Structural and dynamical basis for the interaction of HSP70-EEVD with JDP Sis1

Carolina O. Matos[1], Glaucia M.S. Pinheiro[1], Icaro P. Caruso[2], Gisele C. Amorim[3,4], Fabio C. L. Almeida[3,5,6*] and Carlos H. I. Ramos[1,6*]

[1]Institute of Chemistry, University of Campinas UNICAMP, Campinas SP, Brazil.

[2]Multiuser Center for Biomolecular Innovation (CMIB), Department of Physics, São Paulo State University (UNESP), São Jose do Rio Preto, São Paulo, Brazil.

[3]National Center of Nuclear Magnetic Resonance (CNRMN), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil;

[4]Multidisciplinary Center for Research in Biology (NUMPEX), Campus Duque de Caxias Federal University of Rio de Janeiro, Duque de Caxias, Brazil.

[5]National Institute of Science and Technology for Bioimage and Structural Biology INBEB, Brazil.

[6]Institute of Medical Biochemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil,

***** Correspondence should be addressed to FCLA, falmeida@bioqmed.ufrj.br and CHR, cramos@unicamp.br

Keywords: HSP40; Sis1; J-domain; protein protein-interaction; HSP70; NMR

Abbreviations: CSP, chemical shift perturbation; CTD, C-Terminal Domain; HSP, heat shock protein; HSP70, 70 kDa heat shock protein; IDR, intrinsically disordered region; JDP, J-domain protein; NBD, nucleotide-binding domain; R_1, longitudinal relaxation rate; PRE, Paramagnetic relaxation enhancement; R_2, transverse relaxation rate; SBD, substrate binding domain; RMSD, root-mean-square deviation; Sis1, type B JDP from *Saccharomyces cerevisiae*; Ssa1, HSP70 from *Saccharomyces cerevisiae*;

Abstract

Protein folding, refolding and disaggregation are facilitated by the binding and releasing activity of HSP70 proteins. Such activity is aided by J-domain proteins (JDPs, DNAJs or HSP40s) that stimulate the ATPase activity of HSP70 and stabilize complexes between HSP70 and nonnative proteins. The C-terminus EEVD motif of HSP70 interacts with client proteins and with JDPs. Deletion of the EEVD disrupts the ability of HSP70 to associate with class B JDPs. A new understanding of the EEVD interaction with a JDP arises from the results of this work, which gives a detailed NMR characterization of the dynamics and interaction between a class B JDP, Sis1, and the HSP70-EEVD. The EEVD motif binds to multiple sites in Sis1. The interactions at CTDI contribute to the anchoring of HSP70 to Sis1, at site I, and the displacement of the client protein at site II. The competing interaction between the J-domain and the N-terminal boundary of α-helix 6 (at the GF-region) contributes to the interaction of J-domain to HSP70. These results support a mechanism involved in the regulation of the interaction with HSP70 that is autoinhibitory. Finally, a detailed model of the interaction of the EEVD with the J-domain of Sis1 is proposed.

Introduction

Chaperones of the 70 kDa heat shock protein (HSP70) superfamily are key components of the cellular proteostasis system which are composed of a nucleotide-binding domain (NBD), a substrate binding domain (SBD), and a C-terminal disordered region ending in an EEVD motif (CTD)[1–3]. The SBD is composed of the SBDβ subdomain, also known as the base subdomain, and the SBDα subdomain, the lid subdomain. The CTD is involved in the intramolecular regulation of HSP70 function and intermolecular interactions with J-domain proteins (JDPs, also known as HSP40s or DNAJs), a family of cochaperones involved in the delivery of client-proteins. JDPs bind to the interface between the NBD and the SBD of HSP70 and it is necessary for stimulation of its ATPase activity[4,5].

JDPs are classified into three classes, which are all characterized by the presence of a J-domain, N-terminal in classes A and B, and anywhere in class C. A disordered region follows the J-domain, which is composed of a glycine-phenylalanine-rich (GF) in class A and a GF region followed by a glycine/methionine-rich (GM) in class B. Both classes A and B have homologous C-terminal β-barrel domains (CTDI and CTDII) and a dimerization domain (DD). The CTDI of class A contains a zinc-finger-like region (ZFLR) which is not present in class B (Fig. 1)[6–9].

Two cytosolic and dimeric JDPs from *Saccharomyces cerevisiae*, the class A Ydj1 and the class B Sis1, stimulate the ATPase activity of HSP70 and its function[10]. Remarkably, *sis1* is an essential gene while *ydj1* is not. Cells knocked out of the *ydj1* gene show a phenotype of severe growth defect and a decrease in stress tolerance, while the overexpression of Sis1 can partially suppress its slow growth phenotype. Nevertheless, overexpression of Ydj1 does not restore the viability of cells knocked out of the *sis1* gene[11,12]. It is well-established that the EEVD motif of yeast HSP70/Ssa1 can bind to Sis1 by electrostatic interactions, which is functionally important for its chaperone activity[13]. Deletion of the EEVD ($HSP70_{\Delta EEVD}$) disrupts the ability of Sis1 to associate with HSP70 but does not affect the interaction between Ydj1 and HSP70. Substitution of the J-domain of Ydj1 or Xdj1 (a paralog of Ydj1) for that of Sis1, results in the modified Sis1 associating with $HSP70_{\Delta EEVD}$. Sis1-EEVD interaction is also important for Ssa1 (*S. cerevisiae* HSP70) recruitment to bind in aggregated substrates[14].

An autoinhibitory mechanism for HSP70 binding was shown for the class B JDPs DNAJB1 and DNAJB6[15,16], in which the binding site located at the J-domain is occluded

by a conserved α-helix 5 at the GF region (Fig. 1c). The occlusion can be reverted by the interaction with the HSP70-EEVD motif. This self-regulation is unique for class B JDPs and is essential for the disaggregation of amyloid fibers by HSP70/DNAJB1[15]. Our group recently showed the key importance of the transient interaction of Sis1 to HSP70[17]. Sis1 J-domain binds to HSP70 through a hydrophobic and positively charged patch mainly at α-helices 2 and 3 with the participation of the conserved HPD motif, located between these helices, and is released by the addition of ATP[17]. In the context of the full-length Sis1 the J-domain is in a transient intermediate conformation, in which the HSP70-interacting patch is protected by internal transient interactions[17].

Structural information on the interaction between JDPs and the EEVD-motif is still scarce, and there are many open questions about the biological effects caused by the association of these chaperones. A better understanding of the EEVD interaction with JDP arises from the results of this work, which gives a detailed NMR characterization of the dynamics and interaction between full-length Sis1 (residues 1-352, herein referred to as $Sis1_{1\text{-}352}$) and the HSP70-EEVD motif. Characterization of the dynamics and interaction between Sis1 J-domain (residues 1-81, herein referred to as $Sis1_{1\text{-}81}$) and the HSP70-EEVD motif was also performed to reveal more details about the interaction.

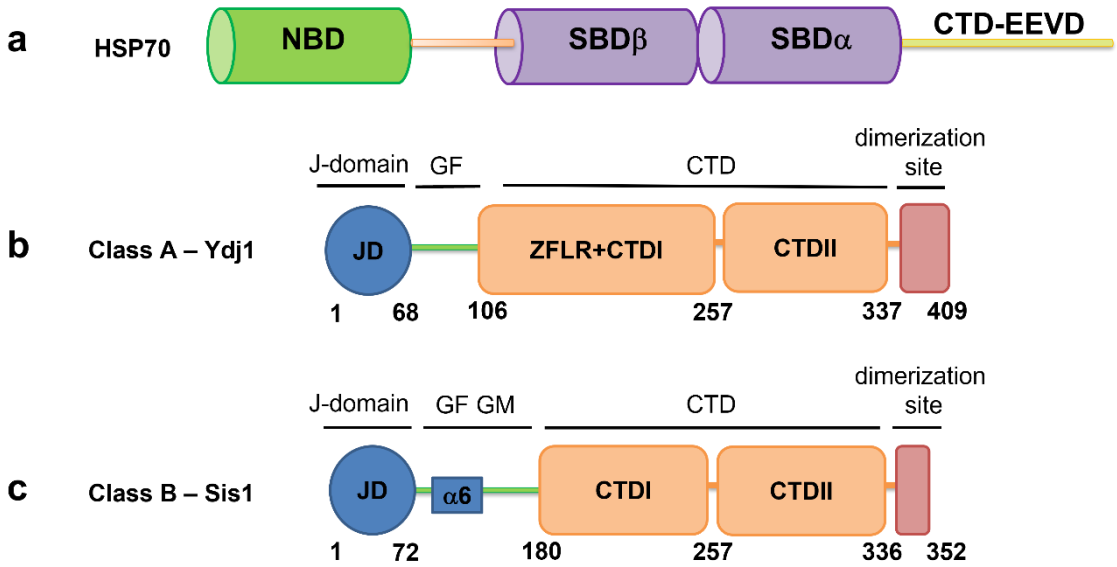

Fig. 1 | Domain architecture of HSP70 and JDPs. a Schematic representation of HSP70 chaperones, composed of a nucleotide-binding domain (NBD, which binds ATP) in green, substrate binding domain (SBD) in purple, divided into SBDβ (base) and SBDα (lid) subdomains, that are connected by a flexible linker (orange) and a C-terminal disordered region terminating in a high conserved EEVD motif (yellow). **b** Schematic representation of the domain organization of Ydj1, class A, and **c** Sis1, class, B from yeast. The different domains are marked as follows:

JD: J-domain (blue, contains α-helix 1-5); GF: Gly-Phe rich region and GM: Gly-Met rich region (green). The GF region of Sis1contains the α-helix 6, which is conserved in class B JDPs and is described as α-helix 5 for DNAJB1[15] and others; ZFLR: zinc-finger like region; CTD: C-terminal domain (orange) and dimerization domain (DD, red).

Results

The interdomain dynamics of Sis1$_{1\text{-}352}$ is affected by EEVD

The understanding of the interaction between the HSP70 C-terminus EEVD motif with the cochaperone Sis1 is important because the deletion of this motif abolishes the ability of Sis1 to partner with HSP70 and pursue its chaperone activity[14]. We first studied the dynamics of free Sis1 (full-length protein, Sis1$_{1\text{-}352}$) and the changes elicited by the interaction of a GPTIEEVD peptide (representing the EEVD motif). The ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates of Sis1$_{1\text{-}352}$ free and bound to the EEVD peptide were measured. Table 1 describes the ^{15}N R_2/R_1 ratio and the apparent rotational diffusion (τ_c^{app}) of each domain in the context of the full-length protein[18]. τ_c^{app} reflects the rotational freedom of each domain in the free and EEVD-bound states and was calculated by filtering out residues in conformational exchange or thermal flexibility using the average ^{15}N R_2/R_1 of each region considering only the values within the average plus or minus one standard deviation (sd).

