Intrinsically Disordered Proteins 2:1, e970499; October 1, 2014; © 2014 Published with license by Taylor & Francis Group, LLC

Unreported intrinsic disorder in proteins: Building connections to the literature on IDPs

Vladimir N Uversky[1,2,3,*]

[1]Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute; Morsani College of Medicine; University of South Florida; Tampa, FL USA; [2]Institute for Biological Instrumentation; Russian Academy of Sciences; Pushchino, Russia; [3]Biology Department; Faculty of Science; King Abdulaziz University; Jeddah, Kingdom of Saudi Arabia

Keywords: intrinsically disordered protein, protein function, protein-protein interaction, molecular recognition, post-translational modification

Abbreviations: AHR, aryl hydrocarbon receptor; AiBS,ANCHOR-identified binding site; AIP, AHR interacting protein; AIPL1, AIP like 1; AMSH, associated molecule with a Src homology 3 domain of]; APC, activated protein C; ApoA-I, apolipoprotein A-I; ARVD, arrhythmogenic right ventricular dysplasia type 2; CBP, CREB binding protein; CDK, cyclin-dependent kinase; Cdk2, cyclin-dependent kinase 2; Ci-VSP, Ciona intestinalis voltage sensitive phosphatase; CREB, cAMP response element binding protein; CR, C-terminal arginine; Csp, cold shock protein; CTD, C-terminal domain; CTR, C-terminal regulatory segment; DENV, dengue virus; DHPR, dihydropyridine receptor; dN, deoxynucleoside; DNAJA1, DnaJ homologue subfamily A member 1; dNTP, deoxynucleoside triphosphate; DUB, deubiquitinating enzyme; EBOV, Ebola virus; eIF4A, eukaryotic translation initiation factor 4A; eIF4B, eukaryotic translation initiation factor 4B; eIF4G, eukaryotic translation initiation factor 4G; ELM, eukaryotic linear motif; ESCRT, endosomal sorting complexes required for transport machinery; Frmpd1, FERM and PDZ domain containing 1; FVIII, factor VIII; Gla-OCN, γ-carboxyglutamic acid; GLTSCR2, glioma tumor suppressor candidate region 2 gene product; GPCR, G protein coupled receptor; GPI, glycosylphosphatidylinositol; GRK, G-protein dependent kinase; HCMV, human cytomegalovirus; HCV, hepatitis C virus; HD, histidine-aspartic domain; HDL, high density lipoprotein; HD-MS/MS, hydrogen/deuterium exchange tandem mass spectrometry; HDX, hydrogen/deuterium exchange; HIV-1, human immunodeficiency virus type 1; HRSV, human respiratory syncytial virus; HSF, heat shock transcription factor; Hsp, heat shock protein; IAV, influenza A virus; IDP, intrinsically disordered protein; IDPR, intrinsically disordered protein region; IFITM, interferon-inducible transmembrane protein; ILT7, immunoglobulin-like transcript 7; ISG, interferon-stimulated gene; ITAM, immunoreceptor tyrosine-based activation motif; ITIM, immunoreceptor tyrosine-based inhibitory motif; MARV, Marburg virus; MAVS, mitochondrial antiviral signaling protein; MIC-CAP, microcephaly capillary malformation; MLCP, myosin light chain phosphatase; Mpa, mycobacterium proteasome ATPase; MSC, "molecular sandwich" complex; NES, nuclear export signal; NF-Y, nuclear transcription factor Y; Nrp, neuropilins; NRTK, nonreceptor tyrosine kinase; NTD, N-terminal domain; PAI-1, plasminogen activator inhibitor 1; PafA, proteasome accessory factor A; PAS, primosome assembly site; PECAM-1, platelet endothelial cell adhesion molecule-1; Pcl, Pho85 cyclin; PD, phosphatase domain; PICT-1, protein interacting with carboxyl terminus-1; PIP, phosphatidylinositol phosphate; PKA, protein kinase A; PP1, protein phosphatase-1; PPIase, prolyl peptidyl isomerase; PRD, proline-rich domain; PSF, protein-associated splicing factor; PTEN, phosphatase and tensin homologue; Pup, prokaryotic ubiquitin-like protein; PUR, purine-rich; Purb, purine-rich element binding protein B; Rb, retinoblastoma protein; RCL, reactive center loop; RING, Really Interesting New Gene; RyR, Ryanodine receptor; SAM, sterile alpha motif; SAMHD1, sterile alpha motif and HD containing protein 1; SANS, small angle neutron scattering; SARS CoV, severe acute respiratory syndrome coronavirus; Sema3, semaphoring; Sirt1, sirtuin-1; SRSF1, serine/arginine-rich splicing factor 1; ssDNA, single-stranded DNA; STAM, signal transducing adaptor molecule; T4P, Type IV pili; TRIM, tripartite motif; VKD, vitamin K-dependent; Vn, vitronectin

*Correspondence to: Vladimir N Uversky; Email: vuversky@health.usf.edu

Submitted: 11/12/2013

Accepted: 09/08/2014

http://dx.doi.org/10.4161/21690693.2014.970499

This review opens a new series entitled "Unreported intrinsic disorder in proteins." The goal of this series is to bring attention of researchers to an interesting phenomenon of missed (or overlooked, or ignored, or unreported)

disorder. This series serves as a companion to "Digested Disorder" which provides a quarterly review of papers on intrinsically disordered proteins (IDPs) found by standard literature searches. The need for this alternative series results from the observation that there are numerous publications that describe IDPs (or hybrid proteins with ordered and disordered regions) yet fail to recognize many of the key discoveries and publications in the IDP field. By ignoring the body of work on IDPs, such publications often fail to relate their findings to prior discoveries or fail to explore the obvious implications of their work. Thus, the goal of this series is not only to review these very interesting and important papers, but also to point out how each paper relates to the IDP field and show how common tools in the IDP field can readily take the findings in new directions or provide a broader context for the reported findings.

Introduction

Recent years evidence an exponential increase in the number of papers on intrinsically disordered proteins (IDPs), indicating that the phenomenon of protein intrinsic disorder is becoming a popular research topic. This point is also reflected in the articles of the "Digested Disorder" series published in this journal.[1-3] To make this point stronger some interesting numbers are provided below. PubMed search (as of August 03, 2014) for ((intrinsically disordered) OR (natively unfolded) OR (intrinsically unstructured) OR (intrinsically unfolded) OR (intrinsically flexible protein)) returned 2,490 hits. Restricting this search to the past one and a half years only (i.e., searching for ((intrinsically disordered) OR (natively unfolded) OR (intrinsically unstructured) OR (intrinsically unfolded) OR (intrinsically flexible protein)) AND ("2013/01/01"[Date - Publication]: "3000"[Date - Publication])) gives 659 hits. Curiously, although papers on IDPs (2,490) constitute just 0.045% of all the protein-related papers (5,569,481) published during the past 115 years, the fraction of IDP-related papers increases to 0.18% if only publications of the past year and a half are taken into account. It is of interest to compare these trends with the related tendencies of the research on proteins in general. Protein-related papers (5,569,481) constitute ~23% of all the papers published during the past 115 y and annotated in PubMed (24,068,783). However, this number decreases to ~21% if only papers published during the past year and a half are taken into account (in 2013–2014, the overall number of published papers is 1,719,568; the number of papers dealing with proteins is 362,612). This suggests that an IDP is a new rising star in the field of protein science.

Despite this steady and sure increase in the appreciation of protein intrinsic disorder, there are still numerous instances when this concept is overlooked, or missed or simply ignored. Unfortunately, such "missed disorder" phenomenon continues to be rather common in the modern literature. The original goal of the first article in the new "Unreported disorder in proteins" series was to find several papers published in different journals that talk about IDPs (or hybrid proteins containing both ordered and disordered regions) not recognizing that they are talking about such proteins, to show how consideration of intrinsic disorder can be used to strengthen conclusions and draw relationships to prior discoveries in the field. However, analysis of recent publications in just 2 journals revealed that the "missed disorder" happened to be essentially more abundant than it was originally expected. In fact, as of August 3, 2014, of 677 papers published during 2013–2014 in the *Journal of Molecular Biology*, 609 were dealing with proteins, with 87 papers being dedicated to the IDPs or hybrid proteins. Papers were considered to be dealing with intrinsic disorder if their texts contained at least a single mention of disorder, flexibility, conformational flexibility, or unfoldedness in relation to the protein of interest or any of its regions. Obviously, this is a very relaxed and inclusive approach, since simple mentioning of disorder or conformational flexibility is not sufficient for the detailed elucidation of the functional role of disorder/flexibility. The direct PubMed search for the IDP-related papers published in the *Journal of Molecular Biology* using the search criteria ((intrinsic disorder) OR (intrinsically disordered) OR (natively unfolded) OR (intrinsically unstructured) OR (intrinsically unfolded) OR (intrinsically flexible protein)) AND ("2013/01/01"[Date - Publication]: "3000"[Date - Publication]) AND ("Journal of Molecular Biology"[Journal]) generated just 17 hits. Analysis of the table of content of this journal combined with a brief computational analysis of related proteins revealed that there are at least 22 more papers, research subject of which are IDPs. Analogous analysis of the publications in *Biochemistry* during the 2013–2014 provided comparable data: of 1455 published papers, 25 were "officially" dedicated to intrinsic disorder or to intrinsically disordered/natively unfolded/intrinsically unstructured/intrinsically unfolded/intrinsically flexible proteins. Mining of papers published in *Biochemistry* during 2013–2014 revealed that some 189 papers contained at least a single mention of "conformational flexibility" OR "intrinsic disorder" OR "disordered protein" OR "disordered peptide" OR "disordered region" OR "natively unfolded" in relation to the protein of interest or any of its regions. Furthermore, based on the combination of literature mining and a brief computational analysis 20 "hidden gems" were found; i.e., papers that missed protein disorder.

The mentioned 42 papers dealing with the unreported intrinsic disorder and their related IDPs or hybrid proteins containing ordered and intrinsically disordered regions are briefly outlined below. **Table 1** contains all the proteins considered in this review. Proteins are arranged according to the increase in the extent of their disorder evaluated as percentage of the residues predicted to be disordered (i.e., possessing disorder scores above 0.5) by PONDR® VSL2, which is among the more accurate disorder predictors.[4] **Table 1** shows that the extent of unreported intrinsic disorder in 143 proteins ranges from 10% to 100%, indicating that all these proteins belong to the category of moderately or highly disordered proteins; i.e., proteins

Table 1. The extent of intrinsic disorder in all the proteins discussed in this paper

Protein name	UniProt ID	Number of residues	PONDR VSL2 (% disordered residues)	Reference
eIF4A	P10081	395	10.1	Andreou & Klostermeier[5]
Human PAI-1	P05121	402	13.2	Florova et al.[17]
T4 UvsW helicase	P20703	587	14.1	Perumal et al.[18]
Pseudomonas aeruginosa PilF	Q51385	379	14.2	Koo et al. [33]
Human integrin α4	P13612	1,032	15.8	Bonet et al.[36]
Human TRIM73	Q86UV7	250	16.8	Rajsbaum et al.[52]
Human TRIM74	Q86UV6	250	16.8	Rajsbaum et al.[52]
Human viperin	Q8WXG1	361	16.9	Helbig and Beard[58]
Human TRIM23	P36406	574	17.1	Rajsbaum et al.[52]
SAMHD1	Q9Y3Z3	626	17.6	Sze et al.[60]
Human β defensin 1	P60022	64	17.6	Wisons et al.[176]
Human TRIM48	Q8IWZ4	208	17.8	Rajsbaum et al.[52]
E. coli CspC	P0A9Y6	69	20.3	Lenz & Ron[65]
Human TRIM77	I1YAP6	450	21.1	Rajsbaum et al.[52]
Human TRIM13	O60858	407	21.4	Rajsbaum et al.[52]
Human TRIM49	P0CI25	452	21.5	Rajsbaum et al.[52]
Ciona intestinalis Ci-VSP	Q4W8A1	576	21.9	Kalli et al.[73]
Human TRIM50	Q86XT4	487	24.4	Rajsbaum et al.[52]
Pukovnik Xis	B3VGI6	56	25.0	Singh et al.[78]
Human AIP	O00170	330	25.5	Li et al.[84]
Human IFITM1	P13164	125	25.6	Perreira et al.[102]
Mouse ryanodine receptor RyR2A	E9Q401	4,966	25.7	Amador et al.[113]
Human activated protein C (APC)	P04070	461	26.0	Takeyama et al.[86]
Human TRIM60	Q495X7	471	26.1	Rajsbaum et al.[52]
Neutrophil defensin 3	P59666	64	26.6	Wisons et al.[176]
Human TRIM70	Q309B1	348	26.7	Rajsbaum et al.[52]
Human TRIM2	Q9C040	744	26.9	Rajsbaum et al.[52]
E.coli primosomal DnaT	P0A8J2	179	27.4	Szymanski et al.[91]
Human TRIM43	Q96BQ3	446	27.4	Rajsbaum et al.[52]
Neutrophil defensin 1	P0A6E6	64	27.6	Wisons et al.[176]
Bacillus PS3 Tε	P0A6E6	139	28.8	Rodriguez et al.[118]
Human TRIM51	Q9BSJ1	452	28.8	Rajsbaum et al.[52]
Human TRIM18	O15344	667	29.1	Rajsbaum et al.[52]
Human TRIM35	Q9UPQ4	493	29.4	Rajsbaum et al.[52]
Human TRIM39	Q9HCM9	518	29.9	Rajsbaum et al.[52]
Candida albicans cyclin CaPcl5	Q5AK08	304	30.3	Simon et al.[120]
Human IFITM5	A6NNB3	132	30.3	Perreira et al.[102]
Human sirtuin-2	Q8IXJ6	389	30.4	Rack et al.[127]
Human semaphorin-3	Q13275	785	30.6	Parker et al.[143]
Bovine osteocalcin	P02820	49	30.6	Malashkevich et al.[148]
Human TRIM1	Q9UJV3	735	30.7	Rajsbaum et al.[52]
Human TRIM45	Q9H8W5	580	31.4	Rajsbaum et al.[52]
Human TRIM9	Q9C026	710	31.5	Rajsbaum et al.[52]
Human TRIM62	Q9BVG3	475	31.6	Rajsbaum et al.[52]
Human TRIM68	Q6AZZ1	485	31.8	Rajsbaum et al.[52]
Human PTEN	P60484	403	32.0	Kalli et al.[73]
Pseudomonas aeruginosa PilQ	P34750	714	32.9	Koo et al. [33]
Human TRIM32	Q13049	653	32.9	Rajsbaum et al.[52]
Human TRIM69	Q86WT6	500	33.2	Rajsbaum et al.[52]
Human NF-κB	P19838	968	33.3	Sasi et al.[154]
Human AIPL1	Q9NZN9	384	33.6	Li et al.[84]
Human TRIM36	Q9NQ86	728	34.2	Rajsbaum et al.[52]
Human TRIM46	Q7Z4K8	759	34.5	Rajsbaum et al.[52]
Human TRIM34	Q9BYJ4	488	35.2	Rajsbaum et al.[52]
Mouse melanopsin	Q9QXZ9	521	36.3	Blasic et al.[172]
Human TRIM59/TRIM57	Q8IWR1	403	36.5	Rajsbaum et al.[52]
Human TRIM42	Q8IWZ5	723	36.9	Rajsbaum et al.[52]
Human PECAM-1	P16284	738	37.0	Tourdot et al.[96]
Human TRIM5	Q9C035	493	37.7	Rajsbaum et al.[52]
Human TRIM71	Q2Q1W2	868	38.1	Rajsbaum et al.[52]

(*continued on next page*)

Table 1. The extent of intrinsic disorder in all the proteins discussed in this paper *(Continued)*

Protein name	UniProt ID	Number of residues	PONDR VSL2 (% disordered residues)	Reference
Human TRIM22	Q8IYM9	498	38.6	Rajsbaum et al.[52]
Human TRIM67	Q6ZTA4	783	38.7	Rajsbaum et al.[52]
Human TRIM21	P19474	475	38.9	Rajsbaum et al.[52]
Human TRIM16	O95361	564	39.0	Rajsbaum et al.[52]
Human TRIM4	Q9C037	500	39.2	Rajsbaum et al.[52]
Human TRIM27	P14373	513	39.2	Rajsbaum et al.[52]
Neutrophil defensin 4	P12838	97	40.2	Wisons et al.[176]
Human TRIM64	A6NGJ6	449	40.5	Rajsbaum et al.[52]
Human integrin β2	P05107	769	40.8	Bonet et al.[36]
Human IFITM2	Q01629	132	40.9	Perreira et al.[102]
Human TRIM14	Q14142	442	41.0	Rajsbaum et al.[52]
Human β defensin 125	Q8N687	156	41.0	Wisons et al.[176]
Human β defensin 129	Q9H1M3	183	41.5	Wisons et al.[176]
Retinoblastoma-associated protein	P06400	928	41.9	Burke et al.[179]
Human IFITM3	Q01628	133	42.1	Perreira et al.[102]
Human TRIM10	Q9UDY6	481	42.4	Rajsbaum et al.[52]
Human TRIM6	Q9C030	488	43.2	Rajsbaum et al.[52]
Human TRIM63	Q969Q1	353	43.3	Rajsbaum et al.[52]
Human integrin β3	P05106	788	43.8	Bonet et al.[36]
Human TRIM75	A6NK02	468	43.8	Rajsbaum et al.[52]
Human NF-YB	P25208	207	44.4	Sasi et al.[154]
TPR motif-containing protein LGN	P81274	684	44.9	Pan et al.[188]
Human β defensin 116	Q30KQ4	102	45.1	Wisons et al.[176]
Human TRIM61	Q5EBN2	209	45.5	Rajsbaum et al.[52]
Human TRIM38	O00635	465	46.0	Rajsbaum et al.[52]
Human TRIM15	Q9C019	465	47.3	Rajsbaum et al.[52]
Human TRIM58	Q8NG06	486	47.7	Rajsbaum et al.[52]
Human TRIM3	O75382	744	48.0	Rajsbaum et al.[52]
Human TRIM65	Q6PJ69	517	49.7	Rajsbaum et al.[52]
Bacteriophage T4 gp32	P03695	501	49.8	Perumal et al.[18]
Human TRIM26	Q12899	539	49.9	Rajsbaum et al.[52]
Human TRIM47	Q96LD4	638	50.5	Rajsbaum et al.[52]
Human integrin β1	P05556	798	51.0	Bonet et al.[36]
Human defensin-6	Q01524	100	51.0	Wisons et al.[176]
Human TRIM56	Q9BRZ2	755	52.1	Rajsbaum et al.[52]
Human TRIM25	Q14258	630	52.2	Rajsbaum et al.[52]
Human TRIM40	Q6P9F5	258	52.3	Rajsbaum et al.[52]
Human TRIM19	P29590	882	53.1	Rajsbaum et al.[52]
Human sirtuin-1	Q96EB6	747	53.5	Davenport et al. [203]
Human AMSH	O95630	424	53.8	Davies et al.[210]
Human TRIM54	Q9BYV2	358	55.3	Rajsbaum et al.[52]
Human TRIM17	Q9Y577	477	55.6	Rajsbaum et al.[52]
Human integrin β7	P26010	798	57.0	Bonet et al.[36]
Human TRIM31	Q9BZY9	425	57.2	Rajsbaum et al.[52]
E.coli TonB	P02929	239	59.0	Freed et al.[215]
Frmpd1	Q5SYB0	1,578	59.3	Pan et al.[188]
Human TRIM52	Q96A61	297	59.9	Rajsbaum et al.[52]
Human IFITM10	A6NMD0	228	60.5	Perreira et al.[102]
Human DNAJA1	P31689	397	61.6	Stark et al.[219]
Human TRIM55	Q9BYV6	548	61.9	Rajsbaum et al.[52]
Human TRIM7	Q9C029	511	62.0	Rajsbaum et al.[52]
Human TRIM28	Q13263	835	63.1	Rajsbaum et al.[52]
Human TRIM11	Q96F44	468	63.2	Rajsbaum et al.[52]
Mouse SRSF1	Q6PDM2	248	63.7	Aubol et al.[99]
Human TRIM24	O15164	1,050	63.9	Rajsbaum et al.[52]
Human TRIM33	Q9UPN9	1,127	65.0	Rajsbaum et al.[52]
Human TRIM37	O94972	964	65.5	Rajsbaum et al.[52]
Purβ	O35295	324	66.4	Romora et al. [221]
Human TRIM8	Q9BZR9	551	66.6	Rajsbaum et al.[52]
Human TRIM29	Q14134	588	68.0	Rajsbaum et al.[52]

(*continued on next page*)

Table 1. The extent of intrinsic disorder in all the proteins discussed in this paper *(Continued)*

Protein name	UniProt ID	Number of residues	PONDR VSL2 (% disordered residues)	Reference
Human TRIM41	Q8WV44	650	71.1	Rajsbaum et al.[52]
Human tetherin	Q10589	180	73.3	Hotter et al.[225]
Human HSF-1	Q00613	529	73.3	Sasi et al.[154]
Human TRIM66	O15016	1,216	74.1	Rajsbaum et al.[52]
eIF4B	P34167	436	74.3	Andreou & Klostermeier[5]
eIF4G	P39935	952	74.6	Andreou & Klostermeier[5]
Human scaffolding protein p21	P38936	164	75.6	Jahn et al.[231]
Mouse apolipoprotein A-I	Q00623	264	76.1	Nguen et al.[232]
Human β defensin 112	Q30KQ8	113	78.8	Wisons et al.[176]
MAVS	Q7Z434	540	79.2	Jacobs and Coyne[239]
Human PSF	P23246	707	79.2	Jahn et al.[231]
Human NF-YC	Q13952	458	79.5	Sasi et al.[154]
Human TRIM44	Q96DX7	344	80.5	Rajsbaum et al.[52]
Human MYPT1	O14974	1,030	81.1	Khasnis et al.[240]
Human defensin-5	Q01523	94	83.0	Wisons et al.[176]
Human apolipoprotein A-I	P02647	267	83.1	Nguen et al.[232]
Human NF-YA	P23511	347	86.2	Sasi et al.[154]
Human TRIM76	Q8N3K9	4,069	86.2	Rajsbaum et al.[52]
Geobacillus stearothermophilus GerD	Q5L3Q1	155	88.4	Li et al.[241]
Mycobacterium tuberculosis Pup	P9WHN5	64	89.1	Forer et al.[247]
Human CREB	P16220	341	92.1	Sasi et al.[154]
Human PICT-1/GLTSCR2	Q9NZM5	478	98.7	Borodianskiy-Shteinberg et al.[248]
HIV-1 nucleocapsid protein p7	P04585	55	100.0	Spears et al.[251]

with the disorder content ranging from 10% to 30% and from 30% to 100%, respectively. **Table 1** also shows that a very significant fraction of such proteins with overlooked disorder (>75%) belongs to the "highly disordered" category.

