub-fragment 577–821, SPAG1 was described years ago as a GTPase protein [22]. In our study, we failed to highlight catalytic activity for the sub-fragment 622–926 of SPAG1. Although different in size, these two sub-fragments equally contain a putative P-loop motif that lays between residues 781 and 788. Using NMR, we have demonstrated the stability and the proper folding of sub-fragment 622–926 (Figure 6A,B). In addition, sequence analysis of SPAG1 defines the region 577–621 in SPAG1 as a linker connecting TPR domains 2 and 3 devoided of any specific motif, whereas region 822–926 appears folded (Figure 6A, Supplementary Figure S1). Thus, although the use of the sub-fragment 622–926 appears totally relevant, the P-loop motif remained silent in our experiments. Moreover, the SPAG1 sequence did not reveal the presence of a G domain, which is a conserved domain in GTPases.

However, among partners, SPAG1 was also able to interact with RUVBL1 and RUVBL2 proteins, two other protein chaperones both shown to be able to hydrolyze ATP. To fulfill this function, RUVBLs proteins contain Walker A and B motifs, as well as ADP/ATP sensor motifs. The Walker A/P-loop sequence is a common motif for ATP and GTP binding. Merging these data together leads us to propose that SPAG1 might have a general enhancing effect on nucleotide hydrolysis. Indeed, we could easily assume that SPAG1 might have stimulated GTPases from *E. coli* in the past experiment [22]. In turn, this hypothesis suggests that SPAG1 could be a key activator of Walker nucleotidases, such as RUVBL1/2 proteins, which could facilitate their role in the quaternary assembly of macromolecular complexes. Interestingly enough, the C-terminal domain of SPAG1 is homologous to the C-terminal domain of RPAP3, which has been shown to be in direct contact with the RUVBL1/2 hetero-hexameric ring. Thus, this domain could be an important module in the regulation of the nucleotide hydrolysis function. As a perspective, extensive activity assays might highlight that SPAG1, through its multiple domains, is a fine tuner of the nucleotide hydrolysis activity of non-Walker and Walker nucleotidases (i.e. HSP and RUVBL proteins, respectively) within a large R2SP/HSP complex.

## **Abbreviations**

cryo-EM, electronic cryo-microscopy; CS, Chord and Sgt1; CSP, chemical shift perturbation; DNAAF, dynein axonemal assembly factors; FCC, Fraction of Common Contacts; HSP, heat shock protein; HSQC, heteronuclear single quantum coherence; IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside; ITC, isothermal titration calorimetry; Kd, dissociation constant; MD, molecular dynamics; NMR, nuclear magnetic resonance; PCD, primary ciliary dyskinesia; PIH, protein interacting with HSP; TPR, tetratricopeptide repeat.

## **Data Depositions**

Chemical shifts and 3D coordinates of the free form of SPAG1-TPR3 were deposited in the Biological Magnetic Resonance Data Bank and in the Protein Data Bank under respective accession numbers 34329 and 6I57.

## **Author Contribution**

M.Q. conceived and coordinated the project. M.E.C, S.D. and X.M. were involved in the DNA cloning steps. M.E.C., R.D.S.M., D.L. and M.Q. performed and analyzed biochemistry and biophysics experiments. C.C., F.D. and M.Q. performed and analyzed in silico experiments. M.Q. wrote the original draft. All authors reviewed the manuscript.

## **Funding**

This work was supported by the French National Center for Scientific Research (CNRS), University of Lorraine (France), the 'Région Grand Est' and the 'Agence Nationale de la Recherche' [Grant No. ANR-16-CE11-0032-02].

C.C and F.D. acknowledge the State-Region Plan 'Technological Innovations, Modeling and Personalized Medical Support' (IT2MP) and the European Regional Development Funds (ERDF)

for generous support.

## **Acknowledgements**

We thank IBSLor (UMS 2008 CNRS, INSERM, Université de Lorraine) for ITC, DLS, SEC-SLS and NMR core facilities. We thank Institut Jean Barriol (IJB, FR2843 CNRS) for 31P NMR core facilities.

## **Competing Interests**

The Authors declare that there are no competing interests associated with the manuscript.

