COMMENTARY

Connecting the αα-hubs: same fold, disordered ligands, new functions

Lasse Staby, Katrine Bugge, Rasmus Greve Falbe-Hansen, Edoardo Salladini, Karen Skriver* and Birthe B. Kragelund*

Abstract

Background: Signal fidelity depends on protein–protein interaction–'hubs' integrating cues from large interactomes. Recently, and based on a common secondary structure motif, the αα-hubs were defined, which are small α-helical domains of large, modular proteins binding intrinsically disordered transcriptional regulators.

Methods: Comparative structural biology.

Results: We assign the harmonin-homology-domain (HHD, also named the harmonin N-terminal domain, NTD) present in large proteins such as harmonin, whirlin, cerebral cavernous malformation 2, and regulator of telomere elongation 1 to the αα-hubs. The new member of the αα-hubs expands functionality to include scaffolding of supra-modular complexes mediating sensory perception, neurovascular integrity and telomere regulation, and reveal novel features of the αα-hubs. As a common trait, the αα-hubs bind intrinsically disordered ligands of similar properties integrating similar cellular cues, but without cross-talk.

Conclusion: The inclusion of the HHD in the αα-hubs has uncovered new features, exemplifying the utility of identifying groups of hub domains, whereby discoveries in one member may cross-fertilize discoveries in others. These features make the αα-hubs unique models for decomposing signal specificity and fidelity. Using these as models, together with other suitable hub domain, we may advance the functional understanding of hub proteins and their role in cellular communication and signaling, as well as the role of intrinsically disordered proteins in signaling networks.

Keywords: Interactome, Intrinsically disordered protein, IDP, NCBD, PAH, RST, HHD, Signaling, TAFH, Hub proteins

Introduction

Fast and efficient regulation of cellular signaling is key to cell viability. Fidelity in signaling is mediated by large networks of interacting proteins linked by a few highly connected proteins called hubs (Fig. 1a). As a consequence, protein interaction networks are highly sensitive to the removal of hubs, which may lead to premature fatality [2, 3]. Accordingly, hub proteins are often subject to gene duplication resulting in functional

redundancy that may protect against hub failure [4, 5]. Protein intrinsic disorder (ID) provides adaptability and is critical for hub functionality [6, 7]. Indeed, hub proteins are longer and have higher degree of ID than non-hubs [8], and they can be distinguished by specific sequence features of importance to their network evolution [9]. A detailed functional characterization of hub networks including both hubs and their binding partners revealed augmented disorder-enrichment in hub interactions among disease-associated proteins [10]. Nonetheless, folded hubs also exist, in which case structural disorder is found in the hub-partners. Thus, whereas a disordered hub, such as p53, can use different disordered regions for partner binding, a structured

*Correspondence: KSkriver@bio.ku.dk; bbk@bio.ku.dk
REPIN, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Staby *et al. Cell Commun Signal*   (2021) 19:2

Page 2 of 5

(See figure on next page.)

