

α -Hub domains and intrinsically disordered proteins: A decisive combo

Received for publication, October 7, 2020, and in revised form, December 22, 2020. Published, Papers in Press, December 29, 2020.
<https://doi.org/10.1074/jbc.REV120.012928>

Katrine Bugge^{1,2}, Lasse Staby^{1,2}, Edoardo Salladini¹, Rasmus G. Falbe-Hansen¹, Birthe B. Kragelund^{1,2,*}, and Karen Skriver^{1,*}

From the ¹REPIN and The Linderstrøm-Lang Centre for Protein Science; ²Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Edited by Ronald Wek

Hub proteins are central nodes in protein–protein interaction networks with critical importance to all living organisms. Recently, a new group of folded hub domains, the α -hubs, was defined based on a shared α -hairpin supersecondary structural foundation. The members PAH, RST, TAFH, NCB, and HHD are found in large proteins such as Sin3, RCD1, TAF4, CBP, and harmonin, which organize disordered transcriptional regulators and membrane scaffolds in interactomes of importance to human diseases and plant quality. In this review, studies of structures, functions, and complexes across the α -hubs are described and compared to provide a unified description of the group. This analysis expands the associated molecular concepts of “one domain–one binding site”, motif-based ligand binding, and coupled folding and binding of intrinsically disordered ligands to additional concepts of importance to signal fidelity. These include context, motif reversibility, multivalency, complex heterogeneity, synergistic α -hub:ligand folding, accessory binding sites, and supramodules. We propose that these multifaceted protein–protein interaction properties are made possible by the characteristics of the α -hub fold, including supersite properties, dynamics, variable topologies, accessory helices, and malleability and abetted by adaptability of the disordered ligands. Critically, these features provide additional filters for specificity. With the presentations of new concepts, this review opens for new research questions addressing properties across the group, which are driven from concepts discovered in studies of the individual members. Combined, the members of the α -hubs are ideal models for deconvoluting signal fidelity maintained by folded hubs and their interactions with intrinsically disordered ligands.

Fast and efficient regulation of complex cellular signaling pathways is mediated by highly connected protein nodes called hubs (1–5). These hub proteins are intimately linked to intrinsic structural disorder (ID) (1–5), either containing intrinsically disordered regions (IDRs) themselves (1) or being ordered proteins binding IDRs (6–8). IDRs exist in ensembles of interconverting and dynamic states, endowing them with

adaptability, multivalency, and high chemical modification potential (9–11). Indeed, ID appears to be a prerequisite for fidelity in signaling pathways through exploitation of the many different mechanisms encoded in these properties (12, 13).

Recognition of IDRs in and by hubs depends on short linear motifs (SLiMs), which are stretches of 2 to 12 residues with only a few highly conserved positions (14, 15). It has been proposed that the eukaryotic SLiMome consists of up to 1 million different SLiMs (16), but SLiMs active in hub interactions have very similar features (17). Thus, it remains enigmatic how signal fidelity is orchestrated by hubs. Several folded domains present in large scaffolding proteins act as hubs, binding IDRs, typically transcription factor (TF) regulatory domains (TRDs), *via* SLiMs. These include the TAZ (18, 19), the KIX (19), the GACKIX (20), and the α -hub domains (21). Despite belonging to different families, the domains share structural traits such as being a relatively short chain of <100 residues that folds into topologies constructed solely by α -helices (19, 21). All the domains are also part of multidomain hub proteins carrying both order and disorder (19, 21), increasing the valency of their interactions. As an additional layer of regulation, protein cofactors bring the transcriptional machinery to target genes through interactions with TFs (22) and aid in scaffolding of the transcriptional machinery (23). How specificity and regulation within the associated multi-component complexes are controlled is far from understood.

The α -hub domains have only recently emerged as a group of folded hub proteins (21, 24), and hence, the full potential for understanding hub proteins from studies across these similar domains has yet to be unfolded. Furthermore, understanding their role in organizing disordered transcriptional regulators and membrane protein scaffolds in interactomes of importance to human diseases and plant quality is of broad interest. This review focuses on the α -hub domains and brings an overview and comparative analysis of their structures, functions, complexes, and mechanisms. The α -hub domains have low sequence identity (4%–15%) (21) and are diversely involved in distinct biological functions, while still binding ligands of similar structural and chemical properties. Thus, this group of hubs constitutes a suitable model for addressing how selectivity and specificity in interactomes are controlled and how fidelity is encoded. Through specific examples, we ask

This article contains [supporting information](#).

* For correspondence: Birthe B. Kragelund, bbk@bio.ku.dk; Karen Skriver, KSkriver@bio.ku.dk.

how the α -hub domains maintain fidelity and highlight concepts and open questions related to the α -hub *modus operandi*.

The α -hubs

The α -hub domains share key fold features

The α -hub domains were recently defined based on a common structural foundation (21) and originally included the following domains: RCD1, SRO, and TAF4 (RST) from Radical-Induced Cell Death1 (RCD1) (21); paired amphipathic helix (PAH)1, PAH2, and PAH3 from Sin3 of the Sin3/histone deacetylase corepressor complex (25); TATA-box associated factor homology (TAFH) (or nervy homology region 1 [NHR1]) from transcription initiation factor TFIID-subunit 4 (TAF4) and eight-twenty-one (ETO) (or MTG8/CBFA2T1) (26); and nuclear coactivator binding domain (NCBD) (or IRF-binding domain) from CREB binding protein (CBP) (27) (Figs. 1–3, A–D). In addition, the harmonin homology domain (HHD) (or harmonin-N-terminal domain) from whirlin (28),

harmonin (29), cerebral cavernous malformation 2 (CCM2) (30), and regulator of telomere elongation helicase 1 (RTEL1) (31) was assigned to the α -hub domain group (24) (Figs. 1E and 2E).

The α -hub domains consist of 3 to 5 α -helices of \sim 10 to 20 residues. Their core defining feature is an α -hairpin supersecondary structure motif (32) constituted by two consecutive antiparallel α -helices with a crossing angle close to 180° , connected by an inflexible loop (L2) (Fig. 1F). For all members, except NCBD, these two helices are helix 2 (H2) and 3 (H3) from the N terminus, and for clarity, helices are for all members numbered relative to these. In the prototypical member, L2 is folded into the five-residue link motif α_L - β_4 (33), with the β_3 -position carrying a well-sized ($>100 \text{ \AA}^3$) hydrophobic side chain that anchors between H2 and H3 (Fig. 1F). For the PAH and TAFH domains β_3 is either Ile, Leu, or Val, and for the RST domain either Ile or Met (Fig. 1, A–C). Connected to the H2-H3 core are typically two, but sometimes one or three, additional α -helices organized on the same side of the hairpin (Fig. 1F) (here referred to as the front). This leaves the other

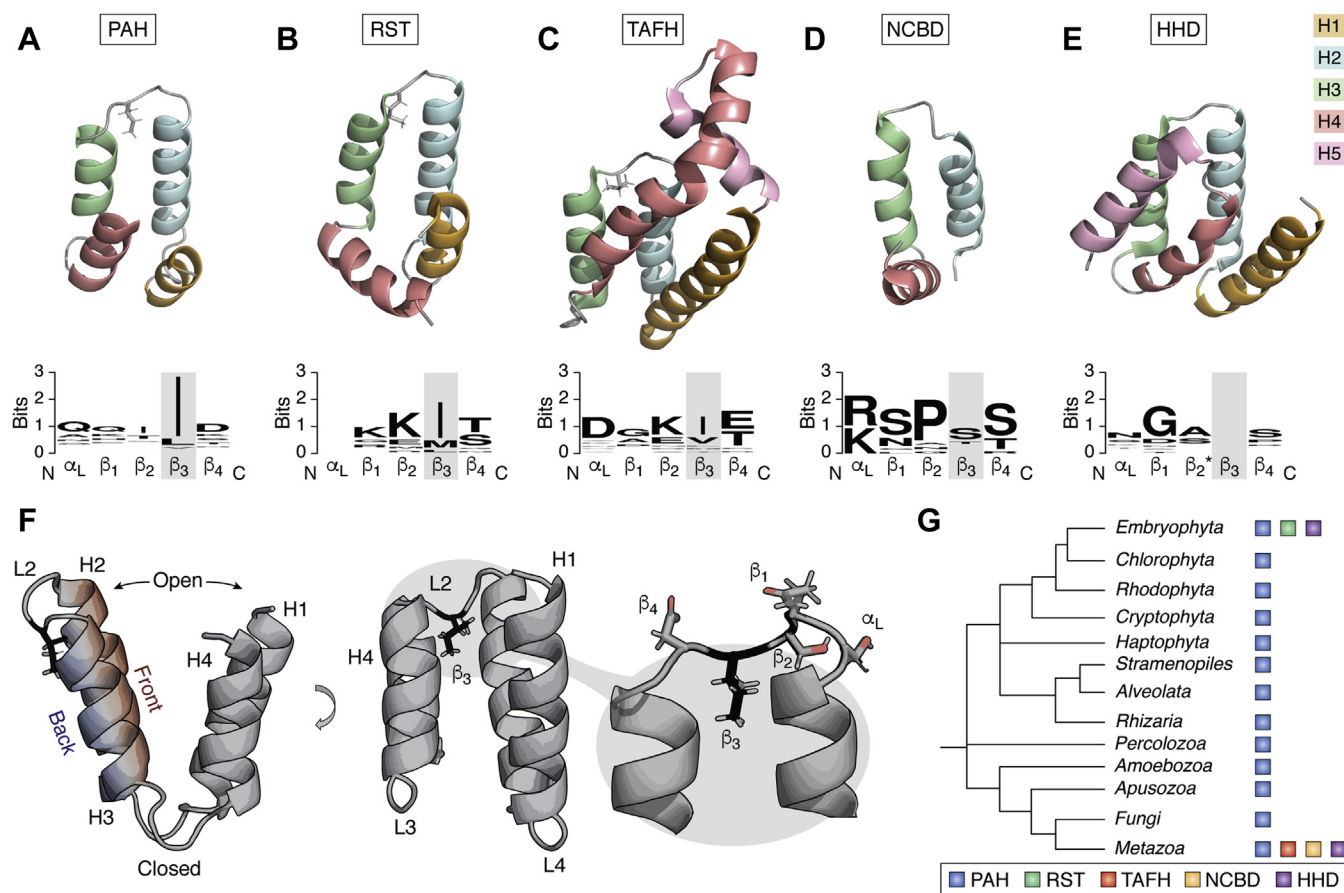


Figure 1. Structures and evolution of α -hub domains. A–E, representative structures of the current α -hub subgroups PAH (2rnr), RST (5n9q), TAFH (2pp4), NCBD (2kkj), and HHD (4fqh), respectively. For each domain, the helices are color coded with H1 (orange), H2 (blue), H3 (green), H4 (red), and H5 (pink). For the domains containing the α_L - β_4 loop, the hydrophobic β_3 -position is shown as gray sticks. Sequence logos below each domain illustrate the conservation of the H2-H3 loop region across phylogenetically representative species with each position named according to the α_L - β_4 loop nomenclature. In the structures of HHD the β_2 -residue (marked with an asterisk) is located in the site normally occupied by the β_3 -residue in the α_L - β_4 (see also Fig. 51C). Empty positions indicate either lack of conservation (for RST) or the presence of a gap in the alignment (HHD). F, compositional features of the prototypical α -hub. Side and front views illustrate the different surfaces, helices, and loops as defined in this review. Zoom shows the configuration of the α_L - β_4 loop with the hydrophobic β_3 -position forming stabilizing interactions with side chains from H2 and H3. G, evolutionary proliferation of α -hubs and relationships between major eukaryotic groups (159). Branch lengths are arbitrary. Blue, PAH; green, RST; red, TAFH; orange, NCBD; purple, HHD.

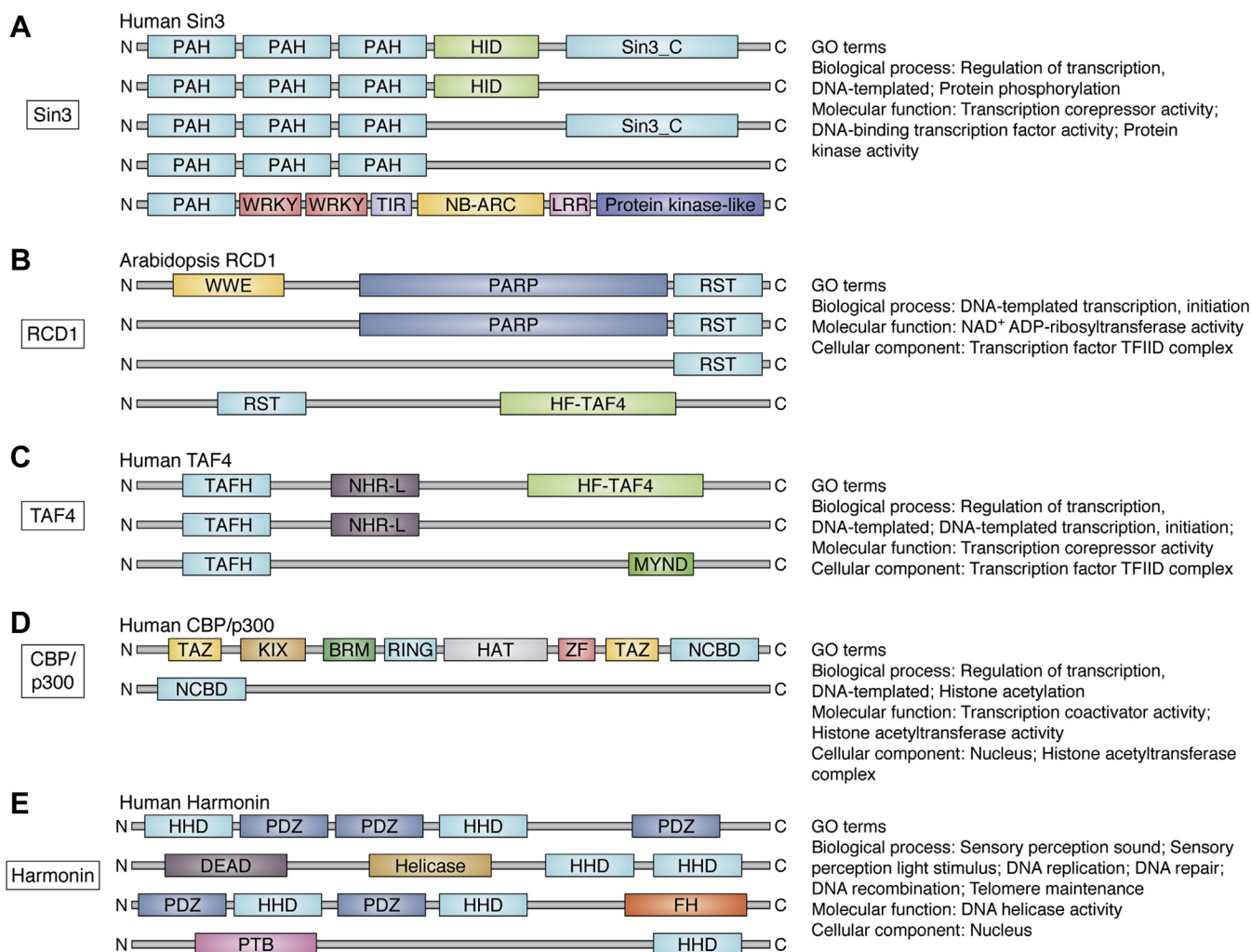


Figure 2. α -Hub protein domain structures. α -hub domain structures present in proteins with known functions or appearing more than 10 times in InterPro. The α -hub domains included are: A, PAH, B, RST, C, TAF4, D, NCBD, E, HHD. The GO terms associated with the different α -hub proteins are shown to the right. The schematics are not drawn to scale, the relative distance between the domains vary, and some of the proteins have more than one copy of the domains shown. GO, gene ontology.