The order of sections in this review crudely follows the order of proteins in **Table 1**. When several proteins are discussed in a paper, the corresponding section within the text is placed at a position ascribed by **Table 1** to a protein with the lowest disorder content. The corresponding sections are organized in the following way: first, a brief description of the paper is provided; then, the biological importance of the related proteins is discussed; next, the results of the computational and bioinformatics analyses of the disorder status of a given protein (or set of proteins) are represented to show how consideration of intrinsic disorder can enhance conclusions of a paper.

Functionality of Unreported Intrinsic Disorder

Eukaryotic translation initiation factors 4A, 4B and 4G (eIF4A, eIF4B, and eIF4G)

Andreou & Klostermeier investigated the peculiarities of the combined action of 3 eukaryotic translation initiation factors, eIF4A, eIF4B, and eIF4G, in RNA unwinding needed to resolve secondary structure elements from the 5′-untranslated region of mRNAs to enable ribosome scanning.[5]

A set of eukaryotic translation initiation factors (eIFs) is involved in the complex, multi-stage process of initiation of mRNA translation in eukaryotes. In the first step of this process, the eIF4F complex binds to the ^{m7}Gppp cap on the 5′-end of the mRNA. This eIF4F complex consists of the cap-binding protein eIF4E, the scaffolding protein eIF4G, and the DEAD-box helicase eIF4A.[6,7] At the next steps of the translation initiation, the eIF4F complex recruits the 43S ribosomal pre-initiation complex which scans the 5′-untranslated region in 5′-to-3′ direction in search for the translation start codon.[8,9] One of the important points in this process is the resolution of secondary and tertiary structures in the mRNA, which is typically done by the DEAD-box helicase eIF4A, which possesses RNA-dependent ATPase and ATP-dependent RNA helicase activities.[10] These 2 activities of eIF4A are enhanced by some auxiliary proteins, such as eIF4B and eIF4G that act synergistically stimulating the eIF4A helicase activity in the mRNA scanning process.[5]

Figure 1 illustrates the disorder propensities of the 3 proteins related to the mRNA unwinding. Disorder was evaluated by a family of PONDR predictors. Here, scores above 0.5 correspond to disordered residues/regions. PONDR® VSL2B is one of the most accurate stand-alone disorder predictors,[11] PONDR® VL3 possesses high accuracy in finding long IDPRs,[12] PONDR® VLXT is not the most accurate predictor but has high sensitivity to local sequence peculiarities which are often associated with disorder-based interaction sites,[13] whereas PONDR-FIT represents a metapredictor which, being moderately more accurate than each of the component predictors, is one of the most accurate disorder predictors.[14] The various predictors often give different predictions of disorder for the same protein, perhaps leading so some confusion when trying to understand the implications of this and the following figures. These differences arise because of different computational methods and especially because different training sets were used. Despite the differences, the overall disorder prediction trends are typically similar for each protein.

In agreement with the notion that catalytic functions require specific and ordered structure, eIF4A, with its ATPase and ATP-dependent RNA helicase activities, is predicted to be mostly ordered protein (see **Fig. 1A**; UniProt ID: P10081). The mostly ordered nature of this enzyme is further supported by the fact that almost entire sequence is seen in the crystal structure (see red structure at **Fig. 1B**, PDB ID: 2VSO), except to the residues 1–11, 126–135, 351–356, and 394–395. On the contrary, the auxiliary proteins eIF4B and eIF4G are predicted to possess long disordered regions (see **Figs. 1C and 1D**; UniProt IDs: P34167 and P39935, respectively). eIF4G is also predicted to have an ordered region that coincides with its middle or eIF4A interacting domain (residues 571–854). This region was co-crystallized with the eIF4A factor (see blue structure at **Fig. 1B**). Curiously, several eIF4G regions were missing in structure of the eIF4A-eIF4G complex (e.g., residues 571–576, 583–596, 686–688, 717–729, 803–811, and 853–854).

Long intrinsically disordered regions are important for functions of eIF4B and eIF4G. In fact, according to ANCHOR,[15,16] eIF4B (UniProt ID: P34167) is expected to have 14 disorder-based binding sites (residues 13–20, 33–40, 48–53, 68–74, 103–109, 159–166, 214–219, 234–244, 259–271, 286–298, 312–324, 338–355, 376–386, and 407–426). Similarly, eIF4G (UniProt ID: P39935) is predicted to have 18 AiBSs, residues 5–24, 32–53, 59–120, 130–158, 198–226, 230–241, 292–307, 329–342, 351–398, 411–433, 450–465, 474–485, 505–515, 570–583, 818–827, 832–844, 893–920, and 933–949.

Human plasminogen activator inhibitor 1 (PAI-1)

Florova et al. investigated the mechanisms of the stabilization of human plasminogen activator inhibitor 1 (PAI-1, also known as serpin E1) during the formation of the transient ternary "molecular sandwich" complex (MSC) containing PAI-1, vitronectin (Vn), and the target enzyme.[17] Major players involved in the formation of this complex are PAI-1, which is a major endogenous inhibitor of plasminogen activators, a cell adhesive glycoprotein Vn, and a proteinase inhibited by PAI-1. An important feature of PAI-1 is its ability to exist in 2 forms, a metastable active conformation with solvent-accessible reactive center loop (RCL) and thermodynamically stable, inactive, latent conformation, where RCL is spontaneously inserted to the middle of the β-sheet B of PAI-1.[17] This "magic" chameleon-like RCL is located in the 331–350 region. **Figure 2A** shows that although human PAI-1 (UniProt ID: P05121) is predicted to be mostly ordered, its RCL is located within grayish area with a mean disorder score close to 0.5, suggesting that this region is characterized by noticeable intrinsic mobility.

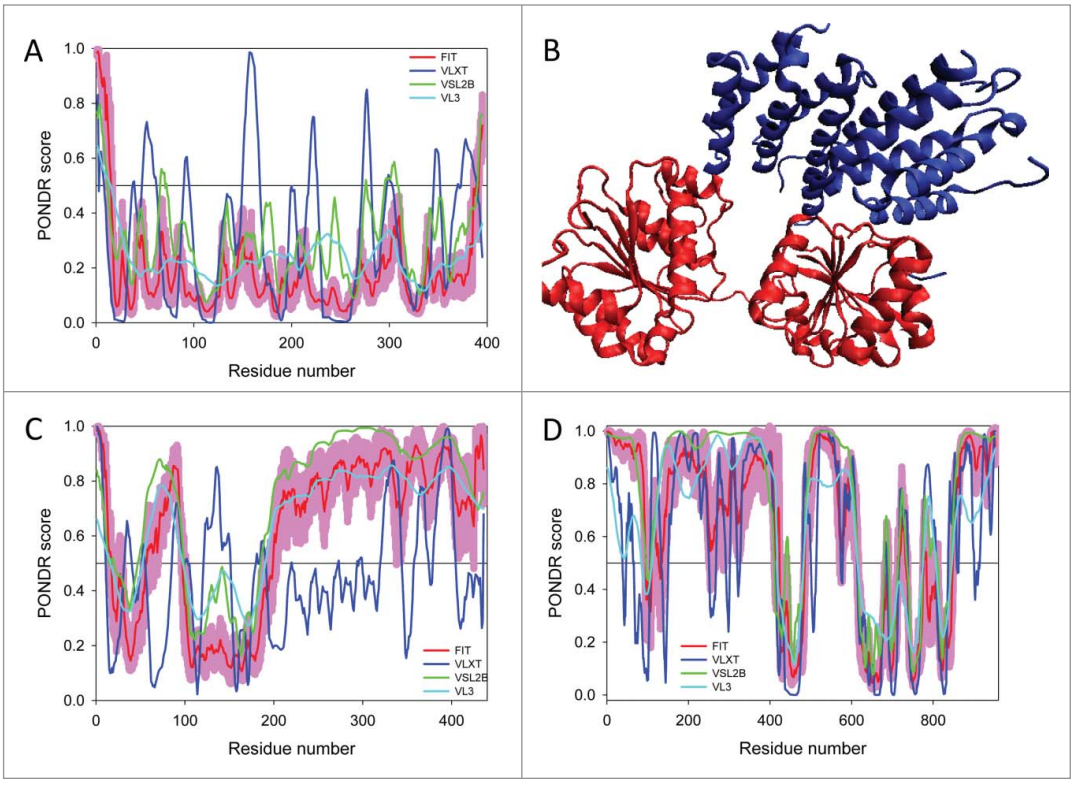

Figure 1. Intrinsic disorder in the eIF4A (UniProt ID: P10081, **A**) eIF4B (UniProt ID: P34167, **C**), and eIF4G (UniProt ID: P39935, **D**) evaluated by PONDR® FIT (red lines), PONDR® VLXT (blue lines), PONDR® VSL2B (green lines), and PONDR® VL3 (cyan lines). (**B**) Crystal structure (PDB ID: 2VSO) of the complex between the eIF4A (red) and the middle domain of eIF4G (residues 572–853, blue).

T4 UvsW helicase and Single-Stranded DNA Binding Protein gp32

Perumal et al. reported that a specific interaction between the bacteriophage T4 UvsW helicase and the T4 single-stranded DNA (ssDNA) binding protein gp32 is required for enhancement of the UvsW DNA unwinding function.[18] Curiously, UvsW interact with gp32, both in the presence and absence of DNA, through the C-terminal acidic tail of the gp32 protein. In the absence of this interaction, the ssDNA annealing and ATP-dependent translocation activities of UvsW are severely inhibited when gp32 coats the ssDNA lattice. However, when UvsW and gp32 do interact, UvsW is able to efficiently displace the gp32 protein from the

ssDNA. The tail-less gp32 inhibits activities of UvsW on DNA in the presence of gp32.[18]

The UvsW helicase, a member of the SF2 superfamily of helicases[19] (i.e., enzymes that use ATP to carry out mechanical work related to the unwinding of duplex DNA structures and reorganization of RNA secondary structures[20]), is one of the 3 helicases encoded by the bacteriophage T4,[21] where it plays a number of roles in a variety of DNA repair and recombination pathways.[22-27] The involvement of this helicase in DNA replication is achieved via the UvsW-mediated unwinding of R-loop structures generated at the replication origin by the host RNA polymerase.[22] UvsW (UniProt ID: P20703) activity involves the generation and consumption of single stranded DNA (ssDNA). In agreement with the overall low disorder predisposition (see **Fig. 3A**), almost entire sequence of UvsW was successfully crystallized (**Fig. 3B**, PDB ID: 2OCA).

Naked ssDNA rarely occurs within T4-infected cells as they are immediately coated with gp32 ssDNA binding protein, which serves to protect it against degradation by endonucleases. The gp32 is also thought to coordinate the many activities necessary to carry out DNA replication, recombination and repair by recruiting and modulating the activity of the enzymes involved in these processes.[28] The bacteriophage T4 gp32 (UniProt ID: P03695) is a 301 residue-long protein. The crystal structure of the gp32 central region (residues 22–239) is known (see **Fig. 3C**).[29] Gp32 binds preferentially to ssDNA and destabilizes double-stranded DNA. It is involved in DNA replication, repair and recombination, dinds ssDNA as the replication fork advances and stimulates the replisome processivity and accuracy. The N-terminal region of gp32 is important for the cooperative binding of gp32 monomers on ssDNA, whereas the C-terminal acidic tail is involved interaction with the replicative DNA polymerase, the primase and helicase loader protein.[30-32] Also, this C-terminal region (residues 254–301) was shown to play a crucial role in controlling the D-loop unwinding capability of UvsW, since the presence of a mutant gp32 protein lacking the acidic

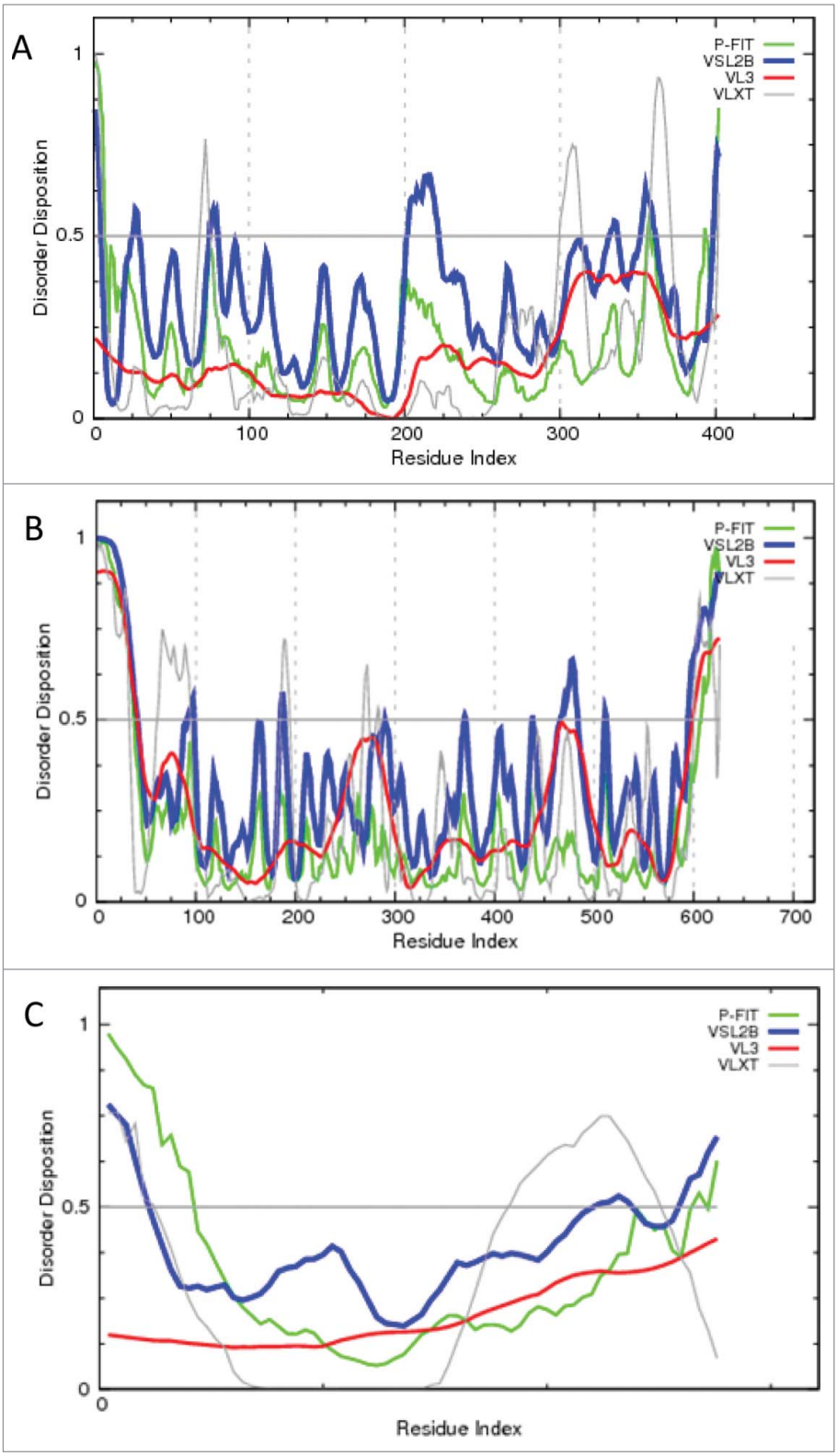

Figure 2. (**A**) Intrinsic disorder in human plasminogen activator inhibitor 1 (PAI-1, UniProt ID: P05121). (**B**) Intrinsic disorder propensities of SAMHD1 (UniProt ID: Q9Y3Z3). (**C**). Intrinsic disorder in CspC from *E. coli* (UniProt ID: P0A9Y6). Intrinsic disorder propensities are evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

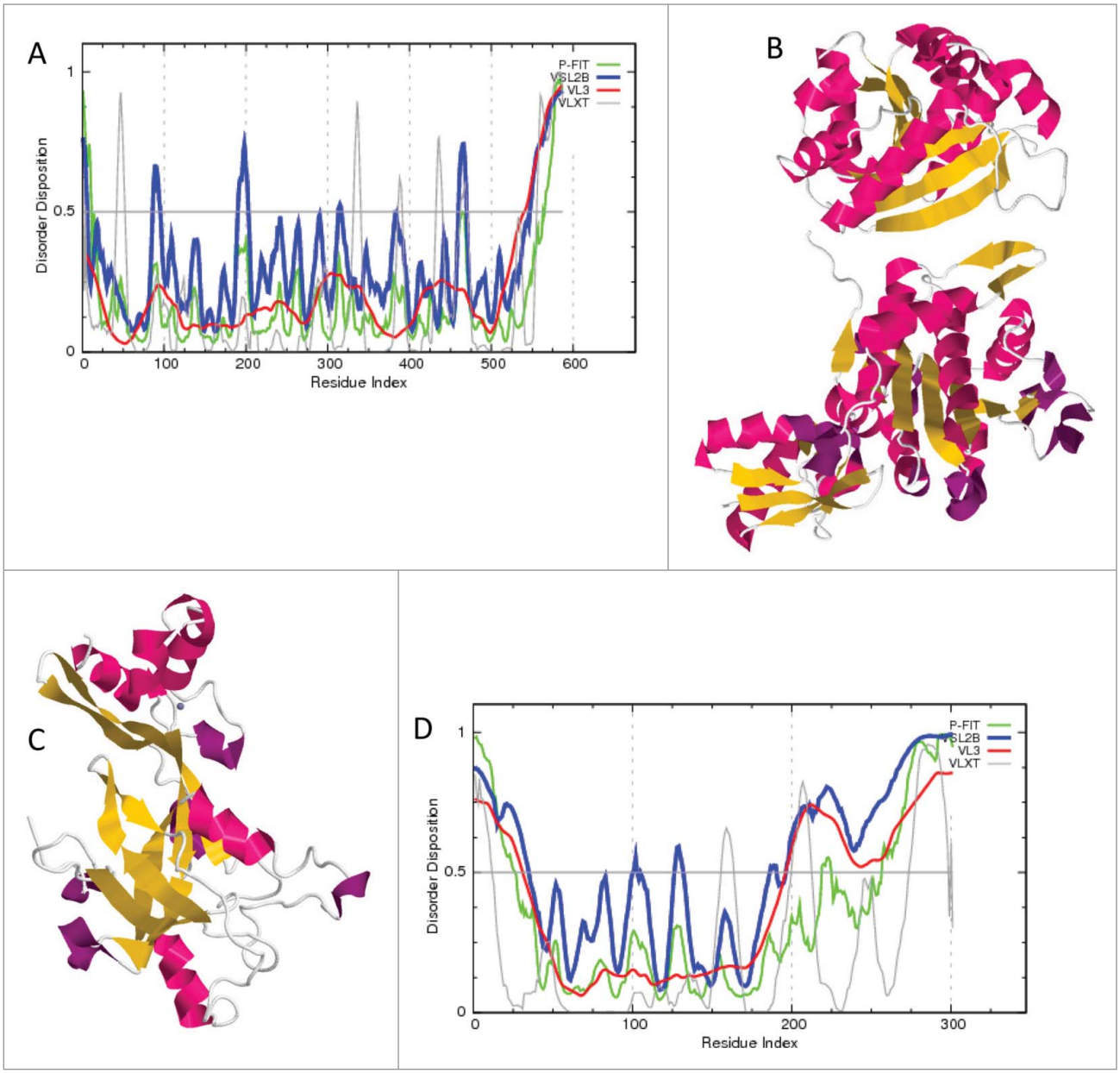

Figure 3. (**A**) Evaluation of the disorder propensity of the UvsW helicase (UniProt ID: P20703) from the bacteriophage T4. Disorder propensity was evaluated by a set of predictors from the PONDR family, PONDR® FIT, VSL2B, VL3, and VLXT. Scores above 0.5 correspond to disordered residues/regions. (**B**) Crystal structure of the bacteriophage T4 UvsW helicase (PDB ID: 2OCA). (**C**). Crystal structure of gp32 (PDB ID: 1GPC). (**D**) Disorder propensity in the gp32 protein (UniProt ID: P03695) evaluated by a set of predictors from the PONDR family, PONDR-FIT, VSL2B, VL3, and VLXT.

tail (delta254–301, gp32-A) led to a significant reduction in the unwinding of the D-loop substrate.[18] **Fig. 3D** shows that a significant portion of the C-terminal half of this protein is predicted to be mostly disordered (residues 200–301). Curiously, this disordered region includes the functionally important C-terminal tail that defines the ability of gp32 to modulate the UvsW activity.