## **References**

- Heydeck, W., Fievet, L., Davis, E.E. and Katsanis, N. (2018) The complexity of the cilium: spatiotemporal diversity of an ancient organelle. *Curr. Opin. Cell Biol.* 55, 139–149 <https://doi.org/10.1016/j.ceb.2018.08.001>
- Wang, L. and Dynlacht, B.D. (2018) The regulation of cilium assembly and disassembly in development and disease. *Development* 145, dev151407 <https://doi.org/10.1242/dev.151407>
- Mitchison, H.M., Schmidts, M., Loges, N.T., Freshour, J., Dritsoula, A., Hirst, R.A. et al. (2012) Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat. Genet.* 44, 381–389. S1–S2 <https://doi.org/10.1038/ng.1106>
- Knowles, M.R., Ostrowski, L.E., Loges, N.T., Hurd, T., Leigh, M.W., Huang, L. et al. (2013) Mutations in SPAG1 cause primary ciliary dyskinesia associated with defective outer and inner dynein arms. *Am. J. Human Genet.* 93, 711–720 <https://doi.org/10.1016/j.ajhg.2013.07.025>
- Maurizy, C., Quinternet, M., Abel, Y., Verheggen, C., Santo, P.E., Bourguet, M. et al. (2018) The RPAP3-Cterminal domain identifies R2TP-like quaternary chaperones. *Nat. Commun.* 9, 2093 <https://doi.org/10.1038/s41467-018-04431-1>
- Dafinger, C., Rinschen, M.M., Borgal, L., Ehrenberg, C., Basten, S.G., Franke, M. et al. (2018) Targeted deletion of the AAA-ATPase Ruvbl1 in mice disrupts ciliary integrity and causes renal

disease and hydrocephalus. *Exp. Mol. Med.* 50, 75 <https://doi.org/10.1038/s12276-018-0108-z>

Kakihara, Y. and Houry, W.A. (2012) The R2TP complex: discovery and functions. *Biochim. Biophys. Acta* 1823, 101–107 <https://doi.org/10.1016/j.bbamcr.2011.08.016>

Houry, W.A., Bertrand, E. and Coulombe, B. (2018) The PAQosome, an R2TP-based chaperone for quaternary structure formation. *Trends Biochem. Sci.* 43, 4–9 <https://doi.org/10.1016/j.tibs.2017.11.001>

Lage P, Z., Stefanopoulou, P., Styczynska-Soczka, K., Quinn, N., Mali, G., von Kriegsheim, A. et al. (2018) Ciliary dynein motor preassembly is regulated by Wdr92 in association with HSP90 co-chaperone, R2TP. *J. Cell Biol.* 217, 2583–2598 <https://doi.org/10.1083/jcb.201709026>

Liu, G., Wang, L. and Pan, J. (2018) Chlamydomonas WDR92 in association with R2TP-like complex and multiple DNAAFs to regulate ciliary dynein preassembly. *J. Mol. Cell Biol.* <https://doi.org/10.1093/jmcb/mjy067>

Martino, F., Pal, M., Muñoz-Hernández, H., Rodríguez, C.F., Núñez-Ramírez, R., Gil-Carton, D. et al. (2018) RPAP3 provides a flexible scaffold for coupling HSP90 to the human R2TP co-chaperone complex. *Nat. Commun.* 9, 1501 <https://doi.org/10.1038/s41467-018-03942-1>

Rivera-Calzada, A., Pal, M., Muñoz-Hernández, H., Luque-Ortega, J.R., Gil-Carton, D., Degliesposti, G. et al. (2017) The structure of the R2TP complex defines a platform for recruiting diverse client proteins to the HSP90 molecular chaperone system. *Structure* 25, 1145–1152.e4 <https://doi.org/10.1016/j.str.2017.05.016>

Pal, M., Morgan, M., Phelps, S.E., Roe, S.M., Parry-Morris, S., Downs, J.A. et al. (2014) Structural basis for phosphorylation-dependent recruitment of Tel2 to Hsp90 by Pih1. *Structure* 22, 805–818 <https://doi.org/10.1016/j.str.2014.04.001>

Henri, J., Chagot, M.E., Bourguet, M., Abel, Y., Terral, G., Maurizy, C. et al. (2018) Deep structural analysis of RPAP3 and PIH1D1, two components of the HSP90 co-chaperone R2TP complex. *Structure* 26, 1196–1209.e8 <https://doi.org/10.1016/j.str.2018.06.002>

Quinternet, M., Rothé, B., Barbier, M., Bobo, C., Saliou, J.M., Jacquemin, C. et al. (2015) Structure/function analysis of protein-protein interactions developed by the yeast Pih1 platform protein and its partners in box C/D snoRNP assembly. *J. Mol. Biol.* 427, 2816–2839