The relaxation parameters for free Sis1$_{1\text{-}352}$ (Fig. 2A, Table 1) highlighted the differences in dynamics between its extended J-domain (1-72), GF (73-121), GM (122-178), CTDI (179-257), CTD II (258-335) and dimerization (DD, 336-352) domains. We measured a significant difference of ^{15}N R_2/R_1 for the J-domain (22 ± 2), GF/GM (30 ± 13) and CTD (70 ± 8), indicating that the domains located at the C-terminus, CTD and DD, have the most restricted rotational freedom. The calculated apparent rotational correlation time (τ_c^{app}) showed that all domains had more rotational freedom than expected for a rigid globular 75.2 kDa protein, considering the scissor-like anisotropic structure of Sis1. The CTD showed the most restricted domain motion (τ_c^{app} = 18 ± 2 ns), possibly due to its proximity to the dimerization domain. The J-domain and the GF/GM regions behaved as a highly flexible arm with τ_c^{app} of 9.8 ± 1.0 ns and 11.0 ± 5.0 ns, respectively. Note that the GF/GM region (sd=5) presents a higher dispersion of ^{15}N R_2/R_1 values when compared to the J domain (sd = 1) or CTD (sd = 2).

The chemical shift-based order parameter (S^2) analysis, for $Sis1_{1-352}$ (Fig. 2B) and $Sis1_{1-180}$ (Fig. 2D), was obtained using TALOS-N[19,20] and was informative on the structural order of each of the regions/domains of the protein. We observed high order for the well-folded J-domain, CTDI, CTDII and dimerization domain (DD). On the other hand, GF/GM showed high backbone flexibility, typical of an IDR, except for the residues 104 to 112, which are highly ordered. The region 104-112 contains the N-terminal portion of α-helix 6, which spans from 107 to 119 (Fig. 2E). Although the higher order for the helical segment comes as no surprise, remarkably, there was a mismatch between the ordered region within the GF (104-112, Fig 1D) and α-helix 6 (107-119, Fig 1E).

The orthologous human DNAJB1 also shows an α-helix (α-helix 5) within the GF region that blocks the J-domain interacting site to HSP70, as an autoinhibitory mechanism[15]. Sis1 structural predictions made by Alphafold[21,22] also suggest the interaction of α-helix 6 with the J-domain, indicating that Sis1 is also regulated by an autoinhibitory mechanism. The observed order in this region (residues 104-112, Fig. 2B) supports this interaction, nevertheless, the observed mismatch between the ordered (104-112) and helical residues (107-119), suggests an on-off equilibrium between a J-domain-associated (on-state) and a free, or J-domain-dissociated (off-state) α-helix 6. In this situation, the assigned chemical shift in solution reflects the average between the J-associated and J-dissociated α-helix 6, similar to what was observed for DNAJB1[15] and DNAJB6[16]. Furthermore, the lower structural order observed for the C-terminal portion of the α-helix 6 (which is positioned between residues 113-119) could be attributed to a partial or total unfolding of the helix in the off-state. This agrees with secondary structure analyses by Jpred4[23], which does not predict a well-folded α-helix 6. Jpred4 informs on secondary structure tendency, based solely on the sequence, meaning that the off state would be disordered. Another evidence of an on-off equilibrium is that residues S78, G102, A105, F106 and S107, which are in the GF region and N-terminal to α-helix 6, are in conformational exchange, indicated by their high ^{15}N R_2/R_1 (Fig. 2C).

We evaluated the effect of the EEVD-peptide binding on the dynamics of $Sis1_{1-352}$ (Fig. 2). The presence of the EEVD led to a subtle but significant increase in the rotational freedom of the J-domain. τ_c^{app} decreased from 9.8 ± 1.0 ns to 9.2 ± 1.0 ns, which is statistically significant ($P < 0.05$, Fig. 2F, Table 1), given the high number of measurements (35 measurements of ^{15}N R_2/R_1 in the absence and 37 in the presence of

the EEVD). Note that most of the ^{15}N R_2/R_1 in the presence of EEVD had a decrease in value in Fig. 2F.

The decrease in ^{15}N R_2/R_1 for the residues S78, G102, A105, F106 and S107 (Figs. 2A and 2C) suggested that the on-off equilibrium of α-helix 6 was suppressed by the presence of the EEVD. We also observed a significant decrease in ^{15}N R_2/R_1 for the residues G141, G166, S169 (Figs. 2A and 2C), suggesting a quenching in the conformational exchange of residues that are at the GM region. Residue S172 was the only one with increased ^{15}N R_2/R_1 elicited by the binding of EEVD. However, a significant change in τ_c^{app} of the CTD was not observed when EEVD was present. In conclusion, altogether, the results indicated significant changes in the dynamics of *Sis1*$_{1\text{-}352}$ caused by the binding of EEVD.

Table 1. ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates (R_2/R_1) and the apparent rotational diffusion (τ_c^{app}) of free and EEVD-bound Sis1$_{1\text{-}352}$

Domain (free or bound)	^{15}N-R_2/R_1	τ_c^{app}, s
J-domain (free)	22 ± 2	9.8 ± 1.0 (x10^{-9})*
J-domain (bound)	20 ± 2	9.2 ± 1.0 (x10^{-9})*
GF/GM region (free)	30 ± 13	11 ± 5 (x10^{-9})
GF/GM region (bound)	22 ± 8	10 ± 4 (x10^{-9})
CTD (free)	70 ± 8	18 ± 2 (x10^{-9})
CTD (bound)	71 ± 8	18 ± 2 (x10^{-9})
* τ_c^{app} of the bound domain is lower than that of a free domain at significance level $P < 0.05$)		

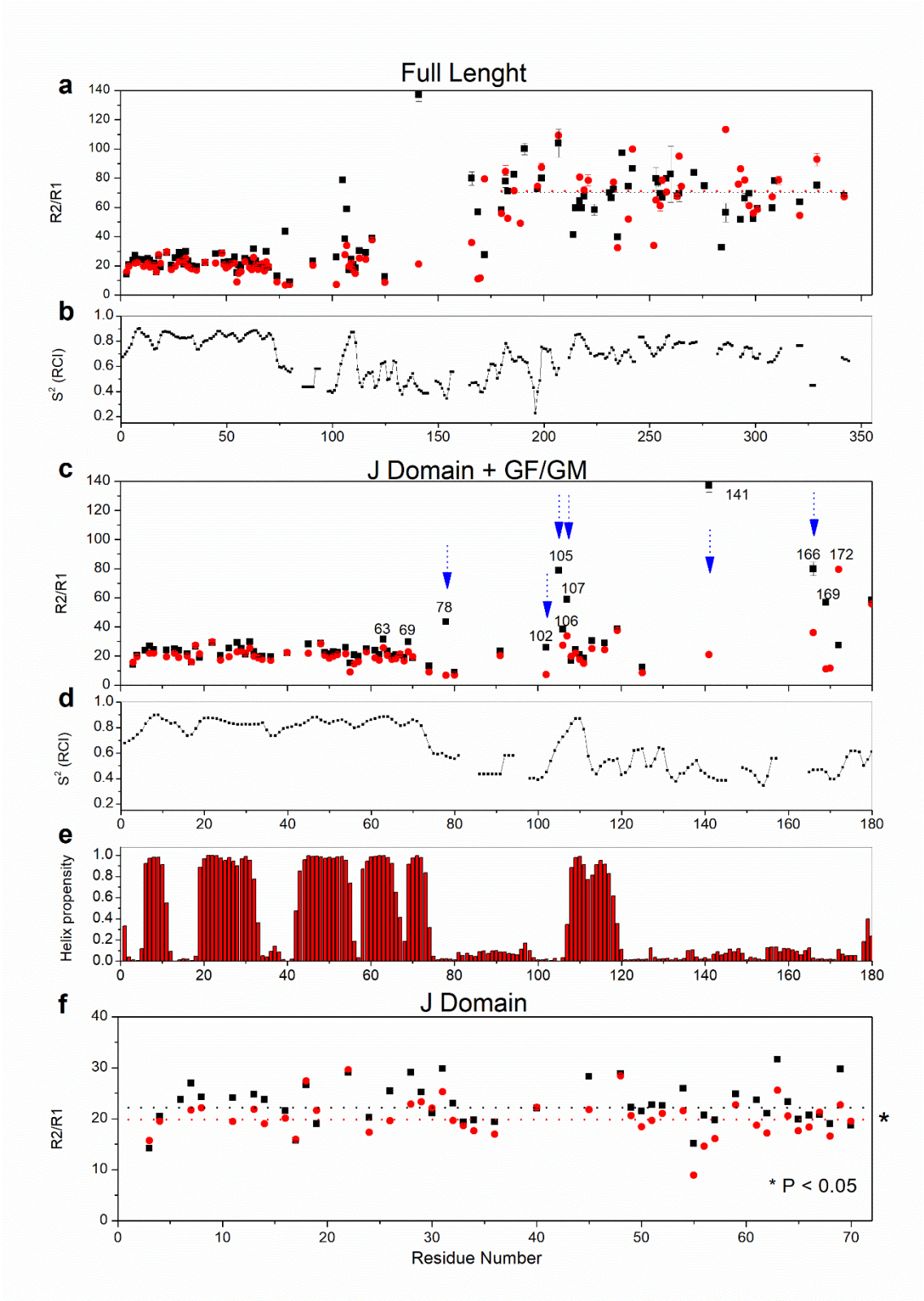

Fig. 2 | Relaxation parameters, chemical shift-based order parameter and Talos-N secondary structure of Sis1$_{1\text{-}352}$. a ^{15}N-R$_2$/R$_1$ for Sis1$_{1\text{-}352}$ (full-length) free (black squares) and bound to EEVD (red circles) for each residue. The dotted lines show the average of ^{15}N-R$_2$/R$_1$ for the CTDI, CTDII and dimerization domain. **b** Chemical shift-based order parameter (S^2) for residues of free Sis1$_{1\text{-}352}$. The lower the value of S^2 the higher is the internal flexibility. **c** Zoomed plot (residues 1 to 180) showing ^{15}N-R$_2$/R$_1$ for J-domain and GF/GM region for the free (black

squares) and EEVD-bound state of $Sis1_{1-352}$ (red circles). Blue arrows highlight the residues for which ^{15}N-R_2/R_1 decreased upon binding to EEVD. **d** Zoomed plot (residues 1 to 180) showing S^2 for J-domain and GF/GM region for the free $Sis1_{1-352}$. **e** Zoomed plot (residues 1 to 180) showing the helical propensity of the free state predicted by Talos-N. **f** Zoomed plot (residues 1 to 72) showing ^{15}N-R_2/R_1 for the J-domain free (black squares) and EEVD-bound (red circles). Dotted lines, average of ^{15}N-R_2/R_1. *, statistically significant for $P < 0.05$ decrease in average ^{15}N-R_2/R_1.

EEVD interacts in multiple sites within $Sis1_{1-352}$

To further understand the effect that the EEVD binding causes in $Sis1_{1-352}$ we used two methods to map the binding sites. Paramagnetic relaxation enhancement (PRE) on $Sis1_{1-352}$ residues elicited by an N-terminal TEMPO-labeled EEVD peptide (Fig. 3A) and the chemical shift perturbation (CSP) caused by the binding to the EEVD peptide (Fig. 3B). We assigned 70% of the backbone of $Sis1_{1-352}$ and, to have more reliability in the conclusions, all the overlapping peaks from the analysis were removed. We mapped two major regions of interaction, guided by the residues presenting PRE < 0.88 or 0.7 (dotted line in Fig 3A) or CSP > 0.018: (i) the GF IDR, spanning from residue 78 to 102, which is between the J-domain and α-helix 6, and (ii) part of GM and the CTDI domain (residues 180-256). Interestingly, both regions in free $Sis1_{1-352}$ contain residues in conformational exchange, as shown by their above-the-average values of ^{15}N-R_2/R_1 (Fig. 2A). In the presence of EEVD the ^{15}N-R_2/R_1 dropped to average values, meaning the mili to microsecond motions (conformational exchange) were suppressed by the binding.