PilF and PilQ of the Pseudomonas aeruginosa Type IV pilus system

Some bacteria and archaea contain specific cell envelope-spanning biomolecular machines, Type IV pili (T4P), which are used by bacteria and archaea to interact with the environment. Pilus biogenesis is controlled by the assembly of the outer membrane PilQ secretin channel through which the pili are extruded. Koo et al. investigated the roles of the *Pseudomonas aeruginosa* type IV pili (T4P) pilotin protein PilF in the outer membrane PilQ secretin channel targeting, oligomerization, and function.[33]

PilF belongs to the class 1 pilotins, which are α-helical proteins containing several tetratricopeptide (TPR) motifs, that are comprised of approximately 34-amino acid and are able to form a

superhelical fold that mediates protein–protein interactions in both prokaryotes and eukaryotes.[33] Secretins (including PilQ sectretin) possess a conserved C-terminal region, containing the secretin domain that is putatively embedded into the outer membrane, and a variable, system-specific N-terminal region.[34,35] **Figure 4A** shows that the *P.aeruginosa* PilF (UniProt ID: Q51385) is predicted to be a hybrid protein possessing both ordered and disordered regions. Curiously, about 2/3 of this protein possesses disorder scores close to 0.5. This grayish area is likely to be characterized by high conformational plasticity and is predicted to contain 4 AiBSs, residues 5–16, 41–46, 76–81 and 111–115. Similarly, PilQ (UniProt ID: P34750) is predicted to be a hybrid protein (see **Fig. 4B**) with several disorder-based binding sites (residues 101–107, 120–122, 245–251, 502–505, and 533–547).

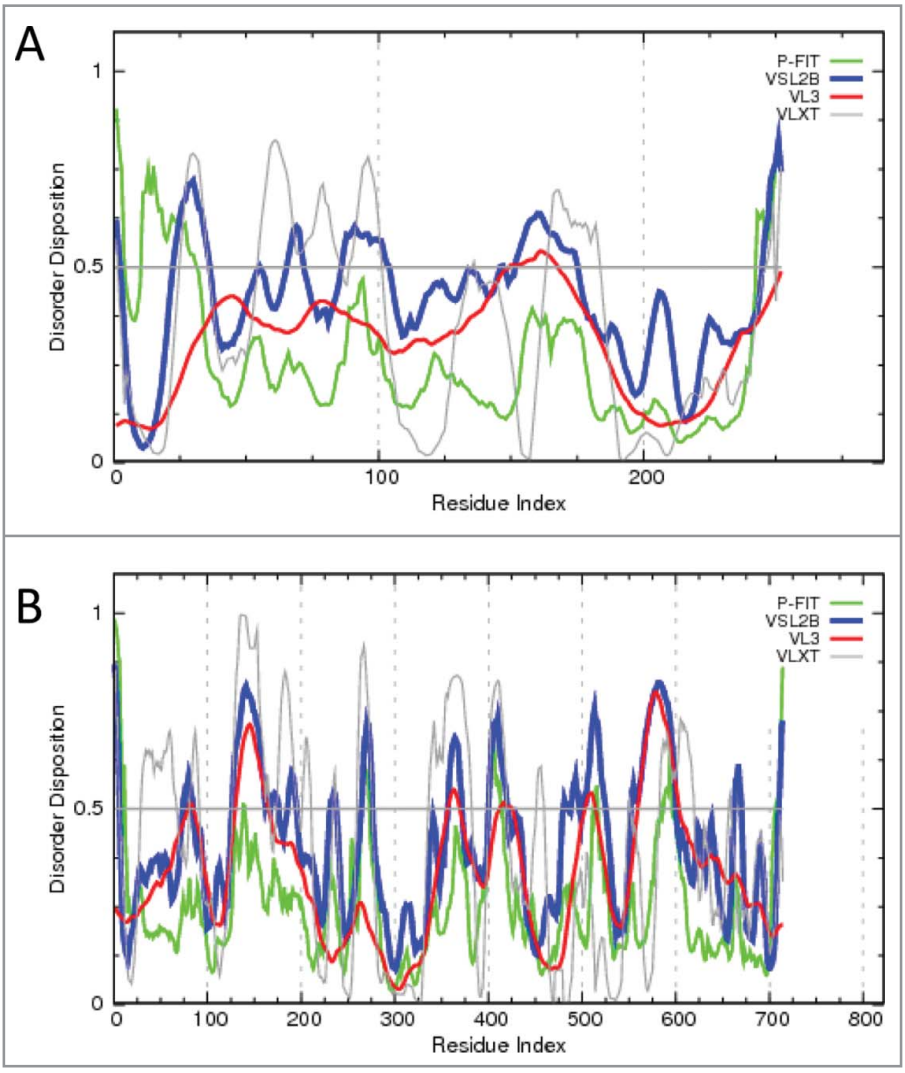

Figure 4. Intrinsic disorder in the type 4 fimbrial biogenesis protein PilF (**A**, UniProt ID: Q9HXJ2); the fimbrial assembly protein PilQ (**B**, UniProt ID: P34750). Intrinsic disorder propensities are evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

Interaction of 14–3–3-ζ with integrin tails

Bonet et al. analyzed the molecular mechanisms of interaction of the 14–3–3ζ protein with several integrin proteins.[36] The authors provided a detailed biophysical characterization of the cytoplasmic tails of α4, β1, β2 and β3 integrins binding to 14–3–3ζ and showed that binding affinities and interaction modes of different integrins with this 14–3–3 are rather different. Furthermore, although many structural features of these interactions are similar to other known 14–3–3 complexes, the binding exhibits specific features involving secondary sites. Particularly, in addition to a canonical binding mode for the α4 phospho-peptide, some residues outside the consensus 14–3–3ζ binding motif of this integrin were shown to be essential for an efficient interaction. Although a short β2 phospho-peptide is sufficient for high-affinity binding to 14–3–3ζ, the authors also found novel 14–3–3ζ/integrin tail interactions that were independent of phosphorylation. [36]

The members of the 14–3–3 family are highly conserved acidic proteins of ~30 kDa that are abundantly expressed in all eukaryotic cells. In human, there are 7 14–3–3 isoforms (β, γ, ε, η, σ, τ and ζ) forming homodimers or heterodimers. These proteins regulate and control many signaling pathways, such as cytoskeletal dynamics programmed cell death, and cell cycle progression, and are involved in the pathogenesis of several human diseases, such as cancer and neurological disorders.[37,38] Although 14–3–3 proteins are considered as phosphor-serine/threonine binding modules possessing 2 consensus recognition motifs, RSXpSXP and RXF/YXpSXP,[39] some other binding modes are known including interaction with some unphosphorylated motifs.[40,41] X-ray structures of various 14–3–3 isoforms and their numerous complexes are known providing considerable knowledge on the various binding modes.[42,43] Curiously, a detailed analysis of more than 200 binding partners of 14–3–3 proteins showed neither structural nor functional relatedness in this group of proteins.[44] However, bioinformatics analysis established that >90 % of the 14–3–3 interactors contain disordered regions and that almost all 14–3–3-binding sites are located inside disordered regions.[44]

Among various partners of 14–3–3 proteins are membrane-spanning receptors, integrins, which are involved in

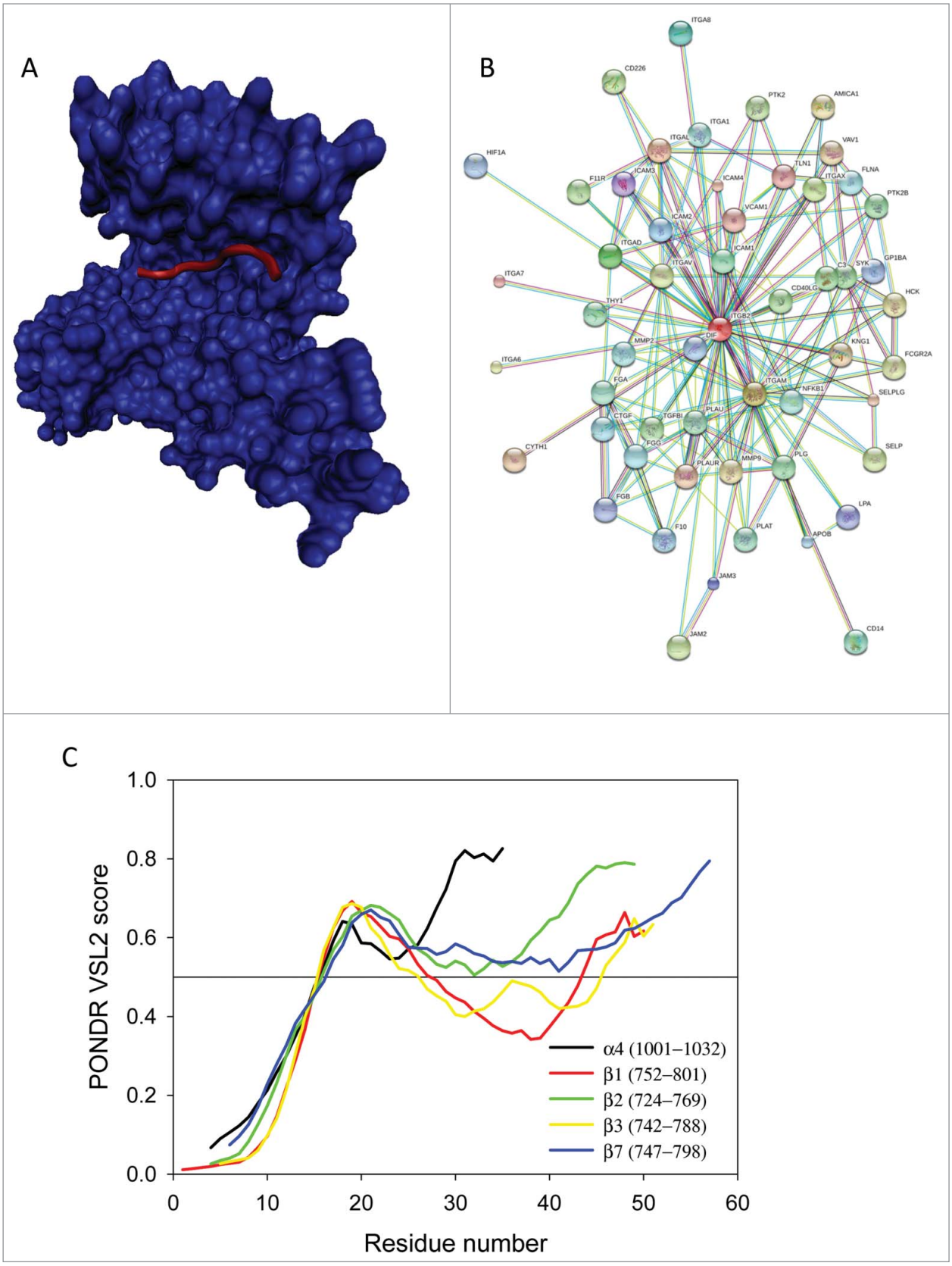

Figure 5. For figure legend, see page 11.

numerous biological functions, such as cell adhesion, migration and differentiation and are associated with a wide-range of diseases.[45,46] Integrins are heterodimers formed by α- and β-subunits composed of large extracellular (ecto) domains, trans-membrane domains and short (13–70 residues) C-terminal cytoplasmic domains,[47] which are flexible tails acting as hubs for numerous integrin-based protein-protein interactions.[48-50] These cytoplasmic tails of the 2 integrin subunits also mediates the bidirectional inside-to-outside and outside-to-inside signaling, where signals transmitted by integrins from outside to inside the cell promote cell survival and proliferation, and where integrin affinity for the extracellular ligands can also be controlled by intracellular factors.[47]

Figure 5A represents a structure of the complex between the 14–3–3ζ protein (blue cloud) and a phosphorylated peptide from the β2 integrin tail (red chain).[51] It is seen that similar to many other 14–3–3 complexes, the β2 integrin tail is bound in a highly extended form. To illustrate interactivity of the β2 integrin, **Figure 5B** shows the results of the analysis of this protein by STRING. Here, settings were chosen to find binding partners with the highest confidence (0.9). Finally, **Figure 3C** shows some disorder-based alignments of the C-terminal tails used in the work of Bonet et al.,[36] namely residues 1001–1032 of human integrin α4 (black line, UniProt ID P13612), residues 752–801 of human integrin β1 (black line, UniProt ID P05556), residues 724–769 of human integrin β2 (black line, UniProt ID P05107), residues 742–788 of human integrin β3 (black line, UniProt ID P05106), and residues 747–798 of human integrin β7 (black line, UniProt ID P26010). **Figure 5C** clearly shows that all these integrin tails involved in the specific interaction with the 14–3–3ζ protein are mostly disordered.

TRIMunity

In the review by Rajsbaum et al., an important family of proteins, tripartite motif (TRIM) proteins is introduced together with the multitude of their functional roles in the innate antiviral immunity.[52] Since there are more than 70 distinct members in the family of human TRIM proteins, it is physically impossible to consider all of them even very superficially. The feature that links all these proteins together is the fact that they share 3 conserved N-terminal domains: a Really Interesting New Gene (RING) domain, one or 2 B-Boxes (B1/B2) and a coiled-coil domain.[53-56] In addition to immune-related functions many TRIM proteins are involved in a wide range of biological activities, such as transcriptional regulation, apoptosis, cell differentiation, development, oncogenesis, ubiquitin E3 ligases, and E3 ligases for other ubiquitin-like molecules such as SUMO and the IFN-inducible protein ISG15.[57] Computational analysis (see **Table 1**) revealed that all members of the human TRIM family are extensively disordered.

Multiple roles of viperin in the innate antiviral response

Helbig and Beard overviewed the ability of viperin to modulate conditions within the cell and to interfere with proviral host proteins in order to create an unfavorable environment for viral replication.[58] Viperin is one of the few products of numerous interferon-stimulated genes (ISGs) possessing direct antiviral activity, being able to limit a broad range of viruses and to play an emerging role in modulating innate immune signaling. This is a highly species conserved protein consisting of 3 distinct domains: a variable N-terminal domain that contains an amphipathic helix and a leucine zipper region, a highly conserved central domain containing a "radical SAM domain," and a conserved C-terminal critical for the antiviral properties of viperin against a number of viruses.[58] The authors showed that viperin plays a role in innate immune signaling, limits different viruses through both direct inhibition of replication and interference with viral budding/release, and disrupts the actin cytoskeleton to increase infectivity of human cytomegalovirus (HCMV) in a known example of evolutionary escape of HCMV from the antiviral properties of viperin.[58] Viperin is able to interact with the 5 host proteins FPPS, TFP, IRAK1, VAP-A and TRAF6 and the 3 viral proteins DENV NS3, HCV NS5A and HCMV vMIA in order to accomplish these diverse biological functions.[58] **Figure 6A** represents the results of the STRING[59]-based evaluation of viperin interactivity. Although it has been emphasized that "It is unusual for viperin to be able to interact with such a divergent range of other proteins and to potentially mediate quite distinct cellular functions",[58] the mystery of such "unusual" polyfunctionality is naturally solved by the presence of extensive disorder in this protein (see **Fig. 6B**; UniProt ID: Q8WXG1).

SAMHD1 host restriction factor

Sze et al. overviewed available information on the sterile α motif and histidine-aspartic domain (HD) containing protein 1 (SAMHD1), which is a member of the unique group of host restriction factors that limit retroviral replication at distinct stages of the viral life cycle.[60] SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase responsible for the degradation of the deoxynucleoside triphosphates (dNTP) into their deoxynucleosides (dN) and inorganic triphosphates, thus depleting the cellular dNTP pool required for cellular DNA polymerase.[61] Activity of this protein is modulated by post-translational modifications, cell-cycle-dependent functions and cytokine-mediated changes, as well as via interaction with the Vpx accessory protein.[60]

SAMHD1is a predominantly nuclear protein composed of 2 functional domains, the sterile α motif (SAM) domain involved

Figure 5 (**see previous page**). (**A**) A crystal structure (PDB ID: 2V7D) of the complex between the 14–3–3ζ protein (blue cloud) and a phosphorylated peptide from the β2 integrin tail (red chain).[51] (**B**) Interactome of the β2 integrin as evaluated by STRING.[59] Here, settings were chosen to find binding partners with the highest confidence (0.9). (**C**). Disorder-based alignments of the C-terminal tails of several human integrins: residues 1001–1032 of human integrin α4 (black line, UniProt ID: P13612), residues 752–801 of human integrin β1 (black line, UniProt ID: P05556), residues 724–769 of human integrin β2 (black line, UniProt ID: P05107), residues 742–788 of human integrin β3 (black line, UniProt ID: P05106), and residues 747–798 of human integrin β7 (black line, UniProt ID: P26010).