<https://doi.org/10.1016/j.jmb.2015.07.012>

Tian, S., Yu, G., He, H., Zhao, Y., Liu, P., Marshall, A.G. et al. (2017) Pih1p-Tah1p puts a lid on hexameric AAA+ ATPases Rvb1/2p. *Structure* 25, 1519–1529.e4 <https://doi.org/10.1016/j.str.2017.08.002>

Blatch, G.L. and Lässle, M. (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 21, 932–939 [https://doi.org/10.1002/\(SICI\)1521-1878\(199911\)21:11<932::AID-BIES5>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N)

Chen, S. and Smith, D.F. (1998) Hop as an adaptor in the heat shock protein 70 (Hsp70) and hsp90 chaperone machinery. *J. Biol. Chem.* 273, 35194–35200 <https://doi.org/10.1074/jbc.273.52.35194>

Schmid, A.B., Lagleder, S., Gräwert, M.A., Röhl, A., Hagn, F., Wandinger, S.K. et al. (2012) The architecture of functional modules in the Hsp90 co-chaperone Sti1/Hop. *EMBO J.* 31, 1506–1517 <https://doi.org/10.1038/emboj.2011.472>

Kryzstofinska, E.M., Evans, N.J., Thapaliya, A., Murray, J.W., Morgan, R.M.L., Martinez-Lumbreras, S. et al. (2017) Structure and interactions of the TPR domain of Sgt2 with yeast chaperones and Ybr137wp. *Front. Mol. Biosci.* 4, 68 <https://doi.org/10.3389/fmolb.2017.00068>

Benbahouche Nel, H., Iliopoulos, I., Török, I., Marhold, J., Henri, J., Kajava, A.V. et al. (2014) *Drosophila* Spag is the homolog of RNA polymerase II-associated protein 3 (RPAP3) and recruits the heat shock proteins 70 and 90 (Hsp70 and Hsp90) during the assembly of cellular machineries. *J. Biol. Chem.* 289, 6236–6247 <https://doi.org/10.1074/jbc.M113.499608>

Lin, W., Zhou, X., Zhang, M., Li, Y., Miao, S., Wang, L. et al. (2001) Expression and function of the HSD-3.8 gene encoding a testis-specific protein.

*Mol. Hum. Reprod.* 7, 811–818 <https://doi.org/10.1093/molehr/7.9.811>

Liu, N., Qiao, Y., Cai, C., Lin, W., Zhang, J., Miao, S. et al. (2006) A sperm component, HSD-3.8 (SPAG1), interacts with G-protein  $\beta$ 1 subunit and activates extracellular signal-regulated kinases (ERK). *Front. Biosci.* 11, 1679–1689 <https://doi.org/10.2741/1913>

Huang, C., Wu, D., Khan, F.A., Jiao, X., Guan, K. and Huo, L. (2016) The GTPase SPAG-1 orchestrates meiotic program by dictating meiotic resumption and cytoskeleton architecture in

mouse oocytes. *Mol. Biol. Cell* 27, 1776–1785 <https://doi.org/10.1091/mbc.e16-02-0132>

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) Distantly related sequences in the  $\alpha$ - and  $\beta$ -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945–951 <https://doi.org/10.1002/j.1460-2075.1982.tb01276.x>

Ewens, C.A., Su, M., Zhao, L., Nano, N., Houry, W.A. and Southworth, D.R. (2016) Architecture and nucleotide-dependent conformational changes of the Rvb1-Rvb2 AAA+ complex revealed by cryoelectron microscopy. *Structure* 24, 657–666 <https://doi.org/10.1016/j.str.2016.03.018>

Diebold, M.L., Fribourg, S., Koch, M., Metzger, T. and Romier, C. (2011) Deciphering correct strategies for multiprotein complex assembly by co-expression: Application to complexes as large as the histone octamer. *J. Struct. Biol.* 175, 178–188 <https://doi.org/10.1016/j.jsb.2011.02.001>

Kriznik, A., Yelehe-Okouma, M., Lec, J.C., Groshenry, G., Le Cordier, H., Charron, C. et al. (2018) CRDSAT generated by pCARGHO: a new efficient lectin-based affinity tag method for safe, simple, and low-cost protein purification. *Biotechnol. J.* 14, e1800214 <https://doi.org/10.1002/biot.201800214>

Anthis, N.J. and Clore, G.M. (2013) Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci.* 22, 851–858 <https://doi.org/10.1002/pro.2253>

Shen, Y. and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks.