The observed PRE and CSP in multiple regions of the protein raise the question of whether EEVD binds in more than one site (majorly GF and CTDI). Supporting the hypothesis of multiple sites was the observation of PRE and CSP for residues in the J-domain: V2, A10, S14, T42 and I48, and for a few residues in the CTDII domain (Fig. 3). Although CSP could arise from an allosteric long-range consequence of a binding, PRE is a direct effect of the presence of the paramagnetic center within ~25 Å from the nuclei[24], meaning that there is likely more than one site of interaction of EEVD, due to the presumed long distance between J-domain/GF to CTDI. To further test the multiple binding site hypothesis, we used NMR data-based docking of the EEVD peptide to $Sis1_{1-352}$ and concluded that only multiple binding sites are plausible.

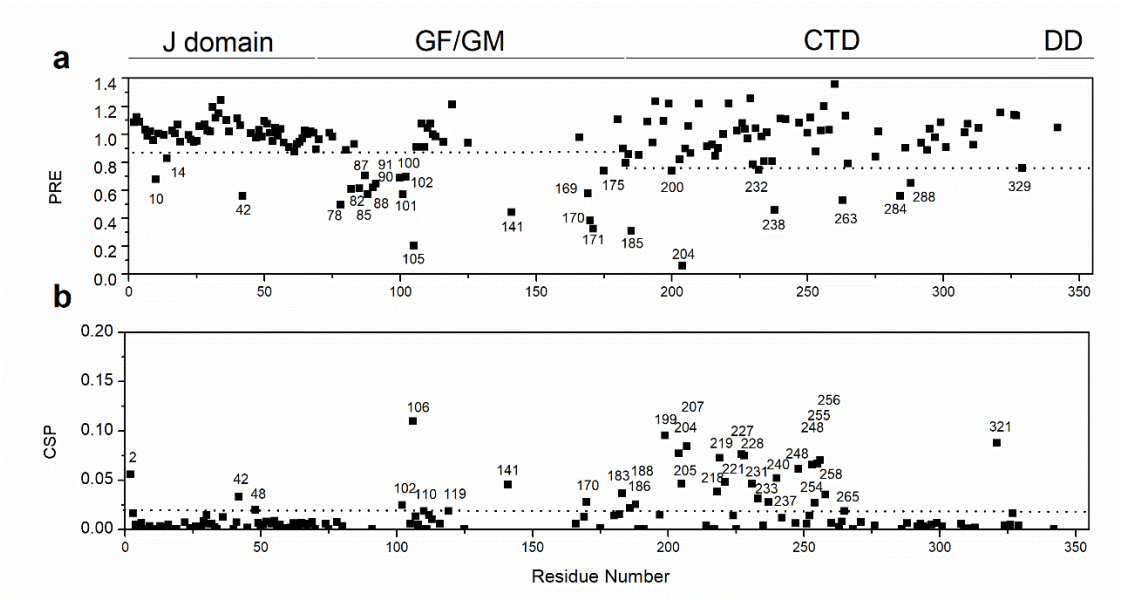

Fig. 3 | EEVD binding to Sis1$_{1\text{-}352}$. a Experimental intensity ratios of backbone amide Sis1$_{1\text{-}352}$ (full-length) in complex with an EEVD paramagnetic peptide (see Material and Methods) for each residue. An intensity ratio of 1 indicates no effect of the spin label on an amide proton. Residues with a significant PRE are indicated. PRE is the normalized intensity ratio (I^{para}/I^{dia}) between the paramagnetic and diamagnetic (reduced with ascorbate) TROSY spectra. **b** Chemical shift perturbation of Sis1$_{1\text{-}352}$ resonances upon the addition of the EEVD peptide. CSPs larger than the mean plus one standard deviation are indicated by dotted lines.

Deriving structural models for the interaction of EEVD with Sis1$_{1\text{-}352}$

Haddock[25] was used to dock the EEVD peptide into the different regions of Sis1$_{1\text{-}352}$. First, the docking targeted the dimeric CTDI (180-257) using only the PRE and CSP data relative to this region. We observed poses of the EEVD-peptide binding to the CTDI in mainly two interaction sites, named here I and II, which are in opposite faces of the domain (Figs. 4 and S1). The interaction space at the CTDI (sites I and II) was well-defined, but not the orientation and position of the EEVD-peptide relative to these sites, varied among different poses. The strategy for analysis focused on describing the interaction space within the CTDI and was based on the many observed poses of the EEVD relative to this domain (Fig. S1C). For that, we selected the 100 lowest binding energy conformers, which varied from -399.8 to -174.0 kcal/mol (Fig. S1A). Similar results would be obtained if we had chosen the 100 lowest energy structures because of the linear correlation observed between the energy of the complex and binding energy

(Fig. S1B). Note that by using the 100 lowest binding energy structures we warrant that all selected structures had negative complex and binding energy (Fig. S1B), representing well-behaved complex models in terms of geometry.

Next, we separated the 100 structures into two clusters: cluster I, containing all the poses that bind to site I (73 structures) and cluster II containing the poses that bind to site II (27 structures) (Fig. S1C). Using the *Volmap 1.9.3*, a structural tool of VMD software[26], we calculated the atomistic (mass) probability density map for both clusters. The atomic probability density map for site I (cluster I) and site II (cluster II) are described by the meshes in cyan (site I) and magenta (site II) in Fig. 4. Note that the position of the density maps matches with the experimental PRE (Fig. 4A) and CSP (Fig. 4B) in the CTDI as measured for $Sis1_{1\text{-}352}$ in the presence of the EEVD peptide. Figs. S1D and S1E show the lowest binding energy poses, which superpose well with the calculated density maps. For site I, the density map also coincides exactly at the position of an HSP70-EEVD peptide complexed with the human DNAJB1[27] (Fig. S2A). One of the poses even forms an antiparallel β-strand with the CTDI β-strand 4 (β4), the same configuration of the EEVD complex at site I of human DNAB1[27]. For site II, the density map is between β1 and β2, while both crystal structures describe the interaction of EEVD peptide with either DNAB1 (3AGY, Fig. S2B) or Sis1 (2B26, Fig. S2C) show interaction with β2, forming and antiparallel β-strand. To assert the stability of the docking presented here, we run 100 ns molecular dynamics (MD) simulations of the 2 lowest energy docking poses for each site (I and II). As a control we also run the MD simulations starting from the crystallographic structure of EEVD complexed with DNAJB1 and Sis1, at sites I and II (Fig. S3 and S4). The docked structures were stable along the MD simulations.

At this point, for the sake of comparison, it is important to analyze the electron density maps of all available crystal structures describing the interaction of the EEVD peptide and the CTDI. In a crystal structure of $Sis1_{171–352}$ complexed with the HSP70-GPTVEEVD peptide[13] (PDB 2B26), the EEVD backbone, which is bound to site II, is well determined but none of the side chains has a well-defined electron density, making it difficult to be assertive on the orientation of the peptide relative to β2 (see Fig. S5A and S5B). It is also important to mention that, in this case, the EEVD-peptide is not involved in any crystal contact.

In a crystal structure of DNAJB1[27] complexed with the EEVD peptide (PDB 3AGY), at site II, the same observations are valid, with low electron density for the EEVD-peptide backbone and none for the side chains (Fig. S5C and S5D). For DNAJB1[27] at site I, the electron density for the backbone is well-defined and sidechains P2, I4, E5, and E6 have reasonably well-defined electron densities, enabling the assertiveness of the relative orientation of the EEVD-peptide and β4, as antiparallel β-strands (Fig. S5C and S5E).

The analysis of the lowest energy docking models generated in this work for site I from Sis1, in which the atomic probability map coincides with the electron density map of the crystal structures, allows for two possible orientations of the peptide relative to β4 (Fig. S6). The EEVD-peptide lowest energy structure (Fig. S6A) has a parallel orientation relative to the CTDI-β4, while the 2nd lowest energy structure has an antiparallel orientation to this domain (Fig. S6B), similar to the EEVD-peptide into DNAJB1 at site I (Fig. S6C). Remarkably, both orientations are stabilized by a considerable number of salt bridges, hydrophobic and hydrogen bond interactions (Fig. S7 and Table 2).

Similar analyses were performed for site II and the atomic probability density maps of the EEVD-peptide (Fig. S2B and S2C) did not coincide with those of the electron densities for the crystal structures of the crystal complexes of either Sis1[13] or DNAJB1[27]. The position and orientation resulting from the docking models suggest that the EEVD-peptide binds between β1 and β2 of CTDI, which is compatible with the CSP and PRE data (Fig. 4). The relative orientation to β2 of the lowest energy complexes is the same as those depicted in the crystal structures (antiparallel to β2, Fig. S6D-G). The lowest energy structures are stabilized by a considerable number of salt bridges (Table 2, Fig. S8), hydrogen bonds and hydrophobic interactions.

Table 2. Orientation and salt bridges from the CTDI:EEVD-peptide complexes at sites I and II

Complex	Orientation	Salt-bridges
Site I		
Lowest energy (Sis1, this work)	Parallel to β4	K234/D8, K230/E5

2nd Lowest energy (Sis1, this work)	Antiparallel to β4	K230/E6, K222/E5, K226/D8
DNAJB1[27] (crystal structure, 3AGY)	Antiparallel to β4*	K217/E5 and K213/D8 (if D8 were present)
Site II		
Lowest energy (Sis1, this work)	Parallel to β1 and anti-parallel to β2	K199/E6, K199/D8, K256/D8
2nd Lowest energy (Sis1, this work)	Parallel to β1 and anti-parallel to β2	K256/E5, K197/D8
DNAJB1[27] (crystal structure, 3AGY)	Antiparallel to β2**	K181/E6, K182/D8
Sis1[13] (crystal structure, 2B26)	Antiparallel to β2**	K214/E6, K202/E6

* Electron density enables assertiveness on the orientation.
** Electron density does not enable assertiveness on the orientation

The interaction of the J-domain with the EEVD-peptide is largely electrostatic, being stabilized by 2 salt bridges. The interaction of J-domain with HSP70 forms 5 salt bridges: R27/D211, D35/R167, R22/E206 and K26/E217, with the NBD and K48/D477 with the SBDβ domains[28], i.e. it is also largely electrostatic. EEVD peptide is negatively charged, being able to compete with the salt bridges stabilizing the J-domain/HSP70 interaction. It is important to mention that the only available structure of the J-domain complexed with HSP70 is that of the *E.coli* DNAJ[28], nevertheless, all the charged residues (R22, K26, R27, D35 and K48) involved in intermolecular salt-bridges with DNAJ are conserved in the J-domain of Sis1 (K23, R27, K28, D36 and K46, respectively)

