**Dock-and-lock binding of SxIP ligands is required for stable and selective EB1 interactions.**

Teresa Almeida[1,*], Eleanor Hargreaves[1,*], Tobias Zech[2], Igor Barsukov[1,#]

1 Department of Biochemistry, Cell Signalling and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

2 Department of Molecular & Clinical Cancer Medicine, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

* Authors contributed equally

# Corresponding author, email: igb2@liverpool.ac.uk

Running title: Dock-and-lock binding of EB1 SxIP ligands

Key words: EB1, microtubules, MACF, cytoskeleton.

## SUMMARY

End Binding protein 1 (EB1) is a key component of the signalling networks located at the plus ends of microtubules. It incorporates an N-terminal microtubule binding CH domain and the C-terminal EBH domain that interacts with the SxIP-containing sequences of other microtubule plus end tracking proteins (+TIPs). By using a series of SxIP containing peptides derived from the microtubule-actin cross-linking factor, MACF, we show that the SxIP motif itself binds to EBH with low affinity, and the full interaction requires contribution of the post-SxIP residues. Based on the solution structure and dynamics of the EBH/MACF complex we proposed a two-step 'dock-and-lock' model for the EBH interaction with targets, where the SxIP motif initially binds to a partially-formed EBH pocket, which subsequently induces folding of the unstructured C-terminus and transition to the stable complex. We dissect contributions from different interactions into the binding and design MACF mutations of the post-SxIP region that enhance the affinity by two orders of magnitude, leading to a nanomolar interaction. We verify the enhanced recruitment of the mutated peptide to the dynamic plus ends of MTs in a live cell experiment. Our model explains EB1's interaction with the SxIP-containing ligands and can be used to design of small molecule inhibitors that can block SxIP interaction with EB1.

## INTRODUCTION

Microtubules (MTs) are fundamental for many cellular processes, such as intracellular transport, cell shape and organization, polarity and division. A large number of proteins are associated with MT functions, stabilizing or destabilizing MTs, guiding their growth or linking them to other cell components (Akhmanova & Steinmetz, 2008; Howard & Hyman, 2003). Plus-end tracking proteins (+TIPs) are a very diverse group of proteins that accumulate at the growing (plus) ends of MTs. They form complex interaction networks that are fundamental for the MT dynamics and signalling properties. Disruption of plus end regulation in diseases like

cancer can lead to abnormal cell division and migration, facilitating disease progression (Aseervatham 2020). Drugs targeting MTs are currently widely use in cancer therapies.

End Binding protein 1 (EB1) is a key member of +TIPs networks that autonomously tracks the microtubule plus ends via its N-terminal CH domain. CH binding to MTs affects MT stability (Zhang, Alushin et al. 2015). The C-terminal EBH domain, on the other hand, regulates +TIPs networks by recruiting a range of other +TIPs, such as cancer related proteins APC and MACF, through the interaction with the SxIP motifs (Gouveia & Akhmanova, 2010; Honnappa et al., 2009; Kumar et al., 2012). These small four residue conserved motifs are usually located in the C-terminal intrinsically disordered regions of +TIPs (Buey et al., 2012; Jiang et al., 2012).

The crystal structure of the EB1 EBH domain in complex with MACF SxIP peptide (Figure 1A) showed that the motif fits into a deep hydrophobic pocket between the C-terminal loop and the long helix of the leucin zipper (Honnappa, Gouveia et al. 2009). The conserved Ser forms a network of hydrophobic interactions, and Ile and Pro are embedded into a deep hydrophobic pocket; these interactions were proposed to be the main determinants of the ligand binding. The additional contributions of the non-SxIP residues into the binding has been shown through the systematic single-residue substation analysis that highlighted the importance of the hydrophobic residues that immediately follow SxIP and the positive charges of the subsequent residues (Buey, Sen et al. 2012). However, it has not been clear to what extent these residues change the overall affinity.

In the solution structure of free EBH the region corresponding to the C-terminal loop is unstructured and dynamic and the binding pocket is only partially formed (Figure 1B) (Almeida, Carnell et al. 2017), leading to the conclusion that the binding of the MACF peptide induces the folding of the C-terminus to complete the binding pocket. However, we found that the 4-residue peptide SKIP fragment of MACF has a very low affinity to EBH and induces only small chemical shift changes in the $^{1}$H,$^{15}$N-HSQC, compared to an 11-residue MACF SxIP-containing peptide, demonstrating that post-SxIP residues have a critical contribution into the complex formation.

Here, we characterize the mechanism and outline the structural features of EBH interactions with the MACF-derived SxIP peptides using biophysical techniques such as NMR and ITC. We show that the post-SxIP residues induce the folding of the EBH C-terminus, thus making a large additional contribution to the peptide binding. To evaluate this contribution, we propose a two-step 'dock-and-lock' model for the peptide interaction with EBH, where the SxIP motif initially docks into the partly formed binding pocket, which subsequently induces the C-terminus folding. Using systematic modifications of EBH and the peptide, we evaluate the thermodynamic and kinetic parameters of the binding. We dissect the contributions of the different peptide regions into the binding and design a peptide with nanomolar affinity to EBH. Our model can be used for the prediction of the EBH affinity to different natural ligands and as a rational for the design of small molecules to inhibit EB1 interactions that have a recognized contribution into human neoplasic diseases such as cancer. (Abiatiari et al., 2009; Liu et al., 2009).

## RESULTS

### SxIP gives specificity in binding to EB1, but flanking residues strengthen the interaction

Our previous analysis suggested that the folding of the C-terminus induced by the SxIP peptide is required for high binding affinity (Almeida, Carnell et al. 2017). To investigate the roles of different peptide regions in the folding of the C-terminus and the connection between folding and affinity we compared the binding of the 4 and 6 residue fragments, SKIP (4MACF) and SKIPTP (6MACF), with the 11-residue MACF peptide KPSKIPTPQRK (11MACF) that was used to solve the crystal structure of the complex (Honnappa, Gouveia et al. 2009). These peptides induced progressively larger changes in the NMR spectra of the $^{15}$N-labelled EBH domain (Figure 1C and S1). For 4MACF and 6MACF we observed fast exchange between the free and bound states, characterised by a progressive linear change of the chemical shifts in the $^{1}$H,$^{15}$N-HSQC spectra of EBH on peptide addition. This corresponds to a weak EBH interaction with the peptides. Fitting the chemical shift changes into a two-state binding model (Figure S1D and E) estimates the dissociation constant $K_d$ of the biding as ~10mM and ~2mM for 4MACF and 6MACF, respectively, similar to the $K_d$ values we measured previously for the SxIP-like molecules (Almeida, Carnell et al. 2017). For the 11MACF peptide many signals are in the slow exchange regime, in agreement with a much stronger interaction ($K_d$ 3.5 µM from ITC, Table 1 and Figure S1).

The chemical shift changes in the $^{1}$H,$^{15}$N-HSQC spectra for the peptides map to increasingly larger regions on the EBH surface (Figure S2). Notably, the changes for the 4 to 6 residue peptides are located predominantly in the structured part of EBH, with relatively small effect on the unstructured C-terminal region. In contrast, the 11-residue peptide induced the largest chemical shift perturbations in the C-terminal region (Figures 1A and 1SA). Additionally, we observed a different effect on the intensity of the $^{1}$H,$^{15}$N-HSQC cross-peaks assigned to the C-terminal region of EBH by the peptides (Figure 1D). In the free EBH the intensity of the signals from C-terminal the region E251-G260 are much higher than the signals of the coiled-coil part of the protein, as expected for the dynamical C-terminus. In the presence of the 4 and 6-residue peptides the intensities of these signals remain much higher than the signals corresponding to the coiled-coil part folded part of EBH, while in the presence of the 11-residue peptide the intensities of many C-terminal signals became similar to those of the folded part. The combination of the chemical shift and intensity changes suggest that the folding of the C-terminal regions seen in the crystal structure of the complex is induced not by SxIP motif, but by the residues that follow it, and that this folding leads to the much higher binding affinity. While not conserved in the ligand sequence and not defining ligand specificity (Honnappa, Gouveia et al. 2009), the post-SxIP residues clearly have a critical role in the interaction.

To define the contribution of the EBH C-terminus, we analysed peptide binding to the EBH-ΔC fragment lacking the C-terminal SxIP-binding region. For the signals of the EBH-ΔC coiled-coil region we observed similar changes on peptide addition as in the EBH spectra, however the amplitudes of the changes were reduced, corresponding to much weaker interactions (Figure S3A-C). For the 4MACF peptide the interactions were too weak to estimate $K_d$, while the affinities to 6MACF and 11MACF are reduced 4 and 10-fold, respectively (Figure S3D and E, and Table 1). The ITC thermodynamic parameters show that for 11MACF the reduction in the affinity is caused by the large decrease of the absolute value of the interaction enthalpy, partly compensated by the reduction of the entropy loss (Table 1, Figure S3F). This agrees with the partial loss of the binding pocket and the lack of the additional ordering of the unstructured C-terminus, as the residues that become immobilised in the complex are missing. Notably, chemical shift perturbations in the coiled-coil part of EBH are less extensive for the 11MACF (Figure S2F), indicating that in the absence of the fully formed binding pocket the C-terminal region of the peptide has a reduced contact with the coil-coiled region.

The above observations support a 'dock-and-lock' model for the SxIP binding where the SxIP-motif itself only provides the initial recognition, while the full affinity of the interaction is controlled by the folding of the EBH unstructured C-terminal, leading to the additional interactions with the peptide residues immediately following the SxIP motif (Figure 1E). To validate this model, we analysed the solution structure of EBH/MACF complex and measured the interaction dynamics for a series of peptides.