side of the hairpin (the back) accessible (Fig. 1F). The anti-parallel organization of H2 and H3 orient the short, but flexible loops connecting to the preceding and proceeding α -helices at the same end of the fold, resulting in the formation of an “open” and “closed” end (Fig. 1F). Together, the helices support a hydrophobic binding cleft at the open end. The prototypical α -hub is thus a domain in a modular protein consisting of four α -helices (H1–H4), of which H2 and H3 make up the α -hairpin supersecondary structure stabilized by the hydrophobic β_3 -loop anchoring residue. The organization of H1 and H4 is the distinctive feature of each α -hub subgroup, resulting in different angles to H2–H3 (Fig. 1, A–E). Based on this difference in topology, the five different subgroups of the α -hubs can be defined: PAH1/2/3, RST, TAFH, NCBD, and HHD.

Phylogenetic proliferation of the α -hub domains

According to the literature and InterPro (34) searches, the α -hubs are exclusive to eukaryotes (Fig. 1G). The PAH

domain is present in most of the major clades of eukaryotes (35), whereas the RST domain has been reported in land plants including mosses and liverworts (36). HHD is present in animals and plants (28), whereas TAFH and NCBD have been found only in animals (37, 38). As orthologous genes have a higher degree of intron position conservation than nonorthologous genes (39), the structure of a gene may provide information about phylogenetic relationships. According to the RefSeq database at NCBI (40), plant RST genes and animal TAFH genes have a conserved intron position right before the α_L - β_4 link motif, which is missing in the remaining α -hubs. Since RST is unique to land plants and TAFH is unique to animals, but both are present in TAF4 proteins (Figs. 2, B–C and 3C) (36, 37) and they have a conserved intron structure, an early evolutionary relationship between these domains is likely. Thus, α -hub domain proteins are dominant in higher eukaryotes with a likely evolutionary link between TAFH and RST dividing them into two different kingdoms of life. The remaining α -hubs have no

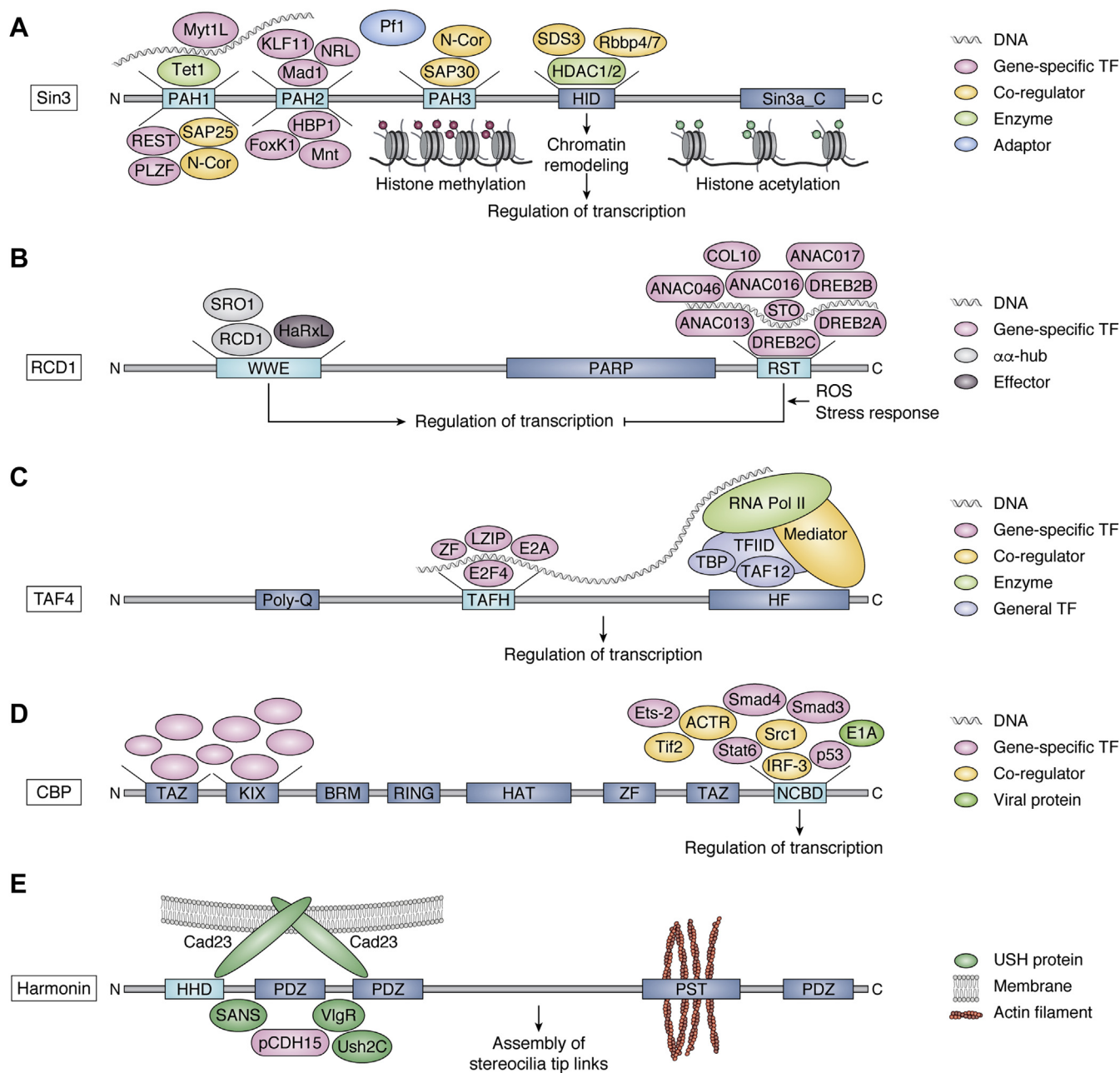


Figure 3. α -Hub protein interactions. Examples of protein complexes and interactions of α -hub proteins, focusing mainly on the interactions of the α -hub domains. *A*, Sin3 as part of a coregulator complex. TFs, coregulators, adapters, and enzymes are shown bound to or as part of the binding pool of their target domains in Sin3. The classic functional role of Sin3 as a scaffold for chromatin remodeling is also shown. *B*, RCD1 in regulation of biotic and abiotic stress responses. The interactions mediate suppression of plant immunity (through HaRxL) and coordination of communication between ROS signals emitted from mitochondria and chloroplasts (through ANAC013 and ANAC017). *C*, TAF4 as part of the RNA Pol II preinitiation complex. TFs binding to the TAFH domain and implicated in embryonic patterning, and programmed cell death are shown. *D*, CBP/p300 is a central node in eukaryotic regulatory networks regulating TFs and chromatin *via* their histone acetyl transferase activity. The TAZ, KIX, and NCBND domains are scaffolds for interactions with IDRs of proteins shown with names for only NCBND ligands. *E*, Harmonin anchored interactions in regulation of hearing and vision. Cad23 assembles the upper part of the tip link, and its cytoplasmic tail is anchored to the actin filament of stereocilia *via* binding to harmonin.

obvious evolutionary links, except for their structural similarities.

The α -hub subgroups have distinctive features and bear characteristics of analogous folds

Proteins with similar folds, such as the α -hub domains, can be divided into three general categories: homologs (derived from a common ancestor), remote homologs (less obvious

sequential similarity because of distant ancestor), and analogs (converged to similar advantageous fold independently) (41, 42). Since the number of ways nature can arrange a few secondary structural elements in a stable manner is limited, analogous folds commonly occur for small and relatively simple protein structures (42). Furthermore, analogous structures are typically similar but with distinct features and key binding site residues (41). As alluded to above, various

deviations from the prototypical features are found among the α -hub subgroups. Some PAH2 domains differ by having extended H2s and L2s (Fig. 4B) (25, 43–47), but with persistent β_3 -anchoring (Fig. 1A) (21). HHD differs by generally having an atypical L2 loop lacking β_3 -anchoring, but having a H5. In CCM2-HHD, the β_2 -residue is located in the site normally occupied by the β_3 -residue and is typically a small side chain residue (Fig. 1E and Fig. S1C), while in structures of harmonin-HHD (Protein Data Bank [PDB] codes 2kbq, 2lsr, 2kbr) (48, 49), β_3 is a Met but does not anchor between H2 and H3. For both, H5 packs between H3 and H4, possibly rescuing any lost stability from lack of β_3 -anchoring. Indeed, in harmonin-HHD the β_3 Met interacts with side chains of H5. NCBD also lacks β_3 -anchoring, but here this coincides with a lack of H1. Of note, NCBD stands out by existing in a molten globule-like state when free (50, 51). Hence, the absence of

prototypical features appears to be counteracted by helices outside the α -hairpin, either intrinsically present in the hub or from binding partners (21). This way, the hubs may maintain stability, while exposing a hydrophobic and solvent accessible binding pocket. Taken together, the α -hub domains have with their similar, but small, simplistic folds with distinct differences between subgroups, the typical characteristics of analogous folds.

To address conservation of the hub topology in terms of fold-defining positions across subgroups, we compared the sequences for each subgroup across phylogenetically representative species (Fig. 4, Figs. S1, and S2). 3D structures of each subgroup were manually inspected and compared with the sequence alignment to identify fold-defining positions (identity >50% and with tertiary side chain contacts). Within subgroups, many residues making up the hydrophobic core, and

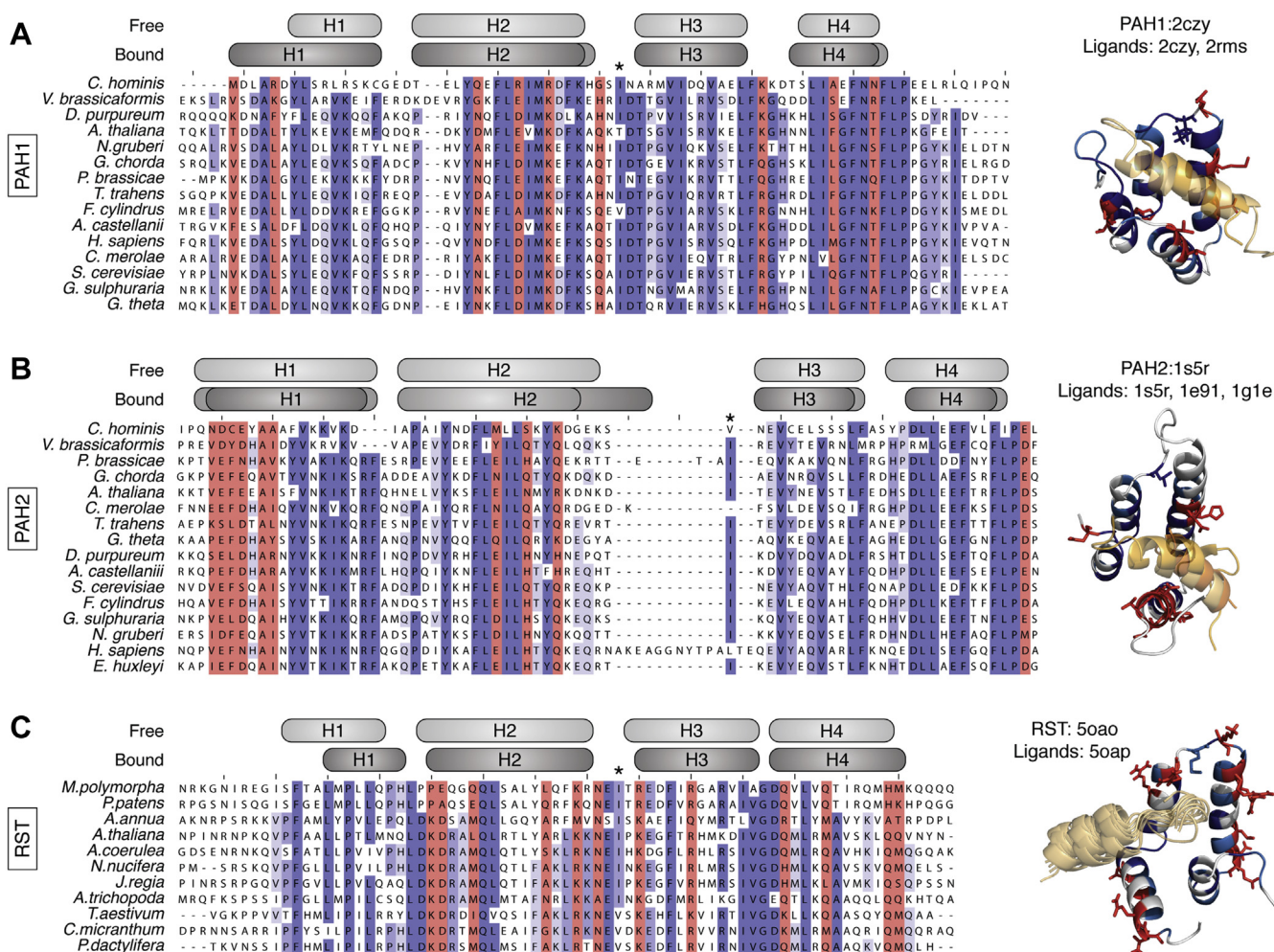


Figure 4. Alignments of sequences of PAH1, PAH2, and RST, respectively, from phylogenetically representative species and comparison with 3D structures. Sequences were aligned with Clustal Omega and visualized in Jalview. Available 3D structures of each subgroup were manually inspected and compared with the conservation alignment, and residues with identity >50% that could not be readily explained by fold-conservation (no tertiary side chain contacts) were highlighted in red (alignments and structures). The fold-defining positions (identity above 50% and tertiary side chain contacts) were colored blue in accordance with percentage identity (darker is higher identity, alignments, and structures). Above each alignment, the β_3 -position is highlighted with “*,” and the gray boxes indicate the helix boundaries in the free (light gray) and complexed (darker gray, variations are different structures) α -hubs. Species are given as four-letter abbreviations, with full names given in Table S2. A, PAH1. Protein Data Bank (PDB) codes 2czy, 2rms. The peptides of the ligands RST (2czy) and SAP25 (2rms) are shown semitransparent in orange variations. B, PAH2. PDB codes 1s5r, 1e91, 1g1e. The peptides of the ligands HBP1 (1s5r) and Mad1 (1e91, 1g1e) are shown semitransparent in orange variations. C, RST. PDB codes 5oap, 5oap. The ligand peptide of DREB2a (5oap) is shown semitransparent in yellow as an ensemble of 10 lowest-energy structures.