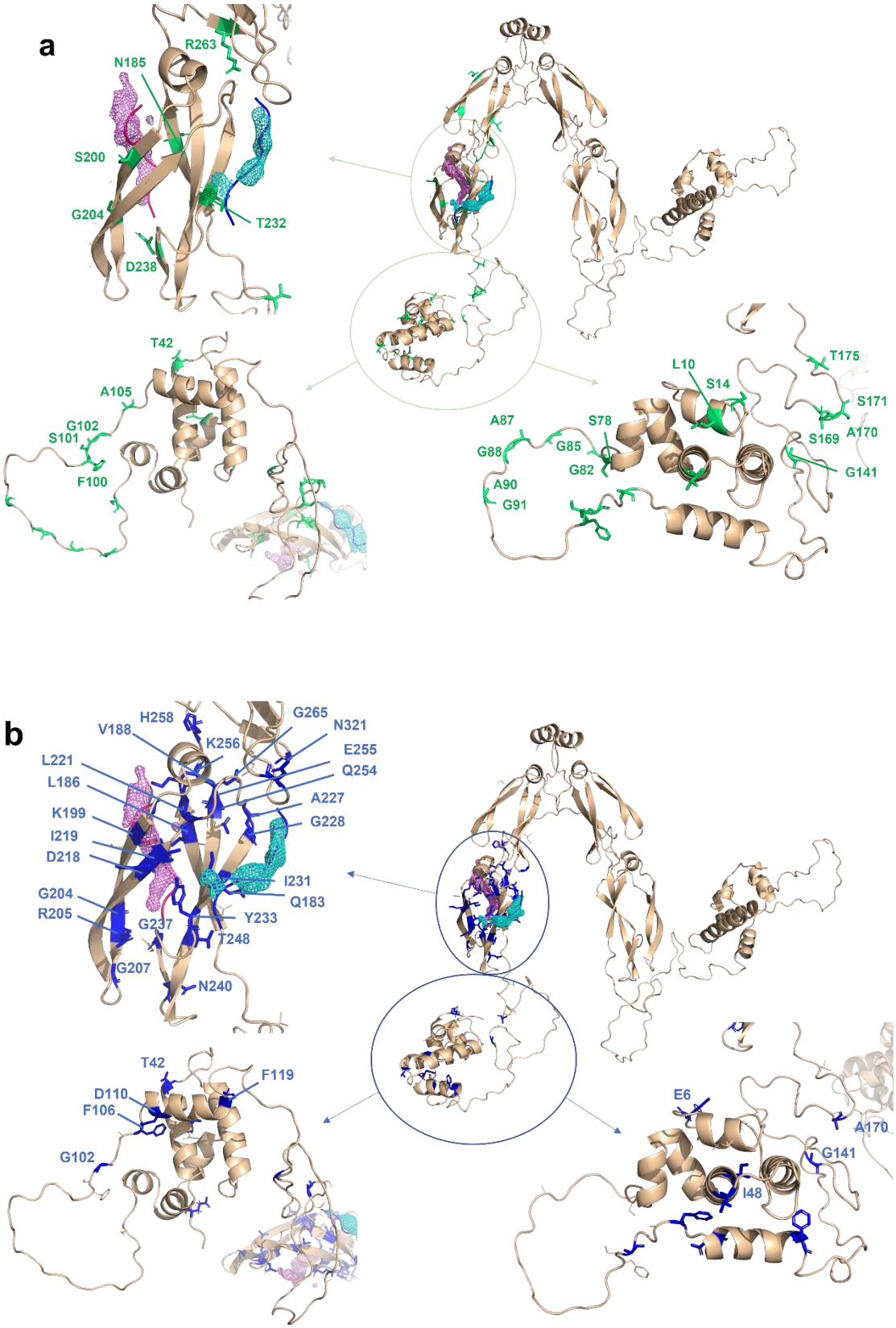

Fig. 4 | EEVD binds to multiple sites of Sis1$_{1\text{-}352}$. Cartoon representation of the Sis1 structure predicted by AlphaFold highlighting the CTDI (top of a and b), in which the atomic probability map density of the EEVD peptide at sites I (cyan) and II (magenta) are shown, and the J-domain and GF region (bottom of a and b). Residues are labeled green for PRE **a** and blue for CSP **b** data.

EEVD interaction with the GF region and the J-domain of Sis1

As discussed, the EEVD-peptide interacts transiently with the Sis1-CTDI in two sites and the analysis of both CSP and PRE data indicated interactions with residues of the J-domain (10, 14 and 42; see Fig. 3) and the GF region, mainly residues 78 to 106, which are N-terminal to α-helix 6 (Fig. 3). These data suggested that both the EEVD and the α-helix 6 competitively interact with the J-domain. The EEVD-peptide interaction with the J-domain is addressed by describing the structure of the complex between this motif and the isolated J-domain of Sis1. For that, we used the $Sis1_{1-81}$ construct which contains the J-domain (1-72) and part of GF (73-81) and had both structure and interaction with HSP70 determined[17]. We first titrated $Sis1_{1-81}$ with the EEVD-peptide by running a series of ^{1}H-^{15}N HSQCs (Fig. S9). The ^{1}H-^{15}N HSQC spectrum of the $Sis1_{1-81}$:EEVD-bound is typical of a well-folded and monomeric protein and the assignment was obtained using a standard combination of triple resonance experiments. To calculate the structure of $Sis1_{1-81}$:EEVD-bound we used ^{15}N and ^{13}C-edited NOESY experiments, which yielded 2750 intramolecular NOEs, enough to fold the $Sis1_{1-81}$ (for statistics, see Table S1). We also acquired ^{13}C-edited half-filtered experiments in the sample containing ^{12}C-EEVD-peptide and ^{13}C-$Sis1_{1-81}$ to obtain unambiguous information on the intermolecular NOEs, but due to the transient low-affinity character of the interaction, we did not get enough intermolecular NOEs to calculate an NOE-based structure of the complex.

Upon structural calculation with Aria2.3/CNS1.21, we obtained a well-defined ensemble of 20 structures of $Sis1_{1-81}$:EEVD-bound (Fig. S10A). The bound conformation of $Sis1_{1-81}$ displays five α-helices (α1 to α5, Fig. S10B), four located in the J-domain (α1 to α4) and one at the beginning of the GF region (α5). The global folded conformations of α1 (residues 6-11), α2 (19-33), α3 (42-56), α4 (58-66), and α5 (69-74) are similar to those found in free $Sis1_{1-81}$ (see reference[17]) with α2 and α3 antiparallel to one another and linked by a loop containing the HPD motif.

The ^{1}H-^{15}N CSP of $Sis1_{1-81}$ upon addition of the EEVD-peptide (1:4 molar ratio) is shown in Fig. 5A. Residues with significant chemical shift differences between the free and the bound conformation are therefore identified as being directly or indirectly related to the binding site. In such a way that several $Sis1_{1-81}$ residues had large chemical shift variations in ^{1}H-^{15}N-HSQC NMR spectra upon EEVD-peptide binding (Fig. 5A),

delimiting a well-defined region located at the α2/α3 open patch located at the J-domain, having the interaction in the fast exchange regime on the NMR time scale (Fig. S9). In the bound state, the binding patch α2/α3 is open, when compared to the free state. α3 became slightly twisted in the bound conformation, probably to accommodate the EEVD-peptide (Fig. 5B). The RMSD between the free and bound $Sis1_{1-81}$ states was 3.22 Å, as measured on Cα atoms. Residues A30, K32, and Y33 located at the α2, H34 and G40 located at the loop between α2 and α3, and E43, K44, F45, and I48, which are located at the α3 exhibited the largest CSP values (Fig. 5C), indicating that these residues are either engaged in direct contacts or are indirectly affected by conformational changes induced by the binding of the EEVD-peptide (Fig. 5G).

The binding of EEVD to $Sis1_{1-81}$ is much more evident and perturbs many more residues than when binding to the full-length Sis1 (Fig. 5). Residues V2 and K3 located at the N-terminal region, A30 located at the α2, and T42, K46, and I48 located at the α3 exhibited the largest CSP values (Fig. 5D). We also compared the CSPs of $Sis1_{1-81}$ with the EEVD-bound with those of $Sis1_{1-81}$ upon addition of HSP70, previously reported by our group[17]. Worth noting, residues A29 located at α2, H34 and T39 located at the loop, F45 and F52 located at α3, and Y67 located at α4 were perturbated by the presence of either HSP70[17] or the EEVD-peptide, indicating that the results with peptide are similar to those with the full-length HSP70.

To get the distance restraints from the EEVD-peptide to the protein, we measured intermolecular PRE rates for the backbone amide groups of $Sis1_{1-81}$ and $Sis1_{1-352}$ (Fig. 3) in complex with the spin-labeled EEVD (TEMPO-EEVD, see Material and Methods). Curiously, the region affected by PRE for isolated $Sis1_{1-81}$ was the same as observed in the CSP analysis (Fig. 5E). PRE data also showed that EEVD-peptide bound to $Sis1_{1-81}$ is much more evident and involved many more residues than when bound to the full-length $Sis1_{1-352}$ (Fig. 5F). Both isolated $Sis1_{1-81}$ and the full-length $Sis1_{1-352}$ had CSP and PRE values in the same region between α2 and α3, even though results for the full-length protein are less pronounced, probably because in the isolated domain ($Sis1_{1-81}$) the EEVD binding site is exposed, while in the full-length $Sis1_{1-352}$, EEVD has to compete with α-helix 6, which occludes the patch containing the EEVD binding site. The mapped interaction site with HSP70 is partially blocked by α-helix 6 in the context of the full-length Sis1, as supported by the observation of the transient interaction of the EEVD-

peptide, mapped by PRE, with many residues in the GF region (78-106). Worth pointing out again, the decrease in ^{15}N R_2/R_1 for the residues S78, G102, A105, F106 and S107 suggests a conformational selection upon EEVD-peptide binding, a condition in which the on-off equilibrium of α-helix 6 was suppressed (Fig. 2C, blue arrows). Altogether, these observations support the hypothesis that the EEVD motif competes with the α-helix 6 for the J-domain.