**Peptide binding changes the structure and dynamics of the EBH C-terminus**.

Following the 11MACF binding, chemical shifts of the EBH C-terminal signals change dramatically and their intensity becomes comparable to the signals on the structured regions of EBH (Figure 1C and D). To relate the changes in the spectra to the structural rearrangement on binding in solution we solved the NMR structure of EBH following the methodology described in (Almeida, Carnell et al. 2017). For the coiled-coil EBH region that is structured in the free form (leucin zipper and the 4-helix bundle) we observed a similar pattern of intra- and inter-molecular contacts in the complex, and similar values of the dihedral angles derived from the $^{13}$C-chemical shifts as for the free EBH. In contrast, a large number of new intra-molecular contacts were detected in the C-terminal regions, as well as the inter-molecular contacts between the 11MACF peptide and EBH. This directly demonstrates that the peptide binding has a limited effect on the overall fold of EBH, and most of the structural changes occur in the C-terminal region.

The solution structure determined from the NMR data shows that EBH in the complex with 11MACF has a coiled-coil structure consisting of a leucine zipper and a four helical bundle in the upper region that is practically identical to the structure of the free form (Figure 2A). The C-terminus folds over the peptide (Figure 2B), forming a loop similar to the loop observed in the crystal form (Honnappa, Gouveia et al. 2009). Only a small number of contacts were detected for the N-terminal KPSK region of the peptide, including Ser5477, that shows a single distance contact to Phe218. This suggests a limited role of this region in the interaction with EBH. Interestingly, an extensive H-bond network involving Ser5477 has been identified in the crystal structure (Honnappa, Gouveia et al. 2009). The lack of NOE contacts of Ser5477 suggest that H-bonds in solution are transient and play less significant structural role than has been expected from the crystal structure. The following residue, Lys5478, is solvent exposed and has no NOE contacts with the helical part of EBH. In contrast, the side-chain Ile5479 shows the largest number of NOE distance contacts to the residues of coiled-coil part of EBH, including Tyr217, Phe218, Leu221, Arg222, Ile224, Glu225, Leu246 and Tyr247. The side-chains of these residues form a deep hydrophobic pocket that matches the shape of Ile. Similarly, Pro5480 shows hydrophobic contacts to Tyr217, Phe218 and further down with Thr249. Both Pro5480 and Thr5481, engage with the lower region of the four-helix bundle, just below the SxIP binding site, with NOEs to the aromatic patch $^{216}$FYF$^{218}$.

Additionally, the $^{5480}$PT$^{5481}$ part of the peptide makes numerous contacts with the folded hydrophobic region $^{253}$FVIP$^{256}$ of the C-terminus (Figure 2B). These interactions position the C-terminus in the linear conformation on top of the peptide, protecting it from the solvent. The aromatic ring of Phe253 in the EBH C-terminus slots between the two Pro residues of the peptide and stacks against the ring of Phe261 of the coiled-coil region, anchoring the folded loop to the EBH core. This is the only direct interaction between the C-terminus and the EBH core; in the absence of the peptide, it would be insufficient to stabilise the structure of the C-terminus. This agrees with the lack of EBH C-terminal folding in the complexes with the shorter MACF peptides (see above).

To assess the dynamic properties of the loop in the free EBH and the peptide complex we measured $^{15}N$ relaxation at 600 and 800 MHz NMR fields, which gives information on both slow and fast exchange processes. At both magnetic fields the relaxations parameters showed matched sequence distribution profiles (Figure S4). The elevated R2/R1 ratios and NOE values, corresponding to the low internal mobility, are located in the helical regions of EBH structure, while the reduced values, corresponding to the high dynamics, in the loop between the helices, and the EBH C-terminus (Figure S4). The order parameter $S^2$ that quantitatively characterises the fast motion, and the exchange contribution into the relaxation $R_{ex}$ that characterises slow structural rearrangements have been derived by the model-free analysis in the Relax software (d'Auvergne and Gooley 2008). Mapping of these values on the EBH structure (Figure 2C) shows that in the free EBH the whole C-terminus is highly dynamic, in agreement with the variations in the NMR structures. Surprisingly, high dynamics is also observed in the loop between the helices that is well-determined in the NMR structures. This discrepancy is likely caused by the multiple conformations of the loop leading to partially contradicting NMR restraints. In such cases the NMR structures represent an averaged conformation. In the EBH/11MACF complex only the last four residues are dynamic, the rest of the C-terminus is fully immobilised. Notably, the $S^2$ values of the C-terminus are very close to the values in the helical parts, demonstrating that the structure of the C-terminus in the complex is rigid despite the high initial dynamics of the C-terminus of free EBH and the free peptide in solution. Surprisingly, the low values of $R_{ex}$ show the lack of slow conformational changes in the immobilised C-terminus that are often observed in similar systems. Once formed, the complex is very rigid, suggesting well-matching structures of the peptide and the EBH C-terminus.

The structural and dynamics analysis supports a two-step model of the SxIP peptide binding (Figure 1E). In the free form only the SxIP-recognition part of the binding pocket is formed that allows the initial docking of the peptide predominantly through the hydrophobic interaction of Ile5479. These interactions are insufficient for strong binding because they involve only a small number of the peptide residues. The initial docking induces the folding of the EBH C-terminus through the complementarity between the hydrophobic region $^{253}$FVIP$^{256}$ of the C-terminus and the peptide residues that immediately follow the SxIP motif. The folded C-terminus is fixed to the coiled-coil region through the aromatic stacking of Phe253 and Phe261. The additional interactions induced by the folding of the C-terminus extend the binding pocket and increase the affinity 20-fold, based on the EBH C-terminal deletion experiments. Next, we used the binding model to design peptides with enhanced binding affinity and to dissect contributions from different peptide regions to the binding.

**Peptide optimisation in the post-SxIP region.**

The folding of the EBH C-terminus after the peptide binding brings the highly hydrophobic region $^{253}$FVIP$^{256}$ of EBH into the contact with a relatively hydrophilic $^{5481}$TPQ$^{5483}$ region of the peptide (Figure 3A). We reasoned that the affinity of the interaction can be enhanced by making this area of the peptide more hydrophobic. To assess a suitable substitution, we docked three versions of the modified peptides into the EBH structure of the complex with MACF: KPSKIP<u>LLL</u>RK (11MACF-LLL), KPSKIP<u>VLL</u>RK (11MACF-VLL) and KPSKP<u>ILL</u>RK (11MACF-ILL), Figure S5. All variants showed a high complementarity to the binding site. The highest fitness score was obtained for 11MACF-LLL, followed by 11MACF-VLL. The score of 11MACF-ILL was similar to that of the wild-type. Based on the docking, KPSKIP<u>LLL</u>RK and KPSKIP<u>VLL</u>RK were selected for the experimental analysis.

The ITC data (Figure 3B and C, Table 1) show a 10-fold increase in binding affinity for 11MACF-LLL and a further 3-fold increase for 11MACF-VLL, when compared with the wild-type 11MACF. Notably, the enthalpy of the interaction of the 11MACF-LLL variant is lower than for the WT and 11MACF-VLL, suggesting a less optimal match to the binding pocket. This is compensated by a favourable change in the entropy, implying an increased hydrophobic effect, as predicted by the modelling. The binding enthalpy for 11MACF-VLL was similar to the WT, indicating a good fit to the binding pocket. The change in the entropy is less favourable than for 11MACF-LLL, agreeing with the reduced hydrophobicity. However, the net change in the negative free energy is larger for 11MACF-VLL, explaining the affinity increase.

The positively charged RK region at the C-terminus of the peptide does not make any stable contacts with EBH. However, this part of the peptide is close to the negatively charged patch on the EBH surface (Figure 3B). We speculated that this proximity creates a favourable electrostatic interaction, enhancing the affinity of the peptide. To test this prediction, we duplicated the RK sequence at the C-terminus in the extended peptide KPSKIPVLLRKRK (11MACF-VLLRK). In agreement with the prediction ITC showed further a 5-fold increase in affinity, leading to $K_d$ of 16 nM (Table 1). Overall, through the rational sequence modification we enhance the affinity by nearly 2 orders of magnitude, changing the binding from the micromolar to nanomolar range.

Comparison of the $^{1}H,^{15}N$-HSQC spectra of EBH in the free form and different EBH complexes reflects the structural changes associated with the changes in affinity (Figure 4A and S6). The binding of the WT 11MACF leads to the large changes in the positions of the signals of the residues in the binding site through the direct contact. Additionally, the peptide induces the folding of the C-terminus, introducing even larger changes of the corresponding signals. Because of the extensive interaction between EBH and the peptide, nearly all $^{1}H,^{15}N$-HSQC signals change position in the complex. Modification from TPQ to LLL results in further large chemical shift changes, primarily in the signals of the folded hydrophobic part of the C-terminus ($^{253}$FVIP$^{256}$) that interacts with the modified peptide region, and of Phe218 that forms an aromatic stacking contact with Phe253. No significant changes are observed in the signals of the residues forming the SxIP binding site. This suggests structural rearrangement of the C-terminus to match the modified, more hydrophobic peptide sequence. On the LLL to VLL change in 11MACF-VLL the spectral changes are much smaller and in the same C-terminal EBH region, demonstrating further, small structural rearrangements. The addition of the positively charged RK residues in 11MACF-VLLRK have minimal effect on the spectra, indicating that the contacts of the charged residues are transient.

Taking ITC and NMR data together, the described peptide modifications optimise the property of the post-SxIP region for high affinity. The residues immediately following SxIP are engaged in the hydrophobic interactions and have to match the complementary FVIP region of EBH. This region is initially unstructured, allowing it to fit different peptide sequences. The closer fit would increase the affinity; however, the specificity of these interactions is limited because the flexible C-terminus can adapt to different sequence. This is reflected in the low sequence conservation beyond the SxIP motif (Figure S7). Additional, smaller contribution into the affinity comes from the non-specific charge interactions between the positive C-terminus and the negative surface of EBH. The separation of the specificity and affinity of the peptide binding shows that different sets of interaction control each stage of the dock-and-lock model (Figure 1E). Additionally, transition to the full complex (lock) depends on the kinetics of C-terminus folding induced by the post-SxIP region of the peptide. To evaluate the kinetic parameters and

to develop the binding model further we investigated the kinetics of the EBH interaction with WT and modified MACF peptides.