hence defining the fold, are highly conserved (Fig. 4, Figs. S1, and S2). Most distinctively, PAH1 and PAH2 have highly conserved cores, sharing many of the conserved core residues across all four helices, whereas PAH3 is the least conserved of all the α -hubs (Fig. 4, A–B, and Fig. S1A). Across the subgroups, however, no clear conservation pattern of even H2-H3 core residues is evident, consistent with their low sequence identity (21), and despite a high degree of core residues within each subgroup. However, despite the identified evolutionary relationship described in the previous section, a sequence-relationship between RST and TAFH could not be established through this analysis (Fig. 4C and Fig. S1B). Hence, the structural similarity between the α -hubs cannot be traced from any recognizable sequential relatedness, and besides the conserved intron structure between RST and TAFH, we found no evidence to support emergence from a common ancestor. Rather, the α -hub folds should be considered analogous folds (52), although more extensive analysis would be required to rule out remote homology. As a consequence, the possibilities for identification of new α -hubs directly from sequence alone is currently limited. An alternative will be searches through 3D-structure alignment using, e.g., PDBeFold (53), as done in the defining work on the α -hub group (21). However, this approach is naturally limited to targets with described 3D structures, and the identification of potential additional α -hub group members is therefore a challenge.

The functions of α -hub domains

The α -hub domains are linked to different domains of diverse functions

To obtain an overview of the domain compositions of the α -hub domain proteins, we searched the literature and InterPro (34). Most PAH-domain proteins, including Sin3, also contain a histone deacetylase (HDAC) interacting domain (HID) and a Sin3 C-terminal domain (Fig. 2A) (54), but numerous PAH-domain proteins contain only some of these domains. The PAH domain is also present in the plant protein WRKY19, which additionally contains a WRKY DNA-binding domain, a kinase domain, and a central TIR-NB-ARC-LRR module implicated in plant immunity (55). The gene ontology (GO) terms for the PAH domain proteins suggest a function in transcriptional regulation. The RST domain is found in RCD1 and is responsible for most RCD1 interactions (56–58). In addition, RCD1 contains an N-terminal WWE domain followed by a poly(ADP-ribose) polymerase (PARP) domain (Fig. 2B). The RST domain is also present in proteins lacking either the WWE or the PARP domain, or both, and in combination with a histone-fold (HF) domain in plant TAF4, which is reflected in GO terms related to transcription. Human TAF4 consists of a TAFH domain followed by a HF domain and is crucial for structural integrity of the TFIID complex (37, 59) (Fig. 2C). TAFH-domains are also found in conjunction with NHR-like domains and in ETO proteins, in which a MYND zinc finger for corepressor recruitment is also found (60). Overall, GO terms reveal a function of TAFH domain proteins in transcription. The multidomain proteins, CBP and its paralog p300 (19), both have histone acetyltransferase activity

(Fig. 2D), as reflected in the GO terms associated with the NCBD-containing proteins suggesting functions within transcription. HHD is present in proteins with several PDZ domains, as in the case of whirlin (28) and harmonin (61) (Fig. 2E). In addition, HHD is found in combination with DEAD-, phosphotyrosine binding-, and formin homology domains (28). The GO terms for the HHD proteins suggest functions in sensory perception and telomere maintenance. Thus, both similar and versatile functions and domains are linked to α -hub domains.

Orchestration of function from networks by α -hubs

As hubs, the α -hub domains serve to organize larger networks and multicomponent complexes. Sin3 proteins are coregulators of gene expression and implicated in processes such as cell cycle regulation, energy metabolism, senescence, and organ development (for recent reviews see (54, 62, 63)). Early studies showed that Sin3 is associated with HDAC1 and HDAC2 in multiprotein complexes, with its central domains, PAH3 and HID (Fig. 3A), interacting with the core complex components HDAC1, HDAC2, Rbbp4/7, SAP30, SAP18, and SDS3 (64–68). The Sin3–HDAC complex mediates histone deacetylation, which together with methylation, leads to gene repression (69), but Sin3 also interacts with the DNA demethylase Tet1 to regulate transcription epigenetically (70). The PAH1 and PAH2 domains bind numerous TFs, as shown using various different methods including biochemical methods such as pulldown assays (44, 71–75) and fluorescence anisotropy (76), genetic methods such as yeast two-hybrid assays (75, 77) and biophysical methods such as NMR spectroscopy (44, 76, 78). The TF ligands include REST/NRSF (71), PLZF (72), Mad1/Mdx1 (45), NRL (79), HBP1 (44), FoxK1/MNF- β (75), Mnt/Rox (73), KLF11/TIEG2 (74), and Myt1L (76), which recruit the Sin3 complex to target genes to regulate expression (54). The importance of the PAH domain:ligand interactions is apparent from several studies. For example, Tet1 depends on interactions with Sin3a-PAH1 for repression of transcription in cells (80) and for PLZF, which interacts with Sin3a-PAH1, histone deacetylation inhibition interferes with its ability to mediate transcriptional repression (72).

In agreement with RCD1 being a hub (56, 81), *rcd1* knockout mutants, which have premature stop codons in the region encoding the PARP domain (Fig. 2B), thus affecting the RST and PARP domains, display pleiotropic phenotypes in gene expression, stress responses, and developmental processes. More specifically, *rcd1* shows increased ozone and salt sensitivities, changed leaf morphologies and early flowering times, as well as altered stomatal regulation (56, 81–83). The WWE domain interacts with itself and the RCD1 paralog SRO1, and with the downy mildew effector HaRxL106 to suppress plant immunity (84) (Fig. 3B). The interactions of RCD1-RST with TFs, which have been studied using both yeast two-hybrid assays and biophysical techniques (21, 56, 58), play important roles in plant biology. Thus, the RST domain of RCD1 affects stress responses *via* interactions with DREB2a (85) and ANAC013 and ANAC017 (86, 87), the latter two of which

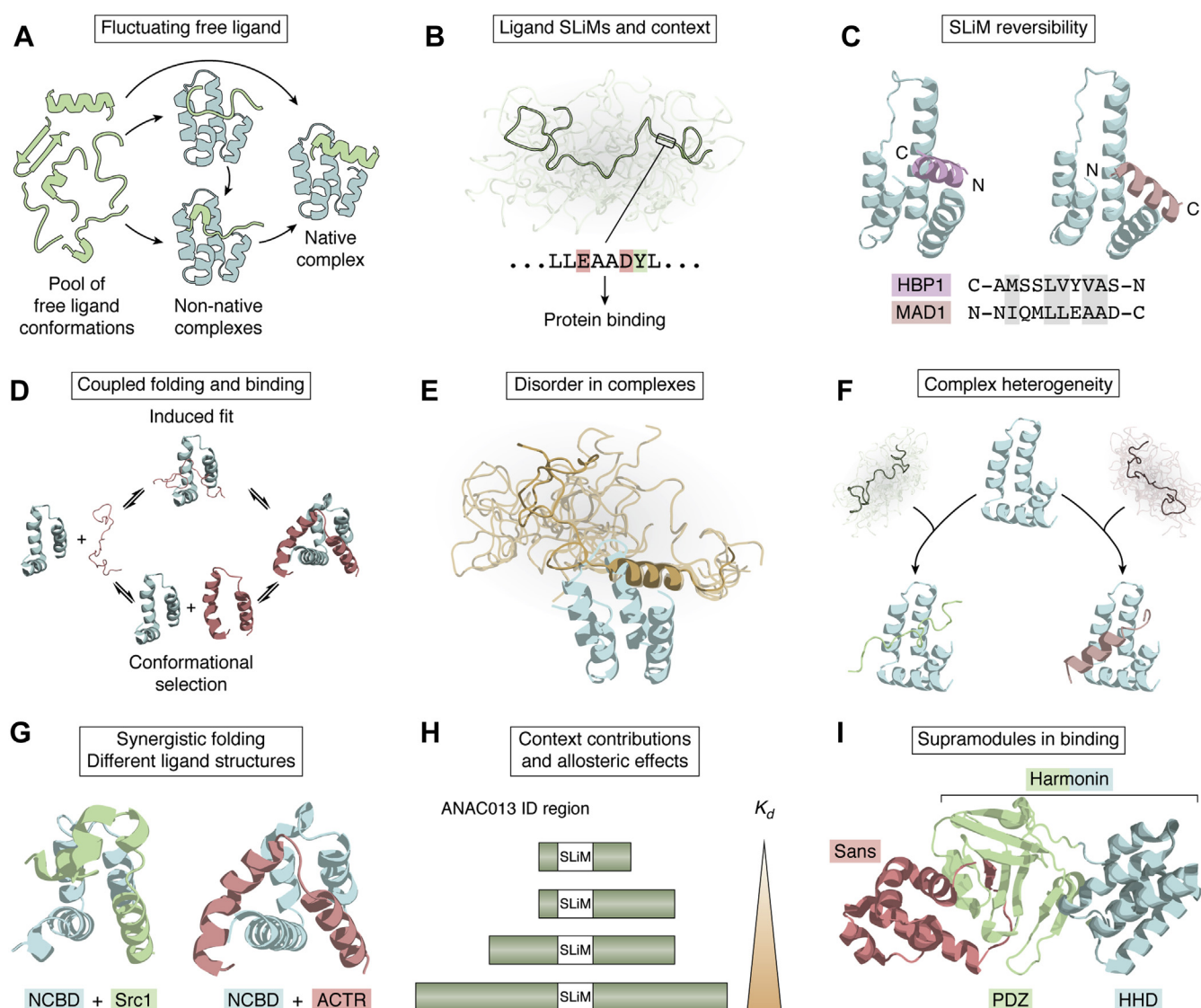


Figure 5. The modus operandi of α -hubs. *A*, the α -hub-binding region of free protein ligand may fluctuate between hairpins, helices, and bent structures as in the case of the Sin3b-PAH1-binding SLiM of REST (107). *B*, protein ligand using a SLiM with hydrophobic and acidic residues for α -hub binding as in the Sin3-PAH2-binding SLiM of Mad1 (108). The SLiM is often part of a larger intrinsically disordered context. *C*, protein ligands may use SLiM reversibility for governing specificity as in the case of Sap25 and REST binding to Sin3-PAH1 (Protein Data Bank [PDB] codes 1s5q and 1s5r) (44). *D*, ligands using coupled folding and binding, through conformational selection and/or induced fit, as in the case of ACTR association with NCBD (based on PDB codes 2kkj and 1kbh) (132). *E*, α -Hub:ligand complexes may retain some disorder as in the case of the Sin3a-PAH1:SAP25 complex (PDB code 2rms) (103). *F*, structural heterogeneity in an α -hub:ligand complex as in the case of RCD1-RST complexes with NAC and DREB2a transcription factors (PDB codes 5oao and 5oap) (21, 36, 58). *G*, α -hub domains may fold synergistically with a disordered protein ligand to form different bound ligand structures as in the case of NCBD complexes with Src1 (left) and ACTR (right), respectively (PDB codes 2c52 and 1kbh) (121). *H*, allosteric effects of the SLiM context on ligand association with α -hubs as in the case of RCD1-RST association with ANAC013 (58). *I*, α -hubs may be part of supramodules as in the case of the harmonin:sans complex (PDB code 3k1r) (48).

contribute to coordination of reactive oxygen species signals emitted from mitochondria and chloroplasts (86).

The TAFH protein TAF4 is crucial for structural integrity of the TFIID complex (59), which contains 13 additional TAF subunits and TATA binding protein (TBP) that by binding to genes triggers formation of the transcriptional preinitiation complex (37, 88, 89) (Fig. 3C). This, in addition, contains RNA polymerase II, general TFs, and the large Mediator complex. Through HF domains, TAF4 interacts with TAF12 to stabilize the TFIID complex (90). The TAFH domain contributes to the regulation of the expression of approximately 400 genes (91) and has been experimentally shown to interact directly with

TFs such as ZF and LZIP (26). TAFH binding of the E-protein TFs, HEB and E2A, implicated in embryonic patterning and programmed cell death (92), is critical to gene activation by enhancing TFIID promoter binding (91).

CBP is a central node in eukaryotic regulatory networks (19) and regulates TFs and chromatin *via* its histone acetyl transferase activity (93). The TAZ, KIX, and NCBD domains form the scaffold for the interactions of CBP with IDRs of regulatory proteins (19). NCBD alone has multiple experimentally identified interaction partners, including IRF-3 (27), p160 nuclear receptor coactivator 1 (NCOA1;Src1), NCOA2 (Tif2), and NCOA3 (ACTR) (27, 94, 95), tumor suppressor p53 (96), Ets-2

(27), Smad3 and 4 (97), Stat6 (98), and the adenoviral protein E1A (27) (Fig. 3D). Also, in these cases, the ligands may depend on interactions with the α -hub domain, both *in vivo* and *in vitro*, as in the transcriptional activation by IRF-3 (27).