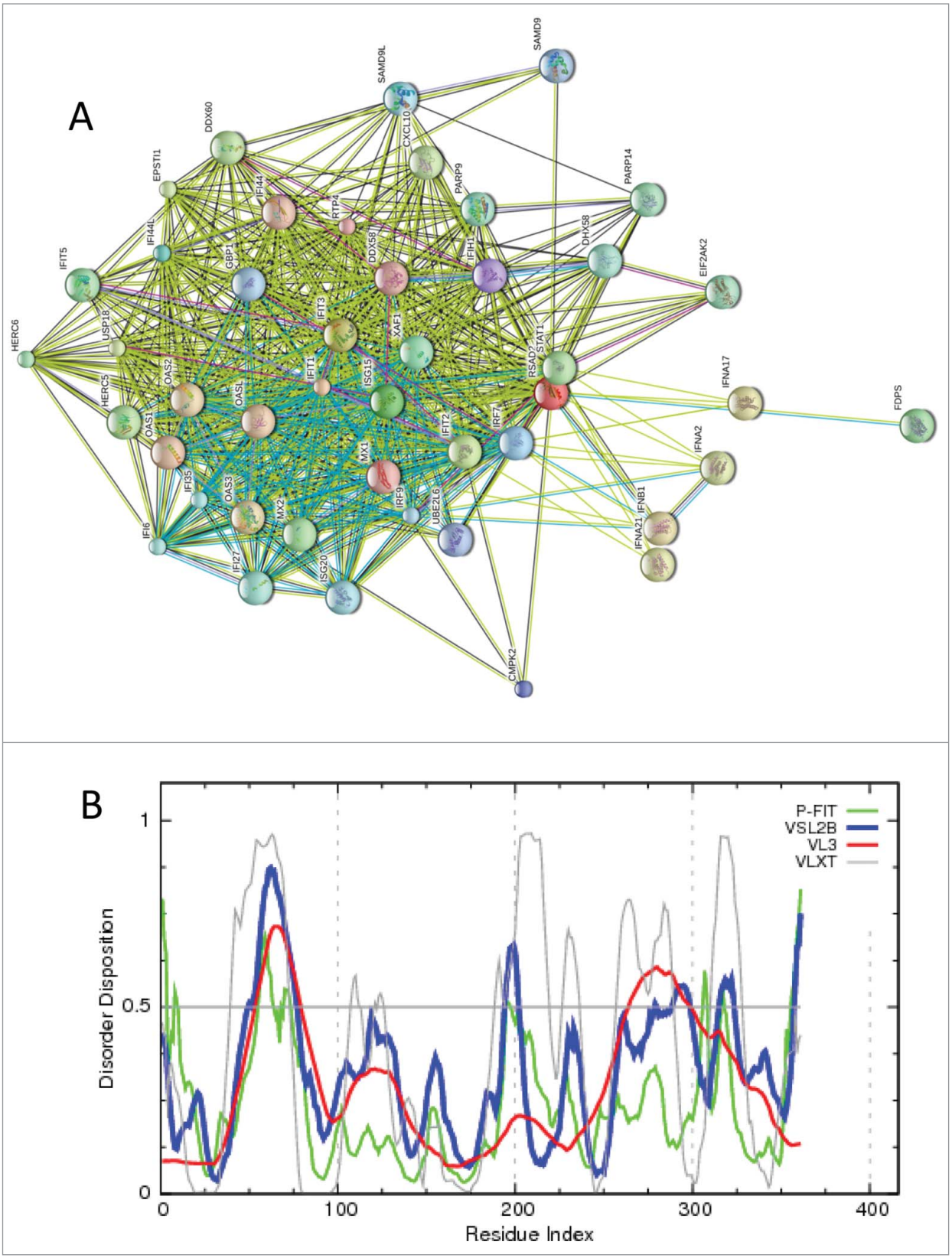

Figure 6. For figure legend, see page 13.

in protein–protein and SAMHD1-nucleic acid interactions,[62] and the HD domain containing the enzymatic sites crucial for its triphosphohydrolase activity, RNA binding and nuclease activity.[63] Another functionally important region is located at the C-terminus of SAMHD1, where a V-domain capable of interaction with the HIV-2/SIVsm Vpx accessory protein is located (residues 595–626).[64] Based on the intrinsic disorder propensity analysis (**Fig. 2B**), SAMHD1 (UniProt ID: Q9Y3Z3) is expected to be a predominantly ordered protein possessing disordered N- and C-terminal tails (residues 1–100 and 580–626). Obviously, intrinsic disorder in the C-terminal tail is important for the SAMHD1 interaction with the HIV-2/SIVsm Vpx accessory protein.

Interaction between the major bacterial heat shock chaperone GroESL and an RNA chaperone CspC

In their recent paper, Lenz & Ron described a novel interaction between the major heat shock chaperone GroESL of *E. coli* (Hsp60) with an RNA chaperone (CspC) leading to the CspC proteolysis needed for the transient nature of the heat shock response.[65]

Protein chaperones and proteases have multiple functions in controlling the well-being of a cell, acting as protein quality control that enables the cells to cope with the unfolding and aggregation of proteins.[66,67] CspC is a member of the cold shock protein (Csp) family that consists of 7 small homologues with high affinity to single-strand nucleic acid and that serve as RNA chaperones.[68-71] It has been shown that CspC is degraded during heat shock and that interaction of CspC with the major protein chaperone GroESL plays a role in the enhanced, temperature-dependent proteolysis of CspC.[65]

The fact that CspC is able to interact with GroESL (which is known to bind partially unfolded and misfolded protein species) is a clear indication that CspC does not possess rigid structure, at least under the conditions favoring such interaction. The likely explanation for this binding is in the potential intrinsically disordered nature of CspC. In fact, many protein and RNA chaperones were shown to be intrinsically disordered or hybrid proteins possessing both ordered and intrinsically disordered regions/domains.[72] In agreement with this observation, **Figure 2C** represents the PONDR plots for the CspC from *E. coli* (UniProt ID: P0A9Y6) and shows that a significant portion of this small proteins (up to 35%) is predicted to be disordered.

Interaction of PTEN with phosphatidylinositol phosphate

Kalli et al. used multiscale molecular dynamics simulations to define the interaction mechanisms of phosphatase and tensin homolog (PTEN; UniProt ID: P60484) and of the PTEN domain of *Ciona intestinalis* voltage sensitive phosphatase (Ci-VSP; UniProt ID: Q4W8A1) with phosphatidylinositol phosphate (PIP)-containing lipid bilayers.[73] The authors revealed that the association of the PTEN with such bilayers involves the formation of an initial electrostatics-driven encounter complex between the protein and bilayer followed by reorientation of the protein to optimize its interactions with PIP molecules in the membrane.[73] PTEN is a cytosolic enzyme that can interact with the inner leaflet of the plasma membrane and, being bound to membrane, catalyzes dephosphorylation of $PI(3,4,5)P_3$ to $PtdIns(4,5)P_2$.[74] PTEN has 4 domains, an N-terminal PIP_2-binding module, a phosphatase domain (PD), a C2 domain, and a C-terminal tail. Several recent studies indicated that both N- and C-tails of this protein are intrinsically disordered.[75-77] Furthermore, it has been emphasized that post-translational modifications, conserved eukaryotic linear motifs, and molecular recognition features are present in the disordered C-tail of PTEN, enhancing protein-protein interactions of this protein needed for the various cellular functions of PTEN.[75]

Domain swapped Pukovnik Xis

Singh et al. determined the structure of Pukovnik Xis by X-ray crystallography.[78] Xis is the recombination directionality factor that serves as a DNA bending machine that difines the outcome of integrase-mediated site-specific recombination by redesign of higher-order protein–DNA architectures.[79] Mycobacterium phage Pukovnik (which is a relative of phage L5) contains a small Xis protein (56 residues) that binds cooperatively to *attR* DNA at specific X1–X4 binding sequences. Similar to L5, Pukovnik Xis stimulates integrase-mediated excision and inhibits integration.[80,81] The presence of both Xis and intergrase is needed for the formation of an *attR* intasome.[78] The cooperative binding of several Xis proteins to DNA is driven by a winged-helix motif and relies on the use of contacts between a central loop and the wing motif to mediate interactions between Xis proteins when bound to DNA leading to the formation of a micronucleoprotein filament with a modest bend.[78,82,83]

Singh et al. showed that in the DNA-bound form, 5 individual Pukovnik Xis subunits stack onto each other through an extensive array of protein–protein interactions forming a filament with left-handed superhelical twist.[78] Stacking of Xis subunits within the filament is rather regular, and each monomer exhibits a twist of 40° and an ~60° bending angle. Within the asymmetric unit, the filament region contains 4 Xis protomers whereas the fifth protomer is positioned adjacent to the filament. This "outside" protomer forms the domain-swapped complex with the Xis protomer 3. This domain-swapped Xis

Figure 6 (**see previous page**). (**A**) Evaluation of the interactivity of the human viperin (UniProt ID: Q8WXG1) by STRING, which is the online database resource Search Tool for the Retrieval of Interacting Genes that provides both experimental and predicted interaction information.[59] STRING produces the network of predicted associations for a particular group of proteins. The network nodes are proteins. The edges represent the predicted functional associations. An edge may be drawn with up to 7 differently colored lines - these lines represent the existence of the 7 types of evidence used in predicting the associations. A red line indicates the presence of fusion evidence; a green line - neighborhood evidence; a blue line – co-occurrence evidence; a purple line - experimental evidence; a yellow line – text mining evidence; a light blue line - database evidence; a black line – co-expression evidence.[59] (**B**) Intrinsic disorder propensity of the human viperin. Disorder propensity was evaluated by a set of predictors from the PONDR family, PONDR® FIT, VSL2B, VL3, and VLXT. Scores above 0.5 correspond to disordered residues/regions.

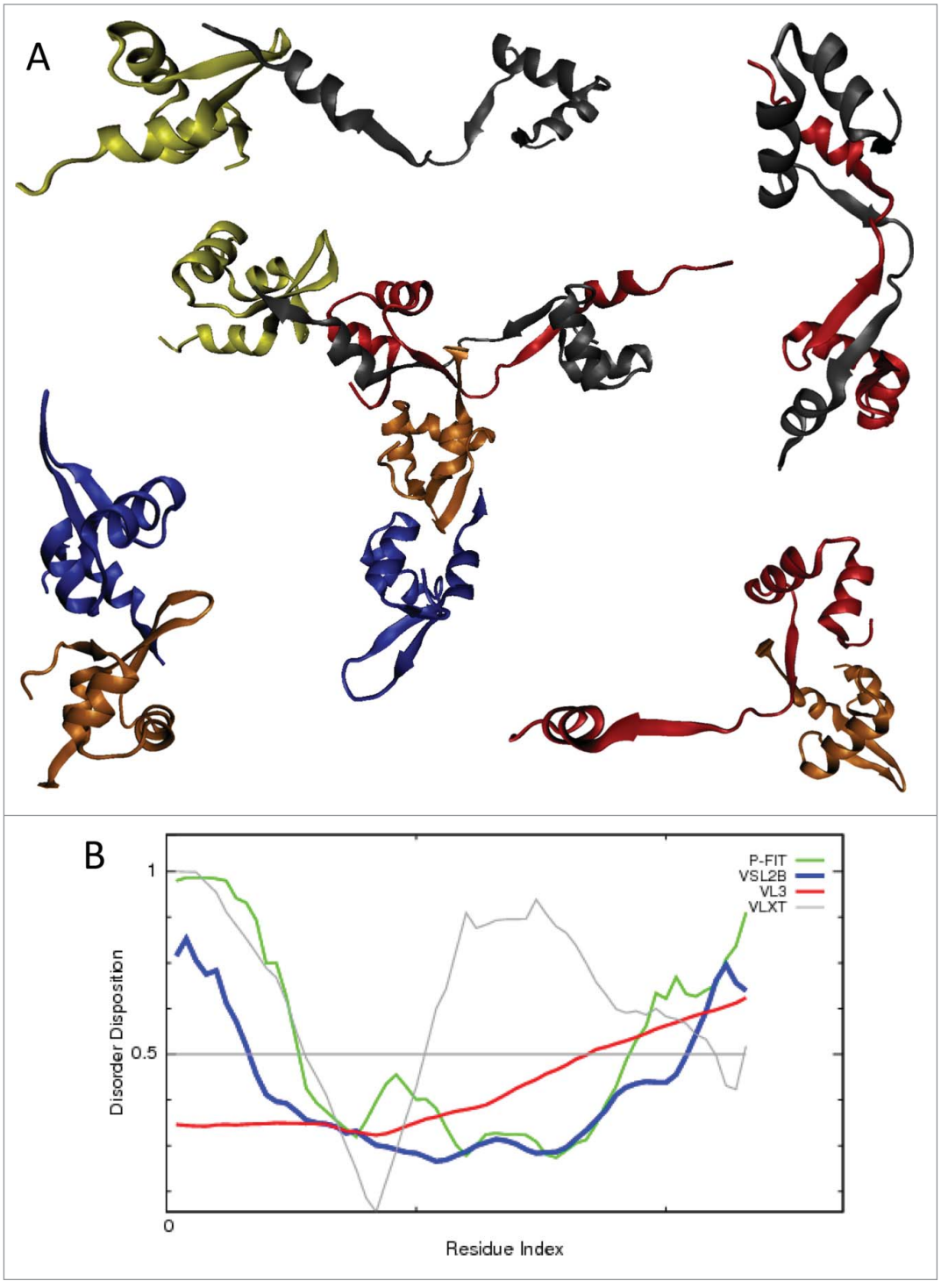

Figure 7. (**A**) Crystal structure of the Pukovnik Xis pentamer (PDB ID: 4J2N) and 4 pairs of different dimers found within this pentameric filament. (**B**) Intrinsic disorder propensity of Pukovnik Xis (UniProt ID: B3VGI6) evaluated by by a set of predictors from the PONDR family, PONDR® FIT, VSL2B, VL3, and VLXT. Scores above 0.5 correspond to disordered residues/regions.

pair is created due to a dramatic rearrangement within the loop connecting β1 and β2 (residues 36–39) allowing the C-terminal residues 40–56 of one subunit to interact with the N-terminal 35 residues of a neighboring Xis monomer.[78] It has been emphasized that the most important protein-protein interactions mediating the filament formation are formed by residues 51–54 of Pukovnik Xis that make extensive contacts with the wing of the adjacent monomer and deletion of which greatly diminishes DNA binding affinity and abolishes excision.[78] Therefore, although the Xis filament is composed of identical subunits, the individual Xis protomers are involved in very different interactions and form very different interfaces. To further illustrate this point, **Figure 7A** represents a crystal structure of the Pukovnik Xis pentamer (PDB ID: 4J2N) and 4 pairs of different dimers found within this pentameric filament. **Figure 7B** shows that Pukovnik Xis (UniProt ID: B3VGI6) is predicted to contain substantial amount of intrinsic disorder (note scale of the Y-axis), including disordered N- and C-terminal tails. Therefore, it is likely that the intrinsically disordered nature determines the ability of this small protein to be involved in a complex net of protein-protein and protein-DNA interactions, including its ability to form a wide array of regular and domain-swapped dimers.

Proline-rich domain of the AIPL1 chaperone

Li et al. provided a detailed characterization of the functional mechanisms of interesting Hsp90 co-chaperones, human aryl hydrocarbon receptor (AHR) interacting protein (AIP) and AIP like 1 (AIPL1).[84] AIP and AIPL1 share 49% sequence identity, contain an N-terminal FKBP-like prolyl peptidyl

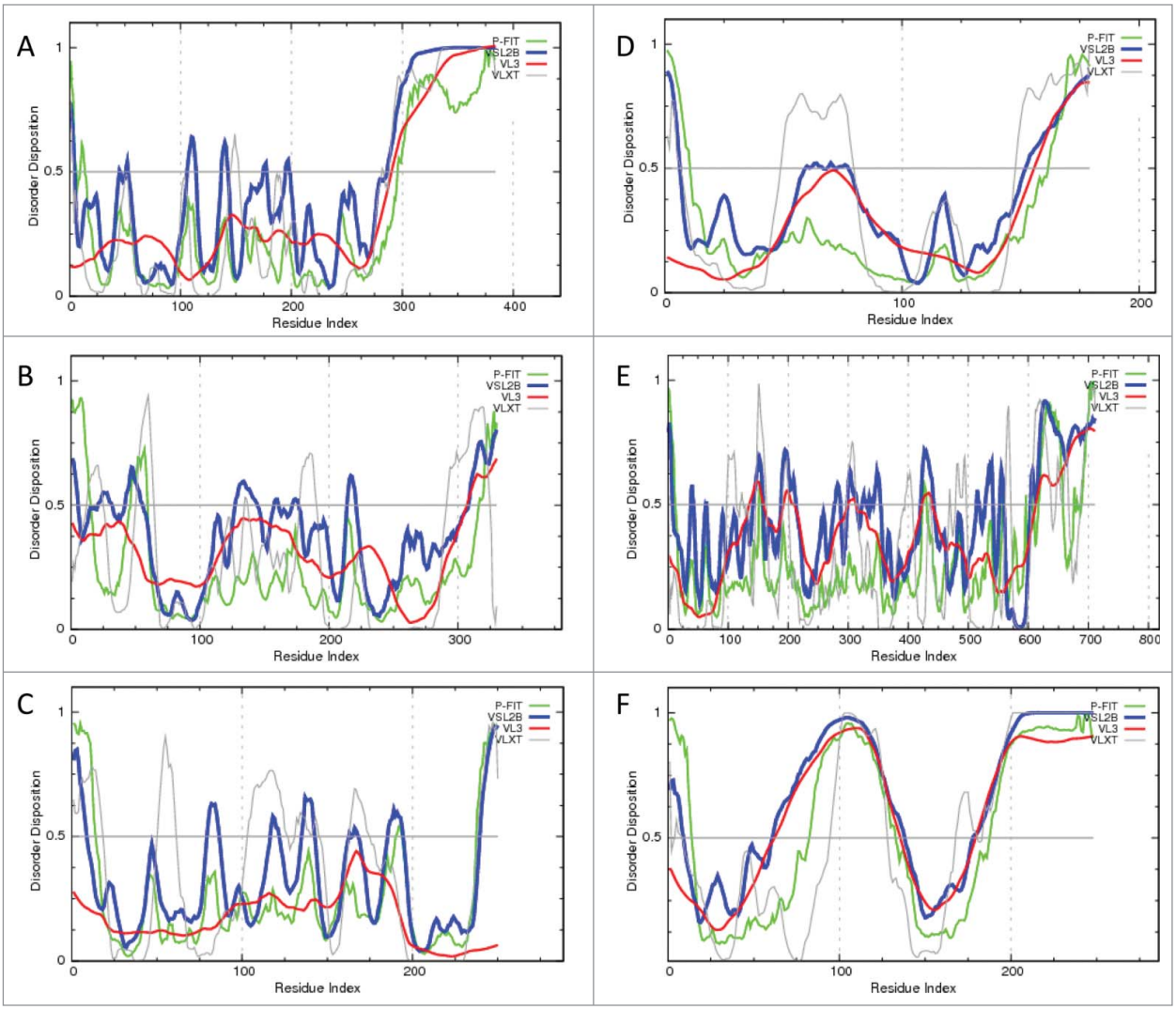

Figure 8. Evaluation of intrinsic disorder in Human aryl-hydrocarbon-interacting protein-like 1 (**A**, AIPL1, UniProt ID: Q9NZN9); Human AIP (**B**, UniProt ID: O00170); Human APC (**C**, UniProt ID: P04070, residues 16–266); *E.coli* DnaT protein (**D**, UniProt ID: P0A8J2); Human PECAM-1 (**E**, UniProt ID: P16284); Mouse Serine/arginine-rich splicing factor 1 (**F**, UniProt ID: Q6PDM2). Intrinsic disorder propensities are evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

isomerase (PPIase) domain (which is inactive in both proteins) followed by a tetratricopeptide repeat (TPR) domain. In addition, AIPL1 harbors a unique C-terminal proline-rich domain (PRD).[84] The authors showed that AIP is inactive as a chaperone, whereas AIPL1 exhibits chaperone activity and prevents the aggregation of non-native proteins, suggesting that PRD is crucial for the chaperone function of this protein providing a means for efficient binding of AIPL1 to non-native proteins.[84] Since AIPL1 possesses decreased affinity to Hsp90, the C-terminal PRD plays a role of a negative regulator of the AIPL1-Hsp90 interaction.[84]

Figure 8A shows that the major portion of the human AIPL1 (N-terminal 75%) containing FKBP-like prolyl peptidyl isomerase and TPR domains is predicted to be mostly ordered (UniProt ID: Q9NZN9), whereas the C-terminal domain is predicted to be highly disordered. On the other hand, **Figure 8B** indicates that the human AIP (UniProt ID: O00170) is expected to be mostly ordered. In agreement with these disorder predictions, 3D structures are known for the PPIase FKBP-type domain of human AIP (residues 2–166, PDB ID: 2LKN) and for its TPR domain (residues 173–330, PDB ID: 4AIF). The intrinsically disordered nature of the C-terminal PRD of the human AIPL1 possessing chaperone activity is in accord with earlier observations that chaperones are often either entirely disordered or contain long disordered regions.[72,85]

Basic residues in the activated protein C (APC) exosite

Takeyama et al. provided a detailed description of the interaction between the activated protein C (APC) and factor (F) VIIIa.[86] The authors showed that the basic residues located within the 39-, 60-, and 70–80-loops of APC constitutes an exosite that contributes to the binding of FVIII and therefore are important for the

subsequent proteolytic inactivation of FVIII.[86] APC, which is also known as vitamin
K-dependent protein C, anticoagulant protein C, autoprothrombin IIA, and blood coagulation factor XIV, is a single chain vitamin K-dependent zymogen for a plasma serine protease that upon activation by the thrombin–thrombomodulin complex down regulates the coagulation cascade by limited proteolysis of FVa and FVIIIa.[87-89] Analysis of the crystal structure of human APC (PDB ID: 1AUT) revealed that there are 3 surface loops (39, 60, and 70–80), rich in basic residues, located in the protease domain of APC near the active site pocket.[86,90] FVIIIa contains acidic C-terminal sequences that may potentially provide interactive sites for the basic exosite of APC.[86] **Figure 8C** represents the results of the computational disorder analysis in human APC (UniProt ID: P04070) and shows that mentioned loops enriched in basic residues are predicted to be disordered or very flexible, thereby providing an interesting mechanistic plane for the molecular basis of APC recognition and binding of FVIII.