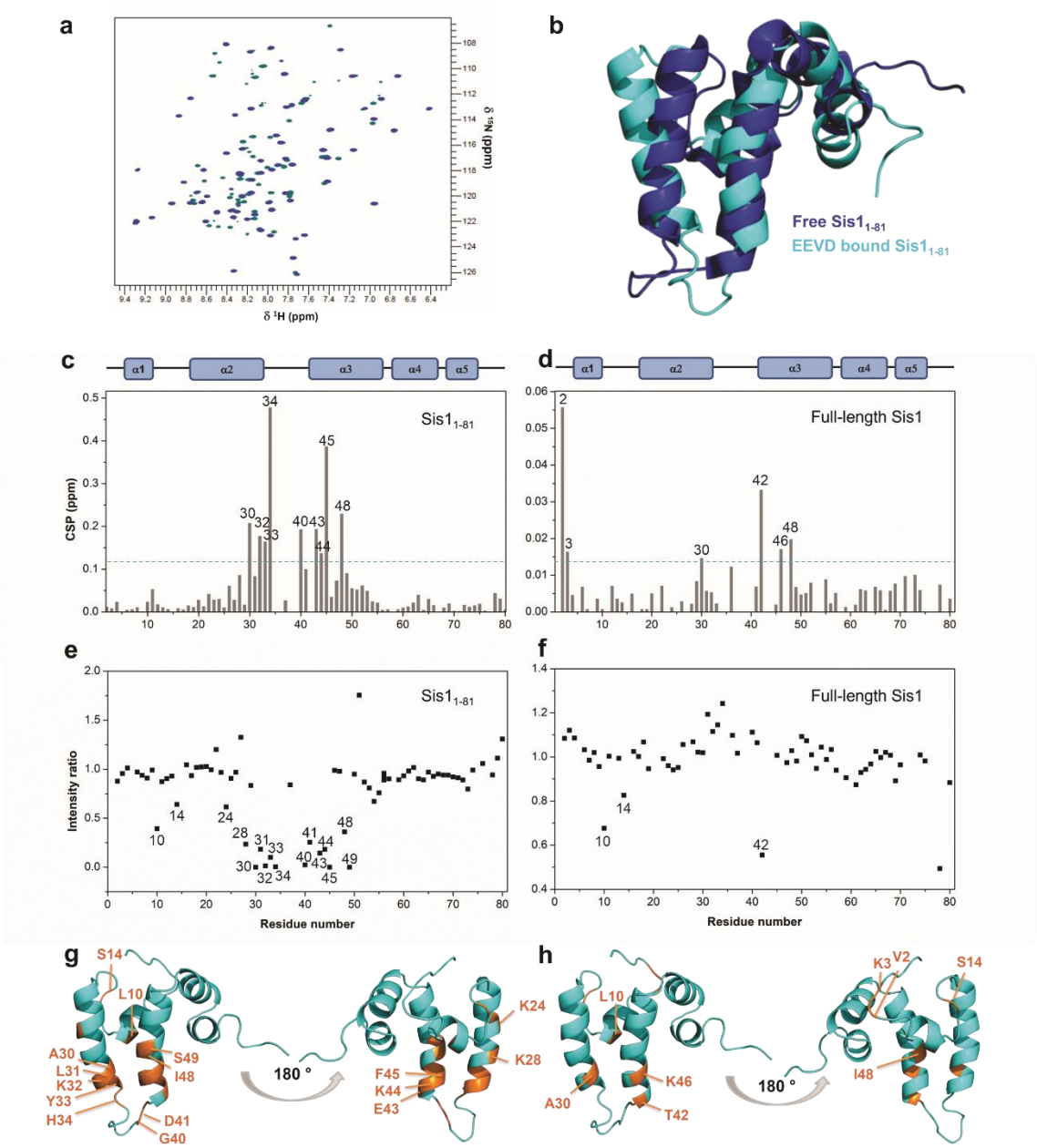

Fig. 5 | NMR analysis of Sis1$_{1-81}$ bound to the EEVD-peptide. a Overlay of 2D ^{1}H-^{15}N HSQC spectra collected for free (blue) and EEVD-bound (4-fold molar excess; cyan) Sis1$_{1-81}$. **b**

Superposition of the NMR structures of free $Sis1_{1-81}$ (blue, PDB 6D6X) and $Sis1_{1-81}$:EEVD-bound (cyan, this work). **c** Chemical shift perturbation of $Sis1_{1-81}$ resonances upon the addition of the EEVD-peptide. CSP index is shown as vertical bars for each residue. CSPs larger than the mean plus one standard deviation are those above the blue line (also shown in orange in Fig 5g). **d** CSPs of $Sis1_{1-81}$ resonances of the full-length Sis1 ($Sis1_{1-352}$) upon addition of the EEVD-peptide are those above the blue line (also shown in orange in Fig. 5h). **e** Experimental intensity ratios of backbone amide $Sis1_{1-81}$ in complex with the paramagnetic EEVD-peptide. **f** Experimental intensity ratios of backbone amide of the full-length Sis1 ($Sis1_{1-352}$) in complex with the paramagnetic EEVD-peptide, however only residues 1-81 are shown to facilitate comparison with Fig. 5e). An intensity ratio of 1 indicates no effect of the spin label on an amide proton. CSP and PRE are mapped (orange) on the $Sis1_{1-81}$ **g** and the $Sis1_{1-352}$ **h** EEVD-bound conformations. In H, only residues 1-81 are shown.

PRE and CSP data and HADDOCK[25,29] were used to produce NMR data-based structural models of $Sis1_{1-81}$:EEVD-peptide complex (see Fig. S11 for the lowest energy structural model). Ten clusters were presented for docking and the best-docked cluster was selected based on a low Haddock score and a low RMSD value (Table S2). In the $Sis1_{1-81}$:EEVD-peptide complex model, the peptide extends between α2 and α3 with its N-terminus at the HPD-loop (Fig. S11), in good agreement with both CSP and PRE data. Molecular dynamics (MD) simulations were used to generate the $Sis1_{1-81}$:EEVD-peptide complex and the RMSD values of the backbone atoms of the $Sis1_{1-81}$ and EEVD-peptide from the starting structure (HADDOCK model, Fig. S11). The RMSD of the $Sis1_{1-81}$ was stable all over the 1 μs simulation, while the RMSD of the EEVD peptide increased subtly in the first 20 ns and became stable throughout the MD simulation (Fig. 6A). The average RMSD was 4.0 ± 0.1 Å and 8.1 ± 0.2 Å for $Sis1_{1-81}$ and EEVD-peptide, respectively (Fig. 6B). The number of contacts (distance < 0.6 nm) formed between atoms of $Sis1_{1-81}$ and the EEVD-peptide revealed that they interacted throughout all the MD simulations with 2027 ± 339 contacts (Fig. 6B, top).

The structural model showed that $Sis1_{1-81}$:EEVD-peptide complex is formed with an average value of 4 (3.7 ± 1.5) intermolecular hydrogen bonds present throughout all the MD simulations (Fig. 6B, bottom). Percentage of persistence higher than 10% were considered and are among the residues Q20-V7, K23-D8, R27-E6, H34-P2, and E26-Y26 (Table S3). The structural model also showed that the conformation of the EEVD peptide, which is extended between α2 and α3, is in very good agreement with NMR and HADDOCK data (Fig. 6C). In addition, two salt bridges were observed between K23-D8 and R27-E6, respectively (Fig. 6D, Table S3).

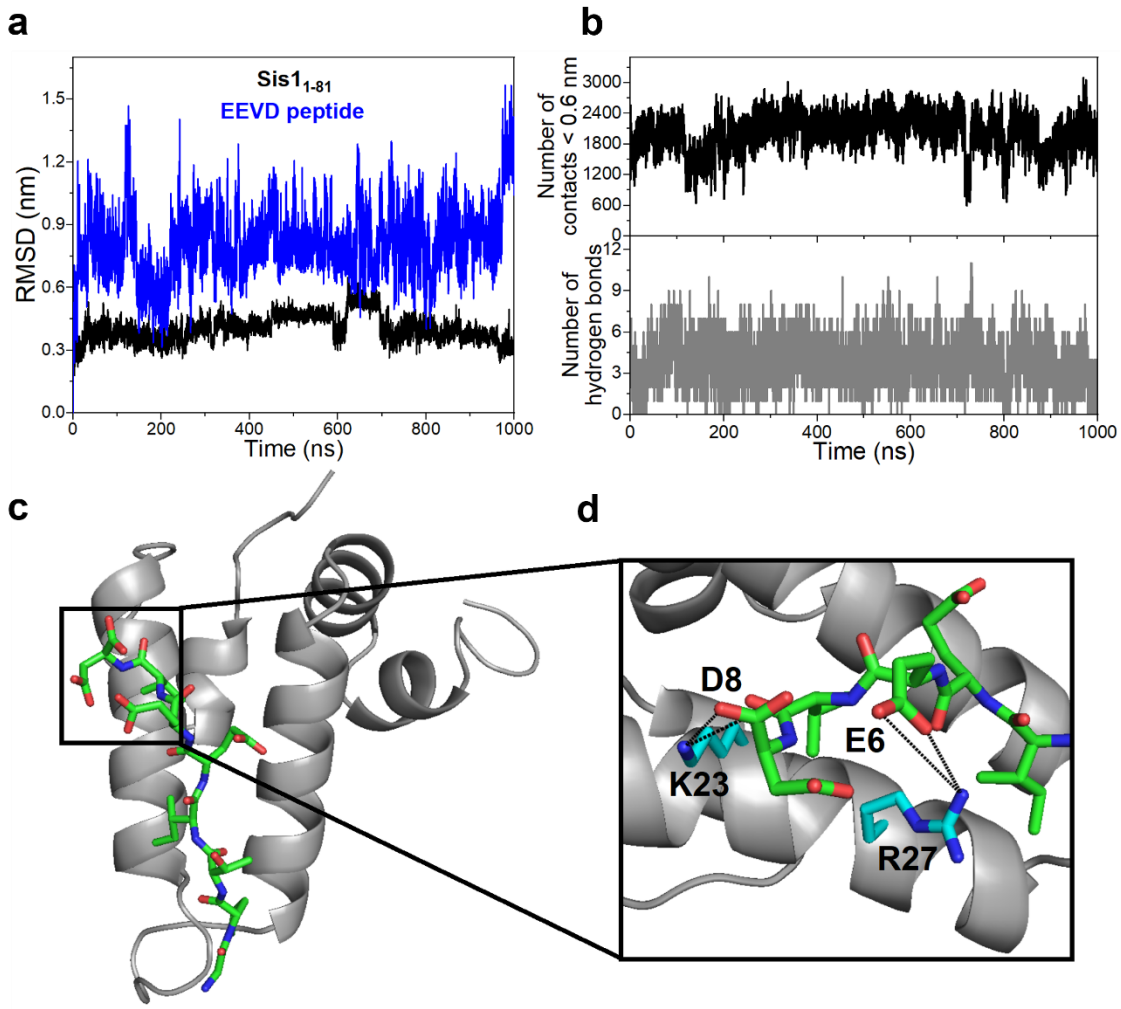

Fig. 6 | Analysis of the parameters from 1.0 µs molecular dynamics simulations of the complex between Sis1$_{1-81}$ and EEVD peptide. a Values of RMSDs of the backbone atoms of Sis1$_{1-81}$ (black) and EEVD (blue) from the starting structure (docking structure, Fig. S11). **b** number of contacts < 0.6 nm formed between atoms (Top) and a number of hydrogen bonds (Bottom) of Sis1$_{1-81}$ and EEVD peptide. **c** Representative structure of the Sis1$_{1-81}$:EEVD complex obtained from cluster analysis of the trajectory. Sis1$_{1:81}$ is denoted as a gray cartoon and EEVD as sticks with carbon in green, oxygen in red, and nitrogen in blue. **d** Detail of the intermolecular salt bridges (K23-D8 and R27-E6) formed between the Sis1$_{1-81}$ and the EEVD peptide.

Discussion

J-domain proteins (JDPs) are responsible for transferring partially unfolded client proteins (also known as substrates) to HSP70, but the mechanism involved in this transfer is not completely understood. JDPs are responsible for the recognition of the client proteins, and for this reason, there are many more JDP genes in the eukaryotes than HSP70 genes. Sis1 is a class B JDP from *S. cerevisiae* that has an essential interaction with the HSP70 C-terminus which is an IDR that ends in an EEVD motif [30–32]. To generate

a more complete understanding of this interaction, this work provided a thorough investigation of the structure and the role of the dynamics of the full-length $Sis1_{1-352}$, and on the $Sis1_{1-81}$ construct, bound to an EEVD-peptide. Most importantly, these studies were carried out in solution, being complementary to available structural information on the interaction of EEVD with class B JDPs.