The HHD proteins whirlin and harmonin are implicated in Usher syndrome causing hearing-vision loss (99). Usher syndrome proteins are organized in interactomes with harmonin, whirlin, and sans as scaffolds and cadherin23 (Cad23), protocad15, sans, VlgR, and Ush2C binding to harmonin (Fig. 3E) (100, 101). Structural and thermodynamic analyses have shown that HHD and PDZ1 of harmonin form a supramodule that binds sans with high affinity (48), and harmonin-HHD also binds Cad23 (61). This interaction, together with the harmonin-PDZ2: Cad23 interaction, represents multidentate binding *via* supramodule exploitation (48), providing a structural platform for the tip link complex of stereocilia (61). Furthermore, the tail of Cad23 promotes Cad23:harmonin polymer formation by binding to harmonin-HHD or by self-dimerization (49). Harmonin thus connects tip link complexes with the actin cytoskeleton (102). For an α -hub protein, harmonin has an atypical biological function and sensory perception, but typical molecular function in scaffolding.

The α -hub domains as protein–protein interaction hubs

Disordered α -hub ligands have SLiMs of similar characteristics that maintain specificity

Many α -hub ligands use IDRs for binding, but identification of most α -hub ligands dates back before the general appreciation of ID. Still, ID has often been mentioned as a feature of the free state of the hub-binding regions (46, 78, 103) or has been computationally predicted (24), whereas experimental characterization of the IDRs has mostly appeared in studies of RST (36, 57, 104, 105) and NCBD ligands (94, 96, 106). Molecular dynamics (MD) simulations have also been used for characterizing free α -hub ligands. Thus, the Sin3b-PAH1-binding region of REST was suggested to fluctuate between hairpins, helices, and bent structures with population shifts and induced folding working cooperatively in coupled folding and binding (Fig. 5A) (107). ID-associated flexibility provides the structural adaptability needed for REST to function as a hub itself, and for the α -hub ligands, ID is in general a prerequisite for adaptable SLiM-based interactions.

Although different intrinsically disordered ligands use different SLiMs for α -hub binding, most are simple and depend on hydrophobic residues for contacts with the hydrophobic α -hub cleft (Fig. 5B). Initial work to identify a PAH2-binding SLiM based on screenings, sequence comparisons, ligand affinity measurements, as well as structural analysis revealed the motif ϕ ZZ ϕ ϕ XAAXX ϕ nXXn (X, nonproline residue; ϕ , bulky hydrophobic residue; Z, aliphatic side chain; n, negatively charged) (25, 43, 45, 79). Later, structural work identified two orientations of PAH-bound SLiMs, types I and II, as exemplified by the PAH1-binding SLiMs from REST (ϕ X ϕ ϕ SX ϕ S) (71, 103) and Sap25 (103) (S ϕ XS ϕ ϕ X ϕ) (S, short side chain) (Table 1), respectively. In PAH2-complexes, the SLiMs of Pfl and Mad1 (ϕ ZZ ϕ ϕ XAAXX ϕ n) and of HBP1 (A(A/V)X ϕ ϕ XX ϕ) also adapt different orientations (Fig. 5C) (43, 44, 46). Despite fold similarities and SLiM simplicities, the α -hubs show remarkable selectivity. The \sim 40 times difference in affinities of Sin3-PAH2 for Mad1 ($K_d \sim$ 50 nM) (45, 103, 108) and Pfl ($K_d \sim$ 2 μ M) (46) was explained by a phenylalanine in the first position of the Pfl-SLiM constituting a steric disadvantage (46). The minimal Mad1-SLiM consists of eight residues, with only three being essential for the interactions with Sin3-PAH2. One of these, L12, inserts into the hydrophobic cleft of PAH2 and is important for affinity, whereas the other two, A15 and A16, determine specificity for PAH2, owing to their proximity to bulky side chains of PAH2 in the complex. Thus, hydrophobic residues are implicated in both affinity and specificity of PAH:SLiM interactions (108).

A combined bioinformatics and experimental approach, including substitution analysis, was used to identify the RST-binding SLiM (D/E)X(1,2)(Y/F)X(1,4)(D/E)L (where X(1,2) denotes 1 or 2 Xs) (Table 1), which has essential binding contributions from aromatic, acidic, and leucine residues (57, 58). The RST domain and the RST-binding SLiM were traced back 480 million years to the emergence of land plants, and SLiM variants, identified from the evolutionary analysis, suggested numerous additional RCD1-interactome members (36). Among the few known TAFH ligands (Fig. 3D), the TFs HEB, cMyb, and STAT6 and the corepressor N-Cor use the SLiM (D/E) ϕ X ϕ ϕ for binding ETO-TAFH (109, 110). Using phage display, D ϕ ϕ XX ϕ ϕ was identified as the TAF4-TAFH-binding SLiM present in ZF, LZIP, and E2A (26). The lack of a common NCBD-binding SLiM likely reflects partner-templated modulation of the NCBD structures. However, similar

Table 1
 α -Hub-interacting SLiMs

Ligand	α -Hub	SLiM	Reference
ACTR	CBP-NCBD	ϕ ϕ XX ϕ and ϕ XX ϕ ϕ ^a	(111)
ANAC013/016/017/046, bZIP23, COL10, DREB2a/b/c, STO	RCD1-RST	(D/E)X(1,2)(Y/F)X(1,4)(D/E)L	(58)
ANAC087	RCD1-RST	(Y/F)X(1,4)(D/E)(LI)	(36)
cMyb, HEB, N-Cor, STAT6	ETO-TAFH	(D/E) ϕ X ϕ ϕ	(109, 110)
E2A, LZIP, ZF	TAF4-TAFH	D ϕ ϕ XX ϕ ϕ	(26)
HBP1	Sin3-PAH2	A(A/V)X ϕ ϕ XX ϕ (type II)	(44)
Mad1, Pfl	Sin3-PAH2	ϕ ZZ ϕ ϕ XAAXX ϕ n (type I)	(43, 46, 108)
p53	CBP-NCBD	ϕ ϕ XX ϕ and ϕ XX ϕ ϕ	(96)
REST	Sin3-PAH1	ϕ X ϕ ϕ SX ϕ S (type I)	(71, 103)
SAP25	Sin3-PAH1	S ϕ XS ϕ ϕ X ϕ (type II)	(103)

^a X, nonproline residue; ϕ , bulky hydrophobic residues; Z, aliphatic component in the side chain; S, short side chain; n, negatively charged.

Table 2
Thermodynamic parameters of the α -hub

Domain	ΔG (kcal mol ⁻¹)	ΔH (kcal mol ⁻¹)	ΔC_p (kcal mol ⁻¹)	M (kcal mol ⁻¹ M ⁻¹)	T_m (°C)	Lack of rigidity ^a					Reference
						H1	H2	H3	H4	H5	
mSin3a-PAH1						X					(103)
mSin3a-PAH2						X					(44, 103)
mSin3b-PAH2						X ^b					(25, 43)
hSin3b-PAH1	12.9 ± 0.5	140 ± 1	1.53 ^c		65.7 ± 0.4						(140)
hSin3b-PAH2	10.2 ± 0.5	119 ± 1	1.53 ± 0.02		64.4 ± 0.4						(140)
hSin3b-PAH3	9.9 ± 0.5	117 ± 1	1.54 ± 0.02		63.7 ± 0.4						(140)
A/RCD1-RST	3.8 ± 0.2					X					(21)
hETO-TAFH						X				X	(158)
hCCM2-HHD					75	X					(30)
whirlin-HHD						X				X	(28)
hCBP-NCBD	1.5 ± 0.1			0.69 ± 0.02				X			(126, 129)

^a The X corresponds to lack of rigidity experimentally obtained by X-ray crystallography or NMR spectroscopy of the different α -helices (H).

^b The observation has been evaluated based on the structure of the complex and not on the free state.

^c The value has not been experimentally determined but derived from fitting.

SLiMs, $\phi\phi XX\phi$ or $\phi XX\phi\phi$, mediate the interactions between NCBD and ACTR, and the TRD regions activation domain (AD)1 and AD2 of p53 (96, 111). HHDs have only a few identified ligands and no known SLiMs. Still, similar to other α -hub ligands, hydrophobic residues are prominent in the HHD-binding ligand region (49, 61).

In summary, the simple generic $\phi XX\phi\phi$, recurring in TF:coregulator interactions (17), is also dominant among the α -hub-interacting SLiMs, which use both hydrophobic and charged residues for securing binding affinity and specificity. Furthermore, PAH1 and PAH2 may use SLiM reversibility for governing specificity.

The affinities and thermodynamic profiles of α -hub interactions vary

The affinities of the α -hub:ligand interactions have been determined using a number of different methods including stopped-flow fluorescence spectroscopy, fluorescence titration, NMR spectroscopy, surface plasmon resonance, and isothermal titration calorimetry (ITC), with ITC being the most frequently used (43, 58, 61, 96, 106, 111, 112) (Table S1). In addition to providing information about affinities, ITC also allows determination of changes in binding enthalpy and entropy. It is generally assumed that IDRs pay an entropic cost upon binding owing to conformational restrictions (113, 114). However, IDRs may also use entropy for binding through counter-ion release (115), increased conformational flexibility (116), or expansion of the surrounding IDRs (117). For the α -hubs, complexes form with K_d s ranging from low nanomolar to mid micromolar, with most affinities in the low micromolar range (Table S1). In the high-affinity end, the Sin3a-PAH3:Sap30 complex has a K_d of 9 nM, resulting from cooperative recognition of two discrete Sin3a-PAH3 surfaces by the tripartite binding region in SAP30 (66). The high affinity may reflect constitutive Sin3:SAP30 association (68). A similar high affinity (K_d 9 nM) was measured for the RCD1-RST:ANAC013 complex, notwithstanding the lack of demonstrated induced structure in ANAC013 upon binding (58). The Sin3a-PAH2:HBP1 complex has a $K_d \sim 2$ orders of magnitude larger than that of the Sin3a-PAH2:Mad1 complex (44, 45, 103, 108), possibly reflecting the biological functions of the two ligands with Mad1 replacing HBP1 in Sin3a complexes during differentiation (44). Large differences in affinities (40-fold) have also been detected for Sin3a-PAH1 interactions with SAP25 and Myt1L, explained by the Myt1L-SLiM diverting from the canonical SLiM (76).

The thermodynamic profiles for ligand binding vary among the hubs, even for the same α -hub under the same experimental conditions. Some complexes are entropy driven, as exemplified by RCD1-RST complexes with Col10 (ΔH -9.2 kJ mol⁻¹, $-T\Delta S$ -27.2 kJ mol⁻¹; K_d 418 nM), STO (ΔH -3.8 kJ mol⁻¹, $-T\Delta S$ -36.3 kJ mol⁻¹, K_d 90 nM), and ANAC087 (ΔH -15.9 kJ mol⁻¹, $-T\Delta S$ -16.9 kJ mol⁻¹, K_d 1.8 μ M) (42). Other, such as the RCD1-RST:DREB2a (ΔH -63.3 kJ mol⁻¹, $-T\Delta S$ 18.7 kJ mol⁻¹, K_d 16.0 nM) and the CBP-NCBD:ACTR (ΔH -132.6 kJ mol⁻¹, $-T\Delta S$ 89.1 kJ mol⁻¹, K_d 34 nM) complexes are driven by enthalpy (36, 111). In a

study addressing binding of Sin3 isoforms, Sin3a-PAH2 and Sin3b-PAH2 bound Pfl with comparable affinities but apparently different thermodynamic profiles. This likely reflects that apo-Sin3a-PAH2 samples both folded and partially folded conformations and forms a monomer–dimer equilibrium and that apo-Sin3b-PAH2 is monomeric and mostly folded (see below) (43, 46, 118, 119). Accordingly, Sin3a-PAH2 and HBP1 undergo mutual coupled conformational transitions upon association (44). The thermodynamics of the α -hub interactions with ligands thus appears diverse, ranging from highly entropically to highly enthalpically driven. However, since the different ligands can be folding to different degrees upon binding, and since both ΔH and ΔS vary with temperature, a comparison of the profiles across the different hubs is complex. Thus, it would be relevant to include more in-depth analyses under varying temperatures, which will allow determination of ΔC_p , and through that infer on differences in binding-induced folding.

Properties of α -hub–ligand complexes

The α -hubs share a common superset with topological variations

The majority of α -hub complex structures have been solved with ligand peptide fragments, entailing an amphipathic α -helix bound through coupled folding and binding (Fig. 5D) in the hydrophobic cleft (21, 25, 43–46, 71, 80, 103, 105, 120). For all the α -hubs, this occurs without substantial changes to the backbone structure of the α -hub, and thus while maintaining the relative helix orientations. For PAH1/2/3, RST, and HHD, binding of these ligands occurs in a shared superset (52), consisting of the hydrophobic cleft formed at the open end of the fold (Figs. 1, A–B, E–F and 4, A–C and Fig. S1, A and C). The shared location of a binding site within apparent analogous domains suggests that it has arisen because it is a particularly advantageous structural motif (52). In this case, the open-end hydrophobic cleft seems particularly well suited for versatile binding of IDRs forming amphipathic α -helices upon binding. NCBD only fully populates the α -hub fold upon complex formation with some ligands (111, 121) and hence does not have the superset in a traditional sense. For PAH1/2/3 and HHD, the cleft is primarily located between H1 and H2. Here α -helices engage in a mostly hydrophobic contact surface of 650 to 750 Å² (49, 61, 103), which is close to the average protein–protein interface size of 800 Å² (12). For RST, the cleft opening is primarily located between H3 and H4 (21). TAFH deviates from the prototypical α -hub traits by its hydrophobic open-end cleft (*i.e.*, the superset) being occupied by a repositioned H4 (Fig. 1, C and F and Fig. S1B). Structures of TAFH complexes revealed binding of ligands in the interfaces between H1 and H4 (Fig. S1B), resulting in a mostly hydrophobic contact surface of 700 Å² (109). This different relative orientation of H1, H4, and H5 in the α -hubs, resulting in different positioning of side chains and geometry of the binding site, may be an additional filter for specificity tuning (21).