Fig. 1 Connecting the αα-hubs. **a** Hub proteins are involved in diverse functions such as transcription, post-translational modification and organization of cell structure. Hub proteins are shown in the center of protein–protein interaction networks where they organize relevant molecular components. The drawings are based on αα-hub protein functions such as transcriptional regulation (Sin3a) and cell structure organization (harmonin). **b** Domain structure of αα-hub proteins. Representative parent protein (not drawn to scale) for each of the founding αα-hubs PAH, RST, TAFH and NCBD, as well as the new member HHD. Experimentally-based interactome sizes obtained from the STRING database are shown for each αα-hub. Hp: small hairpin extension of the HHD-PDZ supramodule. **c** Superimposition of the αα-hairpin super-secondary structure motif. Representatives of each domain type of the αα-hubs (pdb codes 2CZY (PAH1), 2LD7 (PAH3), 5ECJ (TAFH), 2L14 (NCBD), 2KBQ (HHD), 5N9Q (RST)). The zoom illustrates the $α_L$-$β_4$ link motif found in the prototypical αα-hubs [1], where the highlighted hydrophobic $β_3$-anchor residue forms stabilizing interactions between the two hairpin helices. **d** Cartoon structures comparing αα-hubs. The PAH1 domain of Sin3a (purple) is compared to the HHDs of CCM2 (red), whirlin (dark orange) and harmonin (bright orange). Top row illustrates the two domains in their free form whereas the second row shows overlays of free PAH1 with whirlin HHD (left) and CCM2 HHD (right) (pdb codes 2RMR (PAH1), 6FDD (whirlin-HHD), 4FQN (CCM2-HHD)). The third row illustrates complexes of the two domains: PAH1 in complex with SAP25 and CCM2 HHD in complex with MEKK3 (pdb codes 2RMS and 4Y5O, respectively). The ligands are shown in grey. The bottom panel displays the harmonin PDZ-HHD supramodule (PDZ in blue) in complex with sans (grey). The short hairpin extension responsible for tethering the PDZ and HHD is shown in light blue. **e** Disorder profiles of HHD binding ligands cadherin-23 intracellular domain (ICD) and MEKK3. The disorder propensity ranging from 0 to 1 was predicted using IUPred2A (black) [32] and PONDR VSL2 (blue) [33], whereas MoRFs were predicted by MoRFpred (red) [34]. The regions responsible for binding to HHD based on available structures (pdb codes 4Y5O, 2LSR and 2KBR) are highlighted with orange background, whereas the regions binding as α-helix in the hydrophobic cleft are highlighted with white boxes

hub, such as 14-3-3, can associate with many different intrinsically disordered partners [11]. Folded hubs can also exist in large regulatory and modularly built proteins, and similarly, they bind numerous intrinsically disordered partners [12] using intriguing allosteric mechanisms [13], and constitute widespread cellular hubs of key biological relevance.

Bugge, Staby et al. [1] recently determined the structure of the RCD1, SRO and TAF4 (RST) domain from the plant protein Radical Induced Cell Death1 (RCD1), revealing a previously undescribed fold of four helices (H1–H4) forming an exposed hydrophobic binding cleft. However, similar folds have been determined for the PAH (paired amphipathic helix), TAFH (TATA-box-associated factor homology), and NCBD (nuclear coactivator binding domain) domains of the important human transcriptional regulators Sin3, transcription initiation factor TFIID-subunit 4 (TAF4) and CREB binding protein (CBP), respectively (Fig. 1b). These proteins have large interactomes and function as hubs. They have a common structural foundation consisting of an αα-hairpin super-secondary motif linked by an $α_L β_4$-loop, which serves as an organizing platform for malleable helices of varying length and topology (Fig. 1c). Hence, they were designated the αα-hubs. The prototypical αα-hub consists of four α-helices (although an additional helix can be present as in e.g. TAFH), of which two are organized in the αα-hairpin stabilized by the $β_3$-loop residue acting as a hydrophobic anchor (Fig. 1c). As hubs, they bind multiple intrinsically disordered partners with pleiotropic functional outcomes, and their parent proteins are large modular proteins implicated in transcription (Fig. 1b).

In addition to the initially identified αα-hubs, we find, based on structural analyses (Fig. 1d), that the

harmonin-homology domain (HHD), present in the proteins whirlin [14], harmonin [15], cerebral cavernous malformations 2 (CCM2) [16] and regulator of telomere elongation 1 (RTEL1) [17], also has the prototypical traits of αα-hubs [1, 17]. Therefore, we here connect the HHD to the αα-hubs. However, similar to NCBD, it lacks $β_3$-anchoring between H2 and H3, and instead has an additional α-helix H5 positioned between H3 and H4, similar to TAFH. Still, its αα-hairpin superimpose perfectly with those of the other αα-hubs (Fig. 1c), while its overall architecture is most similar to the PAH1 domain (RMSD of 0.85 Å for H1-H4 of whirlin HHD; Fig. 1d). Like the other αα-hub-harboring proteins, the HHD-containing proteins are large, modular proteins, but they have entirely different biological functions, related to hearing-vision perception for whirlin and harmonin [14, 15], neurovascular integrity for CCM2 [16] and telomere length regulation for RTEL1 [17] (Fig. 1a, b).