*J. Biomol. NMR* 56, 227–241 <https://doi.org/10.1007/s10858-013-9741-y>

Berjanskii, M.V., Neal, S. and Wishart, D.S. (2006) PREDITOR: a web server for predicting protein torsion angle restraints. *Nucleic Acids Res.* 34, W63–W69 <https://doi.org/10.1093/nar/gkl341>

Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* 278, 353–378 <https://doi.org/10.1385/1-59259-809-9:353>

Nederveen, A.J., Doreleijers, J.F., Vranken, W., Miller, Z., Spronk, C.A., Nabuurs, S.B. et al. (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* 59, 662–672 <https://doi.org/10.1002/prot.20408>

de Vries, S.J., van Dijk, A.D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V. et al. (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* 69, 726–733 <https://doi.org/10.1002/prot.21723>

Gazda, L., Pokrzywa, W., Hellerschmied, D., Löwe, T., Forné, I., Mueller-Planitz, F. et al. (2013) The myosin chaperone UNC-45 is organized in tandem modules to support myofilament formation in *C. elegans*. *Cell* 152, 183–195 <https://doi.org/10.1016/j.cell.2012.12.025>

Wang, L., Liu, Y.T., Hao, R., Chen, L., Chang, Z., Wang, H.R. et al. (2011) Molecular mechanism of the negative regulation of Smad1/5 protein by carboxyl terminus of Hsc70-interacting protein (CHIP). *J. Biol. Chem.* 286, 15883–15894 <https://doi.org/10.1074/jbc.M110.201814>

Rodrigues, J.P., Trellet, M., Schmitz, C., Kastiris, P., Karaca, E., Melquiond, A.S. et al. (2012) Clustering biomolecular complexes by residue contacts similarity. *Proteins* 80, 1810–1817 <https://doi.org/10.1002/prot.24078>

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J. et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102, 3586–3616 <https://doi.org/10.1021/jp973084f>

MacKerell, Jr, A.D., Feig, M. and Brooks, 3rd, C.L. (2004) Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* 126, 698–699 <https://doi.org/10.1021/ja036959e>

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935 <https://doi.org/10.1063/1.445869>

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E. et al. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802 <https://doi.org/10.1002/jcc.20289>

Tuckerman, M., Berne, B.J. and Martyna, G.J. (1992) Reversible multiple time scale molecular

dynamics. *J. Chem. Phys.* 97, 1990–2001 <https://doi.org/10.1063/1.463137>

Andersen, H.C. (1983) Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* 52, 24–34 [https://doi.org/10.1016/0021-9991\(83\)90014-1](https://doi.org/10.1016/0021-9991(83)90014-1)

Martyna, G.J., Tobias, D.J. and Klein, M.L. (1994) Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101, 4177–4189 <https://doi.org/10.1063/1.467468>

Feller, S.E., Zhang, Y., Pastor, R.W. and Brooks, B.R. (1995) Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103, 4613–4621 <https://doi.org/10.1063/1.470648>

Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092 <https://doi.org/10.1063/1.464397>

Fiorin, G., Klein, M.L. and Hénin, J. (2013) Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* 111, 3345–3362 <https://doi.org/10.1080/00268976.2013.813594>

Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. 27–28 [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)

Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H. et al. (2000) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* 101, 199–210 [https://doi.org/10.1016/S0092-8674\(00\)80830-2](https://doi.org/10.1016/S0092-8674(00)80830-2)

Dominguez, C., Boelens, R. and Bonvin, A.M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125, 1731–1737 <https://doi.org/10.1021/ja026939x>

Blundell, K.L., Pal, M., Roe, S.M., Pearl, L.H. and Prodromou, C. (2017) The structure of FKBP38 in complex with the MEEVD tetratricopeptide binding-motif of Hsp90. *PLoS ONE* 12, e0173543 <https://doi.org/10.1371/journal.pone.0173543>

Tripathi, A., Mandon, E.C., Gilmore, R. and Rapoport, T.A. (2017) Two alternative binding mechanisms connect the protein translocation Sec71-Sec72 complex with heat shock proteins. *J. Biol. Chem.* 292, 8007–8018 <https://doi.org/10.1074/jbc.M116.761122>



Trcka, F., Durech, M., Vankova, P., Chmelik, J., Martinkova, V., Hausner, J., et al. (2019) Human stress-inducible Hsp70 has a high propensity to form ATP-dependent antiparallel dimers that are differentially regulated by cochaperone binding. *Mol. Cell. Proteomics* 18, 320–337 <https://doi.org/10.1074/mcp.RA118.001044>