The positions of binding residues are not always conserved across subgroups

The sequence alignments of the individual α -hub subgroups presented above together with manual inspection of 3D structures allowed identification of fold-defining residues (Fig. 4, Figs. S1 and S2). However, each subgroup also revealed between 4 to 13 conserved residues that cannot be explained by apparent fold-conservation (>50% identity, lack of tertiary contacts) (Fig. 4, Figs. S1 and S2). These are likely conserved because they are crucial components of interaction sites. In, *e.g.*, PAH2, TAFH, and NCBD, 7 of 10, 6 of 7, and 10 of 13, respectively, of the suggested binding residues are in known complex structures indeed in contact with ligands (25, 43–46, 96, 109, 111, 122).

In the α -hub complex structures, the majority of ligand contacts are through the open-end hydrophobic superset (21, 25, 43–46, 48, 49, 71, 96, 103, 109, 111, 122–124). Since the analysis does not pick up residues of the hydrophobic superset that are also part of the core fold, the identified potential binding residues are primarily solvent exposed and, hence, the majority are hydrophilic and charged (Fig. 4, Figs. S1 and S2). All the α -hub domains have conserved binding residues in both H1 and H2 (except for H1 of RST), whereas their presence in H3, H4, and H5 varies between subgroups. The relative position of the conserved binding residues is, however, not consistently conserved throughout the domains, supporting that binding discrimination may be partially encoded in the position of key residues. Even for PAH1 and PAH2, which as described above have many common conserved core residues, the pattern of conserved binding residues is entirely different (Fig. 4, A–B). For PAH1, the 9 identified residues are distributed throughout the domain, whereas for PAH2, 6 of 10 residues are in H1 and none is in H3 and H4. Nonetheless, when inspecting their positions in available structures (Fig. 4, A–B), it is clear that they cluster around the open-end binding pocket between H1 and H2 in both PAH1 and PAH2. This difference is consistent with previous studies showing that conservative replacements of PAH2 residues with equivalent PAH1 residues were sufficient to alter affinity as well as specificity. Thus, substitution of Sin3-PAH2-Leu332, positioned in H2 of the ligand-binding cleft of PAH1/2, with Met, present in the corresponding position in PAH1, resulted in a 7-fold decrease in the affinity for Mad1 (103). For all the α -hubs, particularly the last or second-to-last turn of H2 situated between the core and solvent-exposed side of H2 almost always has a conserved binding residue, which is in contact with ligands in known structures (Fig. 4 and Fig. S1).

NCBD is an outlier, only substantially populating the α -hub fold with certain ligands. Here, the ligand takes the position of H1 in the complex, resulting in many conserved residues engaging in intermolecular interactions. For this reason, it is omitted from the cross comparisons. From the sequences and α -hub-like structures of NCBD (111, 125), 13 conserved binding residues were identified, and 10 of these can be recognized as partaking in complexes (Fig. S2, (111, 125)).

Table 3
Outstanding questions

- Is the α -hub group complete?
- What roles do SLiM contexts play in α -hub:ligand interactions?
- To what extent is disorder and heterogeneity present in α -hub:ligand complexes?
- To what extent is avidity and allosteric regulation used in α -hub:ligand interactions?
- To what extent do α -hubs exploit more than one binding site/surface?
- Can ligand binding occur without involving the hydrophobic supersite?
- Does communication between different α -hub-binding sites allow allosteric regulation?
- What role does α -hub context play in selectivity?
- Is cooperativity between α -hub domains and other (α -hub) domains and supramodules common in binding?
- Is the structural malleability a prerequisite for binding folded partners?
- Does hub stability and flexibility determine selectivity and interactome size?
- How do posttranslational modifications affect α -hub interactions and stability?

The remainder are on the backside of H2-H3 (two residues) or in H2 (one residue) and may engage in complexes with other ligands.

Conserved binding residues across subgroups suggest expanded binding

Many of the identified binding-conserved residues are at the rim of the hydrophobic supersite. However, a subset has geometrically distant locations. Especially noteworthy is that all α -hubs have conserved residues positioned at the backside of H2-H3, and a few also on the solvent-exposed side of H1 (PAH2, CCM2-HHD, RST), H4 (PAH1, RST), or H5 (CCM2-HHD). This suggests these to constitute one or more accessory binding (super)sites. A few complex structures solved with relatively large intrinsically disordered ligand fragments of ~60 to 90 residues (Sin3-PAH1: PDB 2rms (103), Sin3-PAH3: PDB 2ld7 (66)) or folded partners (CCM2-HHD: PDB 4y5o (124), Mtgr1-TAFH: PDB 5ecj (122)) are available. In the Sin3-PAH1 complex, the additional ~25 disordered residues do not engage with the α -hub (Fig. 5E), whereas in the Sin3-PAH3 complex, the additional ~60 residues, intrinsically disordered in the free ligand, form two α -helices engaging with conserved contact residues on the backside of H2-H3 (Fig. S1A). A similar pattern is observed in the CCM2-HHD complex with a folded partner (PDB 4y5o (124)), where a helix of the ligand interacts through the hydrophobic supersite, while additional contacts are made to the backside of H2-H3 (Fig. S1C), again including a conserved contact residue. For the complex of Mtgr1-TAFH with a larger folded partner (PDB 5ecj (122)), the α -hub is almost completely buried within the partner, with H1 as the main anchor and the backside of H3 fully exposed (Fig. S1B). Hence, several structures of α -hub complexes confirm that ligand binding is not limited to the hydrophobic supersite. This, together with the shared pattern of conserved binding residues, especially on the H2-H3 backside, suggests that α -hub binding is more complex than a single supersite cleft, and that this is a shared property across the hubs.

Structural heterogeneity and hub flexibility in α -hub complexes

Protein complexes that involve IDRs may maintain varying degrees of disorder. Indeed, disorder is also preserved in some

α -hub complexes, such as the Sin3a-PAH1:SAP25 complex, in which the SAP25 peptide is largely unstructured (Fig. 5E) (103), although any functional implications of this remains to be determined. For complexes of RCD1-RST with different TFs, ligand plasticity is also likely. Although the α -helical structure was induced in DREB2a, no signs of helical structure were observed for ANAC046 and ANAC013 upon complex formation (Fig. 5F) (58). Thus, for RST different structures may form in the hydrophobic supersite, a trait so far not observed for other α -hub interactomes.

NCBD:ligand interactions are also diverse. CBP-NCBD binds both ACTR and Src1, with the hub and the ligands existing in a molten globule-like and disordered form, respectively (111, 126), and together they cooperatively fold to form helical entities with similar CBP-NCBD folds, but different ligand (helical) topologies (Fig. 5G) (111, 121). CBP-NCBD and the p53-TRD also fold synergistically, with p53 forming a pair of helices docking into a hydrophobic groove of NCBD, in this case separated by a flexible chain (125).

As they have many partners (Fig. 3), α -hub domains must maintain interdomain selectivity, and mechanisms used by other hubs may also be relevant for the α -hubs. Here, flexibility and adaptability by the hub itself is a mechanism highly exploited by the hub's calmodulin (127) and TAZ1 (128). However, to date, studies on the thermodynamic stability and the dynamics and malleability of the α -hub domains remain scarce and limited by studying isolated domains. Helix dynamics of the α -hubs from NMR- and X-ray crystallography data and thermodynamics data of their folding and unfolding (Table 2) suggests that the H2-H3 α -hairpin is rigid and stable, whereas H1, particularly its N-terminal end, is the most dynamic part in all subgroups (21, 25, 28, 30, 44, 103). Some examples of increased dynamics in the C-terminal region of H4 also exist, primarily for α -hubs having a fifth helix (HHD and ETO-TAFH) (28, 43) and for NCBD (126, 129), whose fold is partner dependent. For TAFH, however, it is unclear if the C-terminal dynamics has been affected by premature termination of H5. Although we cannot rule out that dynamics in the hubs arise from domain excision, the dynamics of H1 appears to be independent of the length of the N-terminal tails, and thus suggests this to be an inherent property of the fold. Some of the C-terminal regions are stabilized by folding upon binding with ligands, suggesting the formation of a localized folding transition after binding (103), whereas the lack of rigidity in the N terminus is more common after binding (25, 43, 44). Thus, the α -hairpin super-secondary motif constitutes a structurally stable platform onto which dynamic α -helices can be organized, allowing for flexibility and subtle adaptations to binding partners.

Mechanisms and concepts of α -hub *modus operandi*

The α -hub topology supports complex binding kinetics, cooperativity, avidity, multivalency, and supramodules

Mechanistic dissection of the α -hub:ligand interactions is likely to contribute new conceptual understandings of hub:ID interactions, as identified in (18, 130). The interactions

between CBP-NCBD and IDRs from Src1, Tif2, ACTR, and p53 showed high association rate constants (approximately $1 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$) and ionic strength dependence, reflecting the importance of electrostatics in ID-based α -hub interactions (112). The kinetics of CBP-NCBD:ACTR association is complex, with a heterogeneous transition state reflecting an encounter complex with large structural variations (112, 131). Increased helicity in free ACTR resulted in increased k_{on} and decreased k_{off} for CBP-NCBD:ACTR (132), hinting at binding through conformational selection (Fig. 5D). However, MD simulations suggested that the pre-existing structure in ACTR accelerates association kinetics by promoting folding upon encounter (133). In this case, reduced k_{off} due to increased helicity in ACTR was explained by smaller entropic costs of forming the bound state. In another MD simulation study, CBP-NCBD was proposed to fold through global conformational selection and local induced fit upon p53 binding (134). Contributing to the complexity, two subpopulations of CBP-NCBD, *trans* and *cis*-Pro populations, bind ACTR with different affinities and kinetics, possibly representing a mechanism for interaction-based regulation of signaling (135). So far, mechanistic studies of the other α -hubs are lacking.

Additional biochemical concepts are emerging from α -hub:ligand studies. Thus, full-length Sin3a and a Sin3a fragment consisting of PAH1 and PAH2 interacted more strongly with Mad1 than the individual domains. PAH1 unlikely functions independently in Mad1 binding but instead cooperates with PAH2 (108). A similar pattern was seen for Sin3a:Tet1 association (80). Pfl functions as an adaptor through multivalency by interacting with MRG15 and PAH1/2/3 and HID of Sin3 (Fig. 3A) (46, 136). Furthermore, two regions of Ikaros bound Sin3a and b independently (137), and N-Cor exploited two regions for interactions with Sin3b-PAH1 and -3, respectively (138). Jointly, these studies exemplify avidity in α -hub interactions. In harmonin, the N-terminal HHD is part of a supramodule, harboring a PDZ domain and a hairpin linker, which together form a platform for strong binding of sans (Fig. 5I) (48). Harmonin and Cad23 interact in a tripartite manner, with two different Cad23-tail regions engaging with harmonin-HHD. The three sites of Cad23 do not display synergistic effects in harmonin binding. Instead, these multivalent interactions facilitate the formation of polymeric Cad23:harmonin complexes (Fig. 3E), forming a stable anchorage structure at the tip link of stereocilia (49). Together, these examples demonstrate the diversity in ways α -hub:ligand interactions can be regulated.

Does context matter for α -hub interactions?

The context of α -hub interactions is multifacial, including the disordered SLiM context of the partners, the modular protein scaffold of the hub, as well as the environment (139). However, with the exception of the PAH domains (45, 140), responses to environmental changes have not been investigated (Table 2). The Sin3b-PAH domains are sensitive to pH, being more stable at lower pH (140), where Sin3-PAH2 displays conformational heterogeneity over the pH range 4.5 to 6.0 and the temperature range 15 to 35 °C (at pH 6.0) (45).

Arabidopsis RCD1-RST is less stable than the human Sin3b-PAH domains and has a lower melting temperature than whirlin-HHD (28). Together, this suggests that the α -hub domains can react to different external conditions using lower stability and higher flexibility in a way that may be subgroup and species dependent. This feature could provide a correlation between regulation and environment.

For disordered proteins, both SLiM flanking regions and the remainder of the protein context are emerging as important for affinity and specificity (12, 141). For interactions such as Sin3a-PAH2 with Mad1, HBP1, and Pfl, the α -helical SLiM was regarded as the prime PAH-domain contact (44–46). However, a conserved region flanking the Pfl-SLiM may regulate competitive binding of Sin3 and MRG15 to Pfl (46). Extension of the PAH2-binding SLiM of Mad1 with the charged region ${}_{21}\text{RRER}{}_{24}$ increased affinity ~ 3 -fold, suggestive of important electrostatic interactions mediated by the SLiM-flanking region (43). Furthermore, SAP30 matches the additional binding site in Sin3-PAH3 (see above) by using two α -helices in addition to the helix targeting the supersite in binding (66) (Fig. S1A). Thus, not only the SLiM but also flanking regions and contexts make important contributions to PAH binding.

For the RCD1-RST:TF interactions, SLiM context affected affinity as well as thermodynamic profiles. Removal of the ${}_{269}\text{PEPEPT}{}_{274}$ sequence from ANAC013(254–274) resulted in a 66-fold affinity decrease mainly from enthalpy loss, and a negative and positive allosteric effect on binding was detected for the SLiM contexts of ANAC013 (Fig. 5H) and DREB2a, respectively, when disordered regions outside the SLiM, which are not part of the binding site, were included in the binding experiments. For ANAC013, truncation reversed the entropic contribution to binding from negative to positive (57, 58). Thus, for the RST domain, context can be more deterministic than the SLiM itself. The isolated AD1(14–28) and AD2(38–61) regions of the p53-TRD bound CBP-NCBD approximately 180- and 8-fold weaker, respectively, than when combined in p53(1–61) (96). Similarly, the adenoviral E1A proteins, Adv5 and Adv12, contain three CBP-NCBD-binding sites, which are all required for competition with p53 in binding (106). Thus, binding is complex, likely involving allovalency (142) targeting the same site and exploiting increased local concentration effects (143). For the CBP-NCBD:ACTR association, a positive effect on binding affinity of the NCBD domain context was also demonstrated (144), highlighting the intricate regulatory potential of context, also involving the α -hubs. The additional binding site in CCM2-HHD (see above) is reflected in bound MEKK3, where the N-terminal α -helix targeting the CCM2-HHD supersite is followed by a PB1 domain, also contributing to binding (30, 124). For harmonin-HHD binding to the Cad23 Exon68 region, the affinity decreased when Exon68 was integrated in the Cad23 protein instead of an isolated peptide, because the context of Exon68 allows for a self-dimerization, incompatible with HHD binding (49). Finally, in RTEL1, which has tandem HHDs, a PCNA interacting protein motif resides in the linker context and separates the HHDs, whereby context expands the

organizational platform for the HHDs in RTEL1 (31). These examples all demonstrate that SLiM flanking regions and context play important roles also in α -hub functionality.