The results show that the EEVD-peptide interacts transiently and with low affinity with Sis1 in multiple sites, inducing important changes in the dynamics of the protein. We addressed the effect of these interactions on the protein dynamics by measuring relaxation parameters and pin-pointing the EEVD-peptide interacting sites using CSP and PRE data combined with an experimental-driven computational docking. The EEVD-peptide was able to bind to multiple sites in the full-length Sis1, two of them at the CTDI (sites I and II), one at the boundary of α-helix 6, in the GF-region and the other at the J-domain. The interactions with sites I and II were unambiguous and in good agreement with previously described crystal structures of the complex with an EEVD-peptide for both human DNAJB1[27] and yeast Sis1[13] cochaperones. The observed interactions contributed to a better understanding of the mechanism of class B JDP and HSP70 interaction, which is summarized in Fig. 7 and further explained below.

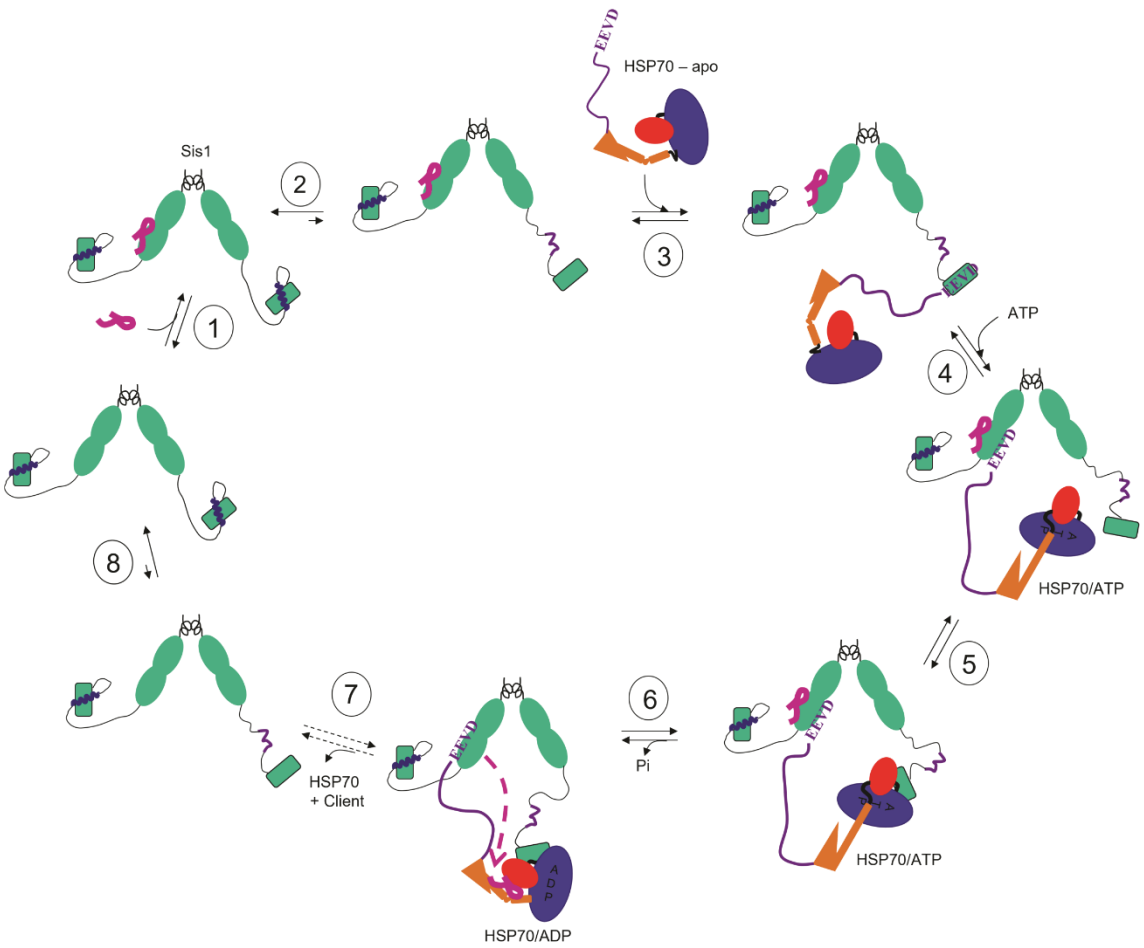

Fig. 7 | Mechanistic chaperone cycle model of HSP70 and class B JDP. From the knowledge generated by the interactions described in this work and from established parameters of JPDs/HSP70 interaction cycle[7,33–35] a mechanistic cycle with 8 steps is proposed (please see Fig. S12 for domain arrangements in the several conformational states of HSP70). The mechanism takes into consideration all the mapped sites of the interaction of the EEVD peptide with Sis1: site I of CTDI, site II of CTDI, J-domain and GF region. Step 1, JDP Sis1 recognizes and binds to a nascent or misfolded protein (client protein, purple) via site II of its C-terminal domain I (CTDI). Steps (2; client bound) and (8, unbound) represent the on-off equilibrium between the J-associated (on-state) and J-dissociated (off-state) α-helix 6. In the absence of HSP70, the equilibrium is shifted towards the on-state, where the J-domain recognition patch (α-helix 2, α-helix 3 and loop containing HPD motif) becomes unavailable to bind HSP70. Step 3 illustrates the transient binding of the C-terminal EEVD of the apo state of HSP70 with the GF region and J-domain promoting an equilibrium shift toward the off-state, where the J-domain recognition patch becomes available to interact with the HSP70 NBD. In this step, the client protein is bound to Sis1, while HSP70 is client-free. It is expected for the client-free apo HSP70 to have a low affinity with the client protein[33] and to be majorly in the closed SBD conformation[5]. In step 4 the EEVD binds to site I of CTDI and ATP binds to NBD leading to a conformational change of the SBD to the open state. In this step, the J-domain becomes available to bind to the open conformation of HSP70/ATP. Step 5 shows the binding of the J-domain at the interface between NBD (blue) and SBDβ (red). Step 6 illustrates SBD interaction at ultra-affinity to the client protein (purple), with the concomitant aid of the EEVD dissociating the client protein by binding to site II of CTDI. The purple arrow illustrates the transfer of the client protein from CTDI to the closed SBD/ADP since the ATPase activity of the NBD is activated to its maximum by the

combined action of the J-domain association and the binding of the client protein. The HSP70/ADP has a high affinity to the client protein (K_D ~ 0.1 μM), whereas the HSP70/ATP has low affinity (K_D ~ 1 μM)[28,33]. The combined action of the J-domain associated with HSP70/ADP closed conformation leads to an ultra-affinity state for the client protein, at the nanomolar range. Step 7 illustrates the dissociation of HSP70:client, making both HSP70 and Sis1 available for another productive cycle. This step is simplified (dotted arrows) because it involves several steps of the HSP70 cycle. Other important information used to build the mechanistic chaperone cycle: The dissociation of the J-domain from the HSP70 is triggered by the presence of ATP[17]. ATP is also the trigger for client release[5]. The schematic view of both Sis1 and HSP70 follows the conformational arrangement of the domains as experimentally determined.

The interaction of the EEVD-peptide with the J-domain and GF region appears to be more complex than with the CTDI. The interpretation of our results, along with the reported structure of human $DNAJB1_{1-111}$, in which the J-domain is complexed with the α-helix 5[15] (α-helix 6 for Sis1), and predictions of the structure of Sis1 made by Alphafold[21], suggested that the interaction of the EEVD-peptide with Sis1 is modulated by an on-off equilibrium between the J-domain-associated (on-state) or J-domain-dissociated (off-state) α-helix 6. The on-off equilibrium is illustrated in paths 2 and 8 of the interaction cycle (Fig. 7). The off-state exposes the interaction patch of the J-domain that is composed of α-helix 2 and 3 and the loop between them, which contains the HPD motif. Noteworthy, this patch binds to the EEVD, to α-helix 6[15], and to an interface between the NBD and SBDβ of HSP70[17,28]. The proposed mechanism in Fig. 7 contemplates the interaction of this patch with the EEVD (step 3), with the NBD/SBDβ (steps 5 and 6, and with α-helix 6 for most steps, where at least one of the J-domain is in the on-state)

Although the J-domain interacting patch is exposed in the Sis_{1-81}, it is mostly occluded in the full-length Sis_{1-352}, as illustrated in steps 2 and 8 of Fig. 7. The PRE and CSP data for this patch was much more pronounced when analyzing the experiment with Sis_{1-81} than with Sis_{1-352} (Fig. 5), but even for the Sis_{1-352}, CSP and PRE were present for the J-domain recognition patch: A30 at α-helix 2, T42, K46 and I48, at α-helix 3. Additionally, the CSP and PRE were much more pronounced for the adjacencies of α-helix 6, located at the GF domain, when Sis_{1-352} was bound to the EEVD-peptide. We interpreted these data as 1) the interaction of EEVD with the on-state occurring majorly at the GF region and 2) only when α-helix 6 is dissociated from the J-domain (off-state) the EEVD binds to the interaction patch of the J-domain. These interpretations are supported by the suppression of the residues in conformational exchange in the GF region

upon binding to the EEVD peptide. Furthermore, there was a subtle increase in the rotational freedom of the J-domain in the presence of EEVD. The high rotational freedom of the J-domain contributes to the biological mechanism of the JDPs, and the binding of EEVD makes it more available for interactions. Further studies are necessary to understand the interplay between the EEVD binding to the J-domain interacting patch and the GF region.

The EEVD/J-domain association is mostly electrostatic, with two salt bridges. On the other hand, the GF region (78-105) has no charged residues, and EEVD probably interacts (presence of PRE, Fig. 3) via hydrophobic contacts involving residues G1, P2, T3, I4 and V7. Since the α-helix 6 association with the J-domain patch is also mediated by hydrophobic interaction, through residues A111, I114, F115, F118 and F119, we suggest that the EEVD competes for these hydrophobic interactions with α-helix 6 while at the same time is attracted by the positive electrostatic potential of the J-domain interacting patch. To sum up, the competing interaction of EEVD with the J-domain patch shifts the equilibrium to the off-state (step 3 in Fig. 7), exposing the J-domain binding site to interact with the NBD and SBDβ of the apo/closed conformation of HSP70.

Finally, all interactions of the EEVD with sites I and II of CTDI are of low affinity and largely electrostatic (salt bridges and hydrogen bonds) raising the hypothesis that the EEVD has two functions: (i) binding to site I of CTDI, keeping HSP70 anchored to Sis1 even without the J-domain association with NBD/SBDβ (step 4 and 5, Fig. 7), and (ii) competing with the client protein for site II (Fig. S13), helping in the dissociation of the client protein, which is transferred to the ultra-affinity state of HSP70 (step 6, Fig. 7). The ternary complex JDP:HSP70:client dissociates (step 7, Fig. 7) from the JDP, consequently releasing HSP70:client complex. From this step on, Sis1 is released to bind to another client protein. Eventually, HSP70 dissociates from the client protein,

Altogether, the EEVD interacts with multiple sites of $Sis1_{1-352}$ acting as a modulating agent or a factor of disorder, facilitating the exposure and availability of the J-domain and CTDI to the interaction with HSP70.