**Kinetics of the target peptide recognition by EBH**

NMR spectroscopy provides a diversity of methods to measure exchange processes at a wide range of timescales (Furukawa, Konuma et al. 2016, Camacho-Zarco, Schnapka et al. 2022). The $^{15}$N relaxation analysis (see above) demonstrates the lack of fast and slow conformational changes in the folded EBH regions both in the free form and in the complex with 11MACF (Figure 2C). This suggests low population of the intermediate “dock” state and/or fast transition to the final “lock” state of the binding model (Figure 1E). This conclusion is further supported by the lack of the exchange contribution in the relaxation dispersion experiments even for the residues with the largest chemical shift changes of ~2000 Hz (data not shown). This prevents us from the direct evaluation of the kinetic parameters for the exchange between the “lock” and the “dock” stages. However, the parameters for the initial peptide docking can be measured with the truncated EBH-ΔC, where the C-terminal hydrophobic residues that contact post-SxIP region of the peptide are removed. Then the kinetic parameters of the second “lock” stage can be estimated from the overall exchange rate in the presence of the C-terminus.

The effect of the chemical exchange is clearly manifested in the exchange broadening of the EBH signals in the $^{1}$H,$^{15}$N-HSQC spectra in the course of the titrations (Figure S1). To evaluate the exchange rates we fitted the spectral changes into the exchange model using TITAN software (Waudby, Ramos et al. 2016). Large changes of the EBH chemical shifts that are observed in many signals on binding allow probing a wide range of the exchange rates for different regions of EBH (see supplementary for more information on the TITAN analysis).

For the quantitative binding analysis in the TITAN software we selected non-overlapped signals with clear changes in the peak amplitude and position in the course of the titration (Figure 4B and S7). When spectral changes were fitted independently for each residue, the $K_d$ values and exchange rates were similar, indicating that binding kinetics is similar for different regions of the binding site. We therefore used a global fitting of all signals to evaluate the overall kinetic parameters of the binding. The $K_d$ values obtained independently from the TITAN lineshape analysis were similar to those measured by ITC, validating the NMR analysis (Table 1).

The dissociation rate of 130 $s^{-1}$ evaluated for the EBH/11MACF interaction with TITAN is optimal for the chemical exchange saturation transfer (CEST) NMR experiments, where saturation is transferred from the high-population free state to the low-population complex (ref). These experiments measure the exchange rate directly and are conducted at the low ligand concentrations, where population of the fully bound EBH dimer with two peptides is much lower than that of a single peptide complex. This makes the indirect allosteric effect on the binding negligible and simplifies the binding model, thus providing direct information on the peptide interaction with the free protein. The large change of the $^{15}$N chemical shifts for the residues in both the pre-formed part of the binding pocket (residues 220, 228, 247, 248 and 249) and the dynamic C-terminus that folds on the peptide binding (residues 250, 252, 255 and 255) allows independent measurement of the kinetic parameters for these two regions. As expected, we detected a strong CEST effect for all the residues with sufficiently large changes in the $^{15}$N chemical shifts (Figure 4C and S8). Independent fitting of the CEST profiles for the individual residues (Figure S9) gave very similar dissociation rates for all the residues. We therefore obtained the overall exchange rates by simultaneous global fitting of all CEST profiles (Figure 4C). The calculated dissociation rate of 120 s-1 was in excellent

agreement with the line-shape TITAN analysis (Table 1). For the weaker EBH-ΔC/11MACF interaction we did not detect any CEST effect, in agreement with much faster dissociation rate of 1900 $s^{-1}$ evaluated by the lineshape analysis. For the stronger EBH/11MACF-VLL interaction the intensity of the CEST effect was significantly lower that for EBH/11MACF at the same protein:peptide ratio, the exchange rates from the CEST were significantly higher than from the line-shape analysis, and protein:peptide ratio estimated by CEST was several times lower than used in the experiment. This indicates that the exchange rates are too slow for the accurate CEST analysis, in agreement with the low rate of 15.6 $s^{-1}$ obtained by the lineshape fitting.

**Model of the target recognition by EBH**

Based on the NMR structural and binding analysis we proposed a two-step 'dock-and-lock' binding model (see above). We can now evaluate the binding parameters for each stage using the exchange parameters derived from the line-shape changes and CEST experiments. The C-terminal deletion of the EBH-ΔC fragment eliminates the 'lock' step and can be used to calculate the parameters for the docking. The fitting of the NMR line-shape changes to a two-state model gives $K_d$ of 27 uM and the dissociations rate of 1900 $s^{-1}$ (Table 1). The dissociation constant from the NMR analysis is close to the value of 41 uM from the ITC measurements. Since ITC is generally a more accurate method for the binding constant measurement, we use the ITC value as the dissociation constant for the dock stage. The on-rate for the docking can then be evaluated as $4.6 \times 10^7$ $s^{-1}M^{-1}$, which is similar to the diffusion-controlled binding (summarised in (Shammas, Crabtree et al. 2016)). This agrees with the easy access to the open, partly formed binding pocket of EBH.

When analysing the line-shape changes for the full EBH domain interactions with 11MACF and the mutated peptides, we did not detect separate signals from the intermediate 'dock' stage despite the large chemical shift difference between the free and the bound states and slow-exchange regime. In addition, we did not observe any exchange broadening for EBH complex formed at the high access of the peptide, where the population of the free state is negligibly low. (Under these conditions the complex is in the equilibrium between the 'dock' state where C-terminal region unfolded with the chemical shifts similar to the free state, and the locked complex with the fully folded C-terminus.) The lack of signals from the 'dock' state and the absence of the exchange contribution from the exchange between the 'dock' and the 'lock' states suggests that the exchange between these states is very fast compared to the chemical shift differences, or the population of the intermediate 'dock' state is very low.

Under the conditions of the fast exchange between the two bound states ('dock' and the 'lock') they can be represented by an average state, and the binding approximated by a two-state model (Brezina, Hanykova et al. 2022) with the overall dissociation constant $K_D'$ and rate $k_{off}'$ represented by:

$$K_D' = K_D^D \frac{K_D^L}{K_D^L + 1} \tag{1}$$

$$k'_{off} = k^D_{off} \frac{K^L_D}{K^L_D + 1} \quad (2)$$

where $K^D_D$ and $k^D_{off}$ are the dissociation constant and rate of the dock step, and $K^L_D$ is the dissociation constant of the lock step. These equations show that the overall dissociation constant $K'_D$ and rate $k'_{off}$ are scaled down proportionally to the fraction of the protein in the intermediate state relative to all bound population. Effectively, the 'lock' step depletes the population of the protein where ligand can dissociate, thus decreasing the overall dissociation constant and dissociation rate. If $K^L_D \ll 1$, corresponding to the case when most of the bound protein is in the final locked state, the overall constants decrease in proportion to the dissociation constant of the lock state.

The parameters of the 'dock' step can be estimated from the peptide interaction with EBH-ΔC that lacks the C-terminal and cannot transfer to the locked state (Figure 1E). The dissociation constant of the 'lock' step can be then calculated from Equations 1 and 2. Thus for 11MACF $K^D_D = 41.5$ μM and $k^D_{off} = 1900$ s$^{-1}$, and $K'_D = 3.5$ μM and rate $k'_{off} = 130$ s$^{-1}$, (Table 1), giving $K^L_D = 0.08 - 0.07$. This corresponds to only ~8% of the bound population is in the intermediate dock state.

The low population of the 'dock' state could potentially still be sufficient to cause exchange line-broadening under fully bound condition (large ligand excess) if the exchange rates for the lock process are sufficiently slow. The absence of the broadening suggests fast exchange on NMR time-scale, with $k^L_{ex} = k^L_{on} + k^L_{off} \gg 2000\ s^{-1}$ (from the maximum differences in the chemical shifts between the free and the bound states). With $K^L_{eq} \ll 1$, the exchange is dominated by the on-rate, giving $k^D_{on} \gg 2000\ s^{-1}$. This rate is much faster that the dissociation rate of the dock stage $k^D_{off} = 1900\ s^{-1}$, supporting the fast exchange approximation.

When the population of the intermediate state is low, its contribution into the NMR signals becomes negligibly small and the measured line-shape changes and CEST saturation transfer are only associated with the exchange between the free and the fully bound 'lock' forms (see supplementary for more information). It can then be shown that an additional 'dock' step has a generally larger effect on the overall dissociation rate than on the overall dissociation constant, with the decrease of both parameters becoming equal in fast exchange (see supplementary for more information):

$$k'_{off}/k^D_{off} \leq K'_D/K^D_D \quad (3)$$

Using the parameters EBH-ΔC/11MACF binding as the approximation for the first 'dock' step and the EBH/11MACF binding for the overall parameters of the two-step model (Table 1), we estimate the ratios $k'_{off}/k^D_{off} = 0.068$ and $K'_D/K^D_D = 0.084$ (12 and 15-fold decrease in the dissociation constant and rate, respectively), in agreement Equation 3. The similarity of these ratios further supports the fast exchange rate of the 'lock' step. The estimated kinetic parameters for the EBH/11MACF interaction are summarised in Figure 5A.

For the EBH/11MACF-VLL interaction the effective overall constants are $K'_D = 0.08$ μM and rate $k'_{off} = 15.6$ s$^{-1}$, showing 500 and 122-fold decrease, respectively, compared to the 'dock' stage constants from the EBH-ΔC/11MACF binding (Table 1). The much larger decrease in

the dissociation constant compared to the decrease in the dissociation rate clearly disagrees with the two-step model prediction (Equation 3). A likely explanation for this is an additional interaction of the hydrophobic VLL region with the coiled-coil, which would enhance affinity to EBH-ΔC the interaction and decrease $K_D^D$ value compared to the 11MACF peptide. Assuming the fast exchange, when $K_D'/K_D^D = k_{off}'/k_{off}^D$, we can calculate the expected dissociation constant for 'dock' step as $K_D^D = k_{off}^D/k_{off}' * K_D' = 10$ uM and from Equation 1 $K_D^L = 0.008$.