The *Escherichia coli* primosomal DnaT protein

Molecular mechanisms of the oligomerization of *E.coli* primosomal DnaT protein are uncovered in the study by Szymanski et al.[91] The authors showed that the removal of the short C-terminal region dramatically affects the oligomerization process and instead of the trimer, the isolated N-terminal domain of DnaT forms a dimer.[91]

Priming of the DNA strand during the replication process is catalyzed by a multiprotein-DNA complex known as the primosome.[92,93] The translocation of this complex along the DNA is fueled by NTP hydrolysis. Besides being responsible for the synthesis of short oligoribonucleotide primers used to initiate synthesis of the cDNA strand, primosome plays a role in the restarting of the stalled replication fork at the damaged DNA sites.[93,94] The assembly of the primosome is driven by an essential replication protein in *Escherichia coli*, the DnaT protein, where, the primosome assembly is initiated by recognition of a specific primosome assembly site (PAS) of the replicating DNA or the damaged DNA site by the PriA protein, or the PriB protein–PriA complex, followed by the association of the DnaT and the PriC protein.[92-94] This primary DNA-protein complex is a scaffold, specifically recognized by the DnaB helicase–DnaC protein complex, which results in formation of the preprimosome. Next, the preprimosome is recognized by the primase, and a functional primosome is formed. The DnaT protein is crucial for the specific entry of the DnaB helicase into the primosome complex. [92-94]

The DnaT monomer consists of the large, N-terminal core domain that includes the first 161 residues of the protein and a small C-terminal region containing the 18 remaining amino acids.[91] Disorder analysis of this protein (UniProt ID: P0A8J2) revealed that although the majority of the DnaT is predicted to be ordered, the crucial for trimerization C-terminal tail is expected to be completely disordered (see **Fig. 8D**). This observation is in agreement with recent notion that disordered protein tails are commonly involved in a wide array of important functions.[95]

Platelet endothelial cell adhesion molecule 1 (PECAM-1)

Tourdot et al. dedicated their research to the analysis of the effect of the sequential phosphorylation of immunoreceptor tyrosine-based inhibitory motifs (ITIMs) on function of the platelet endothelial cell adhesion molecule-1 (PECAM-1).[96] PECAM-1 is a dual ITIM-containing receptor of the Ig superfamily that is capable of inhibiting immunoreceptor tyrosine-based activation motif (ITAM)-induced activation of B cells, T cells, and mast cells, as well as GPVI/FcRγ chain-mediated platelet activation.[97] The inhibitory properties of PECAM-1 require phosphorylation of both of its ITIMs located in the vicinity of tyrosine residues at positions 663 ($VQY_{663}TEV$) and 686 ($TVY_{686}SEV$).[96] Curiously, the ITIMs of PECAM-1 are not phosphorylated in resting cells but are phosphorylated upon cellular activation. Furthermore, these PECAM-1 ITIMs are phosphorylated sequentially, with phosphorylation of Y_{663} depending on prior phosphorylation of Y_{686}.[98] This sequential process of the PECAM-1 phosphorylation starts with the phosphorylation of the C-terminal ITIM by Src family kinases, which enables phosphorylation of the N-terminal ITIM of PECAM-1 by other Src homology 2 domain-containing nonreceptor tyrosine kinases (NRTKs).[96] **Figure 8E** shows that although human PECAM-1 (UniProt ID: P16284) is predicted to be mostly ordered, the residues, phosphorylation of which is crucial for its function (Y_{663} and Y_{686}), are located within the highly disordered C-terminal tail.

Mouse serine/arginine-rich splicing factor 1 (SRSF1)

Aubol et al. studied the molecular mechanisms of the phosphorylation of SR splicing factors, which are the proteins possessing regions enriched in Arg–Ser dipeptide repeats known as RS domains, by a specific serine kinase SRPK1.[99] Phosphorylation of RS domains is crucial for the regulation of the activities of the SR splicing factors. RS domains of SR proteins can range from 50 to over 300 residues in length, and the Arg–Ser dipeptide repeats can vary in both length and position. SRPK1 phosphorylates the prototype SR protein, serine/arginine-rich splicing factor 1 (SRSF1), via a directional mechanism where 11 serines flanked by arginines are sequentially fed from a docking groove in the large lobe of the kinase domain to the active site.[99] **Figure 8F** shows that SRSF1 (UniProt ID: Q6PDM2) is predicted to be a highly disordered protein. The disordered nature of this splicing factor is in agreement with recent bioinformatics studies, which showed that human[100] and yeast spliceosomes[101] are enriched in intrinsic disorder, and abundant disordered regions are crucial for functions of numerous auxiliary protein factors involved in the formation of this ribonucleoprotein complex.[100,101] The potential functionality of the disordered regions found in mouse SRSF1 was further supported by ANCHOR-based analysis,[15,16] which revealed that this protein contains 6 AiBSs, residues 34–43, 75–87, 123–129, 144–177, 186–197, and 221–245.

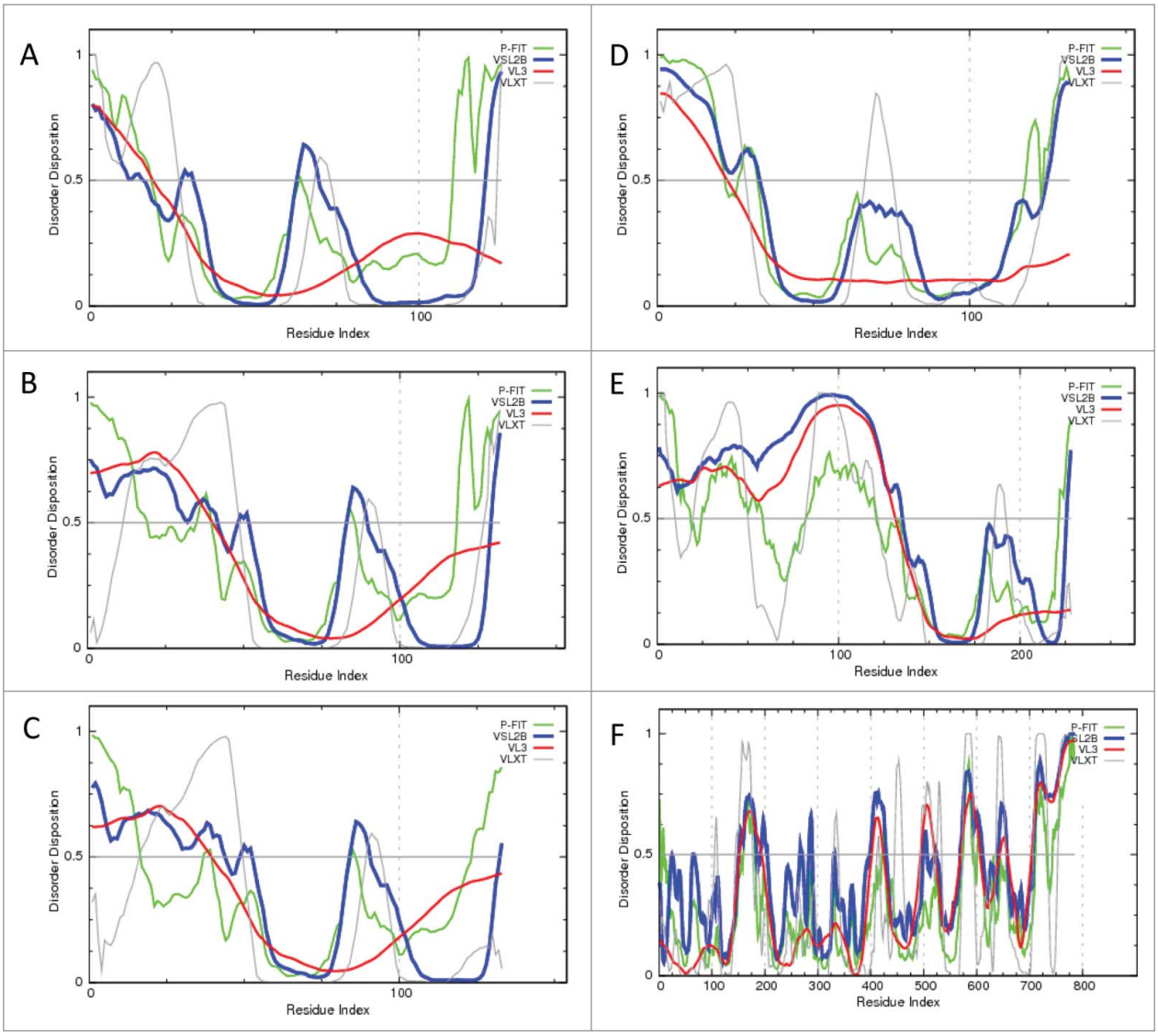

Figure 9. Intrinsic disorder in: (**A**) Human IFITM1 (UniProt ID: P13164); (**B**) Human IFITM2 (UniProt ID: Q01629); (**C**) Human IFITM3 (UniProt ID: Q01628); (**D**) Human IFITM5 (UniProt ID: A6NNB3); (**E**) Human IFITM10 (UniProt ID: A6NMD0); (**F**) Human semaphorin-3 (UniProt ID: Q13275). Intrinsic disorder propensities are evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

The members of human interferon-inducible transmembrane protein (IFITM) family

The roles of human interferon-inducible transmembrane proteins (IFITMs) in antagonizing viral infection were reviewed by Perreira et al.[102] IFITMs are representatives of a diverse group of host restriction factors fighting viral replication at different steps of the viral life cycle. The expression of many restriction factors is transcriptionally controlled by the antiviral cytokine, interferon. The interferon-inducible transmembrane proteins (IFITM) 1, 2, and 3 that inhibit the replication of multiple pathogenic viruses are among the products of such interferon-stimulated genes.[103-105] Among the variety of pathogenic viruses inhibited by IFITMs are influenza A virus (IAV) and influenza B virus, West Nile virus, dengue virus (DENV), severe acute respiratory syndrome coronavirus (SARS CoV), hepatitis C virus (HCV), Ebola virus (EBOV), Marburg virus (MARV), and multitude of the filoviruses, the flaviviruses, the bunyaviruses, and the rhabdoviruses.[102]

Replication of viruses requires breaching the plasma membrane of the host cell. The entry of the enveloped viruses relies on the specialized fusion proteins. There are 3 major classes of viral fusion proteins, all containing a fusion peptide that inserts into the cytolemma, thereby anchoring the 2 membranes side by side.[106] Human IFITMs block the entry of viruses from each of the 3 classes of viral fusion proteins,[102] by preventing viral-host membrane fusion subsequent to viral binding and endocytosis.[104,107] In humans, there are 5 IFITMs (IFITM1, 2, 3, 5, and 10; corresponding UniProt IDs: P13164, Q01629, Q01628, A6NNB3, and A6NMD0). These proteins each contain 2 hydrophobic membrane-associated

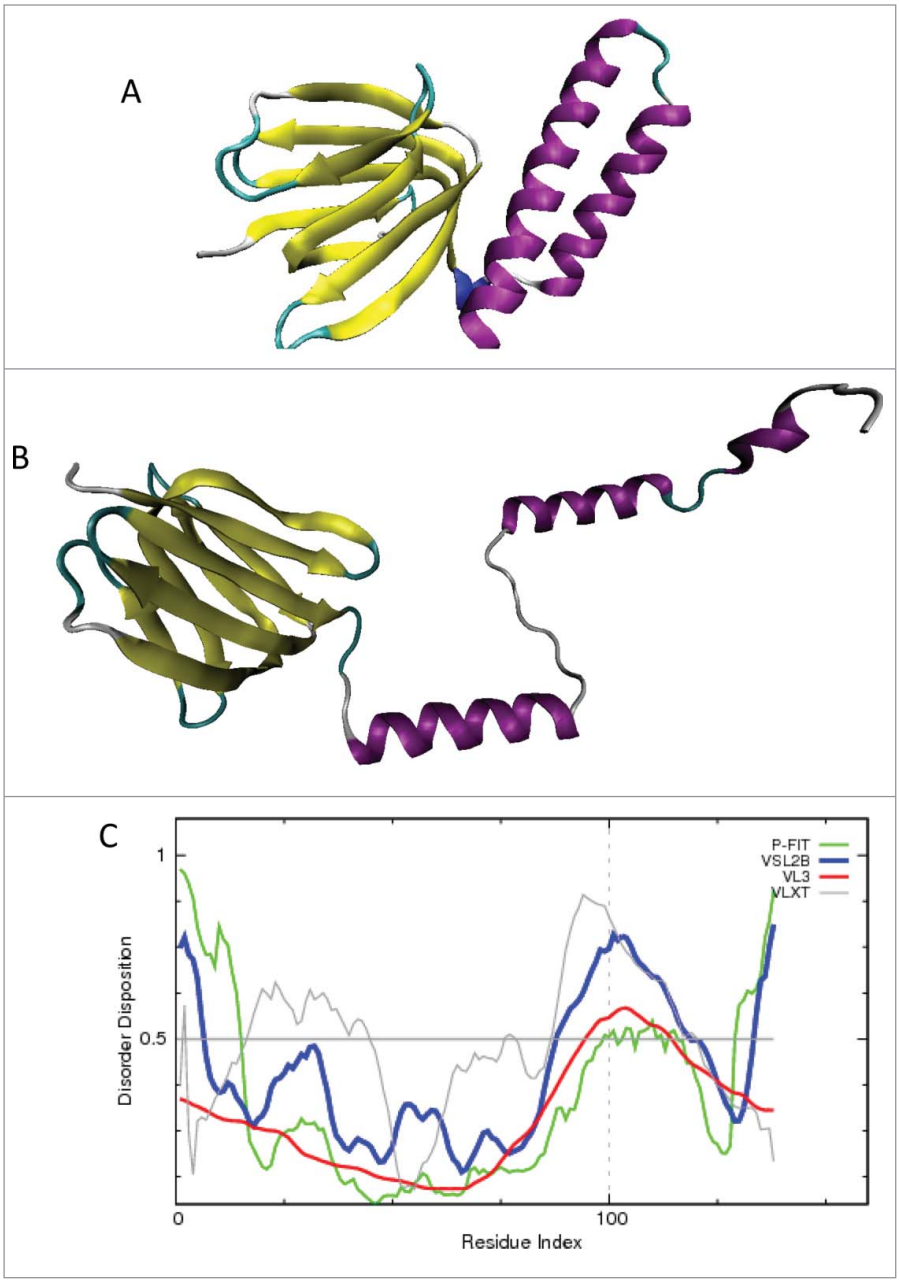

Figure 10. (**A**) Crystal structure of the isolated ε subunit of the F_0F_1 ATP synthase from the thermophile *Bacillus* PS3 (Tε) in a compact structure in the non-bound form (PDB ID: 1AQT). (**B**) Crystal structure of the Tε subunit in a highly extended open form in a complex with other subunits of the bacterial F_1 complex (PDB ID: 3OAA). (**C**). Disorder propensity of the Tε (UniProt ID: P0A6E6) evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

domains separated by a conserved intracellular loop. All of them also contain the cytosolically-located N-terminal domain (NTD) of various length.[102] **Figure 9** shows that the NTDs of all 5 human IFITMs are predicted to be intrinsically disordered. The likely functional role of these disordered tails is reflected in the fact that the longest NTD of IFITM10 contains 3 AiBSs, residues 8–29, 51–64, and 71–82. Shorter tails of other human IFITMs might also be involved in protein-protein interactions since all of them contain characteristic "dips," which are commonly associated with the existence of specific molecular recognition features, MoRFs, which are short folding-prone fragments of disordered regions that are able to fold (at least partially) at interaction with specific binding partners.[108-112]

Chameleon region undergoing α-helix to β-strand transition in the mutant form of the type 2 ryanodine receptor domain A (RyR2A) linked with a heritable cardiomyopathy

Using the solution NMR spectroscopy and X-ray crystallography, Amador et al. studied the effects of mutations associated with arrhythmogenic right ventricular dysplasia type 2 (ARVD2) on the structure and dynamics of the N-terminal domain of the mouse ryanodine receptor RyR2A (UniProt ID: E9Q401).[113] Ryanodine receptors (RyRs) are the largest tetrameric Ca^{2+} release channels (~2.2 MDa) of the sarcoplasmic reticulum in the electrically excitable cells. Here, RyRs resides in close proximity to plasma membrane voltage-gated dihydropyridine receptors (DHPRs) and respond to the DHPR activity through a direct interaction and/or indirect Ca^{2+} sensitivity, propagating sarcoplasmic reticulum luminal Ca^{2+} release, thereby playing an integral role in excitation–contraction coupling in skeletal muscle cells and cardiomyocytes.[114,115] Mutations in the *RyR* genes are associated with several heritable human diseases, with most disease-associated mutations making RyRs hypersensitive to activating stimuli such as Ca^{2+} on either the luminal or the cytosolic side of the receptor. In one of such mutations, the removal of exon 3 in *RyR2* (RyR2AΔ3) results in a 35-residue deletion (N57-G91) that does not destabilize the protein.[116,117] Earlier structural analysis revealed that the β-trefoil fold of this domain is rescued by the transformation of the part of the preceding region of unknown structure (residues 88–109) into the missing β-strand leading to the noticeable increase in the thermal stability of the RyR2AΔ3 domain.[117] Using solution NMR analysis, Amador et al. showed that this "part of the preceding region of unknown structure"

corresponds to a previously unresolved α_2-helix (residues 95–104) in the wild-type RyR2A.[113] Curiously, this α-helix undergoes structural transformation to a β-strand, thereby rescuing the β-strand deleted in the RyR2AΔ3 due to the removal of 3 in *RyR2* gene. Therefore, the authors revealed that this newly discovered α_2-helix corresponds to the region of missing electron density in the earlier crystal structures of the protein, is rather dynamic, exhibits a greater backbone RMSD compared to the remaining secondary structure elements, and is able to undergo an α-to-β transition after the deletion of residues N57-G91 in RyR2AΔ3.[113] In other words, applying NMR to look at the crystallographically invisible region, an important chameleon region was discovered. Although this chameleon region is relatively small, it is really mighty, with at least some mightiness being attributed to its intrinsically disordered nature.

ATP-induced dimerization of the F_0F_1 ε subunit from Bacillus PS3

Rodriguez et al. used complementary biophysical techniques for examining the ATP-induced conformational switching of the isolated ε subunit of the F_0F_1 ATP synthase from the thermophile *Bacillus* PS3 (Tε).[118] The authors showed that ATP binding induces large-scale conformational transition consistent with formation of stable helical structure in the isolated Tε.[118] Based on the complimentary hydrogen/deuterium exchange (HDX) mass spectrometry analysis it has been concluded that this transition is accompanied by a pronounced stabilization in the vicinity of the ATP-binding pocket.[118]

Structurally, the ε subunit consists of a β-sandwich and 2 C-terminal helices, α1 and α2.[119] Curiously, this protein is able to undergo dramatic structural rearrangements from a compact structure in the non-bound form (see **Fig. 10A**; PDB ID: 1AQT)[119] to a highly extended open form in a complex with other subunits of the bacterial F_1 complex (see **Fig. 10B**; PDB ID: 3OAA). The ability of the ε subunit to undergo large conformational change associated with binding to the F_1 complex is determined by the presence of long IDPR at the C-terminal region (see **Fig. 10C**; UniProt ID: P0A6E6).