The known αα-hub ligands are mostly disordered in their free state [18, 19], typically interacting via short linear motifs [20–22]. These are of similar chemical compositions with hydrophobic and acidic residues as key hotspots for the interactions, and typically undergo coupled folding and binding to form an α-helix when bound in the hydrophobic cleft of the αα-hub [1]. Like the other αα-hubs, the currently characterized HHD complexes involving MEKK3 and cadherin-23 primarily use the common hydrophobic cleft for binding of an amphipathic α-helix [16]. Of note, the binding regions of HHD ligands are located in longer predicted disordered regions. As an example, the region of MEKK3 that directly binds to the typical αα-hub binding cleft of the HHD of CCM2 has been predicted to be intrinsically disordered in the free state (7) (Fig. 1e). ID also characterizes the αα-hub

Staby *et al. Cell Commun Signal* (2021) 19:2

Page 3 of 5
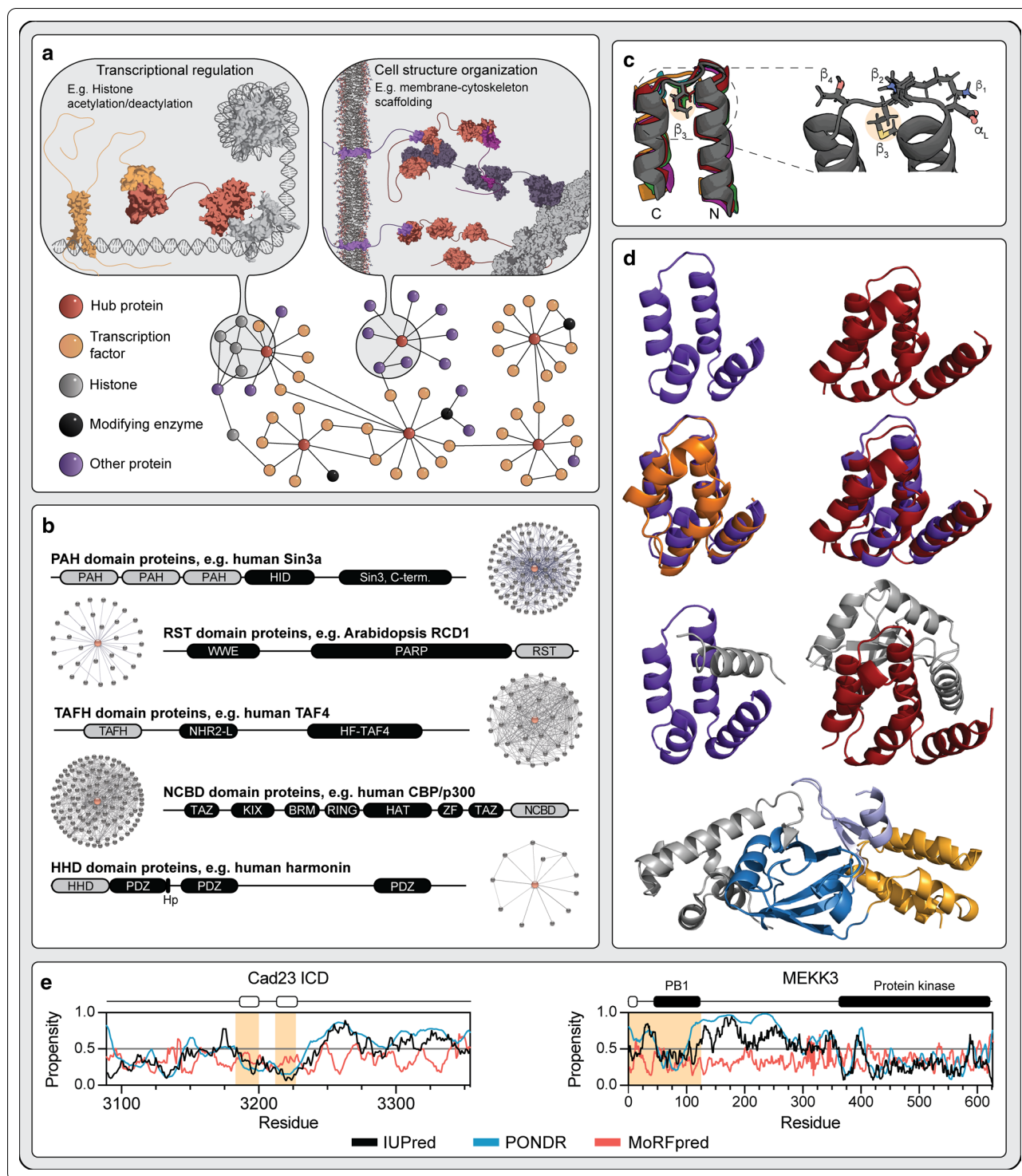
binding regions of the well-characterized PAH domain ligands repressor element 1 silencing transcription factor (REST) [23] and Mad1 [24], and RST ligands DREB2A and ANAC013 [20]. Similar to PAH, the ligand-binding cleft is in HHD located between helix H1 and H2.