The decrease in $K_D^L$ for 11MACF-VLL compared to 11MACF corresponds to decrease in the population of the intermediate state from 0.08 to 0.008. This is expected to cause a small increase in the difference in the chemical shift values between the free and the bound forms for the signals of the C-terminal residues of EBH. In agreement with this prediction, we observed a general systematic increase of the chemical shift differences for the residues at the C-terminus most affected by the peptide binding for 11MACF, 11MACF-VLL and 11MACF-VLL (Figure 4A and S8). This change supports the fast exchange of the C-terminus between the free and the bound state in the saturated complex, although non-linear shift changes between the complexes also indicate some conformational difference in the complex.

The dissociation constants measured for the different peptides (Table 1) allow us to define contributions from different regions into the free energy of the binding (Figure 4E). These contributions can be assigned to the 'dock' and 'lock' stages. The short fragments 4MACF and 6MACF do not cause folding of the EBH C-terminus and, together with the EBH-ΔC/11MACF binding show a steady decrease of the free energy on the increase of the peptide size, from ΔG = -11.2 kJ/mol for 4MACF to -25.2 kJ/mol for the full 11MACF. This demonstrates that approximately half of the free energy change for the initial stage is defined by the recognition of the SxIP motif, and the rest by the non-specific binding of the variable sequence that immediately follows SxIP. The high contribution from SxIP makes the dock stage very sensitive to single-residue mutations in the recognition part of the peptide. The folding of the C-terminus adds -6 kJ/mol, released in the lock stage of the interaction, bringing the overall $K_D'$ from high to low-micromolar range. The TPQ sequence of 11MACF is not optimal for the interaction with the hydrophobic EBH C-terminus. Replacing this with a hydrophobic VLL sequence improves the match and reduces free energy by further -6.1 kJ/mol. Further reduction of the free energy of ~ -4.6 kJ/mol is possible by adding positively charged residues to the peptide C-terminus that interact non-specifically with the negatively-charge surface of EBH. The variation of the sequence in the post-SxIP region observed in the EBH ligands (Figure S7) could lead to the change of free energy that is comparable to the free energy of the SxIP binding. Overall, the change in the free energy on complex formation consists of three similar contribution: SxIP binding to the partially formed binding pocket, non-specific interactions between the post-SxIP region and the folded part of EBH, and folding-induced interaction between the unstructured EBH C-terminus and post-SxIP region of the peptide.

### High affinity peptides localise into comet-like structures in cells

Through the rational optimisations of the peptide sequences, we generated SxIP peptides with the enhanced affinity to EB1. To relate their affinity to the EB1-dependent recruitment to the plus ends of MTs, we inserted them into an mTFP containing mammalian vector and followed their ability to localise into mCherry-EB1 comets in HeLa cells. For the WT peptide ($K_d$ = 3.5 µM) we observed a ubiquitous cytosolic localisation with little to no peptide localised into comets (Figure 5A), similar to that of the GFP control. Close inspection of images and movies

at high contrast revealed small number of comets enriched in the peptide that were too faint for the quantitative analysis, comparable to the reported previously (Honnappa, Gouveia et al. 2009). A measurable localisation to comets could be detected with the MACF-VLLRK peptide ($K_d$ = 16 nM). However, the ratio of peptide to EB1 comets was extremely low at around 0.02 (Figure 5B), showing that only a small number of comets contained detectable amount of the peptide. This ratio dramatically increased to ~0.8 with the dimeric MACF-LZ-VLLRK peptide, showing that only dimeric localised to almost every plus end comet we traced. EB1 comet number remained the same with all peptide transfections (Figure 5C). These experiments suggest that the enhanced affinity increase the localisation of ligands to MTs, but even nanomolar affinity is still insufficient for recruitment of the monomeric ligand.

## DISCUSSION

The interaction of EB1 with the SxIP motif containing proteins is fundamental for the assembly of signalling complexes at the plus ends of MTs. These complexes regulate the assembly and stability of MTs themselves, guide MT growth, and link MTs to other cellular processes, such as cell division or assembly of adhesion complexes (Akhmanova & Steinmetz, 2008; Howard & Hyman, 2003). In this paper we demonstrated that while SxIP motif is the only conserved feature of the SxIP-containing proteins, it alone is insufficient for the high-affinity interaction. By comparing the interaction of the EBH domain of EB1 with a range of SxIP-containing peptides we identified the critical role of the post-SxIP region in the complex formation. These residues induce folding of the EBH unstructured C-terminus following the initial docking via the SxIP motif, which can increase affinity by several orders of magnitude. Thus, the overall binding of the SxIP-containing ligands to EBH can be described as a two-step dock-and-lock process, with the initial docking through SxIP motif providing the specificity of the interaction, and the subsequent folding of the EBH C-terminus defining the full binding affinity (Figure 4D and E). We evaluated the kinetic parameters of the binding and the contributions from different regions into the binding energy, and used the model to increase the affinity of the SxIP peptide ~100 fold through mutations of the post-SxIP region. The mutated peptide showed enhanced dynamic recruitment to the plus ends of MTs in cells, supporting the functional role of the post-SxIP region. Overall, our results agree and extend the reported analysis of the substitutions of the individual residues in the SxIP ligands (Buey, Sen et al. 2012), and provide a dynamic model for the target recognition by EB1.

Our dock-and-lock model is a variant of the induced fit that is a dominant binding mechanism of IDPs (Sugase, Dyson et al. 2007, Wright and Dyson 2009, Mollica, Bessa et al. 2016). We used a separate name for our model to highlight the clear structural definitions of each step of the binding and the lack of the internal motions in the complex. Our previously reported structure (Almeida, Carnell et al. 2017) and the relaxation analysis reported here (Figure 2C) demonstrate that in the absence of the ligand, the EBH domain has only partially formed SxIP binding site and a dynamic C-terminus. The SxIP motif itself has an extremely low affinity to EBH as it fails to induce the formation of the full binding pocket that requires the folding of the EBH C-terminus. This folding requires post-SxIP residues that immediately follow the SxIP motif. The folding is driven by the hydrophobic interaction of the FVIP region in the dynamic EBH C-terminus that folds on top of the post-SxIP region and locks it in the position through the stacking of two Phe aromatic rings (Figure 3A). These interactions completely immobilise the EBH C-terminus and the peptide ligand, making the completed

binding pocket as rigid as the rest of the EBH coiled-coil structure (Figure 2C).

This type of binding resembles the interaction of HIV-1 protease with group-specific oncogene (Gag), where the protease has dynamic flaps that fold over the bound substrate and traps it in the catalytic site, thus facilitation the catalysis (Deshmukh, Tugarinov et al. 2017). Similar loop folding has also been observed in other enzymes, such as some classes of sulfatrahsferases and lipases, where the closed loops shield the substrates from the bulk solvent and prevent non-productive hydrolysis (Cook, Wang et al. 2013, Khan, Lan et al. 2017). However, this is rather unusual in the adaptor proteins with similar to EB1 function. In the majority of the reported cases the folding is limited to the initially unstructured ligand (Morris, Torpey et al. 2021). Part of the ligand often remains dynamic in the complex, forming a set of fuzzy interactions with the folded partner (Sugase, Dyson et al. 2007).

The dock-and-lock mechanism of the EBH binding has a number of functional benefits. The open, partly formed binding pocket is easily accessible, leading to the fast on-rate. The initial, partly formed, binding pocket is very small, primarily fitting the Ile side-chain. This makes the pocket highly suitable for the recognition of the short SxIP motif. Despite the small size, the pocket has a very distinct shape, leading to a very specific recognition. However, because of the small size, the binding energy is low, resulting in the high off-rate. This explains the high specificity, but low affinity of the SxIP binding.

The additional residues that follow the SxIP motif contribute in two different ways. First, they enhance the affinity through the interaction with the folded EBH coiled-coil, with the increase of $K_d$ from 2 mM for 6MACF to 41 uM for the 11MACF interaction with EBH-ΔC (Table 1). The fast off-rate of 1900 $s^{-1}$ determined for the 11MACF/EBH-ΔC interaction from the chemical shift changes in the $^{1}H,^{15}N$-HSQC spectra corresponds to a very high on-rate of $4.7*10^{7}$ $M^{-1}s^{-1}$ calculated from the $K_d$ value. This is similar to the reported diffusion-controlled rates enhanced by the electrostatic interactions (Sugase, Dyson et al. 2007, Shammas, Travis et al. 2013), which agrees with the open binding pocket and contributions of the C-terminal positive charges of the peptide identified with the mutations and strong dependence of the affinity on the salt concentration (Buey, Sen et al. 2012). Limited chemical shift perturbations outside of the SxIP pocket for the 11MACF/EBH-ΔC interaction suggest that in the albescence of the C-terminus the interactions of the post-SxIP residues with the folded EBH domain are transient, and have more effect of the on- than the off-rate.

The second effect of the post-SxIP residues is to induce the folding of the EBH C-terminus, which locks the peptide into the fully folded state. While not available from the direct measurements, the rate of this step can be evaluated from the rates of the dock step (above) and the overall rate for the complex formation using the 2-step model. Both chemical shift analysis and CEST measurement give consistent values of for the overall off-rate of ~130 $s^{-1}$ for the 11MACF/EBH interaction. This corresponds to ~15-fold decrease from the rate 1900 $s^{-1}$ of the dock step, which is similar to the differences in the overall $K_d$ values for the 11MACF/EBH-ΔC (only dock step) and 11MACF/EBH (both dock and lock steps). The proportional decrease corresponds to the fast exchange condition for the lock step (Brezina, Hanykova et al. 2022), with forward rate >> 2000 $s^{-1}$, estimated from the chemical shift differences. This rate is much higher than the dissociation rate of the dock step. As the result, the binding is highly productive, with practically all encounter complexes proceeding to the fully bound state.

Overall, the binding mechanism of EBH maximises the on-rate through the open binding pocket optimised for the short SxIP domain recognition, making this motif the only feature

required for the target recognition. Insufficient binding affinity of the small SxIP-recognition pocket is strongly enhanced by the folding induced by the post-SxIP region. The main requirement for this region is to support the hydrophobic interactions with the EBH C-terminus. These interactions are much less specific as the dynamic C-terminus can adjust and accommodate the changes. This allows for significant sequence variation, and explains the high conservation of the SxIP motif and lack of the conservation in the post-SxIP region (Figure S7). Notably, many IDR ligands have a small number of key hydrophobic residues that are critical for the formation of the encounter complex (Sugase, Dyson et al. 2007). It is not clear yet whether these residues define small recognition motifs for all the IDR ligands, or only some IDR, such as SxIP ligands, have well-defined small recognition motifs.