In the full modular proteins, α -hub context constitutes either two folded domains linked *via* disordered linkers or one folded domain and a disordered tail (Fig. 2). However, structural and functional studies of the α -hubs in their intact protein contexts are lacking. The different lengths and chemical properties of the linkers and tails may effect compaction and dynamics of relevance to the accessibility of the hub (145, 146). Indeed, some of the linkers are short, which support domain cooperation as observed for PAH domains in Sin3 proteins (108), and HHD and PDZ domains in harmonin (48). Furthermore, modifications of the context, such as by phosphorylations and ubiquitylations, may enable involvement of α -hubs in a variety of signaling processes (147). Thirteen phosphosites were recently identified for RCD1 (148), some located in the disordered N- and C-terminal contexts of the RST domain. How these affect ligand binding remains to be revealed, although not all phosphorylated residues may directly impact function.

Why α -hubs: open questions

The proliferation of α -helical folds in transcriptional regulation suggests this to be an optimal topology for signal integration and fidelity. In this review, we explored and compared a group of domains sharing a particular α -helical fold, the α -hub domains, to understand why, and to pose new important questions (Table 3). The dominating model for α -hub:ligand interactions involves paradigmatic coupled folding and binding (13, 78, 120) to form complexes of an α -helical ligand bound to an α -hub supersite (Fig. 4). However, exceptions are emerging that suggest further advantages of this particular fold. First, SLiM contexts affect binding affinity (58) (Fig. 5H) and competition (106), for reasons yet to be delineated. Contexts may remain disordered in the complexes (Fig. 5E), and complexes may involve structural heterogeneity (Fig. 5F) (58), as in the Med15:Gcn4 complex (149). Avidity, of importance to the function of strong acidic ADs (150, 151), and allosteric regulation by context (Fig. 5H) (58) are also mechanisms of relevance to α -hub:ligand interactions.

Second, in addition to the supersite, α -hubs have surfaces decorated with exposed conserved residues. For many of the hubs, it remains to be established if these sites are functionally exploited and if they can be used in the absence of binding to the supersite. Furthermore, it remains to be explored if expanded and/or additional binding surfaces can impose allosteric regulation of binding, as in the case of the KIX hub domain, for which binding of one TF allosterically enhances binding of another TF to a different binding site (152, 153). The use of supramodules in α -hub-bearing multidomain proteins (Fig. 5I) also points toward expandable functionality. If and how the adjoining linker contexts partake in binding and how posttranslational modifications of the hubs affect binding and selection have not yet been addressed.

Third, the H2-H3 hairpin platform combined with flexibility and dynamics of the remainder of the α -hub topology ensures malleability and versatility in binding, also allowing binding of folded ligands. Whether stability and rigidity of the different α -hub topologies (Table 2) correlate with specificity remains an open question. From an evolutionary point of view, higher protein–protein interaction specificity, correlating with a smaller interactome, may result in decreased network resilience, and hub domains showing malleability and promiscuity may represent an evolutionary advantage (154, 155). Comparison of ancestral and extant protein–protein interaction complexes of the plant protein SEP3 and its MAD TF ligands (156) showed that SEP3 has lost interaction partners, while increasing its structural stability (155). The fact that folded hub proteins keep flexibility as an important property is considered one of the main differences between hub and nonhub proteins (3, 157). Thus, it is possible that different degrees of flexibility, stability, and dynamics of the different α -hubs can explain different α -hub interactome sizes. With the uncovering of the many shared concepts valid across the hubs, and manifested by this review, it appears that the α -hubs cope with signal fidelity and specificity using several different strategies. The similarities and differences within the hubs highlighted here establish the α -hub domains as advantageous model systems for addressing general properties for maintaining signal fidelity in protein–protein interaction networks.

Acknowledgements—We thank Dr Kathrine Kemplen for some of the initial analyses of domain compositions in α -hubs.

Author contributions—K. B. and L. S. mainly contributed to the structural part, R. G. F. H. mainly contributed to the gene-related evolutionary part, E. S. mainly contributed to the stability part, K. S. and B. B. K. mainly contributed to the ligand interaction and interactome part and directed the study. All the authors worked closely on discussions. K. B., B. B. K., and K. S. wrote the article with input from all the authors. L. S. was mainly responsible for making the figures.

Funding and additional information—This work was supported by REPIN, *rethinking protein interactions*, the Novo Nordisk Foundation Challenge grant #NNF18OC0033926 to B. B. K. and K. S. Support was also achieved from the Novo Nordisk Foundation grant #NNF18OC0052177 to K. S.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AD, activation domain; CBP, CREB binding protein; CCM2, cerebral cavernous malformation 2; ETO, eight-twenty-one; GO, gene ontology; HDAC, histone deacetylase; HF, histone-fold; HHD, harmonin homology domain; HID, HDAC interacting domain; ID, intrinsic structural disorder; IDR, intrinsically disordered region; ITC, isothermal titration calorimetry; MD, molecular dynamics; NCBP, nuclear coactivator binding domain; NHR1, nervy homology region 1; PAH, paired amphipathic helix; PARP, poly(ADP-ribose)polymerase; RCD1, Radical-Induced Cell Death1; RST, RCD1, SRO, and TAF4; RTEL1, regulator of telomere elongation helicase 1; SLiM, short

linear motif; TAF4, transcription initiation factor TFIID-subunit 4; TAFH, TATA-box associated factor homology; TF, transcription factor; TRD, TF regulatory domain.

References

1. Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N., and Dunker, A. K. (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics* **9** Suppl 1, S1
2. Cumberworth, A., Lamour, G., Babu, M. M., and Gsponer, J. (2013) Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. *Biochem. J.* **454**, 361–369
3. Patil, A., Kinoshita, K., and Nakamura, H. (2010) Hub promiscuity in protein-protein interaction networks. *Int. J. Mol. Sci.* **11**, 1930–1943
4. Han, J. J., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J. M., Cusick, M. E., Roth, F. P., and Vidal, M. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**, 88–93
5. Hu, G., Wu, Z., Uversky, V. N., and Kurgan, L. (2017) Functional analysis of human hub proteins and their interactors involved in the intrinsic disorder-enriched interactions. *Int. J. Mol. Sci.* **18**, 2761
6. Moldovan, G.-L., Pfander, B., and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell* **129**, 665–679
7. Prestel, A., Wichmann, N., Martins, J. M., Marabini, R., Kassem, N., Broendum, S. S., Otterlei, M., Nielsen, O., Willemoës, M., Ploug, M., Boomsma, W., and Kragelund, B. B. (2019) The PCNA interaction motifs revisited: thinking outside the PIP-box. *Cell. Mol. Life Sci.* **76**, 4923–4943
8. Bustos, D. M. (2012) The role of protein disorder in the 14-3-3 interaction network. *Mol. Biosyst.* **8**, 178–184
9. Babu, M. M., van der Lee, R., de Groot, N. S., and Gsponer, J. (2011) Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **21**, 432–440
10. Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60
11. Huang, J.-R., Warner, L. R., Sanchez, C., Gabel, F., Madl, T., Mackereth, C. D., Sattler, M., and Blackledge, M. (2014) Transient electrostatic interactions dominate the conformational equilibrium sampled by multi-domain splicing factor U2AF65: a combined NMR and SAXS study. *J. Am. Chem. Soc.* **136**, 7068–7076
12. Dai, W., Wu, A., Ma, L., Li, Y.-X., Jiang, T., and Li, Y.-Y. (2016) A novel index of protein-protein interface propensity improves interface residue recognition. *BMC Syst. Biol.* **10**, 112
13. Wright, P. E., and Dyson, H. J. (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29
14. Neduva, V., Linding, R., Su-Angrand, I., Stark, A., de Masi, F., Gibson, T. J., Lewis, J., Serrano, L., and Russell, R. B. (2005) Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol.* **3**, e405
15. Jespersen, N., and Barbar, E. (2020) Emerging features of linear motif-binding hub proteins. *Trends Biochem. Sci.* **45**, 375–384
16. Tompa, P., Davey, N. E., Gibson, T. J., and Babu, M. M. (2014) A million peptide motifs for the molecular biologist. *Mol. Cell* **55**, 161–169
17. Plevin, M. J., Mills, M. M., and Ikura, M. (2005) The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem. Sci.* **30**, 66–69
18. Berlow, R. B., Dyson, H. J., and Wright, P. E. (2017) Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature* **543**, 447–451
19. Dyson, H. J., and Wright, P. E. (2016) Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *J. Biol. Chem.* **291**, 6714–6722
20. Novatchkova, M., and Eisenhaber, F. (2004) Linking transcriptional mediators via the GACKIX domain super family. *Curr. Biol.* **14**, R54–R55
21. Bugge, K., Staby, L., Kemplen, K. R., O’Shea, C., Bendsen, S. K., Jensen, M. K., Olsen, J. G., Skriver, K., and Kragelund, B. B. (2018) Structure of radical-induced cell Death1 hub domain reveals a common α -scaffold for disorder in transcriptional networks. *Structure* **26**, 734–746.e7
22. Poss, Z. C., Ebmeier, C. C., and Taatjes, D. J. (2013) The mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* **48**, 575–608
23. Allen, B. L., and Taatjes, D. J. (2015) The mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166
24. Staby, L., Bugge, K., Falbe-Hansen, R. G., Saladini, E., Skriver, K., and Kragelund, B. B. (2021) Connecting the α -hubs: same fold, disordered ligands, new functions. *Cell Commun. Signal.* **19**, 2
25. Spronk, C. A. E. M., Tessari, M., Kaan, A. M., Jansen, J. F. A., Vermeulen, M., Stunnenberg, H. G., and Vuister, G. W. (2000) The Mad1-Sin3B interaction involves a novel helical fold. *Nat. Struct. Biol.* **7**, 1100–1104
26. Wang, X., Truckses, D. M., Takada, S., Matsumura, T., Tanese, N., and Jacobson, R. H. (2007) Conserved region I of human coactivator TAF4 binds to a short hydrophobic motif present in transcriptional regulators. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 7839–7844
27. Lin, C. H., Hare, B. J., Wagner, G., Harrison, S. C., Maniatis, T., and Fraenkel, E. (2001) A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Mol. Cell* **8**, 581–590
28. Delhommel, F., Cordier, F., Saul, F., Chataigner, L., Haouz, A., and Wolff, N. (2018) Structural plasticity of the HHD2 domain of whirlin. *FEBS J.* **285**, 3738–3752
29. Bahloul, A., Pepermans, E., Raynal, B., Wolff, N., Cordier, F., England, P., Nouaille, S., Baron, B., El-Amraoui, A., Hardelin, J. P., Durand, D., and Petit, C. (2017) Conformational switch of harmonin, a submembrane scaffold protein of the hair cell mechano-electrical transduction machinery. *FEBS Lett.* **591**, 2299–2310
30. Fisher, O. S., Zhang, R., Li, X., Murphy, J. W., Demeler, B., and Boggon, T. J. (2013) Structural studies of cerebral cavernous malformations 2 (CCM2) reveal a folded helical domain at its C-terminus. *FEBS Lett.* **587**, 272–277
31. Faure, G., Revy, P., Schertzer, M., Londono-Vallejo, A., and Callebaut, I. (2014) The C-terminal extension of human RTEL1, mutated in Hoyerall-Hreidarsson syndrome, contains harmonin-N-like domains. *Proteins* **82**, 897–903
32. Efimov, A. V. (1991) Structure of α - α -hairpins with short connections. *Protein Eng. Des. Sel.* **4**, 245–250
33. Engel, D. E., and DeGrado, W. F. (2005) α - α linking motifs and inter-helical orientations. *Proteins* **61**, 325–337
34. Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., Brown, S. D., Chang, H. Y., El-Gebali, S., Fraser, M. I., Gough, J., Haft, D. R., Huang, H., Letunic, I., Lopez, R., et al. (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360
35. Bowen, A. J., Gonzalez, D., Mullins, J. G. L., Bhatt, A. M., Martinez, A., and Conlan, R. S. (2010) PAH-domain-specific interactions of the arabidopsis transcription coregulator SIN3-LIKE1 (SNL1) with telomere-binding protein 1 and ALWAYS EARLY2 Myb-DNA binding factors. *J. Mol. Biol.* **395**, 937–949
36. Christensen, L. F., Staby, L., Bugge, K., O’Shea, C., Kragelund, B. B., and Skriver, K. (2019) Evolutionary conservation of the intrinsic disorder-based radical-induced cell Death1 hub interactome. *Sci. Rep.* **9**, 18927
37. Kazantseva, J., and Palm, K. (2014) Diversity in TAF proteomics: consequences for cellular differentiation and migration. *Int. J. Mol. Sci.* **15**, 16680–16697
38. Hultqvist, G., Åberg, E., Camilloni, C., Sundell, G. N., Andersson, E., Dogan, J., Chi, C. N., Vendruscolo, M., and Jemth, P. (2017) Emergence and evolution of an interaction between intrinsically disordered proteins. *Elife* **6**, 1–25
39. Henricson, A., Forslund, K., and Sonnhammer, E. L. L. (2010) Orthology confers intron position conservation. *BMC Genomics* **11**, 412
40. O’Leary, N. A., Wright, M. W., Brister, J. R., Ciuffo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badreddin, A., Bao, Y., Blinkova, O., Brover, V., et al. (2016) Reference sequence (RefSeq) database at NCBI: current status,

- taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745
41. Russell, R. B., Saqi, M. A. S., Sayle, R. A., Bates, P. A., and Sternberg, M. J. E. (1997) Recognition of analogous and homologous protein folds: analysis of sequence and structure conservation. *J. Mol. Biol.* **269**, 423–439
 42. Krishna, S. S., and Grishin, N. V. (2004) Structurally analogous proteins do exist! *Structure* **12**, 1125–1127
 43. Van Ingen, H., Lasonder, E., Jansen, J. F. A., Kaan, A. M., Spronk, C. A. E. M., Stunnenberg, H. G., and Vuister, G. W. (2004) Extension of the binding motif of the Sin3 interacting domain of the mad family proteins. *Biochemistry* **43**, 46–54
 44. Swanson, K. A., Knoepfler, P. S., Huang, K., Kang, R. S., Cowley, S. M., Laherty, C. D., Eisenman, R. N., and Radhakrishnan, I. (2004) HBP1 and Mad1 repressors bind the Sin3 corepressor PAH2 domain with opposite helical orientations. *Nat. Struct. Mol. Biol.* **11**, 738–746
 45. Brubaker, K., Cowley, S. M., Huang, K., Loo, L., Yochum, G. S., Ayer, D. E., Eisenman, R. N., and Radhakrishnan, I. (2000) Solution structure of the interacting domains of the mad-sin3 complex: implications for recruitment of a chromatin-modifying complex. *Cell* **103**, 655–665
 46. Kumar, G. S., Xie, T., Zhang, Y., and Radhakrishnan, I. (2011) Solution structure of the mSin3A PAH2-Pfl SID1 complex: a Mad1/Mxd1-like interaction disrupted by MRG15 in the Rpd3S/Sin3S complex. *J. Mol. Biol.* **408**, 987–1000
 47. van Ingen, H., Baltussen, M. A. H., Aelen, J., and Vuister, G. W. (2006) Role of structural and dynamical plasticity in Sin3: the free PAH2 domain is a folded module in mSin3B. *J. Mol. Biol.* **358**, 485–497
 48. Yan, J., Pan, L., Chen, X., Wu, L., and Zhang, M. (2010) The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4040–4045
 49. Wu, L., Pan, L., Zhang, C., and Zhang, M. (2012) Large protein assemblies formed by multivalent interactions between cadherin23 and harmonin suggest a stable anchorage structure at the tip link of stereocilia. *J. Biol. Chem.* **287**, 33460–33471
 50. Kjaergaard, M., Andersen, L., Nielsen, L. D., and Teilum, K. (2013) A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition. *Biochemistry* **52**, 1686–1693
 51. Papaleo, E., Camilloni, C., Teilum, K., Vendruscolo, M., and Lindorff-Larsen, K. (2018) Molecular dynamics ensemble refinement of the heterogeneous native state of NCBD using chemical shifts and NOEs. *PeerJ* **6**, e5125
 52. Russell, R. B., Sasieni, P. D., and Sternberg, M. J. E. (1998) Supersites within superfolds. Binding site similarity in the absence of homology. *J. Mol. Biol.* **282**, 903–918
 53. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268
 54. Adams, G. E., Chandru, A., and Cowley, S. M. (2018) Co-repressor, co-activator and general transcription factor: the many faces of the Sin3 histone deacetylase (HDAC) complex. *Biochem. J.* **475**, 3921–3932
 55. Warmerdam, S., Sterken, M. G., Sukarta, O. C. A., van Schaik, C. C., Oortwijn, M. E. P., Lozano-Torres, J. L., Bakker, J., Smant, G., and Goverse, A. (2020) The TIR-NB-LRR pair DSC1 and WRKY19 contributes to basal immunity of Arabidopsis to the root-knot nematode *Meloidogyne incognita*. *BMC Plant Biol.* **20**, 73
 56. Jaspers, P., Blomster, T., Brosché, M., Salojärvi, J., Ahlfors, R., Vainonen, J. P., Reddy, R. A., Immink, R., Angenent, G., Turck, F., Overmyer, K., and Kangasjärvi, J. (2009) Unequally redundant RCD1 and SRO1 mediate stress and developmental responses and interact with transcription factors. *Plant J.* **60**, 268–279
 57. O'Shea, C., Kryger, M., Stender, E. G. P., Kragelund, B. B., Willemoës, M., and Skriver, K. (2015) Protein intrinsic disorder in Arabidopsis NAC transcription factors: transcriptional activation by ANAC013 and ANAC046 and their interactions with RCD1. *Biochem. J.* **465**, 281–294
 58. O'Shea, C., Staby, L., Bendsen, S. K., Tidemand, F. G., Redsted, A., Willemoës, M., Kragelund, B. B., and Skriver, K. (2017) 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.* **292**, 512–527
 59. Wright, K. J., Marr, M. T., and Tjian, R. (2006) TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12347–12352
 60. Lutterbach, B., Sun, D., Schuetz, J., and Hiebert, S. W. (1998) The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein. *Mol. Cell Biol.* **18**, 3604–3611
 61. Pan, L., Yan, J., Wu, L., and Zhang, M. (2009) Assembling stable hair cell tip link complex via multidentate interactions between harmonin and cadherin 23. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5575–5580
 62. Chaubal, A., and Pile, L. A. (2018) Same agent, different messages: insight into transcriptional regulation by SIN3 isoforms. *Epigenetics Chromatin* **11**, 17
 63. Kadamb, R., Mittal, S., Bansal, N., Batra, H., and Saluja, D. (2013) Sin3: insight into its transcription regulatory functions. *Eur. J. Cell Biol.* **92**, 237–246
 64. Laherty, C. D., Yang, W. M., Jian-Min, S., Davie, J. R., Seto, E., and Eisenman, R. N. (1997) Histone deacetylases associated with the mSin3 corepressor mediate Mad transcriptional repression. *Cell* **89**, 349–356
 65. Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (1997) Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell* **89**, 357–364
 66. Xie, T., He, Y., Korkeamaki, H., Zhang, Y., Imhoff, R., Lohi, O., and Radhakrishnan, I. (2011) 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.* **286**, 27814–27824
 67. Clark, M. D., Marcum, R., Graveline, R., Chan, C. W., Xie, T., Chen, Z., Ding, Y., Zhang, Y., Mondragón, A., David, G., and Radhakrishnan, I. (2015) Structural insights into the assembly of the histone deacetylase-associated Sin3L/Rpd3L corepressor complex. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3669–E3678
 68. Streubel, G., Fitzpatrick, D. J., Oliviero, G., Scelfo, A., Moran, B., Das, S., Munawar, N., Watson, A., Wynne, K., Negri, G. L., Dillon, E. T., Jammula, S., Hokamp, K., O'Connor, D. P., Pasini, D., et al. (2017) Fam60a defines a variant Sin3a-Hdac complex in embryonic stem cells required for self-renewal. *EMBO J.* **36**, 2216–2232
 69. van Oevelen, C., Wang, J., Asp, P., Yan, Q., Kaelin, W. G., Kluger, Y., and Dynlacht, B. D. (2008) A role for mammalian Sin3 in permanent gene silencing. *Mol. Cell* **32**, 359–370
 70. Williams, K., Christensen, J., Pedersen, M. T., Johansen, J. V., Cloos, P. A. C., Rappsilber, J., and Helin, K. (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–349
 71. Nomura, M., Uda-Tochio, H., Murai, K., Mori, N., and Nishimura, Y. (2005) The neural repressor NRSF/REST binds the PAH1 domain of the Sin3 corepressor by using its distinct short hydrophobic helix. *J. Mol. Biol.* **354**, 903–915
 72. David, G., Alland, L., Hong, S. H., Wong, C. W., DePinho, R. A., and Dejean, A. (1998) Histone deacetylase associated with mSin3A mediates repression by the acute promyelocytic leukemia-associated PLZF protein. *Oncogene* **16**, 2549–2556
 73. Hurlin, P. J., Quéva, C., and Eisenman, R. N. (1997) Mnt, a novel Max-interacting protein is coexpressed with Myc in proliferating cells and mediates repression at Myc binding sites. *Genes Dev.* **11**, 44–58
 74. Zhang, J.-S., Moncrieffe, M. C., Kaczynski, J., Ellenrieder, V., Prendergast, F. G., and Urrutia, R. (2001) A conserved α -helical motif mediates the interaction of sp1-like transcriptional repressors with the corepressor mSin3A. *Mol. Cell Biol.* **21**, 5041–5049
 75. Yang, Q., Kong, Y., Rothermel, B., Garry, D. J., Bassel-Duby, R., and Williams, R. S. (2000) The winged-helix/forkhead protein myocyte nuclear factor beta (MNF-beta) forms a co-repressor complex with mammalian sin3B. *Biochem. J.* **345 Pt 2**, 335–343

76. Marcum, R. D., and Radhakrishnan, I. (2020) The neuronal transcription factor Myt1L interacts via a conserved motif with the PAH1 domain of Sin3 to recruit the Sin3L/Rpd3L histone deacetylase complex. *FEBS Lett.* **594**, 2322–2330
77. Le Guezennec, X., Vriend, G., and Stunnenberg, H. G. (2004) Molecular determinants of the interaction of Mad with the PAH2 domain of mSin3. *J. Biol. Chem.* **279**, 25823–25829
78. Brubaker, K., Cowley, S. M., Huang, K., Loo, L., Yochum, G. S., Ayer, D. E., Eisenman, R. N., and Radhakrishnan, I. (2000) Solution structure of the interacting domains of the mad–sin3 complex. *Cell* **103**, 655–665
79. Le Guezennec, X., Vermeulen, M., and Stunnenberg, H. G. (2006) Molecular characterization of Sin3 PAH-domain interactor specificity and identification of PAH partners. *Nucleic Acids Res.* **34**, 3929–3937
80. Chandru, A., Bate, N., Vuister, G. W., and Cowley, S. M. (2018) Sin3A recruits Tet1 to the PAH1 domain via a highly conserved Sin3-interaction domain. *Sci. Rep.* **8**, 1–10
81. Kragelund, B. B., Jensen, M. K., and Skriver, K. (2012) Order by disorder in plant signaling. *Trends Plant Sci.* **17**, 625–632
82. Overmyer, K., Tuominen, H., Kettunen, R., Betz, C., Langebartels, C., Sandermann, H.J., and Kangasjärvi, J. (2000) Ozone-sensitive Arabidopsis rcd1 mutant reveals opposite roles for ethylene and jasmonate signaling pathways in regulating superoxide-dependent cell death. *Plant Cell* **12**, 1849–1862
83. Ahlfors, R., Lång, S., Overmyer, K., Jaspers, P., Brosché, M., Tauriainen, A., Kollist, H., Tuominen, H., Belles-Boix, E., Piippo, M., Inzé, D., Palva, E. T., and Kangasjärvi, J. (2004) Arabidopsis radical-induced cell Death1 belongs to the WWE protein–protein interaction domain protein family and modulates abscisic acid, ethylene, and methyl jasmonate responses. *Plant Cell* **16**, 1925–1937
84. Wirthmueller, L., Asai, S., Rallapalli, G., Sklenar, J., Fabro, G., Kim, D. S., Lintermann, R., Jaspers, P., Wrzaczek, M., Kangasjärvi, J., MacLean, D., Menke, F. L. H., Banfield, M. J., and Jones, J. D. G. (2018) Arabidopsis downy mildew effector HaRxL106 suppresses plant immunity by binding to radical-induced cell Death1. *New Phytol.* **220**, 232–248
85. Vainonen, J. P., Jaspers, P., Wrzaczek, M., Lamminmäki, A., Reddy, R. a, Vaahtera, L., Brosché, M., and Kangasjärvi, J. (2012) RCD1-DREB2A interaction in leaf senescence and stress responses in Arabidopsis thaliana. *Biochem. J.* **442**, 573–581
86. Shapiguzov, A., Vainonen, J. P., Hunter, K., Tossavainen, H., Tiwari, A., Järvi, S., Hellman, M., Aarabi, F., Alseekh, S., Wybouw, B., Van Der Kelen, K., Nikkanen, L., Krasensky-Wrzaczek, J., Sipari, N., Keinänen, M., *et al.* (2019) Arabidopsis RCD1 coordinates chloroplast and mitochondrial functions through interaction with ANAC transcription factors. *Elife* **8**, e43284
87. De Clercq, I., Vermeirssen, V., Van Aken, O., Vandepoele, K., Murcha, M. W., Law, S. R., Inzé, A., Ng, S., Ivanova, A., Rombaut, D., van de Cotte, B., Jaspers, P., Van de Peer, Y., Kangasjärvi, J., Whelan, J., *et al.* (2013) The membrane-bound NAC transcription factor ANAC013 functions in mitochondrial retrograde regulation of the oxidative stress response in Arabidopsis. *Plant Cell* **25**, 3472–3490
88. Papai, G., Weil, P. A., and Schultz, P. (2011) New insights into the function of transcription factor TFIID from recent structural studies. *Curr. Opin. Genet. Dev.* **21**, 219–224
89. Thuault, S., Gangloff, Y. G., Kirchner, J., Sanders, S., Werten, S., Romier, C., Weil, P. A., and Davidson, I. (2002) Functional analysis of the TFIID-specific yeast TAF4 (yTAFII48) reveals an unexpected organization of its histone-fold domain. *J. Biol. Chem.* **277**, 45510–45517
90. Gazit, K., Moshonov, S., Elfakess, R., Sharon, M., Mengus, G., Davidson, I., and Dikstein, R. (2009) TAF4/4b-TAF12 displays a unique mode of DNA binding and is required for core promoter function of a subset of genes. *J. Biol. Chem.* **284**, 26286–26296
91. Chen, W. Y., Zhang, J., Geng, H., Du, Z., Nakadai, T., and Roeder, R. G. (2013) A TAF4 coactivator function for E proteins that involves enhanced TFIID binding. *Genes Dev.* **27**, 1596–1609
92. Quong, M. W., Romanow, W. J., and Murre, C. (2002) E protein function in lymphocyte development. *Annu. Rev. Immunol.* **20**, 301–322
93. Janknecht, R., and Hunter, T. (1996) A growing coactivator network. *Nature* **383**, 22–23
94. Demarest, S. J. (2004) Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci.* **13**, 203–210
95. Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S. C., Heyman, R. A., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1996) A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403–414
96. Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2010) Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* **49**, 9964–9971
97. Janknecht, R., Wells, N. J., and Hunter, T. (1998) TGF- β -stimulated cooperation of Smad proteins with the coactivators CBP/p300. *Genes Dev.* **12**, 2114–2119
98. Gingras, S., Simard, J., Groner, B., and Pfltzner, E. (1999) p300/CBP is required for transcriptional induction by interleukin-4 and interacts with Stat6. *Nucleic Acids Res.* **27**, 2722–2729
99. Saihan, Z., Webster, A. R., Luxon, L., and Bitner-Glindzicz, M. (2009) Update on Usher syndrome. *Curr. Opin. Neurol.* **22**, 19–27
100. Sorusch, N., Wunderlich, K., Baus, K., Nagel-Wolfrum, K., and Wolfrum, U. (2014) Usher syndrome protein network functions in the retina and their relation to other retinal ciliopathies. *Adv. Exp. Med. Biol.* **801**, 527–533
101. Siemens, J., Kazmierczak, P., Reynolds, A., Sticker, M., Littlewood-Evans, A., and Müller, U. (2002) The Usher syndrome proteins cadherin 23 and harmonin form a complex by means of PDZ-domain interactions. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14946–14951
102. Boëda, B., El-Amraoui, A., Bahloul, A., Goodyear, R., Daviet, L., Blanchard, S., Perfettini, I., Fath, K. R., Shorte, S., Reiners, J., Houdusse, A., Legrain, P., Wolfrum, U., Richardson, G., and Petit, C. (2002) Myosin VIIa, harmonin and cadherin 23, three Usher I gene products that cooperate to shape the sensory hair cell bundle. *EMBO J.* **21**, 6689–6699
103. Sahu, S. C., Swanson, K. A., Kang, R. S., Huang, K., Brubaker, K., Ratcliff, K., and Radhakrishnan, I. (2008) Conserved themes in target recognition by the PAH1 and PAH2 domains of the Sin3 transcriptional corepressor. *J. Mol. Biol.* **375**, 1444–1456
104. Kjaersgaard, T., Jensen, M. K., Christiansen, M. W., Gregersen, P., Kragelund, B. B., and Skriver, K. (2011) Senescence-associated barley NAC (NAM, ATAF1,2, CUC) transcription factor interacts with radical-induced cell death 1 through a disordered regulatory domain. *J. Biol. Chem.* **286**, 35418–35429
105. Staby, L., O’Shea, C., Willemoës, M., Theisen, F., Kragelund, B. B., and Skriver, K. (2017) Eukaryotic transcription factors: paradigms of protein intrinsic disorder. *Biochem. J.* **474**, 2509–2532
106. Haberz, P., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2016) Mapping the interactions of adenoviral E1A proteins with the p160 nuclear receptor coactivator binding domain of CBP. *Protein Sci.* **25**, 2256–2267
107. Higo, J., Nishimura, Y., and Nakamura, H. (2011) 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.* **133**, 10448–10458
108. Cowley, S. M., Kang, R. S., Frangioni, J. V., Yada, J. J., DeGrand, A. M., Radhakrishnan, I., and Eisenman, R. N. (2004) Functional analysis of the Mad1-mSin3A repressor-corepressor interaction reveals determinants of specificity, affinity, and transcriptional response. *Mol. Cell. Biol.* **24**, 2698–2709
109. Park, S., Chen, W., Cierpicki, T., Tonelli, M., Cai, X., Speck, N. A., and Bushweller, J. H. (2009) Structure of the AML1-ETO eTAFH domain-HEB peptide complex and its contribution to AML1-ETO activity. *Blood* **113**, 3558–3567
110. Zhang, J., Kalkum, M., Yamamura, S., Chait, B. T., and Roeder, R. G. (2004) E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* **305**, 1286–1289

111. Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M., and Wright, P. E. (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553
112. Dogan, J., Jonasson, J., Andersson, E., and Jemth, P. (2015) Binding rate constants reveal distinct features of disordered protein domains. *Biochemistry* **54**, 4741–4750
113. Flock, T., Weatheritt, R. J., Latysheva, N. S., and Babu, M. M. (2014) Controlling entropy to tune the functions of intrinsically disordered regions. *Curr. Opin. Struct. Biol.* **26**, 62–72
114. Teilum, K., Olsen, J. G., and Kragelund, B. B. (2015) Globular and disordered—the non-identical twins in protein-protein interactions. *Front. Mol. Biosci.* **2**, 1–6
115. Borgia, A., Borgia, M. B., Bugge, K., Kissling, V. M., Heidarsson, P. O., Fernandes, C. B., Sottini, A., Soranno, A., Buholzer, K. J., Nettels, D., Kragelund, B. B., Best, R. B., and Schuler, B. (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66
116. Heller, G. T., Sormanni, P., and Vendruscolo, M. (2015) Targeting disordered proteins with small molecules using entropy. *Trends Biochem. Sci.* **40**, 491–496
117. Das, R. K., Huang, Y., Phillips, A. H., Kriwacki, R. W., and Pappu, R. V. (2016) Cryptic sequence features within the disordered protein p27^{Kip1} regulate cell cycle signaling. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5616–5621
118. Zhang, Y., Zhang, Z., Demeler, B., and Radhakrishnan, I. (2006) Coupled unfolding and dimerization by the PAH2 domain of the mammalian Sin3A corepressor. *J. Mol. Biol.* **360**, 7–14
119. He, Y., and Radhakrishnan, I. (2007) Solution NMR studies of apo-mSin3A and -mSin3B reveal that the PAH1 and PAH2 domains are structurally independent. *Protein Sci.* **17**, 171–175
120. Shammas, S. L., Crabtree, M. D., Dahal, L., Wicky, B. I. M., and Clarke, J. (2016) Insights into coupled folding and binding mechanisms from kinetic studies. *J. Biol. Chem.* **291**, 6689–6695
121. Waters, L., Yue, B., Veverka, V., Renshaw, P., Bramham, J., Matsuda, S., Frenkiel, T., Kelly, G., Muskett, F., Carr, M., and Heery, D. M. (2006) Structural diversity in p160/CREB-binding protein coactivator complexes. *J. Biol. Chem.* **281**, 14787–14795
122. Nady, N., Gupta, A., Ma, Z., Swigut, T., Koide, A., Koide, S., and Wysocka, J. (2015) ETO family protein Mtgr1 mediates Prdm14 functions in stem cell maintenance and primordial germ cell formation. *Elife* **4**, 1–28
123. Wang, X., Hou, Y., Deng, K., Zhang, Y., Wang, D. C., and Ding, J. (2015) Structural insights into the molecular recognition between cerebral cavernous malformation 2 and mitogen-activated protein kinase kinase 3. *Structure* **23**, 1087–1096
124. Fisher, O. S., Deng, H., Liu, D., Zhang, Y., Wei, R., Deng, Y., Zhang, F., Louvi, A., Turk, B. E., Boggon, T. J., and Su, B. (2015) Structure and vascular function of MEKK3-cerebral cavernous malformations 2 complex. *Nat. Commun.* **6**, 7937
125. Lee, C. W., Ferreon, J. C., Ferreon, A. C. M., Arai, M., and Wright, P. E. (2010) Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19290–19295
126. Kjaergaard, M., Teilum, K., and Poulsen, F. M. (2010) Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12535–12540
127. Villalobo, A., González-Muñoz, M., and Berchtold, M. W. (2019) Proteins with calmodulin-like domains: structures and functional roles. *Cell. Mol. Life Sci.* **76**, 2299–2328
128. Berlow, R. B., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2019) Role of backbone dynamics in modulating the interactions of disordered ligands with the TAZ1 domain of the CREB-binding protein. *Biochemistry* **58**, 1354–1362
129. Jemth, P., Karlsson, E., Vögeli, B., Guzovsky, B., Andersson, E., Hultqvist, G., Dogan, J., Güntert, P., Riek, R., and Chi, C. N. (2018) Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins. *Sci. Adv.* **4**, eaau4130
130. Li, J., White, J. T., Saavedra, H., Wrabl, J. O., Motlagh, H. N., Liu, K., Sowers, J., Schroer, T. A., Thompson, E. B., and Hilser, V. J. (2018) Erratum: correction: genetically tunable frustration controls allostery in an intrinsically disordered transcription factor (eLife (2017) 6 PII: e35768). *Elife* **7**, 1–17
131. Karlsson, E., Andersson, E., Dogan, J., Gianni, S., Jemth, P., and Camilloni, C. (2019) A structurally heterogeneous transition state underlies coupled binding and folding of disordered proteins. *J. Biol. Chem.* **294**, 1230–1239
132. Iešmantavičius, V., Dogan, J., Jemth, P., Teilum, K., and Kjaergaard, M. (2014) Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew. Chem. Int. Ed.* **53**, 1548–1551
133. Liu, X., Chen, J., and Chen, J. (2019) Residual structure accelerates binding of intrinsically disordered ACTR by promoting efficient folding upon encounter. *J. Mol. Biol.* **431**, 422–432
134. Yu, Q., Ye, W., Wang, W., and Chen, H. F. (2013) Global conformational selection and local induced fit for the recognition between intrinsic disordered p53 and CBP. *PLoS One* **8**, e59627
135. Zosel, F., Mercadante, D., Nettels, D., and Schuler, B. (2018) A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction. *Nat. Commun.* **9**, 3332
136. Yochum, G. S., and Ayer, D. E. (2001) Pfl1, a novel PHD zinc finger protein that links the TLE corepressor to the mSin3A-histone deacetylase complex. *Mol. Cell. Biol.* **21**, 4110–4118
137. Koipally, J., and Georgopoulos, K. (2002) A molecular dissection of the repression circuitry of Ikaros. *J. Biol. Chem.* **277**, 27697–27705
138. Heinzel, T., Lavinsky, R. M., Mullen, T. M., Söderstrom, M., Laherty, C. D., Torchia, J., Yang, W. M., Brard, G., Ngo, S. D., Davie, J. R., Seto, E., Eisenman, R. N., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1997) A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. *Nature* **387**, 43–48
139. Bugge, K., Brakti, I., Fernandes, C. B., Dreier, J. E., Lundsgaard, J. E., Olsen, J. G., Skriver, K., and Kragelund, B. B. (2020) Interactions by disorder – a matter of context. *Front. Mol. Biosci.* **7**, 110
140. Hasan, T., Ali, M., Saluja, D., and Singh, L. R. (2015) pH Might play a role in regulating the function of paired amphipathic helices domains of human Sin3B by altering structure and thermodynamic stability. *Biochemistry (Mosc)* **80**, 424–432
141. Henley, M. J., Linhares, B. M., Morgan, B. S., Cierpicki, T., Fierke, C. A., and Mapp, A. K. (2020) Unexpected specificity within dynamic transcriptional protein-protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 27346–27353
142. Olsen, J. G., Teilum, K., and Kragelund, B. B. (2017) Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness. *Cell. Mol. Life Sci.* **74**, 3175–3183
143. Sørensen, C. S., and Kjaergaard, M. (2019) Effective concentrations enforced by intrinsically disordered linkers are governed by polymer physics. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 23124–23131
144. Karlsson, E., Lindberg, A., Andersson, E., and Jemth, P. (2020) High affinity between CREBBP/p300 and NCOA evolved in vertebrates. *Protein Sci.* **29**, 1687–1691
145. Lyle, N., Das, R. K., and Pappu, R. V. (2013) A quantitative measure for protein conformational heterogeneity. *J. Chem. Phys.* **139**, 121907
146. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G., and Pappu, R. V. (2017) CIDR: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112**, 16–21
147. Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter, S., Shikata, H., Messerer, M., Lang, D., Altmann, S., Cyprys, P., Zolg, D. P., Mathieson, T., et al. (2020) Mass-spectrometry-based draft of the Arabidopsis proteome. *Nature* **579**, 409–414
148. Vainonen, J., Shapiguzov, A., Krasensky-Wrzaczek, J., De Masi, R., Gossens, R., Danciu, I., Battchikova, N., Jonak, C., Wirthmueller, L., Wrzaczek, M., and Kangasjärvi, J. (2020) Arabidopsis poly(ADP-ribose)-binding protein RCD1 interacts with photoregulatory protein kinases in nuclear bodies. *bioRxiv*. <https://doi.org/10.1101/2020.07.02.184937>
149. Warfield, L., Tuttle, L. M., Pacheco, D., Klevit, R. E., and Hahn, S. (2014) A sequence-specific transcription activator motif and powerful synthetic

- variants that bind mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E3506–E3513
150. Brodsky, S., Jana, T., Mittelman, K., Chapal, M., Kumar, D. K., Carmi, M., and Barkai, N. (2020) Intrinsically disordered regions direct transcription factor in vivo binding specificity. *Mol. Cell* **79**, 459–471.e4
151. Erijman, A., Kozłowski, L., Sohrabi-Jahromi, S., Fishburn, J., Warfield, L., Schreiber, J., Noble, W. S., Söding, J., and Hahn, S. (2020) A high-throughput screen for transcription activation domains reveals their sequence features and permits prediction by deep learning. *Mol. Cell* **78**, 890–902.e6
152. Yang, K., Stanfield, R. L., Martinez-Yamout, M. A., Dyson, H. J., Wilson, I. A., and Wright, P. E. (2018) Structural basis for cooperative regulation of KIX-mediated transcription pathways by the HTLV-1 HBZ activation domain. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 10040–10045
153. De Guzman, R. N., Goto, N. K., Dyson, H. J., and Wright, P. E. (2006) Structural basis for cooperative transcription factor binding to the CBP coactivator. *J. Mol. Biol.* **355**, 1005–1013
154. Ferreira, R. M., Rybarczyk-Filho, J. L., Dalmolin, R. J. S., Castro, M. A. A., Moreira, J. C. F., Brunnet, L. G., and de Almeida, R. M. C. (2013) Preferential duplication of intermodular hub genes: an evolutionary signature in eukaryotes genome networks. *PLoS One* **8**, e56579
155. Alhindi, T., Zhang, Z., Ruelens, P., Coenen, H., Degroote, H., Iraci, N., and Geuten, K. (2017) Protein interaction evolution from promiscuity to specificity with reduced flexibility in an increasingly complex network. *Sci. Rep.* **7**, 44948
156. de Folter, S., Immink, R. G. H., Kieffer, M., Pařenicová, L., Henz, S. R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M. M., Davies, B., and Angenent, G. C. (2005) Comprehensive interaction map of the arabidopsis MADS box transcription factors. *Plant Cell* **17**, 1424–1433
157. Patil, A., and Nakamura, H. (2006) Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* **580**, 2041–2045
158. Plevin, M. J., Zhang, J., Guo, C., Roeder, R. G., and Ikura, M. (2006) The acute myeloid leukemia fusion protein AML1-ETO targets E proteins via a paired amphipathic helix-like TBP-associated factor homology domain. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10242–10247
159. Burki, F., Roger, A. J., Brown, M. W., and Simpson, A. G. B. (2020) The new tree of eukaryotes. *Trends Ecol. Evol.* **35**, 43–55