Methods

Sample Preparation. $Sis1_{1-352}$ (UNIPROT P25294) and $Sis1_{1-81}$ expression and purification were carried out according to a previous work[36,37]. GPTIEEVD and CGPTIEEVD peptides referring to

the C-terminal motif of HSP70 were synthesized and purified by GenOne Biotechnologies (Rio de Janeiro, Brazil).

NMR spectroscopy. The NMR experiments to assign $Sis1_{1-352}$ were collected in a Bruker Avance III HD 950 MHz Spectrometer equipped with ^{13}C, ^{15}N, ^{1}H, ^{2}H TXI cryoprobe. $Sis1_{1-352}$:EEVD-bound, $Sis1_{1-81}$ and $Sis1_{1-81}$:EEVD-bound were collected on a Bruker Avance III HD 900 MHz spectrometer equipped with an inverse-detection triple resonance z-gradient TXI probe. All NMR samples contained 300 μM of $Sis1_{1-352}$ or 1 mM of $Sis1_{1-81}$ in 25 mM Tris-HCl pH 7.5, 200 mM NaCl and 10 % D_2O at 303K or 298 K, respectively. NMR data were processed and analyzed with nmrPipe[38] and CcpNmr Analysis[39] available on the NMRbox platform[40].

Backbone dynamics. ^{15}N backbone amide relaxation parameters (R_1 and R_2) were acquired on Bruker 900 MHz for a ^{15}N-$Sis1_{1-352}$ at 303 K. The TROSY-based ^{15}N longitudinal (R_1) relaxation rates were measured at 20, 300 and 600 ms and transverse (R_2) relaxation rates were measured at 8.65, 16.96 and 33.92 ms. Relaxation parameters were determined by fitting T_1 and T_2 peak intensities to a single exponential decay. ^{15}N R_1 and R_2 values were determined from the fit of the peak intensities using a mono-exponential equation. Experimental errors of the relaxation parameters were evaluated based on the signal/noise ratio as described previously[41]. All relaxation experiments were acquired as pseudo-3D spectra and converted to 2D data sets. NMR spectra were processed and analyzed with NMRPipe[38]. The overall rotational correlation time (τ_c) of $Sis1_{1-81}$ was estimated from the mean values of R_1 and R_2 measured for each domain, as follows: $\tau c \approx \frac{1}{4\,\pi \upsilon_N}\sqrt{6\,\frac{R_2}{R_1} - 7}$, where υ_N is the ^{15}N resonance frequency (Hz), R_1 and R_2 are the mean values of the ^{15}N relaxation rates[42].

Chemical shift assignment. The backbone resonance assignments of $Sis1_{1-352}$, data were acquired from the following two-dimensional (2D) and three-dimensional (3D) NMR experiments: all BEST-TROSY versions[43]: BT-HSQC, BT-^{13}C-HSQC, BT-HNCACB, BT-HNCA, BT-HNCACO, BT-HNCO, BT-HNCOCA, BT-HNCOCACB, BT-N-NOESY and BT-HNCANH. The assigned chemical shift values of backbone ^{15}N, $^{13}C\alpha$, and $^{13}C'$ of $Sis1_{1-352}$ were used as the input for the TALOS-N program[19,20] to predict secondary structures. The backbone and side chain assignments of $Sis1_{1-81}$:EEVD-bound were deposited in the BMRB (BMRB 51187).

Site-direct Spin-Labeling experiments. A 20-fold excess of 4-maleimide-TEMPO (Sigma-Aldrich) dissolved in acetonitrile was added to CGPTIEEVD peptide at 1 mM in 25 mM Tris-HCl pH 7.5, 200 mM NaCl and incubation for 2 h at room temperature. The final product was the TEMPO spin label attached to the reactive thiol functional group of Cys1. To remove unreacted and excess maleimide spin label, the reaction mixture was injected in a C18 column and HPLC chromatography (Cytiva) in acetonitrile and water and eluted with a 30-70 % (water:acetonitrile 0.1 % TFA) of gradient

Paramagnetic relaxation enhancement (PRE). The direct interaction of $Sis1_{1-352}$ and $Sis1_{1-81}$ with EEVD motif was investigated by analyzing the PRE with the TEMPO-labeled EEVD peptide. ^{15}N-$Sis1_{1-81}$ was titrated with TEMPO-labeled CGPTIEEVD peptide (paramagnetic sample) and ^{1}H-^{15}N TROSY (for $Sis1_{1-352}$) or HSQC (for $Sis1_{1-81}$) spectra were recorded at 900 MHz. The diamagnetic samples were obtained after the reduction with 3 mM ascorbic acid, at room temperature for 4 h. PRE was obtained by the intensity ratio between the paramagnetic and diamagnetic spectrum ($PRE = \alpha I^{para}/I^{dia}$). We observed PRE effects higher than one because of the larger longitudinal relaxation time for the diamagnetic sample when compared to the paramagnetic and a short relaxation time used in the TROSY or HSQC spectra (1 and 1.3 s, respectively). Because of that, we introduced a correction factor (α) to make the average <PRE> =1 for the residues distant from the paramagnetic center.

To calibrate the PRE effect into semiquantitative distance ranges we considered a linear change of the PRE effect as a function of the distance from each NH to the paramagnetic center (d) between 12 and 30 Å[24,44]. We considered d >30 Å for PRE =1 and d < 12 Å for PRE = 0. Distances between 12 and 30 Å were calculated (d_{calc}) assuming a linear change proportional to PRE. The distances were used semiquantitatively for the Haddock calculation: For PRE = 0, d < 12 Å and for PRE between 0.8 and 0, d was in the range between d_{calc} and 30 Å.

Chemical shift mapping. To identify the interaction of $Sis1_{1-352}$ or $Sis1_{1-81}$ with EEVD peptide, ^{1}H-^{15}N-TROSY spectra were collected for free ^{15}N-labeled $Sis1_{1-352}$ at 0.3 mM and upon addition of GPTIEEVD- (from now on referred as EEVD) (1:4 protein: peptide). For ^{15}N-$Sis1_{1-81}$ at 0.5 mM, ^{1}H-^{15}N HSQC spectra were collected of free and upon addition of EEVD (1:4 protein: peptide). Chemical shift perturbations (CSP) were calculated for each backbone amide group, as follows: $CSP = \sqrt{(\Delta\delta^{15N}/10)^2 + (\Delta\delta^{1H})^2}$, where $\Delta\delta^{15N}$ and $\Delta\delta^{1H}$ are the chemical shift differences between the free and bound states of $Sis1_{1-81}$:EEVD or $Sis1_{1-352}$:EEVD. CSP greater than 1 standard deviation from the mean was considered significant.

Structure calculation of Sis1$_{1-81}$:EEVD. NMR data for structure calculation of Sis1$_{1-81}$:EEVD were collected on a 1 mM ^{13}C/^{15}N-labeled Sis1$_{1-81}$ in 25 mM Tris-HCl pH 7.5, 200 mM NaCl and 10 % D_2O at 900 MHz. Distance restraints were derived from 3D ^{15}N-NOESY-HSQC and 3D ^{13}C-NOESY-HSQC both aliphatic and aromatic and dihedral angle restraints were derived from ^{1}HN, ^{15}N, ^{13}Cα, ^{1}Hα, ^{13}Cβ, and C' chemical shifts using TALOS-N[20]. The amino acid sequence, the chemical shift lists, the dihedral angle values and the three NOESY data sets were used as input files for structural determination.

Three-dimensional structures of the Sis1$_{1-81}$:EEVD-bound were calculated using the program Aria 2.3[45,46] combined with CNS 1.2[47] available on the NMRbox platform [40]. In the final calculation the 20 lowest-energy structures, derived from the water refinement step, were selected as representative of the ensemble of Sis1$_{1-81}$:EEVD conformations in solution. For quality validation, we used Protein Structure Validation Software suite (PSVS)[48] (https://montelionelab.chem.rpi.edu/PSVS/PSVS/). Structures were visualized with PyMOL[49] and the atomic coordinates of Sis1$_{1-81}$:EEVD were deposited in the *Protein Data Bank* under accession code 8EOD.

Molecular Docking and Molecular Dynamic (MD) Simulations. The software Haddock[25] was used to dock the EEVD peptide to CTDI (residues 180 to 257) of Sis1$_{1-352}$ using the residues with CSP (Fig. 3b) as active residues and distance restraints semi-quantitatively calibrated from the PRE data (Table S3). Passive residues were automatically assigned as those surrounding the active ones. Residues with low order parameters ($S^2 < 0.6$, Fig. 2B) were set as fully flexible. For Sis1$_{1-352}$, we have NMR-based interaction information with many regions but created docking models for CTDI only. It was unfeasible to create an interaction model of one EEVD peptide that was compatible with all the experimental data. The HADDOCK (version 2.4) server (https://wenmr.science.uu.nl/haddock2.4/)[25] was used for the modeling of CTDI:EEVD. The protein structural coordinates of CTDI used as input were obtained from an Alphafold prediction, which is almost identical to the Protein Data Bank (PDB) under access code 2B26[13]. In total, 2000 complex structures of rigid-body docking were calculated by using the standard HADDOCK protocol with an optimized potential for liquid simulation (OPLSX) parameters. The final 200 lowest-energy structures were selected for subsequent explicit solvent (water) and semi-flexible simulated annealing refinement, to optimize side chain constants. The Volmap tool of the Visual Molecular Dynamics (VMD)[26] software was used for the construction of the atomistic probability map of CTDI-Sis1:EEVD. An ensemble of 100 structures lowest energy structures calculated from Haddock was used. The 100 structures were clustered in two groups. Cluster I with all the poses binding to Site I and Cluster II with all the poses binding to Site II (Figure S1).

The software Haddock[25] was used to dock the EEVD peptide to the EEVD-bound conformation of $Sis1_{1-81}$ (Fig.S10, PDB 8EOD) using the residues with CSP (Fig. 5c) as active residues and distance restraints semi-quantitatively calibrated from the PRE data (Table S4). Passive residues were automatically assigned as those surrounding the active ones. The HADDOCK (version 2.4) server (https://wenmr.science.uu.nl/haddock2.4/)[25] was used for the modeling of $Sis1_{1-81}$:EEVD. In total, 2000 complex structures of rigid-body docking were calculated by using the standard HADDOCK protocol with an optimized potential for liquid simulation (OPLSX) parameters. The final 200 lowest-energy structures were selected for subsequent explicit solvent (water) and semi-flexible simulated annealing refinement, to optimize side chain constants. The final structures were clustered using the fraction of common contacts (FCC) with a cutoff of 0.6.