Based on the understanding of the post-SxIP region contribution into the affinity, we designed an MACF mutant 13MACF-VLLRK with a highly increased affinity from 3.5 uM to 16 nM. In this design we increased the hydrophobicity of the immediate post-SxIP region that interacts with the EBH C-terminus, and introduced additional charges to at the peptide C-terminus for enhance the electrostatic steering effect. These substitutions agree with the reported previously systematic single-residue replacement analysis (Buey, Sen et al. 2012), extending it to the multi-residue substitutions. The high affinity of the mutant that has a small number of changes demonstrates the scope of the $K_d$ modulation through the variation in the post-SxIP regions of different EBH ligands shown by the sequence comparison (Figure S7).

Endogenous EB1 forms dynamic condensates via phase separation that track plus ends of MTs in comet-like structures (Song, Yang et al. 2023). These structures recruit many of the SxIP-containing +TIPs proteins such as MACF and APC (Kita *et al.*, 2006; Honnappa *et al.*, 2009) and regulate microtubule polymerisation, catastrophe and rescue (Vaughan, 2005). The SxIP +TIP proteins effectively localise to EB1 comets (Akhmanova and Steinmetz 2015) and condensates (Song, Yang et al. 2023), and enhance condensate formation to a different degree (Song, Yang et al. 2023). We found that the comet localisation depends on the SxIP-peptide affinity, but even the mutant peptide with the nanomolar affinity does not localise effectively (Figure 5A). Only when the peptide was dimerised though the leucine-zipper motif, the number of the SxIP comets became similar to the number of EB1 comets per cell (Figure 5B). This shows that the recruitment of the SxIP ligands requires multi-valent interactions that are characteristic of the condensates (Boeynaems, Alberti et al. 2018, Zumbro and Alexander-Katz 2021, Mohanty, Kapoor et al. 2022) and is partly driven by the condensation. Notably, most of the SxIP proteins contain additional interaction sites with other +TIP proteins, and some of them have multiple SxIP motifs (Kumar and Wittmann 2012, Akhmanova and Steinmetz 2015). Therefore, effective targeting of the SxIP and other interactions in the +TIP microtubule networks may require high affinity multi-valent inhibitors, which has to be considered in the future drug development.

## EXPERIMENTAL PROCEDURES

### Peptides and Protein Preparation

Human EB1 (Uniprot code Q15691), residues 191–260 and 191-252, was purified as previously described (Almeida, Carnell et al. 2017).

Synthetic MACF peptides with purity >95% were purchased from ChinaPeptides (Shanghai). Peptide amino acid sequences are: 4MACF (SKIP), 6MACF (SKIPTP), 11MACF

(KPSKIPTPQRK), 11MACF-VLL (KPSTAKSKIPVLL), 11MACF-LLL (KPSTAKSKIPLLL) and 11MACF-VLLRK (KPSTAKSKIPVLLRK). Peptides were dissolved with milli-Q ultrapure water to make 10-20 mM stock solutions; pH was adjusted to 7 with NaOH.

**NMR Spectroscopy**

NMR spectra were collected on Bruker Avance III 600 and Neo 800 MHz spectrometers equipped with CryoProbes. Spectra were processed with TopSpin (Bruker) and analysed using CCPNmr Analysis v2.4 (Vranken, Boucher et al. 2005). Experiments were performed at 298K in 20 mM phosphate pH 6.5, 50 mM NaCl, 0.5 mM TCEP, 0.02% (w/v) $NaN_3$.

To achieve full saturation of the complex, data for structure calculations was collected with a slight excess of ligand, and 10:11 EB1/11MACF ratio. The backbone resonances were assigned using triple resonance experiments (HNCO, HN(CA)CO, HNCA, HNCACB and CBCACONH) measured for $^{13}C$-$^{15}N$ labelled EB1 using standard assignment protocols in CCPN Analysis.

Side chain resonance assignments were obtained using HBHA(CO)NH, H(C)CH-TOCSY and (H)CCH-TOCSY experiments. Aromatic side-chains were assigned using 2D-NOESY and $^{1}H$-$^{13}C$-resolved-NOESY-HSQC. The resonances of the ligands were assigned using $^{13}C$,$^{15}N$-filtered 2D TOCSY and NOESY experiments. The structures were calculated using ARIA 2.2 integrated with CCPNmr Analysis (Vranken, Boucher et al. 2005), as fully described previously (Almeida, Carnell et al. 2017). Statistics of the structure determination are presented in Supplementary Table S1.

For the NMR lineshape analysis, a series of $^{1}H$,$^{15}N$-HSQC spectra were collected for the $^{15}N$-labelled EBH domain with variable concentrations of the peptide. The range of the peptide concentrations has been selected to ensure maximum saturation at the highest concentration and even distribution of the saturation states throughout the titration. The concentrations used are listed in the supplementary material (see figures S1 and S3). The titration HSQC data were analysed in TITAN software (v1.6) (Waudby, Ramos et al. 2016). For the EBH-ΔC/11MACF the 1:1 binding model was used because chemical shift changes were linear and no effect allosteric effect from binding to the two different binding sites in the dimer was detected. For the EBH-/11MACF and EBH-ΔC/11MACF-VLL we observed non-linear changes of chemical shifts and multiple HSQC peaks corresponding to the non-symmetrical complex where only one binding site in the dimer is occupied. To account for this, we used the dimer binding model implemented in TITAN. For each residue with a significant chemical shift perturbation that can be followed throughout the titration, lineshapes for the free state and the final titration point were initially fitted to evaluate the parameters of the free state and the complex. Then the titration series was fitted separately for each residue to obtain residue-specific exchange parameters. Finally, the lineshapes were fitted for all selected residues simultaneously to evaluate the global exchange rate.

The $^{15}N$ CEST experiments (Vallurupalli, Bouvignies et al. 2012) were conducted for samples containing 0.75 mM $^{15}N$ EB1 and 2.5% (molar) of the MACF peptide at a $^{1}H$ frequency of 800 MHz and 298K. CEST profiles were measured at $^{15}N$ B1 field strength of 12.5, 25 and 50 Hz applied during a constant period of 400 ms using the standard Bruker pulse sequence. For the residues with detected CEST NMR exchange, data at all B1 values were fitted separately for each residue and simultaneously for all the residues using the ChemEx software (https://github.com/gbouvignies/chemex) as described previously (Vallurupalli, Bouvignies et al. 2012).

The $^{15}N$ relaxation rates $R_1$, $R_2$, and $\{^1H\}$-$^{15}N$ heteronuclear Overhauser effects (nOes) were measured at a $^1H$ frequencies of 600 and 800 MHz and 298K using HSQC-based pulse sequences using standard Bruker pulse sequences. EBH concentration on the samples was 0.5 mM; 3 mM MACF peptide was used to ensure full saturation of the complex. To obtain the dynamic parameters, the data were analysed in the Relax software (d'Auvergne and Gooley 2008) using standard protocols.

**Isothermal titration Calorimetry**

ITC experiments were performed using the Malvern MicroCal Automated PEAQ-ITC instrument. Both protein and peptides were dialysed into 50mM Phosphate (pH 6.5), 50mM NaCl, 0.5mM TCEP and 0.02% $NaN_3$ before use. The sample cell and syringe were filled with 20-40 µM EB1 EBH domain and 200-600 µM MACF peptide respectively. 1.5 µL of MACF peptides were injected into the sample cell for a total of 25 injections. All experiments were performed in triplicate and at 25$^O$C. The data were integrated and fitted into the single-site binding model using the Malvern PEAQ-ITC Control software v1.41.

**Cell culture**

HeLa cells were cultured in Dulbecco's Modified Eagle Serum substituted with 10% fetal bovine serum and 1% penicillin/streptavidin. Cells were maintained at 37$^O$C and 5% $CO_2$.

**MACF peptide constructs and transfection**

Peptide sequences WT (GSRPSTAKPSKIPTPQRK), VLLRK (GSRPSTAKPSKIPVLLRKRK) and dimeric LZ-VLLRK peptide containing GCN4p1 leucin zipper (RMKQLEDKVEELLSKNYHLENEVARLKKLVGERGSRPSTAKPSKIPVLLRK) were inserted into a EKAR2G_design1_mTFP_wt_Venus_wt vector courtesy of Oliver Pertz (Addgene plasmid #39813; http://n2t.net/addgene:39813 ; RRID:Addgene_39813).

Peptide constructs were co-transfected into HeLa cells with an mCherry-EB1 construct using Lipofectamine 3000 regent following manufacturer's instructions (Invitrogen #L3000015). mCherry-EB1-8 was a gift from Michael Davidson (Addgene plasmid #55035, http://n2t.net/addgene:55035 ; RRID:Addgene_55035). Imaging was performed 24 hours after transfection.

**Live cell imaging**

Live cell images were obtained using the Marianas spinning disk confocal microscope system (Intelligent Imaging Innovations, Inc.) and Slidebook2022 (reference) capture software. Total internal reflection microscopy (TIRFM) was employed using 100x1.49 N.A. lens, 488 (FF01-525/30-25) and 561 (FF01-617/73-25) lasers with images captured using a Hamamatsu C11440 camera.

**Data Availability**. The coordinates for the solution structure of EBH/MACF complex are deposited in the PDB, accession code 7OLG. The chemical shifts are deposited in the BMRB, accession code 34629.

**Acknowledgements**. TA was funded by Liverpool University PhD studentship. EH was funded by BBSRC DTP NLD studentship. The NMR spectra were measured at the LIV-SRF High-Field NMR Facility, University of Liverpool. The images were collected at the LIV-SRF Centre for Cell Imaging (CCI) Facility, University of Liverpool.