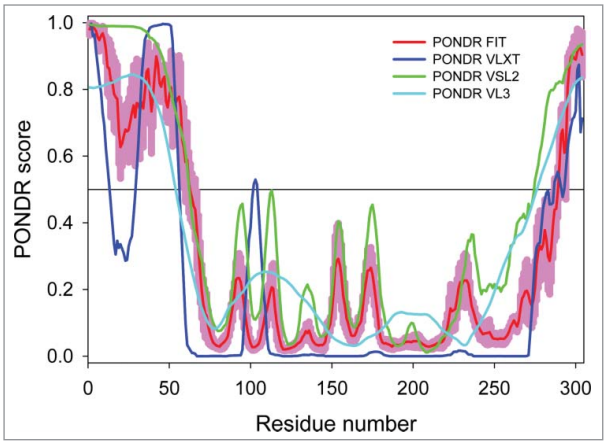

Figure 11. Intrinsic disorder in the CaPcl5 evaluated by PONDR® FIT (red lines), PONDR® VLXT (blue line), PONDR® VSL2B (green line), and PONDR® VL3 (cyan line). Light pink shadow around the corresponding PONDR® FIT curve represents distribution of statistical errors.

Functional implications of the cyclin CaPcl5 phosphorylation

Simon et al. investigated functional peculiarities of the *Candida albicans* cyclin CaPcl5.[120] The authors showed that the activation of the cyclin-dependent kinase Pho85 that induces phosphorylation of the transcription factor CaGcn4 is controlled via the specific binding of CaPcl5 to Pho85. Furthermore, it is shown that CaPcl5 induces its own phosphorylation at 2 adjacent sites in the N-terminal region of the protein and that this phosphorylation causes degradation of the cyclin *in vivo*, whereas *in vitro* studies indicated that this phosphorylation was accompanied by a loss of specific substrate recognition thereby representing a novel mechanism for limiting cyclin activity.[120]

Activity of cyclin-dependent kinases (CDKs) is regulated via binding of ancillary subunits, the cyclins,[121] which also participate in targeting the kinase to specific substrates.[122] Similar to other members of the CDK family, a yeast (*Saccharomyces cerevisiae*) CDK Pho85 can bind up to 10 different Pho85 cyclins (Pcls),[123] which defines the ability of this CDK to phosphorylate different substrates leading to a wide range of Pho85 functions in cellular regulation. Among various regulatory functions of Pho85 complexed with different cyclins, is the Pho85/Pcl5-driven phosphorylation of the bZIP transcription factor Gcn4 leading to its degradation.[124] Curiously, it was shown that cyclin Pcl5 coevolved with its substrate Gcn4, and in the pathogenic yeast *Candida albicans*, degradation of CaGcn4 depends on CaPcl5.[125]

Mutagenic analysis of the CaPcl5 (UniProt ID: Q5AK08) revealed 2 potential CDK phosphorylation sites located in the N-terminal segment, at positions 38 and 43.[120] **Figure 11** shows that both of these phosphorylated sites are located within the disordered region, providing further illustration for an important notion stating that sites of posttranslational modifications in proteins are commonly located within the disordered regions.[126]

Nuclear localization of the alternatively spliced sirtuin-2 isoform

Rack et al. described a discovery of a new alternatively spliced isoform of the human sirtuin-2.[127] The new isoform results from skipping of exons 2–4 in the *SIRT2* gene.[127]

Sirtuin-2 is a member of a large family of protein-modifying enzymes, NAD^+-dependent deacetylases, which are highly conserved throughout bacteria, archaea, and eukaryotes, and are heavily involved in modulation of various cellular processes ranging from metabolic homeostasis, to cell cycle control, to development, and to chromatin organization.[128-133] In their turn, the numerous activities of sirtuins

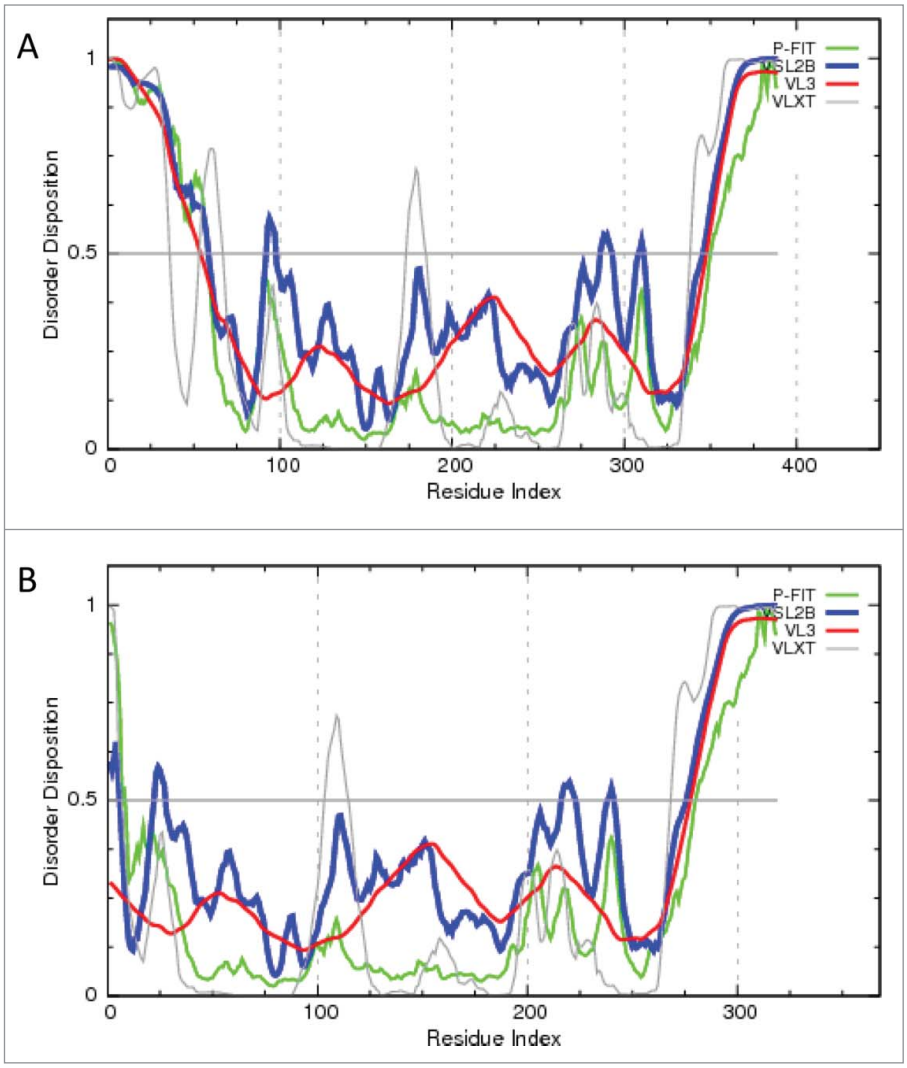

Figure 12. Intrinsic disorder propensities of the isoforms-1 (**A**) and the alternatively spliced isoform-5 (**B**) of human siruin-2 evaluated by PONDR® FIT, PONDR® VSL2B, PONDR® VL3, and PONDR® VLXT.

are regulated by alternative splicing, modulation of expression levels, posttranslational modifications or subcellular compartmentalization.[134-137] In humans, there are 7 sirtuin homologues with distinct cellular locations. For example, Sirt1, Sirt6 and Sirt7 are primarily localized to the nucleus, whereas Sirt3, Sirt4 and Sirt5 are mitochondrial proteins, and Sirt2 is believed to have a predominant cytosolic localization.[138]

Earlier, 4 variants of human *SIRT2* gene were reported, with alternative splicing affecting the N-terminal part of the sirtuin-2 protein in 2 isoforms (residues 1–37 are missing in the isoform-2 and the residues 1–38 (MAEPDPSHPLETQAGKVQEAQDSDSDSEGGAAGGEADM) are substituted by residues MPLAECPSCRCLSSFRSV in the isoform-3), and with the isoform-4 possessing changes at the C-terminus, where residues 266–271 (VQPFAS) are substituted by GRGLAG, and residues 272–389 are missing. These isoforms contain a leucine-rich nuclear export signal (NES) within their N-terminal region which mediate their cytosolic localization[139] where they have numerous functions. A recent finding a new isoform-5 with the predominant nuclear localization helped to resolve an apparent contradiction between the predominant cytosolic localization of this protein and the existence of multiple nuclear functions.[127] The newly discovered alternative splicing event results in the substitution of the codons for amino acids 6–76 of the full-length ORF of isoform-1 by an arginine codon in a new isoform. This isoform-5 is predominantly localized in the nucleus, does not exhibit detectable deacetylase activity, but is properly folded and retains the ability to bind p300.[127]

It has been established that the protein segments affected by alternative splicing are most often intrinsically disordered, suggesting that alternative splicing enables functional and regulatory diversity while avoiding structural complications associated with the deletion of portions of the well-folded proteins.[140] In line with these observations, later studies revealed that the tissue-specific protein segments produced by alternative splicing often contain disordered regions, are enriched in post-translational modification sites, and frequently embed conserved binding motifs.[141] Also, it was proposed that alternative splicing of intrinsically disordered regions containing linear interaction motifs and/or post-translational modification sites results in complete rewiring of protein interactions.[142] **Figure 12** shows that in agreement with these earlier observations the alternatively spliced isoform-5 of human sirtuin-2 differs from the canonical isoform (UniProt ID: Q8IXJ6) by lacking the intrinsically disordered N-terminal tail (cf. **Figs. 12A and 12B**). Curiously, the ANCHOR analysis[15,16] of the human sirtuin-2 revealed that this alternative splicing event removed 2 AiBSs (located at the residues 1–15 and 36–49 of the isoform-1). Therefore, functionality of siruin-2 is modulated by alternative splicing that removes potential binding sites from the disordered region of this protein.

Role of the C-terminal tail in competitive binding of semaphorin-3 to neuropilin-1

Parker et al. investigated the role of the amino acid sequence at the C-terminal region of human semaphorin-3 in interaction of this protein with neuropilin-1.[143] The neuropilin-1 is a member of the family of type I transmembrane receptors that are essential in development, homeostasis, and pathogenesis being involved in the coordination of several important

signaling events in the cardiovascular and nervous system. The members of the class III Semaphorin (Sema3) family of axon guidance molecules are among the ligands neuropilins (Nrps). Ligand binding of Nrps is a subject of the intensive research due to the involvement of these transmembrane receptors in various pathogenic conditions. It has been established that all known Nrp ligands require a C-terminal arginine (C_R) for binding to a conserved pocket in the Nrp b1 domain.[144-147] To better understand the mechanism of interaction between the Nrps and their ligands and to understand the role of residues upstream of the C_R, Parker et al. synthesized a library of semaphorin-3 derived peptides. This study revealed that the C-1 residue (i.e., residue preceding the C_R) serves the critical role of positioning the C_R and C-2 residues to promote concurrent Nrp binding.[143] Bioinformatics analysis revealed that the analyzed sequence (WDQKKPRNRR) is a part of long intrinsically disordered tail that spans over the last 100 residues of human semaphorin-3 (UniProt ID: Q13275, see **Fig. 9F**).

Structure of the unmodified 3 Glu-OCN form of bovine osteocalcin

Malashkevich et al. reported the X-ray crystal structure of the unmodified 3 Glu-OCN form of bovine osteocalcin.[148] This is an important contribution since it adds a crucial structural information on osteocalcin, which is a small, abundant noncollagenous protein synthesized by osteoblasts, and that typically contains 3 γ-carboxyglutamic acid (3 Gla-OCN) residues generated post-translationally in a vitamin K-dependent process[149,150] by the vitamin K-dependent (VKD) carboxylase.[151] Earlier, based on the circular dichroism study it has been concluded that in the presence of Ca^{2+}, the 3 Glu-OCN molecule contained significantly less α-helical structure than the 3 Gla-OCN form.[152] Furthermore, it was shown that Ca^{2+} no longer binds to decarboxylated, unmodified osteocalcin [153] Malashkevich et al. revealed that the crystal structure of the thermally decarboxylated bovine osteocalcin contained residues 17–47, was rather similar to 3 Gla Ca^{2+}-OCN, consisted of 3 α-helices surrounding a hydrophobic core, and contained C23–C29 disulfide bond between 2 of the helices but did not contain bound Ca^{2+}.[148] This is an interesting finding since earlier work clearly showed that in their apo-forms, modified and non-modified osteocalcin from different sources are mostly disordered.[148] The explanation for this apparent contradiction can be derived from the analysis of conditions used for the 3 Glu-OCN crystallization, where protein in 20 mM NaCl and 10 mM $CaCl_2$ (pH 7.0) was mixed with the reservoir solution containing 2.5 M ammonium sulfate and 0.1 M Bis-Tris propane (pH 7.0) in 1:1 ratio.[148] Therefore, high ionic strength solution was used to partially neutralize the high anionic charge in the helical regions of osteocalcin. Although this approach permitted better realization of the α-helix-forming potential of the non-modified osteocalcin, the crystallization conditions are clearly non-physiological.

Coordinated transcriptional regulation of the Hspa1a gene by multiple transcription factors

Sasi et al. investigated the peculiarities of the regulation of expression of the inducible heat shock protein HSPA1A at the transcription level by several transcription factors.[154] HSPA1A, together with HSPA1B are the members of the human inducible heat shock protein family (Hsp70), altered expression of which has been attributed to the pathogenesis of various diseases, such as cancer, cardiovascular diseases, and neurodegenerative disorders.[155-158] It is known that the heat shock response-related proteins (including Hsp70) can be expressed during normal conditions (e.g., during the cell growth and development) or can be induced by various pathological conditions, such as infection, inflammation, and protein conformation diseases. Sasi et al. show that transcription factors HSF-1, CREB, NF-Y, and NF-kB synergistically regulate expression of the heat-shock-induced gene *Hspa1a*.[154]

The initiation of the heat shock response is manifested by the activation of the heat shock transcription factors (HSFs), a family of related transcription factors which, in mammals, is composed of HSF-1, -2, -3, -4, -5, -Y and -X.[159] cAMP response element binding protein (CREB) is the cAMP-regulated transcription factor that has been shown to stimulate target gene expression often via the associating with the coactivator paralogs P300 and CREB binding protein (CBP).[160-162] Complex formation between CREB and CBP/P300 requires protein kinase A (PKA)-mediated phosphorylation of CREB at Ser-133.[160] The nuclear transcription factor Y (NF-Y) is a sequence-specific DNA-binding protein that recognizes the Y box, which is a promoter element common to all major histocompatibility complex class II genes.[163] Also, NF-Y is known to interact with a CCAAT box, which is one of the most common elements in eukaryotic promoters, found in the forward or reverse orientation.[164] Furthermore, NF-Y has been reported to play a key role in the basal expression of many Hsps through this CCAAT box.[165,166] NF-Y is heterotrimeric transcription factor composed of 3 components, NF-YA (subunit α, UniProt ID: P23511), NF-YB (subunit β, UniProt ID: P25208), and NF-YC (subunit γ, UniProt ID: Q13952), where the dimerization of NF-YB and NF-YC is a prerequisite for the NF-YA association and DNA binding.[167] Finally, the transcription factor nuclear factor-kappa B (NF-κB; UniProt ID: 19838) is a key player in an intracellular signaling cascade regulating many inflammatory mediators.[168] NF-κB is a common transcription factor present in almost all cell types and is the endpoint of a series of signal transduction events that are initiated by a vast array of stimuli related to many biological processes such as inflammation, immunity, differentiation, cell growth, tumorigenesis and apoptosis. NF-κB is a homo- or heterodimeric complex formed by the Rel-like domain-containing proteins RELA/p65, RELB, NF-κB1/p105, NF-κB1/p50, REL and NF-κB2/p52, and the heterodimeric p65-p50 complex.

Figure 13 represents the results of the multiparametric evaluation of intrinsic disorder propensities of HSF-1 (**A**, UniProt ID: Q00613), CREB (**B**, UniProt ID: P16220), NF-YA (**C**, UniProt ID: P23511), NF-YB (**D**, UniProt ID: P25208), NF-YC (**E**, UniProt ID: Q13952), and NF-kB (**F**, UniProt ID: P19838) and shows that these proteins are predicted to be highly disordered. The high levels of functional intrinsic disorder is a "family signature" of the transcription factors in general.[169] Furthermore, the recent

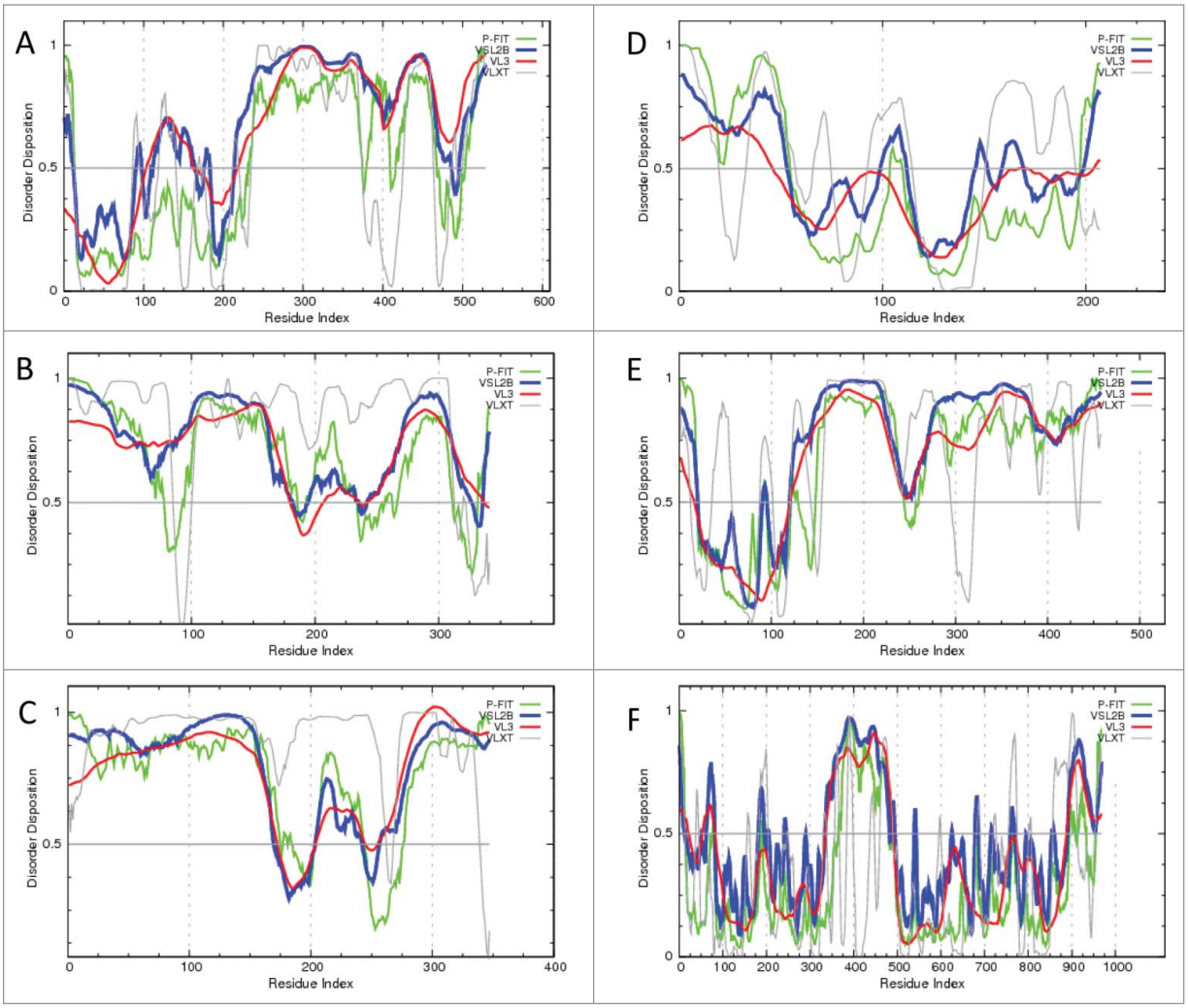

Figure 13. The multiparametric evaluation of intrinsic disorder propensities of HSF-1 (**A**, UniProt ID: Q00613), CREB (**B**, UniProt ID: P16220), NF-YA (**C**, UniProt ID: P23511), NF-YB (**D**, UniProt ID: P25208), NF-YC (**E**, UniProt ID: Q13952), and NF-kB (**F**, UniProt ID: P19838). Intrinsic disorder propensities are evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

analysis of the abundance and importance of IDPRs in functions of various HSFs clearly indicated that the heat shock response requires HSF flexibility to be more efficient.[170] Also, the importance of intrinsic disorder in functions of various proteins involved in the innate immune response (including NF-κB) has been recently analyzed.[171]