Molecular dynamics (MD) calculations for docking models of CTDI at site I and site II and the 2 lowest energy structure of the highest representative cluster of $Sis1_{1-81}$:EEVD complex were performed using GROMACS (version 5.1.4) [50]. The molecular systems were modeled with the corrected AMBER14-OL15 package, including the ff14sb protein force field [51], as well as the TIP3P water model [52]. The structural models of the complexes (from molecular docking) were placed in the center of a cubic box solvated by a solution of 200 mM NaCl in water. Periodic boundary conditions were used, and all simulations were performed in NPT ensemble, keeping the system at 25 °C and 1.0 bar using Nose-Hoover thermostat (τ_T = 2 ps) and Parrinello-Rahman barostat (τ_P = 2 ps and compressibility = $4.5\times10^{-5}\cdot bar^{-1}$). A cutoff of 12 Å for both Lennard-Jones and Coulomb potentials was used. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) algorithm. A conjugate gradient minimization algorithm was used to relax the superposition of atoms generated in the box construction process. Energy minimizations were carried out with the steepest descent integrator and conjugate gradient algorithm, using 1,000 $kJ\cdot mol^{-1}\cdot nm^{-1}$ as the maximum force criterion. Five hundred thousand steps of molecular dynamics were performed for NVT and NPT equilibration, applying force constants of 1,000 $kJ\cdot mol^{-1}\cdot nm^{-2}$ to all heavy atoms of the complex. At the end of preparation, 100 ns MD pulling simulation of the molecular system was carried out using a spring constant of 1,000 $kJ\cdot mol^{-1}\cdot nm^{-2}$ between the protein and the peptide. Next, 1 μs MD simulation was performed for data acquisition. Following dynamic, the trajectories of the complex were firstly concatenated and analyzed according to the RMSD for the backbone atoms of protein and peptide, a number of contacts for distances lower than 0.6 nm between pairs of atoms of CTDI- Sis1/$Sis1_{1-81}$ and EEVD peptide, and a number of protein-peptide hydrogen bonds with cutoff distance (heavy atoms) of 3.5 Å and maximum angle of 30°. The percentages of protein-peptide hydrogen bond persistence were obtained from *plot_hbmap_generic.pl* script[53]. The number of protein-peptide hydrogen bonds with persistence greater than 10% was considered. The tool *g_cluster* of

GROMACS[54] package was used to perform cluster analysis in the 1.0 μs MD trajectory of the complex using cutoff of 3 Å. The structure of the first cluster was used as representative structure of the complex The structural representation of the constructed model was displayed using PyMOL[49].

References

1. Kohler, V. & Andréasson, C. Hsp70-mediated quality control: should I stay or should I go? *Biol. Chem.* **401**, 1233–1248 (2020).
2. Liu, Q., Liang, C. & Zhou, L. Structural and functional analysis of the Hsp70/Hsp40 chaperone system. *Protein Sci.* **29**, 378–390 (2020).
3. O. Tiroli-Cepeda, A. & H.I. Ramos, C. An Overview of the Role of Molecular Chaperones in Protein Homeostasis. *Protein Pept. Lett.* **18**, 101–109 (2011).
4. Kelley, W. L. The J-domain family and the recruitment of chaperone power. *Trends Biochem. Sci.* **40**, 222–227 (1998).
5. Rosenzweig, R., Nillegoda, N. B., Mayer, M. P. & Bukau, B. The Hsp70 chaperone network. *Nat. Rev. Mol. Cell Biol.* **20**, (2019).
6. Ramos, C. H. I., Oliveira, C. L. P., Yang-fan, C., Torriani, I. L. & Cyr, D. M. Conserved central domains control the quaternary structure of Type I and Type II Hsp40 molecular chaperones. *J Mol Biol.* **383**, 155–166 (2009).
7. Kampinga, H. H. & Craig, E. A. The HSP70 chaperone machinery : J proteins as drivers of functional specificity. *Nat. Publ. Gr.* **11**, (2010).
8. María, S., Mariscal, A. & Kirstein, J. J-domain proteins interaction with neurodegenerative disease-related proteins. *Exp. Cell Res.* **399**, 112491 (2021).
9. Aron, R., Lopez, N., Walter, W., Craig, E. A. & Johnson, J. In vivo bipartite interaction between the Hsp40 Sis1 and Hsp70 in Saccharomyces cerevisiae. *Genetics* **169**, 1873–1882 (2005).
10. Lu, Z. & Cyr, D. M. Protein Folding Activity of Hsp70 Is Modified Differentially by the. *J. Biol. Chem.* **273**, 27824–27830 (1998).
11. Caplan, A. J. and & Douglas, M. G. Characterization of YDJI: A Yeast Homologue of the Bacterial dnaJ Protein. *J Cell Biol.* (1991) doi:10.1083/jcb.114.4.609.
12. Luke, M. M., Sutton, A. & Arndt, K. T. Characterization of SISI, a Saccharomyces. *J Cell Biol.* **114**, (1991).
13. Li, J., Wu, Y., Qian, X. & Sha, B. Crystal structure of yeast Sis1 peptide-binding fragment and Hsp70 Ssa1 C-terminal complex. *Biochem. J.* **398**, 353–360 (2006).
14. Yu, H. Y. *et al.* Roles of Intramolecular and Intermolecular Interactions in Functional Regulation of the Hsp70 J-protein Co-Chaperone Sis1. *J. Mol. Biol.* **427**, 1632–1643

(2015).

15. Faust, O. *et al.* HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. *Nature* **587**, 489–494 (2020).
16. Karamanos, T. K., Tugarinov, V. & Clore, G. M. Unraveling the structure and dynamics of the human DNAJB6b chaperone by NMR reveals insights into Hsp40-mediated proteostasis. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21529–21538 (2019).
17. Pinheiro, G. M. S., Amorim, G. C., Iqbal, A., Almeida, F. C. L. & Ramos, C. H. I. Solution NMR investigation on the structure and function of the isolated J-domain from Sis1: Evidence of transient inter-domain interactions in the full-length protein. *Arch. Biochem. Biophys.* **669**, 71–79 (2019).
18. Dosset, P. *et al.* Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J Biomol NMR* 23–28 (2000) doi:https://doi.org/10.1023/A:1008305808620.
19. Berjanskii, M. V. & Wishart, D. S. A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *J. Am. Chem. Soc.* **127**, 14970–14971 (2005).
20. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
21. Ronneberger, O. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, (2021).
22. Mirdita, M., Schütze, K., Moriwaki, Y. & Heo, L. ColabFold : making protein folding accessible to all. **19**, (2022).
23. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: A protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394 (2015).
24. Columbus, L. & Kroncke, B. *Solution NMR Structure Determination of Polytopic α - Helical Membrane Proteins : A Guide to Spin Label Paramagnetic Relaxation Enhancement Restraints*. *Membrane proteins Engineering, Purification and Crystallization* vol. 557 (Elsevier Inc., 2015).
25. Van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
26. Humphrey, W., Dalke, A. & Schulten, K. VMD : Visual Molecular Dynamics. *J. Mol. Graph.* **7855**, 33–38 (1996).
27. Suzuki, H. *et al.* Peptide-binding sites as revealed by the crystal structures of the human Hsp40 Hdj1 c-terminal domain in complex with the octapeptide from human Hsp70. *Biochemistry* **49**, 8577–8584 (2010).
28. Kityk, R., Kopp, J. & Mayer, M. P. Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. *Mol. Cell* **69**, 227-237.e4 (2018).
29. Honorato, R. V. *et al.* Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem.

Front. Mol. Biosci. **8**, 1–7 (2021).

30. Scheufler, C. *et al.* Structure of TPR domain-peptide complexes: Critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* **101**, 199–210 (2000).
31. Alvira, S. *et al.* Structural characterization of the substrate transfer mechanism in Hsp70/Hsp90 folding machinery mediated by Hop. *Nat. Commun.* **5**, (2014).
32. Borges, J. C. *et al.* Identification of Regions Involved in Substrate Binding and Dimer Stabilization within the Central Domains of Yeast Hsp40 Sis1. *PLoS One* **7**, 1–15 (2012).
33. Ambrose, A. J. & Chapman, E. Function, Therapeutic Potential, and Inhibition of Hsp70 Chaperones. *J. Med. Chem.* **64**, 7060–7082 (2021).
34. Kampinga, H. H. *et al.* Function, evolution, and structure of J-domain proteins. *Cell Stress Chaperones* **24**, 7–15 (2019).
35. Schilke, B. A. *et al.* Broadening the functionality of a J-protein/Hsp70 molecular chaperone system. *PLoS Genet.* **13**, 1–29 (2017).
36. Silva, J. C., Borges, J. C., Cyr, D. M., Ramos, C. H. I. & Torriani, I. L. Central domain deletions affect the SAXS solution structure and function of Yeast Hsp40 proteins Sis1 and Ydj1. *BMC Struct. Biol.* **11**, 40 (2011).
37. Pinheiro, G. M. S., Amorim, G. C., Iqbal, A., Ramos, C. H. I. & Almeida, F. C. L. 1H, 15N and 13C resonance assignments of the J-domain of co-chaperone Sis1 from Saccharomyces cerevisiae. *Biomol. NMR Assign.* **12**, 279–281 (2018).
38. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, (1995).
39. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Bioinforma.* **59**, 687–696 (2005).
40. Maciejewski, M. W. *et al.* NMRbox: A Resource for Biomolecular NMR Computation. *Biophys. J.* **112**, (2017).
41. Sanches, K., Caruso, I. P., Almeida, F. C. L. & Melo, F. A. The dynamics of free and phosphopeptide-bound Grb2-SH2 reveals two dynamically independent subdomains and an encounter complex with fuzzy interactions. *Sci. Rep.* **10**, 1–13 (2020).
42. Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).
43. Solyom, Z. *et al.* BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J. Biomol. NMR* **55**, 311–321 (2013).
44. Sjodt, M. & Clubb, R. Nitroxide Labeling of Proteins and the Determination of Paramagnetic Relaxation Derived Distance Restraints for NMR Studies. *Bio-Protocol* **7**, 1–14 (2017).

45. Linge, J. P., O'Donoghue, S. I. & Nilges, M. Automated Assignment of Ambiguous Nuclear Overhauser Effects with ARIA. in 71–90 (2001). doi:10.1016/S0076-6879(01)39310-2.
46. Rieping, W. *et al.* ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382 (2007).
47. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733 (2007).
48. Bhattacharya, A., Tejero, R. & Montelione, G. T. Evaluating protein structures determined by structural genomics consortia. *Proteins Struct. Funct. Bioinforma.* **66**, 778–795 (2006).
49. DeLano, W. L. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. protein Crystallogr.* **40**, 82–92 (2002).
50. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**–**2**, 19–25 (2015).
51. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
52. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
53. Lemkul, A. J. Scripts and Programs. https://osf.io/bafn4/ (2018).
54. Pronk, S. *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).

Acknowledgements

The authors thank Dr. Bernhard Brutsche (responsible scientist), Dr. JDr. Adrien Favier (NMR platform engineer), Dr Jérôme Boisbouvier (Researcher) for help with the acquisition of NMR data. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG ; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), with financial support from the TGIR-RMN-THC Fr3050 CNRS and the National Center of Nuclear Magnetic Resonance (CNRMN/UFRJ, https://www.cenabio.ufrj.br/index.php). This work was supported by FAPESP (2012/50161-8; 2017/26131-5). CHIR and FCLA have research fellowships from CNPq (305148-2019-2 and 313517/2021-5, respectively). FCLA have a research fellowship of FAPERJ (273303, 204432 and 267010). The following authors received research fellowship from FAPESP: COM (2019/16114-1), and GMSP (2017/01074-9; 2018/11948-9).

Author contributions

Data collection, analysis, and interpretation: COM, GMSP, FCLA, GCA, IPC. Data analysis and interpretation and design of the work: CHIR.; FCLA. This work results from the collaboration of two research groups: Laboratory of Biochemistry of the Chaperome (PI CHIR) and Biomolecular NMR Laboratory (PI FCLA). All authors: drafted and critically reviewed the article.

Competing interest statement: The authors declare no competing interests.