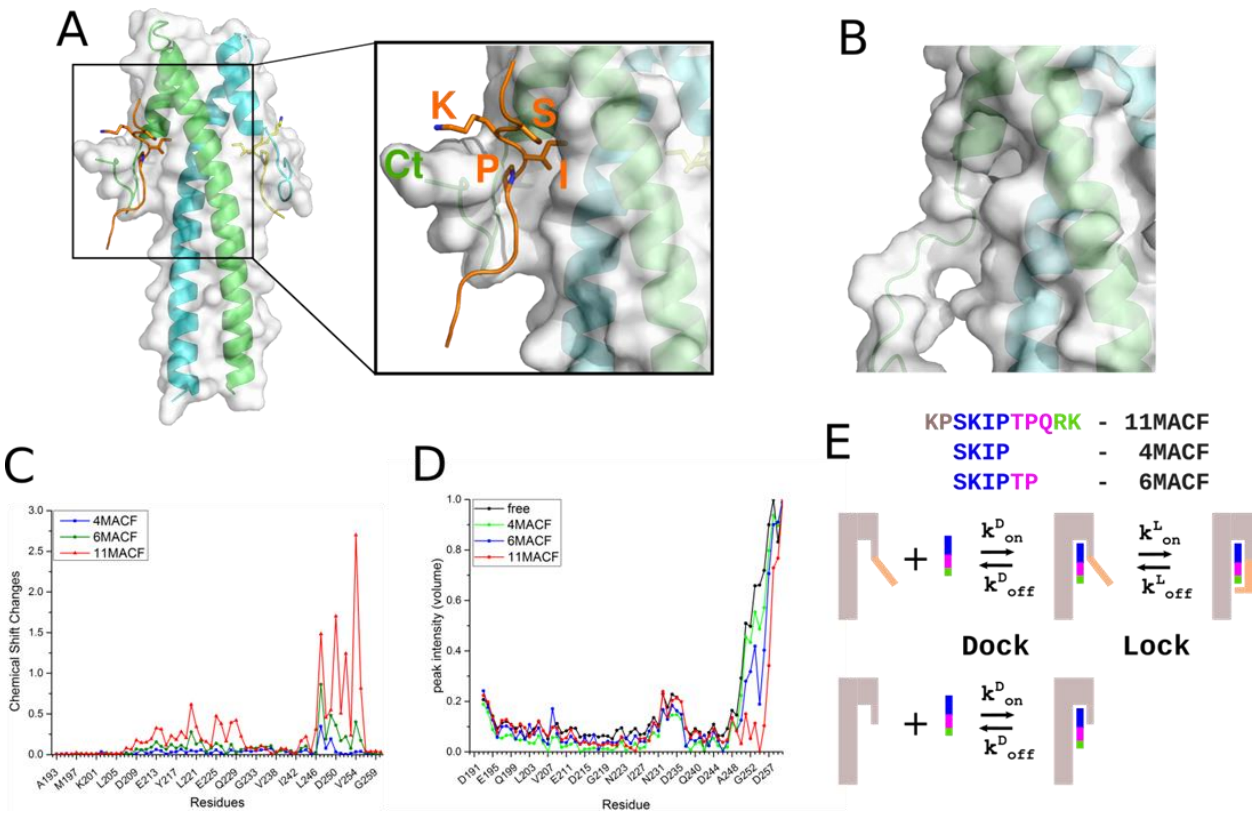


**Figure 1**. **Residues that follow SxIP motif enhance binding by engaging with the C-terminus of the EB1 EBH domain**. (A) Crystal structure of the EBH domain in the complex with MACF peptide (PDB ID 3GJO). EB1 is shown as a cartoon with the subunits of the dimer coloured in green and cyan, and a semi-transparent surface. The MACF peptide is coloured orange, with the SxIP motif chains-chains shown as sticks. The zoomed region highlights the binding pocket of EBH formed by the surface of the coiled-coil and the folded C-terminus. (B) Zoom on the partly formed SxIP binding pocked in the structure EBH domain free in solution (PDB ID 3EVI) in the same orientation as (A). The C-terminal region is unfolded. (C) EB1 chemical shift changes in the $^{1}H,^{15}N$-HSQC spectra induced by 4MACF (blue), 6MACF (green) and 11MACF (red) peptides. (D) Relative intensities of cross-peaks in the $^{1}H,^{15}N$-HSQC spectra of the EBH free in solution (black) and in the presence of 4MACF (blue), 6MACF (green) and 11MACF (red) peptides. (E) “Dock and Lock” binding model that explicitly considers the role of EBH C-terminus in the interaction with the SxIP peptide. Initially, the binding pocket is partially formed and only contains SxIP-recognition region. Following the initial binding (“Dock”), the post-SxIP region of the peptide induces the folding of the of the C-terminus and formation of the full binding pocket (“Lock”). The deletion of the EBH C-terminus removes the “Lock” stage of the binding, thus reducing the affinity of the interaction.

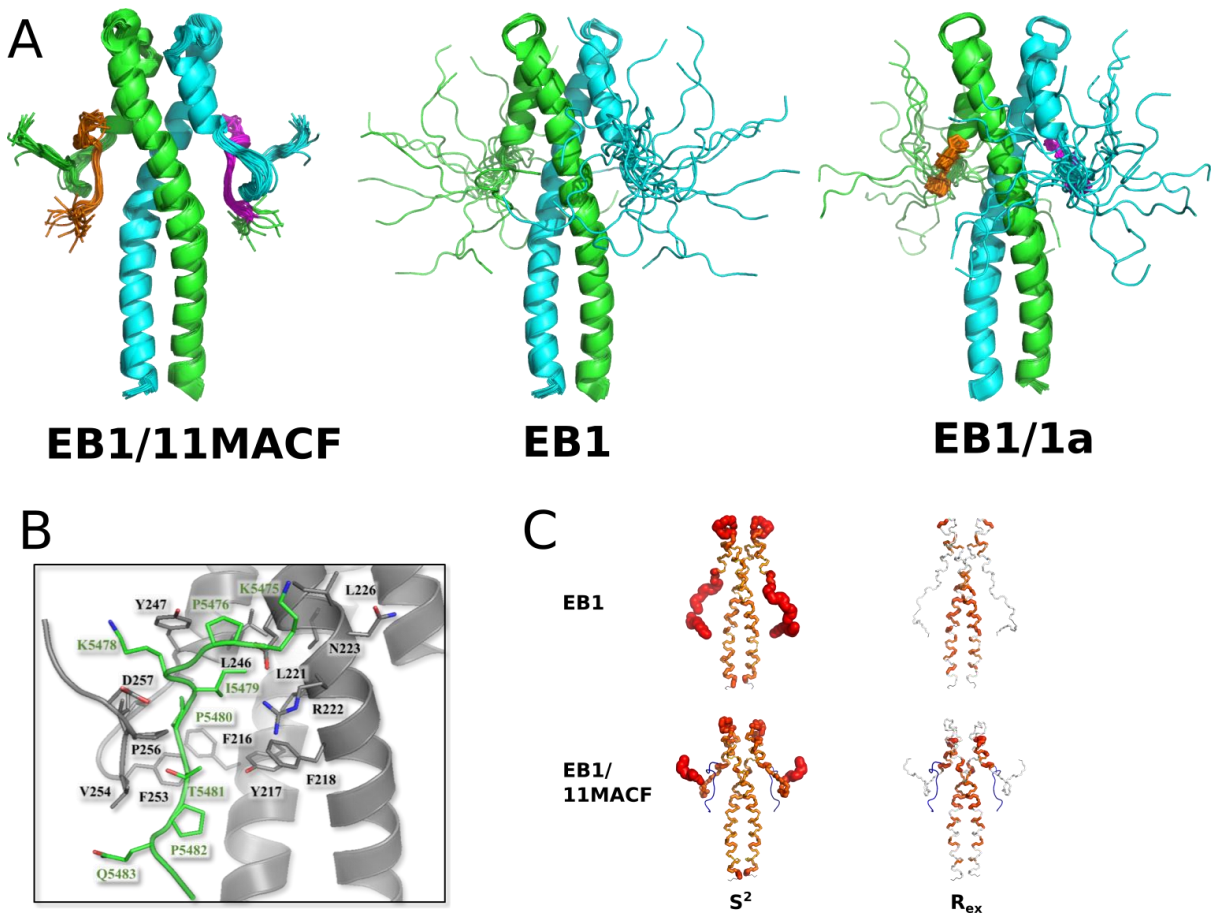


**Figure 2**. **Structure and dynamics of the EB1 EBH domain**. (A) Superposition of 20 NMR structures calculated for EBH in the complex with 11MACF peptide (left). Previously reported NMR structures of the free EBH (PDB ID 6EVI, middle) and the EBH complex with a small SxIP-like molecule 1a (PDB ID 6EVI, right) are shown for comparison. The EBH subunits are coloured in green and cyan, the peptide and the small molecule in orange and magenta. (B) Representation of the residues forming the contact interface between the EBH and 11MACF peptide, both shown as cartoon, with side chain displayed for all the residues involved in the contacts between the two molecules. EB1 is coloured in grey and 11MACF in green, with oxygen shown in red and nitrogen shown in blue. (C) Order parameters $S^2$ (left) and exchange contributions into the relaxation rate $R_{ex}$ (right) calculated from the relaxation parameters for the free EBH (top) and EBH in complex with 11MACF peptide (bottom) mapped on the EBH solution structure. The thickness of the tube is proportional to the value of the corresponding parameter.