Phosphorylation sites in melanopsin

Blasic et al. analyzed the peculiarities of phosphorylation of the C-terminal domain of the photopigment melanopsin possessing 37 serine and threonine sites that are potential sites for phosphorylation by a G-protein dependent kinase (GRK).[172] Melanopsin is a member of the G protein coupled receptor (GPCR) family that undergoes light-dependent phosphorylation that is involved in deactivation of the photoresponse. Blasic et al. revealed that of the 37 phosphorylation sites, a small cluster of 6 or 7 sites in the proximal region of the C-tail is critical for mediating deactivation.[172] **Figure 14A** shows that mouse melanopsin (UniProt ID: Q9QXZ9) is predicted to have long disordered C-tail (residues 380–521). This is a rather expected output since phosphorylation sites are known to be preferentially located within the IDPRs.[173-175]

Human defensins

A recent review of Wisons et al. is dedicated to the antiviral activities of human defensins.[176] These peptides with a wide range of antimicrobial activities are important components of the innate immune system. Among various antiviral mechanisms of human defensins are direct targeting of viral envelopes, glycoproteins, and capsids, inhibition of viral fusion and post-entry neutralization, inhibition of viral replication via disruption of host intracellular signaling and binding and modulation of host cell surface receptors, and augmentation and altering of adaptive immune responses.[176] Of the 2 types of

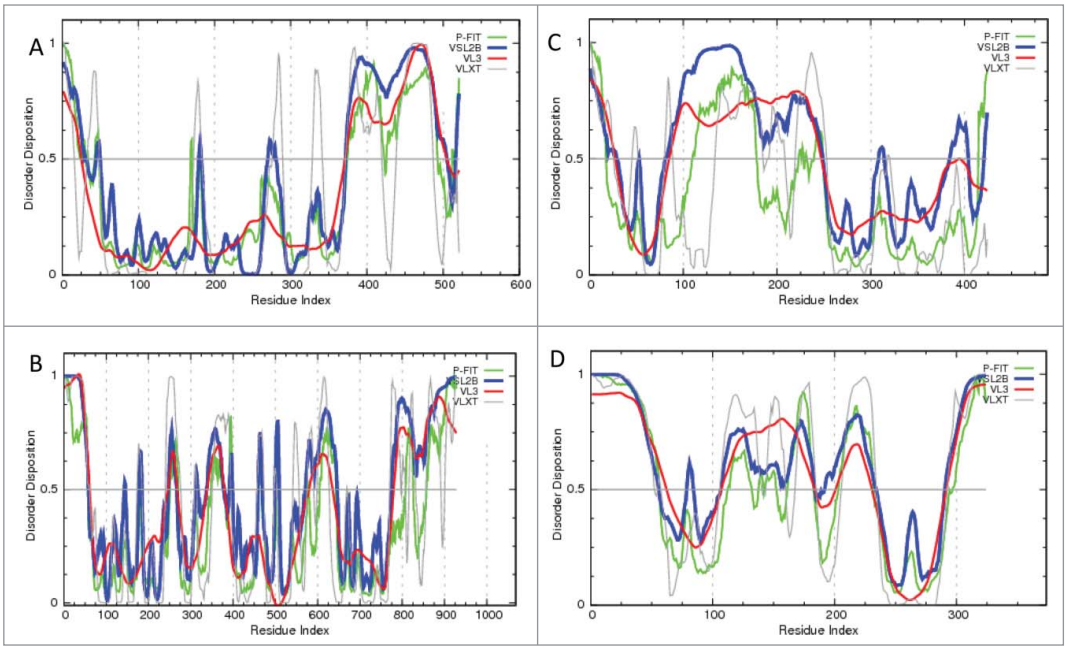

Figure 14. Intrinsic disorder propensities of mouse melanopsin (**A**. UniProt ID: Q9QXZ9); human retinoblastoma-associated protein (**B**, UniProt ID: P06400); human AMSH (**C**, UniProt ID: O95630); mouse transcriptional activator protein Purβ (**D**, UniProt ID: O35295). Intrinsic disorder propensities are evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

defensins, α- and β-defensins, found in human, significant knowledge is accumulated about the molecular mechanisms of the α-defensin action, whereas molecular details on the β-defensin activities are more disperse.[176]

Defensins are small (~29–129 amino acids) cationic, amphipathic polypeptides with a predominantly β-sheet structure stabilized by 3 disulfide bonds.[176] Human α-defensins are typically shorter than β-defensins, with α-defensins staying in a range of 29–39 residues and some β-defensins can being as long as 129 residues. Search of UniProt for human defensins produces about 40 hits, of which 6 entries were α-defensins. Defensins are produced in a form of proprotein containing a signal peptide and/or a propeptide. **Figure 15** represents disorder profile for several representative members of the human defensin family and shows that all proproteins possess noticeable amount of intrinsic disorder, the content of which varies significantly between different defensins. It is likely that the presence of noticeable disordered tails plays a role in the maturation of defensins since sites of the proteolytic attack are commonly located within regions of intrinsic disorder.[177] It is also possible that these disordered tails serve as entropic bristle domains[178] defining solubility of pro-defensins. In β-defensins, long disordered regions can have other functions. For example, the ANCHOR analysis[15,16] of human β-defensin 129 (UniProt ID: Q9H1M3) revealed that this longest member of the human defensin family (there are 129 amino acids in a mature protein and 183 residues in a pro-protein) contains 4 AiBSs, residues 78–87, 89–100, 117–122, and 149–164.

Inhibition of E2F transcription factor by the phosphorylated C-terminal domain of the C-terminal domain of retinoblastoma protein

Burke et al. investigated the roles of the phosphorylation of the C-terminal domain of the retinoblastoma protein (RbC) in interaction of Rb with the E2F transcription factor and related regulation of the Rb activities in growth suppression.[179]

Deregulation of the broad-functioning tumor suppressor retinoblastoma protein (Rb) is associated with several human cancers.[180,181] Among numerous functions ascribed to this important protein is a negative regulation of cell division at the G_1–S transition of the cell cycle,[182] where Rb forms a growth-repressive complex with E2F transcription factors[183] in a phosphorylation-dependent manner. Burke et al. showed that there are at least 2 different mechanisms by which RbC phosphorylation inhibits E2F binding. Here, phosphorylation of S788 and S795 weakens the direct association between the RbC and the marked-box domains of E2F, whereas phosphorylation of S788, S795, S807 and S811 induces an intramolecular association between RbC and the pocket domain, which overlaps with the site of E2F transactivation domain binding.[179]

Figure 14B shows that human retinoblastoma-associated protein (UniProt ID: P06400) is predicted to have numerous disordered regions, including long disordered tails. The longest disordered region is the C-terminal tail (residues 760–928) that contains all the phosphorylation sites discussed above. According to the ANCHOR analysis,[15,16] disordered tails of Rb are enriched in potential binding sites located at residues 7–20, 41–49, 57–64, 830–840, 844–860, 872–881, and 891–919. N-terminal tail of Rb contains 2 regions of compositional bias (poly-Ala and poly-Pro regions, residues 10–18 and 20–29, respectively). Known sites of the Rb interaction with LIMD1 and E4F1 are located within the intrinsically disordered C-terminal tail (residues 763–928 and 771–928, respectively). The fact that phosphorylation sites of the RbC are located within the disordered region is in agreement with the general trend, where sites of many catalytically-induced post-translational modifications including phosphorylation[173] and ubiquitination[184]

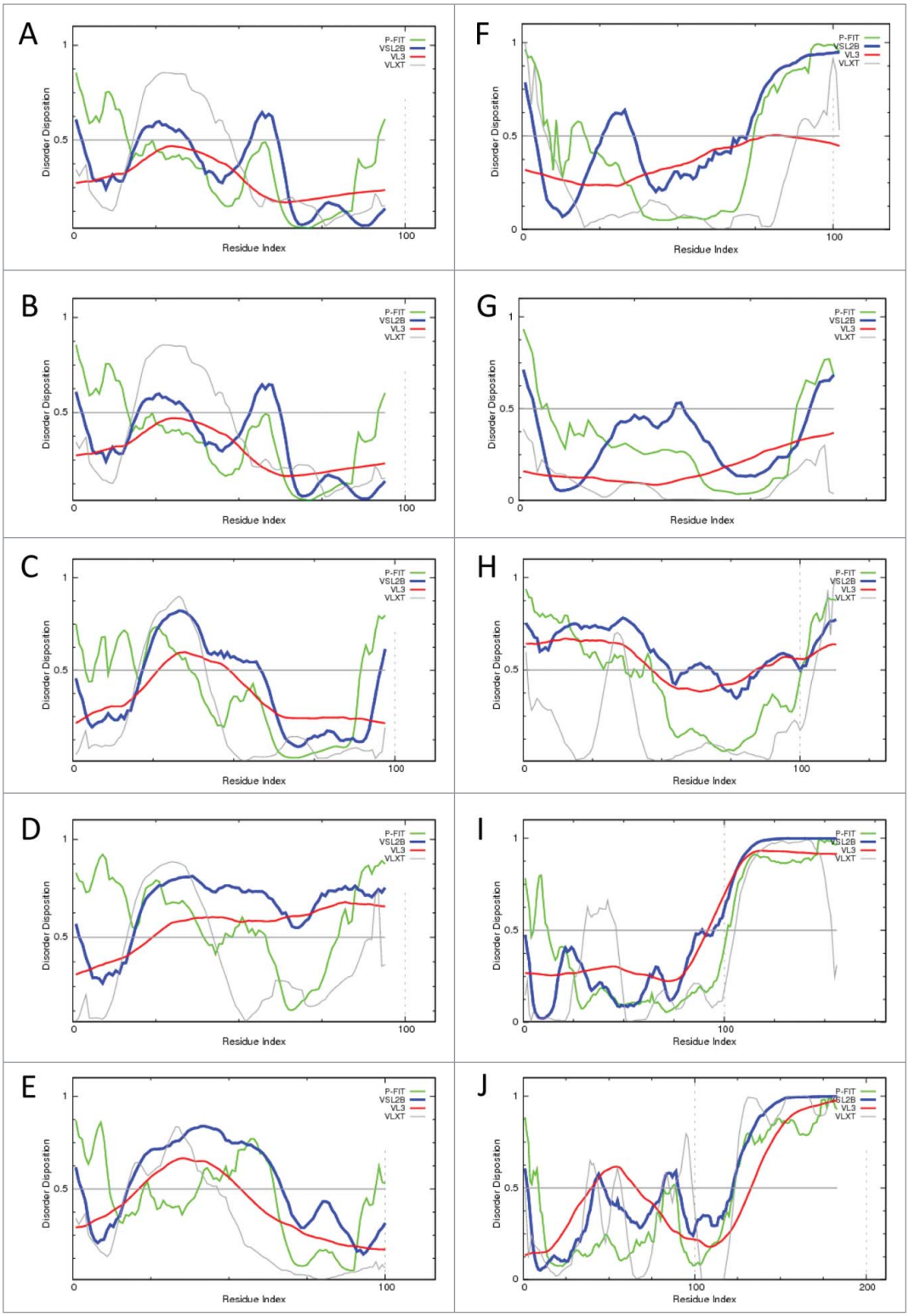

Figure 15. Intrinsic disorder propensity of human defensins evaluated by PONDR® FIT, PONDR® VSL2B, PONDR® VL3, and PONDR® VLXT. (**A**) Neutrophil defensin 1 (P59665); (**B**) Neutrophil defensin 3 (P59666), (**C**) Neutrophil defensin 4 (P12838); (**D**) Defensin-5 (Q01523); (**E**) Defensin-6 (Q01524); (**F**) Beta-defensin 116 (Q30KQ4); (**G**) Beta-defensin 1 (P60022); (**H**) Beta-defensin 112 (Q30KQ8); (**I**) Beta-defensin 125 (Q8N687); (**J**) Beta-defensin 129 (Q9H1M3).

are typically found in regions of intrinsic disorder.[126,174,185-187]

The tetratricopeptide repeat (TPR) motif-containing protein LGN and its binding partner Frmpd1 (FERM and PDZ domain containing 1)

Pan et al. described structural peculiarities of the human LGN-Frmpd1 complex.[188] To this end, a crystal structure of the complex between the 15–350 fragment of human LGN containing 8 tetratricopeptide repeat motifs (UniProt ID: P81274) and the 901–951 fragment of human Frmdp1 (UniProt ID: Q5SYB0) was solved at 2.4 Å resolution.[188] In this LGN-TPR/Frmpd1 complex, almost the entire length of LGN-TPR was well resolved, and most residues of the Frmpd1 fragment were clearly resolved adopting an extended conformation that occupied most of the concave channel formed by the 8 TPR motifs and buried a total of 2541 $Å^2$ surface area.[188] The noticeable exceptions were the 11 residues in the loop connecting the αA and αB of TPR3 (amino acids 153–163) and the 6 residues at the C terminus (amino acids 345–350) in the LGN-TPR, and the N-terminal 8 residues (amino acids 912–919) and the last 2 residues at the C terminus (amino acids 937–938) of the Frmpd1 fragment, all of which were missing in the structure of the LGN-TPR/Frmpd1 complex.[188]

Leu-Gly-Asn repeat-enriched protein LGN/AGS3 in mammals (Pins,

or *p*artner of *ins*cuteable in *Drosophila* neuroblasts) plays a number of crucial roles in regulation of cell polarity and spindle orientation during cellular differentiation and self-renewal in multicellular eukaryotes development.[189-191] Structurally, LGN consist of 8 tetratricopeptide repeat (TPR) motifs in its N-terminal half and 3 or 4 GoLoco motifs (also referred to as G-protein regulatory or GPR motifs) in the C-terminal half.[192-194] Both of these motifs are crucial for protein-protein interactions,[195,196] with each GoLoco motif of Pins/LGN being capable of binding to GDP-bound Gαi[197] leading to stable cortical localization of Pins/LGN,[198] and with the TPR motifs possessing multiple binding partners. The canonical TPR motif is a 34-amino-acid protein–protein interaction module multiple copies of are found in a wide range of proteins with diverse functions such as cell cycle regulations, gene transcription and splicing processes, protein trafficking, and protein folding.[195,199] **Figure 15A** shows that human LGN protein is involved in a wide range of protein-protein interactions, clearly serving as an important hub protein. This interactivity analysis is done by STRING.[59]

Frmpd1 protein (FERM and PDZ domain containing 1) is known to serve as a regulatory binding partner of AGS3, with a short fragment of Frmpd1 (amino acids 901–938) being shown to bind to the 8 TPR motifs of AGS3.[200] Also, a 50-residue fragment (amino acids 901–951) of Frmpd1 was shown to bind to LGN with a K_d~1 μM.[188,194]

Curiously, besides the mentioned above regions of missing electron density in the LGN-TPR/Frmpd1 complex **Figures 15B and 15C** shows that full-length LGN and Frmpd1 proteins both belong to the category of hybrid proteins possessing some ordered domains and significant number of intrinsically disordered protein regions (IDPRs). In fact, the GoLoco motif containing C-terminal half of human LGN (residues 350–684) and the major portion of Frmpd1 (residues 511–1390) are predicted to be mostly disordered by all the computational tools used in this study. Overall, more than 40% of the LGN residues and more than 50% of the Frmpd1 residues are predicted

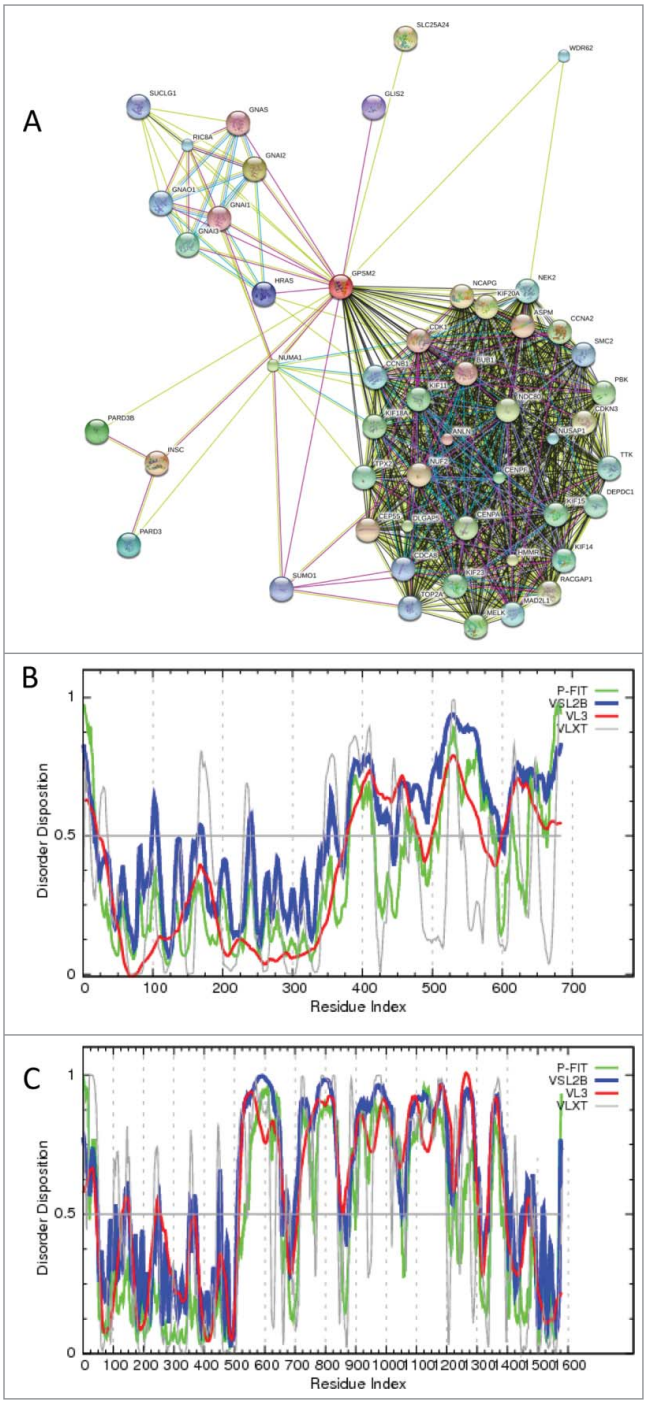

Figure 16. (**A**) Evaluation of the interactivity of the human LGN by STRING, which is the online database resource Search Tool for the Retrieval of Interacting Genes that provides both experimental and predicted interaction information.[59] STRING produces the network of predicted associations for a particular group of proteins. The network nodes are proteins. The edges represent the predicted functional associations. An edge may be drawn with up to 7 differently colored lines - these lines represent the existence of the 7 types of evidence used in predicting the associations. A red line indicates the presence of fusion evidence; a green line - neighborhood evidence; a blue line – co-occurrence evidence; a purple line - experimental evidence; a yellow line – text mining evidence; a light blue line - database evidence; a black line – co-expression evidence.[59] (**B**) Intrinsic disorder propensity of the LGN protein. (**C**) Per-residue disorder distribution in the Frmpd1 protein. Disorder propensity was evaluated by a set of predictors from the PONDR family, PONDR® FIT, VSL2B, VL3, and VLXT. Scores above 0.5 correspond to disordered residues/regions.