With the inclusion of HHD, new features are added to the αα-hubs. The HHD of harmonin tethers the neighboring PDZ domain via a small hairpin extension of PDZ1 to form a functional and structurally stable supramodule responsible for binding the protein Sans as

Staby *et al. Cell Commun Signal*   (2021) 19:2

Page 4 of 5

part of hearing-vision regulation [15] (Fig. 1b, d). Binding of the hairpin extension occurs through a surface different from the α-helix-binding groove, a feature similarly seen for HHD of CCM2 (Fig. 1d). Here, the CCM2-HHD uses both the α-helix-binding cleft between H1 and H2 and a large surface on the backside of H2 and H3 for binding of MEKK3 [16] (Fig. 1d). Despite taking advantage of several surface areas simultaneously for ligand binding, the known harmonin interactome appears much smaller than for the other αα-hubs (e.g. Sin3a has > 100 experimentally identified partners vs. 11 in harmonin) (Fig. 1b).

Some of the new functional features arising from the inclusion of HHD in the αα-hubs can retrospectively be found in the literature for αα-hubs and vice versa. Thus, it is likely that the tandem HHDs of RTEL1, linked by a proliferating cell nuclear antigen (PCNA) interacting protein (PIP) box [25], forms a supramodular function and protein-interaction platform in RTEL1 [17]. Furthermore, the PAH domains may also exist as supramodules, as a Sin3a fragment of PAH1 and PAH2 interacts more strongly with the transcription factor Mad1/Mxd1 than the isolated domains [2]. The PAH3 domain of Sin3a binds the histone deacetylase complex subunit Sin3-associated protein (SAP30) in a high affinity complex resulting from cooperative recognition of two discrete surfaces of PAH3 by the tripartite binding motif in SAP30 [26]. NCBD binds not only disordered proteins in complicated mechanisms [27, 28], but also folded partners [29, 30], and similarly, HHD of CCM2 binds the folded PB1 region of MEKK3 [16] (Figure D). Thus, the αα-hubs are highly versatile and can adapt to many different partners, although intrinsically disordered ligands remain a highly dominant common feature. These findings highlight the importance of identifying groups of similar domains, whereby discoveries in one member may cross-fertilize discoveries in others, advancing their understanding. Furthermore, as a group, the αα-hubs have the potential to foster progress in the understanding of signal fidelity and specificity governed by other small α-helical hub domains, such as KIX and TAZ from CBP [12], for which such highly divergent functions have not yet been observed.