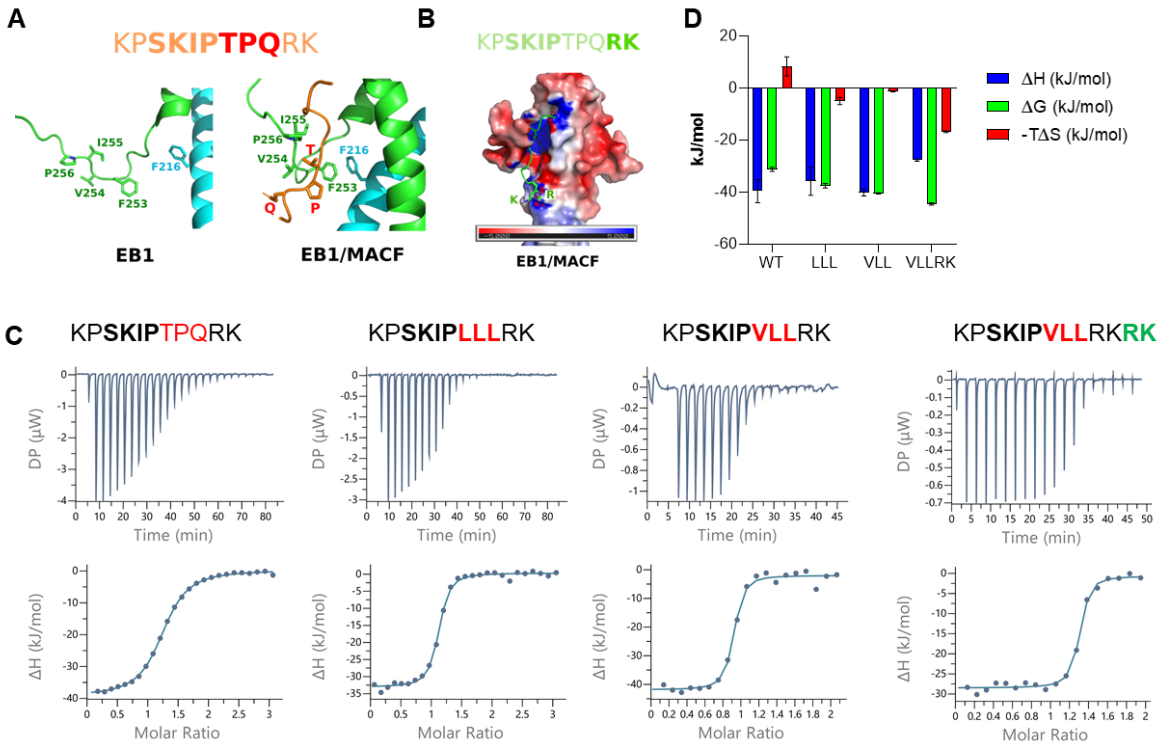


**Figure 3. Enhancement of the peptide binding through the substitution of the post-SxIP residues**. (A) Folding of the C-terminus brings the hydrophobic EBH residues of that region into the contact with the TPQ region of the peptide. Structure of the free EBH (left) and EBH in the complex with 11MACF (right). (B) Proximity of the positively charged RK residues of the peptide to the negatively charged patch on the EBH surface. The peptide is shown as a cartoon, RK side-chains are shown in a stick representation. Positive and negative electrostatic surface potential is represented by blue and red colour, respectively. (C) The ITC titration (top) and the binding isotherm fitted into a single-site binding model (bottom) of the ITC binding experiments for the 11MACF peptide and the mutants 11MACF-LLL, 11MACF-VLL and 11MACF-VLLRK. The peptide sequences are shown above the graphs, with the changed regions highlighted by the red (mutation) and green (insertion) colours. (D) The thermodynamic parameters (ΔG (green), ΔH (blue) and -TΔS (red) calculated from ITC data.

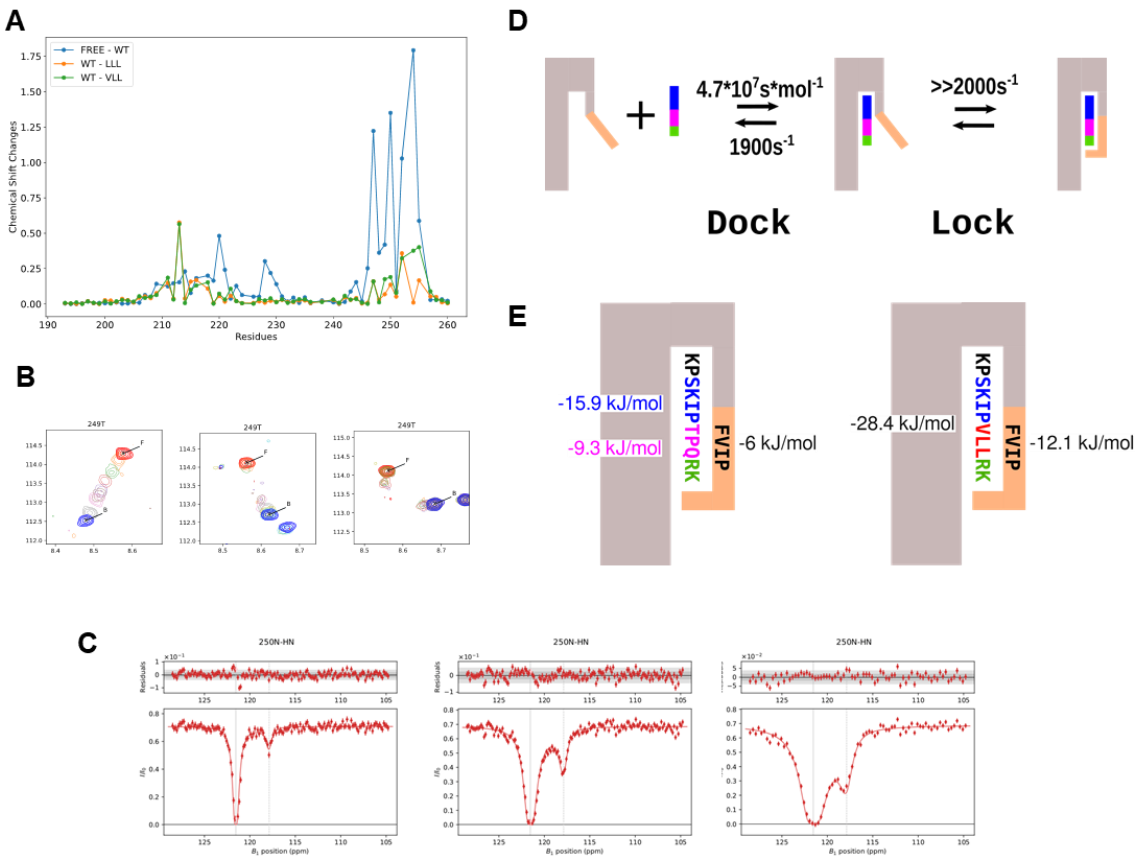


**Figure 4**. **NMR data show large decrease in the exchange rates on the mutations of the post-SxIP residues**. (A) 11MACF-LLL and 11MACF-VLL peptides induce increased chemical shift changes compared to the 11MACF. Chemical shift differences in the $^{1}$H,$^{15}$N-HSQC spectra between the free EBH and the EBH/11MACF complex (blue), EBH/11MACF and EBH/11MACF-LLL (orange), and EBH/11MACF and EBH/11MACF-VLL (green). (B) Chemical shift changes in the $^{1}$H,$^{15}$N-HSQC spectra on peptide addition observed for the EBH interactions with different peptides illustrated for the Thr$^{249}$ signal. Superposition of the spectra for the titration of the EBH-ΔC with 11MACF (left), EBH with 11MACF (middle), and EBH with 11MACF-VLL (right). Signals of the free EBH are shown in red, fully bound EBH in blue, and the intermediate titration points are shown in the pale colours. Notice two additional signals that are observed for the EBH/11MACF-VLL titration at the intermediate concentrations corresponding to the non-symmetrical form where EBH dimer binds a single peptide. These signals can only be observed when the exchange between the different forms is very slow. (C) Example of the CEST profiles measured at the irradiation field strength of 12.5, 25 and 50 Hz (left to right) for Asp$^{250}$. Solid curve represents the fitting of the data into the global exchange model with the dissociation rate 130 s$^{-1}$ calculate with Chemix software. (D) Exchange rates calculated for the EBH interaction with 11MACF peptide from the combination of the NMR data using the two-stage interaction model, where the folding of the EBH C-terminus follows the peptide binding. (E) Free energy contributions into the EBH interaction with the 11MACF peptide (left) and the mutated 11MACF-VLL peptide (right). The SxIP motif itself contributes approximately half of the binding energy (-15.9 kJ/mol), with the second half created by the interaction of the TPQRK region with the coiled-coli folded part of EBH (-9.3 kJ/mol), and the C-terminal EBH region that folds on the peptide binding (-6 kJ/mol). The VLL mutation increases the interaction with both the coiled-coli folded part of EBH energy (-28.4 kJ/mol) and the EBH C-terminus energy (-12.1 kJ/mol).