The simple fold of the αα-hubs and the consistent appearance of small α-helix domains in transcriptional regulation and signaling, and in particular in hubs [1, 12], suggest them to be of optimal architecture for integration via signals embedded in structural disorder. For the αα-hubs it is tempting to hypothesize that the expanded interaction surface and the cooperativity associated with supramodular structures, coupled to structural malleability, form a highly versatile platform for affinity- and specificity-tuning. This could ensure signal fidelity across

networks. Indeed, the harmonin N-terminal HHD-containing supramodule bound a fragment of Sans consisting of the two domains SAM and PBM with a K_d of ~ 1 nM (Fig. 1d), which was three and four orders of magnitude lower than the K_ds for binding the isolated SAM and PBM domains [15]. Whether other αα-hubs also exploit supramodular structures or if they use other interaction surfaces than the hydrophobic cleft, remains to be systematically addressed. So far, the limited focus on "one domain—one binding site" has precluded answering these questions and jointly, the αα-hubs leave more open questions than answers. In fact, many other signaling domains exist, which like the αα-hubs, function as hubs but with different structural features [31]. Connecting hub domains using the approach presented here will have the prospect to cross-fertilize studies beyond the single hub protein, and provide new insight not directly available from one hub on its own.

Conclusion

With the inclusion of HHD in the αα-hubs, we wish to put focus on the αα-hubs as an attractive model system for scrutinizing underlying mechanisms of signal fidelity and specificity, as well as communication by structural disorder within the cell. This is necessary to forward the understanding of key features of cellular hubs, including how interactome sizes are determined. Connecting the hubs, as done here, opens for exciting new research questions addressing properties across the domains, inspired by individual members. The approach used to connect hubs, is directly applicable to other folded hubs. For the αα-hubs, the inclusion of a new member will foster a new focus on structural disorder and its role in HHD linked interactomes.

Supplementary Information

Abbreviations

Acknowledgements

Authors' contributions

Staby *et al. Cell Commun Signal*   (2021) 19:2

Page 5 of 5

References
1. Bugge K, Staby L, Kemplen KR, O'Shea C, Bendsen SK, Jensen MK, et al. Structure of radical-induced cell death1 hub domain reveals a common αα-scaffold for disorder in transcriptional networks. Structure. 2018;26:734-746.e7.
2. Cowley SM, Kang RS, Frangioni JV, Yada JJ, DeGrand AM, Radhakrishnan I, et al. Functional analysis of the Mad1-mSin3A repressor-corepressor interaction reveals determinants of specificity, affinity, and transcriptional response. Mol Cell Biol. 2004;24:2698–709.
3. Dannenberg JH, David G, Zhong S, Van Der Torre J, Wong WH, DePinho RA. mSin3A corepressor regulates diverse transcriptional networks governing normal and neoplastic growth and survival. Genes Dev. 2005;19:1581–95.
4. Vinayagam A, Gibson TE, Lee HJ, Yilmazel B, Roesel C, Hu Y, et al. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. Proc Natl Acad Sci USA. 2016;113:4976–81.
5. Kafri R, Dahan O, Levy J, Pilpel Y. Preferential protection of protein interaction network hubs in yeast: evolved functionality of genetic redundancy. Proc Natl Acad Sci USA. 2008;105:1243–8.
6. Kragelund BB, Jensen MK, Skriver K. Order by disorder in plant signaling. Trends Plant Sci. 2012;17:625–32.
7. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets: the roles of intrinsic disorder in protein interaction networks. FEBS J. 2005;272:5129–48.
8. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, et al. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput Biol. 2006;2:890–901.
9. Dosztányi Z, Chen J, Dunker AK, Simon I, Tompa P. Disorder and sequence repeats in hub proteins and their implications for network evolution. J Proteome Res. 2006;5:2985–95.
10. Hu G, Wu Z, Uversky VN, Kurgan L. Functional analysis of human hub proteins and their interactors involved in the intrinsic disorder-enriched interactions. Int J Mol Sci. 2017;18:2761.
11. Oldfield CJ, Meng J, Yang JY, Qu MQ, Uversky VN, Dunker AK. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genom. 2008;9:S1.
12. Dyson HJ, Wright PE. Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. J Biol Chem. 2016;291:6472–22.
13. Berlow RB, Dyson HJ, Wright PE. Hypersensitive termination of the hypoxic response by a disordered protein switch. Nature. 2017;543:447–51.
14. Delhommel F, Cordier F, Saul F, Chataigner L, Haouz A, Wolff N. Structural plasticity of the HHD2 domain of whirlin. FEBS J. 2018;285:3738–52.
15. Yan J, Pan L, Chen X, Wu L, Zhang M. The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. Proc Natl Acad Sci USA. 2010;107:4040–5.
16. Fisher OS, Deng H, Liu D, Zhang Y, Wei R, Deng Y, et al. Structure and vascular function of MEKK3–cerebral cavernous malformations 2 complex. Nat Commun. 2015;6:7937.
17. Faure G, Revy P, Schertzer M, Londono-Vallejo A, Callebaut I. The C-terminal extension of human RTEL1, mutated in Hoyeraal-Hreidarsson syndrome, contains Harmonin-N-like domains. Proteins Struct Funct Bioinform. 2014;82:897–903.
18. Higo J, Nishimura Y, Nakamura H. A free-energy landscape for coupled folding and binding of an intrinsically disordered protein in explicit solvent from detailed all-atom computations. J Am Chem Soc. 2011;133:10448–58.
19. Iešmantavičius V, Dogan J, Jemth P, Teilum K, Kjaergaard M. Helical propensity in an intrinsically disordered protein accelerates ligand binding. Angew Chem Int Ed. 2014;53:1548–51.
20. O'Shea C, Staby L, Bendsen SK, Tidemand FG, Redsted A, Willemoës M, et al. Structures and short linear motif of disordered transcription factor regions provide clues to the interactome of the cellular hub protein radical-induced cell death 1. J Biol Chem. 2017;292:512–27.
21. Chandru A, Bate N, Vuister GW, Cowley SM. Sin3A recruits Tet1 to the PAH1 domain via a highly conserved Sin3-interaction domain. Sci Rep. 2018;8:1–10.
22. Sahu SC, Swanson KA, Kang RS, Huang K, Brubaker K, Ratcliff K, et al. Conserved themes in target recognition by the PAH1 and PAH2 domains of the Sin3 transcriptional corepressor. J Mol Biol. 2008;375:1444–56.
23. Nomura M, Uda-Tochio H, Murai K, Mori N, Nishimura Y. The neural repressor NRSF/REST binds the PAH1 domain of the Sin3 corepressor by using its distinct short hydrophobic helix. J Mol Biol. 2005;354:903–15.
24. Swanson KA, Knoepfler PS, Huang K, Kang RS, Cowley SM, Laherty CD, et al. HBP1 and Mad1 repressors bind the Sin3 corepressor PAH2 domain with opposite helical orientations. Nat Struct Mol Biol. 2004;11:738–46.
25. Prestel A, Wichmann N, Martins JM, Marabini R, Kassem N, Broendum SS, et al. The PCNA interaction motifs revisited: thinking outside the PIP-box. Cell Mol Life Sci. 2019;76:4923–43.
26. Xie T, He Y, Korkeamaki H, Zhang Y, Imhoff R, Lohi O, et al. Structure of the 30-kDa Sin3-associated protein (SAP30) in complex with the mammalian Sin3A corepressor and its role in nucleic acid binding. J Biol Chem. 2011;286:27814–24.
27. Zosel F, Mercadante D, Nettels D, Schuler B. A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction. Nat Commun. 2018;9:3332.
28. Liu X, Chen J, Chen J. Residual structure accelerates binding of intrinsically disordered ACTR by promoting efficient folding upon encounter. J Mol Biol. 2019;431:422–32.
29. Qin BY, Liu C, Srinath H, Lam SS, Correia JJ, Derynck R, et al. Crystal structure of IRF-3 in complex with CBP. Structure. 2005;13:1269–77.
30. Dogan J, Jonasson J, Andersson E, Jemth P. Binding rate constants reveal distinct features of disordered protein domains. Biochemistry. 2015;54:4741–50.
31. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 2020;43:D257–60.
32. Mészáros B, Erdős G, Dosztányi Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 2018;46:W329–37.
33. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. BMC Bioinform. 2006;7:208.
34. Disfani FM, Hsu WL, Mizianty MJ, Oldfield CJ, Xue B, Dunker AK, et al. MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. Bioinformatics. 2012;28:i75-83.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.