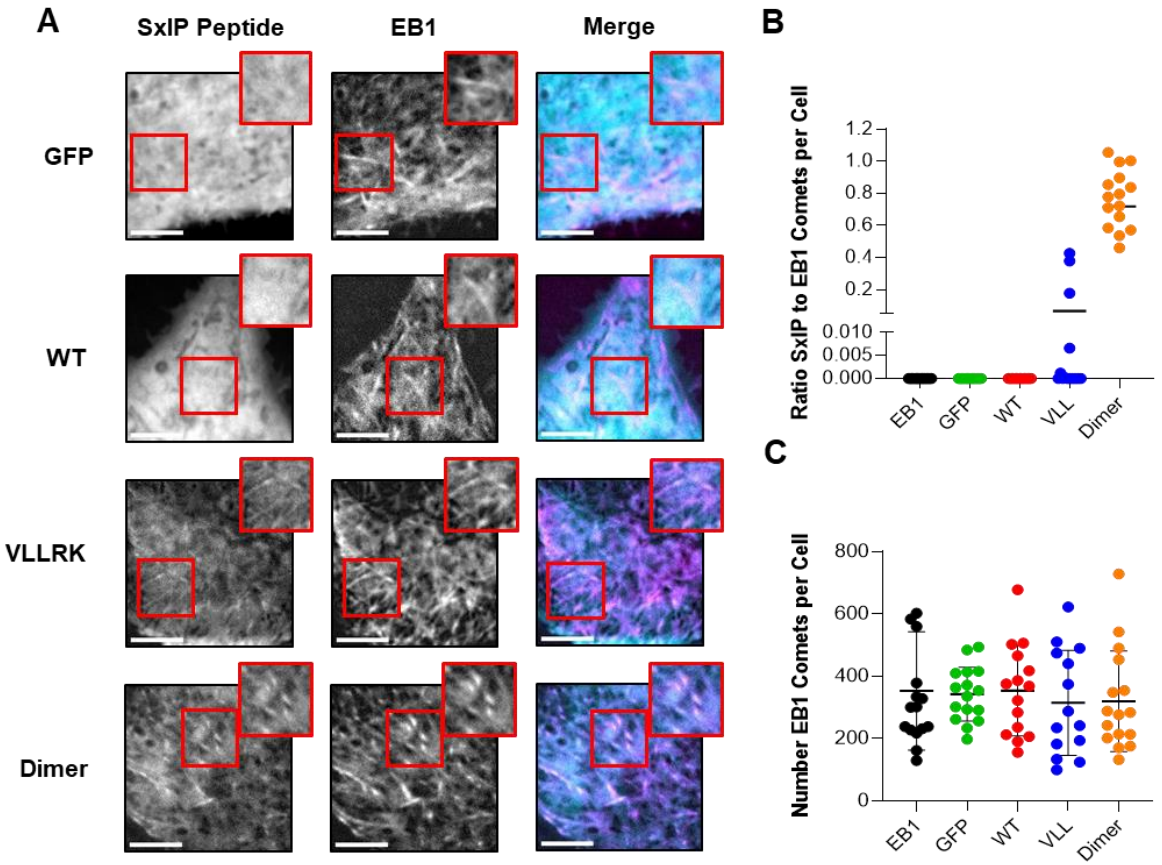


**Figure 5. High affinity peptides form comet-like structures that colocalise with EB1 comets.** (A) Live images of HeLa cells transiently transfected with SxIP peptide constructs and EB1 plasmid. Images represent averaging three frames at 1 second intervals. Inserts represent zoomed-in regions of the cell to depict SxIP peptides at EB1 comets. Scale bar; 5µM. (B) Ratio of SxIP comets to EB1 comets present in a cell. (C) EB1 comet number per cell. Lines represent the mean with error bars representing the standard deviation. Changes in comet number with each peptide were not significant when T-test was used with the EB1 construct only control (EB1). All experiments were performed to an N=3.

**Table 1. Binding parameters of the EBH domain interactions with peptides**

| | $K_d$ (μM) [1] | ΔG (kJ/mol) | ΔH (kJ/mol) | -TΔS (kJ/mol) | $k_{off}$ ($s^{-1}$) [2] | $k_{on}$ ($10^6$ $M^{-1}$ $s^{-1}$) |
|---|---|---|---|---|---|---|
| EBH-ΔC/6MACF | 7000 | -12.3 | | | | |
| EBH-ΔC/11MACF | 41.5 | -25.2 | -29.8 | 1.1 | 1900 | 45 |
| EBH/4MACF | 10400 | -11.3 | | | | |
| EBH/6MACF | 1700 | -15.8 | | | | |
| EBH/11MACF | 3.51 ± 1.038 | -31.2 ± 0.81 | -39.6 ± 4.45 | 8.4 ± 3.65 | 130 (143) | 37 |
| EBH/11MACF-LLL | 0.287 ± 0.088 | -37.5 ± 0.87 | -35.8 ± 5.38 | -4.9 ± 1.41 | 62 | 216 |
| EBH/11MACF-VLL | 0.081 ± 0.009 | -40.5 ± 0.25 | -40.2 ± 1.27 | -1.3 ± 0.09 | 15.6 (57) | 192 |
| EBH/11MACF-VLLRK | 0.016 ± 0.003 | -44.6 ± 0.47 | -27.8 ± 0.30 | -16.8 ± 0.32 | | |

[1] $K_d$ determined by NMR

[2] $k_{off}$ determined by the NMR lineshape analysis and by CEST (values in brackets)

## References


Akhmanova, A. and M. O. Steinmetz (2015). "Control of microtubule organization and dynamics: two ends in the limelight." Nat Rev Mol Cell Biol **16**(12): 711-726.

Almeida, T. B., A. J. Carnell, I. L. Barsukov and N. G. Berry (2017). "Targeting SxIP-EB1 interaction: An integrated approach to the discovery of small molecule modulators of dynamic binding sites." Sci Rep **7**(1): 15533.

Aseervatham, J. (2020). "Cytoskeletal Remodeling in Cancer." Biology (Basel) **9**(11): 385.

Boeynaems, S., S. Alberti, N. L. Fawzi, T. Mittag, M. Polymenidou, F. Rousseau, J. Schymkowitz, J. Shorter, B. Wolozin, L. Van Den Bosch, P. Tompa and M. Fuxreiter (2018). "Protein Phase Separation: A New Phase in Cell Biology." Trends Cell Biol **28**(6): 420-435.

Brezina, V., L. Hanykova, N. Velychkivska, J. P. Hill and J. Labuta (2022). "NMR lineshape analysis using analytical solutions of multi-state chemical exchange with applications to kinetics of host-guest systems." Sci Rep **12**(1): 17369.

Buey, R. M., I. Sen, O. Kortt, R. Mohan, D. Gfeller, D. Veprintsev, I. Kretzschmar, J. Scheuermann, D. Neri, V. Zoete, O. Michielin, J. M. de Pereda, A. Akhmanova, R. Volkmer and M. O. Steinmetz (2012). "Sequence determinants of a microtubule tip localization signal (MtLS)." J Biol Chem **287**(34): 28227-28242.

Camacho-Zarco, A. R., V. Schnapka, S. Guseva, A. Abyzov, W. Adamski, S. Milles, M. R. Jensen, L. Zidek, N. Salvi and M. Blackledge (2022). "NMR Provides Unique Insight into the Functional Dynamics and Interactions of Intrinsically Disordered Proteins." Chem Rev **122**(10): 9331-9356.

Cook, I., T. Wang, S. C. Almo, J. Kim, C. N. Falany and T. S. Leyh (2013). "The gate that governs sulfotransferase selectivity." Biochemistry **52**(2): 415-424.

d'Auvergne, E. J. and P. R. Gooley (2008). "Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor." J Biomol NMR **40**(2): 121-133.

Deshmukh, L., V. Tugarinov, J. M. Louis and G. M. Clore (2017). "Binding kinetics and substrate selectivity in HIV-1 protease-Gag interactions probed at atomic resolution by chemical exchange NMR." Proc Natl Acad Sci U S A **114**(46): E9855-E9862.

Furukawa, A., T. Konuma, S. Yanaka and K. Sugase (2016). "Quantitative analysis of protein-ligand interactions by NMR." Prog Nucl Magn Reson Spectrosc **96**: 47-57.

Honnappa, S., S. M. Gouveia, A. Weisbrich, F. F. Damberger, N. S. Bhavesh, H. Jawhari, I. Grigoriev, F. J. van Rijssel, R. M. Buey, A. Lawera, I. Jelesarov, F. K. Winkler, K. Wuthrich, A. Akhmanova and M. O. Steinmetz (2009). "An EB1-binding motif acts as a microtubule tip localization signal." Cell **138**(2): 366-376.

Khan, F. I., D. Lan, R. Durrani, W. Huan, Z. Zhao and Y. Wang (2017). "The Lid Domain in Lipases: Structural and Functional Determinant of Enzymatic Properties." Front Bioeng Biotechnol **5**: 16.

Kumar, P. and T. Wittmann (2012). "+TIPs: SxIPping along microtubule ends." Trends Cell Biol **22**(8): 418-428.

Mohanty, P., U. Kapoor, D. Sundaravadivelu Devarajan, T. M. Phan, A. Rizuan and J. Mittal (2022). "Principles Governing the Phase Separation of Multidomain Proteins." Biochemistry **61**(22): 2443-2455.

Mollica, L., L. M. Bessa, X. Hanoulle, M. R. Jensen, M. Blackledge and R. Schneider (2016). "Binding Mechanisms of Intrinsically Disordered Proteins: Theory, Simulation, and Experiment." Front Mol Biosci **3**: 52.

Morris, O. M., J. H. Torpey and R. L. Isaacson (2021). "Intrinsically disordered proteins: modes of binding with emphasis on disordered domains." Open Biol **11**(10): 210222.

Shammas, S. L., M. D. Crabtree, L. Dahal, B. I. Wicky and J. Clarke (2016). "Insights into Coupled Folding and Binding Mechanisms from Kinetic Studies." J Biol Chem **291**(13): 6689-6695.

Shammas, S. L., A. J. Travis and J. Clarke (2013). "Remarkably fast coupled folding and binding of the intrinsically disordered transactivation domain of cMyb to CBP KIX." J Phys Chem B **117**(42): 13346-13356.

Song, X., F. Yang, T. Yang, Y. Wang, M. Ding, L. Li, P. Xu, S. Liu, M. Dai, C. Chi, S. Xiang, C. Xu, D. Li, Z. Wang, L. Li, D. L. Hill, C. Fu, K. Yuan, P. Li, J. Zang, Z. Hou, K. Jiang, Y. Shi, X. Liu and X. Yao (2023). "Phase separation of EB1 guides microtubule plus-end dynamics." Nat Cell Biol **25**(1): 79-91.

Sugase, K., H. J. Dyson and P. E. Wright (2007). "Mechanism of coupled folding and binding of an intrinsically disordered protein." Nature **447**(7147): 1021-1025.

Vallurupalli, P., G. Bouvignies and L. E. Kay (2012). "Studying "invisible" excited protein states in slow exchange with a major state conformation." J Am Chem Soc **134**(19): 8148-8161.

Vranken, W. F., W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." Proteins **59**(4): 687-696.

Waudby, C. A., A. Ramos, L. D. Cabrita and J. Christodoulou (2016). "Two-Dimensional NMR Lineshape Analysis." Sci Rep **6**: 24826.

Wright, P. E. and H. J. Dyson (2009). "Linking folding and binding." Curr Opin Struct Biol **19**(1): 31-38.

Zhang, R., G. M. Alushin, A. Brown and E. Nogales (2015). "Mechanistic Origin of Microtubule Dynamic Instability and Its Modulation by EB Proteins." Cell **162**(4): 849-859.

Zumbro, E. and A. Alexander-Katz (2021). "Multivalent polymers can control phase boundary, dynamics, and organization of liquid-liquid phase separation." PLoS One **16**(11): e0245405.