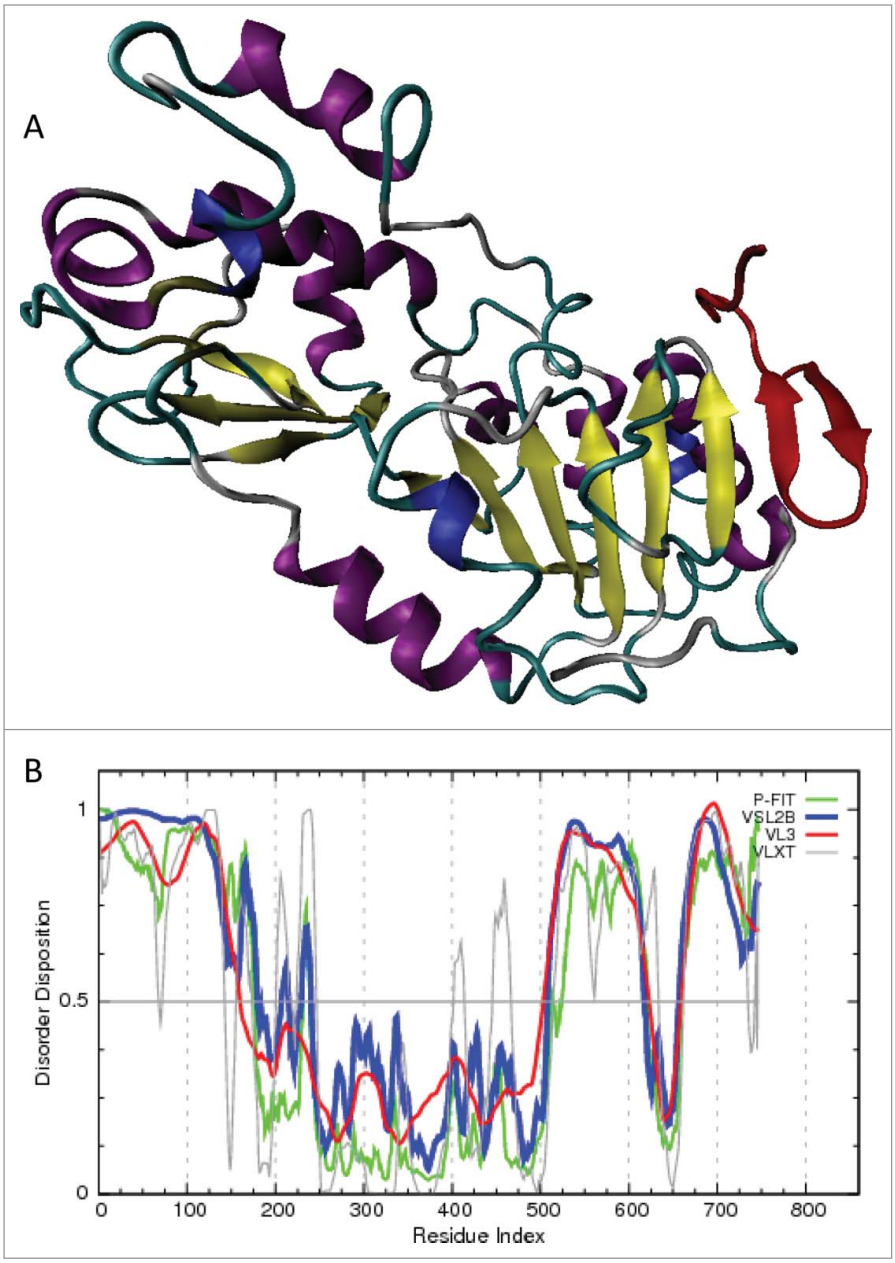

Figure 17. (**A**) Crystal structure of a complex between the catalytic domain of the human sirtuin-1 (residues 234–510) and the C-terminal regulatory segment of this protein (CTR, residues 641–665).[203] In this structure (PDB ID: 4IG9) the catalytic domain is colored based on its secondary structure, whereas CTR is shown as a red β-hairpin. (**B**) Disorder predictions of the human Sirt1 (UniProt ID: Q96EB6) evaluated by PONDR® FIT, PONDR® VSL2B, PONDR® VL3, and PONDR® VLXT.

to be disordered, which clearly places these proteins to the category of highly disordered proteins.

Furthermore, both proteins are expected to have numerous disorder-based interaction sites. These can be identified by ANCHOR,[15,16] a computational tool that relies on the pair-wise energy estimation approach developed for the general disorder prediction method IUPred,[201,202] and is based on the hypothesis that long regions of disorder contain localized potential binding sites that cannot form enough favorable intra-chain interactions to fold on their own, but are likely to gain stabilizing energy by interacting with a globular protein partner.[15,16] Here the term ANCHOR-indicated binding site (AiBS) is used to identify a region of a protein suggested by the ANCHOR algorithm to have significant potential to be a binding site for an appropriate but typically unidentified partner protein. This analysis revealed that there are 10 AiBSs in LGN (residues 370–379, 436–452, 491–499, 508–513, 531–537, 545–553, 566–573, 596–607, 629–638 and 661–671), whereas Frmpd1 contains 23 AiBSs (residues 534–554, 609–617, 622–636, 652–662, 704–715, 747–779, 792–804, 828–878, 896–909, 932–949, 958–967, 983–1019, 1023–1077, 1087–1100, 1106–1114, 1119–1129, 1138–1171, 1192–1220, 1245–1250, 1290–1297, 1312–1319, 1336–1342, and 1424–1431). Of special interest is the fact that the 901–951 fragment of human Frmdp1 used in the mentioned structural analysis[188] is predicted to overlap with 2 AiBSs (residues 896–909 and 932–949). This is a clear indication that intrinsic disorder plays a role in the formation of the LGN-TPR/Frmpd1 complex.

Structural and functional analysis of human sirtuin-1

Davenport et al. reported a crystal structure of human sirtuin-1catalytic domain (residues 234–510) in complex with its C-terminal regulatory segment (CTR, residues 641–665).[203] In this structure (see **Fig. 17A**, PDB ID: 4IG9), the catalytic NAD^+-binding domain adopts the canonical sirtuin fold, whereas CTR forms a β-hairpin structure complementing the β-sheet of the catalytic domain and covering an essentially invariant hydrophobic surface.[203]

Biological significance of the sirtuin family was outlined in the previous section. Human sirtuin-1 (Sirt1) is implicated in a wide range of human diseases due to the fact that this enzyme plays multiple roles in cellular processes ranging from energy metabolism to cell survival via it ability to deacetylate a wide range of substrates, such as p53, NF-κB, FOXO transcription factors, and PGC-1α.[204-206] Deacetylation activity of Sirt1 was shown to be affected by various regions located within the long N- and C-termini that flank the Sirt1 catalytic domain.[207,208]

Figure 17B illustrates that in agreement with a recent bioinformatics study

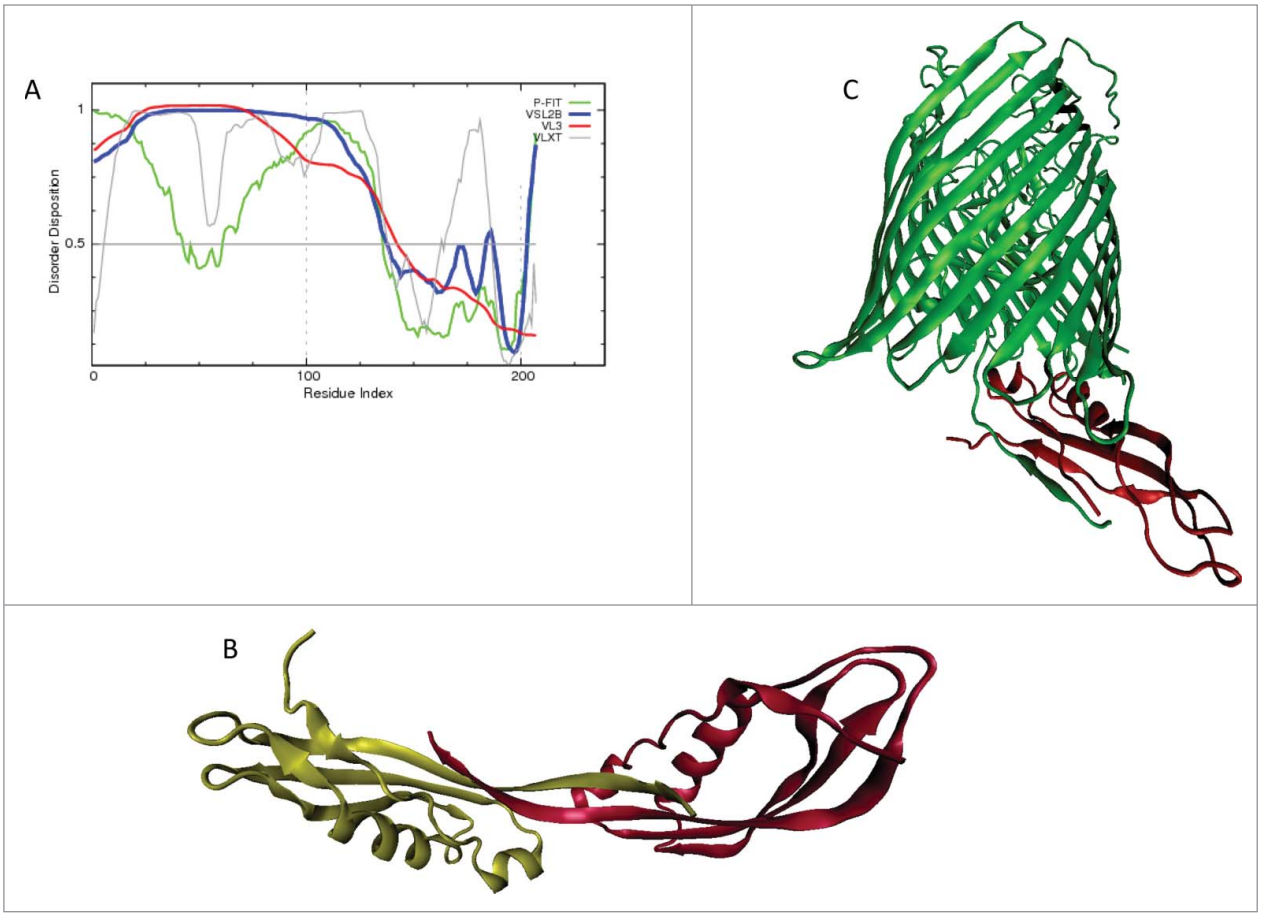

Figure 18. (**A**) Intrinsic disorder in the periplasmic domain of the TolB protein (residues 33–239, UniProt ID: P02929). (**B**) Crystal structure of the dimer containing C-terminal fragment of the *E.coli* TolB protein (residues 150–239, PDB ID: 1U07). (**C**) Crystal structure (PDB ID: 2GRX) of the complex between the TolB (residues 31–239, red structure) and FhuA complex (green structure). Intrinsic disorder propensities are evaluated by PONDR® FIT, PONDR® VSL2B, PONDR® VL3, and PONDR® VLXT.

revealing that N- and C-terminal segments of sirtuins in all known organisms are expected to be intrinsically disordered,[209] both termini of human Sirt1 (UniProt ID: Q96EB6, residues 1–250 and 500–747) are predicted to be mostly disordered, whereas the central catalytic domain is mostly ordered. The ANCHOR analysis[15,16] revealed that human Sirt1 contains 14 AiBSs located at the residues 1–33, 43–49, 53–125, 133–169, 184–193, 220–225, 498–503, 521–528, 549–574, 588–593, 618–624, 637–672, 691–706, and 710–744. It is important to emphasize here that the CTR (residues 641–665) used in the mentioned above structural analysis of human Sirt1[203] is completely embedded within one of the AiBSs (residues 637–672).

The endosome-associated deubiquitinase AMSH

Davies et al. performed a systematic mutational analysis to elucidate the molecular mechanisms of activation of AMSH [associated molecule with a Src homology 3 domain of signal transducing adaptor molecule (STAM)], a deubiquitinating enzyme (DUB) with exquisite specificity for Lys63-linked polyubiquitin chains and recruitment of this protein to the endosomal sorting complexes required for transport (ESCRT) machinery.[210] The fact that the mutations AMSH were implemented in microcephaly capillary malformation (MIC-CAP) syndrome in children[211] further reiterates the importance of this study. AMSH is a member of the JAMM (JAB1/MPN/MOV34) family of DUBs involved in the regulation of ubiquitin signaling by catalyzing the hydrolysis of isopeptide (or peptide) bonds between ubiquitin and target proteins or within polymeric chains of ubiquitin.[212] The catalytic activity of AMSH is stimulated upon binding to STAM, which is a member of the ESCRT-0 complex. Here, the AMSH-STAM interaction is conducted through the binding of the SH3 binding motif of AMSH to the SH3 domain of the STAM.[213]

Davies et al. used in their analysis the catalytic domain of AMSH (residues 219–424).[210] **Figure 14C** represents the results of the disorder analysis of human AMSH (UniProt ID: O95630) and shows that this protein contains a long IDPR (residues 90–250) thereby illustrating that the N-terminal part of the analyzed catalytic domain is predicted to be disordered.

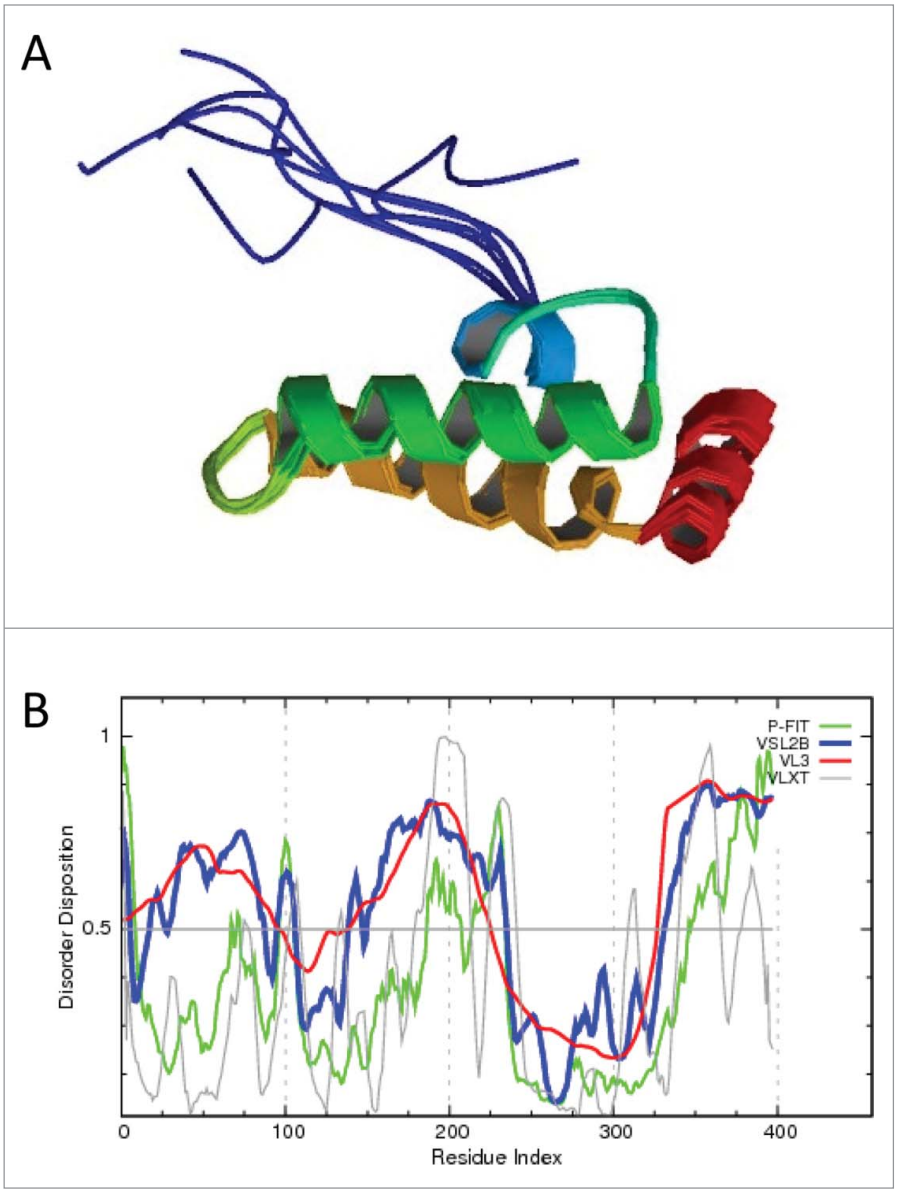

Figure 19. (**A**) NMR solution structure of the human DNAJA1 J-domain (PDB ID: 2M6Y). Intrinsic disorder propensity of the human DNAJA1 (UniProt ID: P31689) evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

According to the ANCHOR analysis, AMSH contains a disorder-based binding site at position 205–213 that is located in the close proximity to the known STAM-binding motif of this protein. Search of the ELM server (http://elm.eu.org/)[214] for functional eukaryotic linear motifs (ELMs) revealed that AMSH has several ELMs in the 200–250 region, such as WDR5 WD40 repeat-binding ligand (residues 199–213), FHA phosphopeptide ligand (residues 208–214), SH3 ligand (residues 224–230), PCSK cleavage site (residues 235–239), cyclin recognition site (residues 235–239), GSK3 phosphorylation sites (residues 2070219 and 240–247) among several other ELMs.

Interactions of the transperiplasmic protein TonB with outer membrane transporters BtuB, FecA, and FhuA

Freed et al. looked at the peculiarities of interaction between the *E.coli* transperiplasmic protein TonB with outer membrane transporters BtuB, FecA, and FhuA.[215]

TonB is the *E.coli* inner membrane protein, which possesses a polyproline motif (residues 70–81) that may span the length of the periplasmic space[216] and a globular C-terminal domain (residues 150–239) that interacts with the transporters.[217,218] **Figure 18A** shows that the N-terminal half of the periplasmic domain of TonB protein (UniProt ID: P02929) is predicted to be highly disordered, whereas its C-terminal domain possesses significant amount of order. **Figure 18B** illustrates an important point that the dimer formed by the truncated periplasmic domain (residues 150–239) is a highly intertwined and elongated structure with noticeable domain swapping. A crystal structure of the TonB-FhuA complex is shown in **Figure 18C**. Curiously, although the entire periplasmic domain of TonB was used in the crystallization experiment, residues 39–157 and 236–239 are not seen in the resulting structure, clearly indicating that these regions are likely to preserve mostly disordered structure even in the bound form of TonB. The disordered N-terminal and C-terminal tails clearly have functional importance since they contains AiBSs, 3 at the N-terminus (residues 39–65, 92–95, and 124–164) and one at the C-terminus (residues 232–239). One more AiBS is predicted at residues 171–188.

Structure and function of human DnaJ homolog subfamily A member 1 (DNAJA1)

Stark et al. performed functional analysis of the important human chaperone, DnaJ homolog subfamily A member 1 (DNAJA1), and solved solution structure of its J-domain (residues 1–67).[219] Human DNAJA1 has multiple important functions and act as protein chaperone, regulator of androgen receptor signaling, and activator of the DnaK protein. Furthermore, levels of the human DNAJA1 serve as a biomarker for pancreatic cancer, since the expression of this protein in pancreatic cancer cells is downregulated fold5-.[220] NMR analysis revealed that the structure of the human DNAJA1 J-domain consists of 4 α-helices, residues 17–21 (α1), 29–42 (α2), 52–65 (α3), and 68–75 (α4) (see **Fig. 19A**; PDB ID: 2M6Y).[219] On the contrary, the evaluation of intrinsic disorder

propensity of the full-length protein suggested that DNAJA1 (UniProt ID: P31689) possesses significant amount of disorder throughout its N- and C-terminal tails (see **Fig. 19B**). Curiously, by ANCHOR analysis, there are 3 disorder-based binding sites in DNAJA1, residues 26–31, 85–93, and 392–397, with binding site #2 being overlapped with α2.

Structure and functions of the purine-rich element binding protein B (Purβ)

In a recent comprehensive study, Romora et al. showed that the purine-rich element binding protein B (Purβ) serves as a suppressor of myofibroblast differentiation and *ACTA2* repression.[221] Purα and Purβ are members of a small family of nucleic acid-binding proteins that interact with purine-rich ssDNA or RNA sequences homologous to the so-called PUR element originally described in eukaryotic gene flanking regions and origins of DNA replication.[222-224] **Figure 14D** shows that human Purβ protein (UniProt ID: O35295) is predicted to possess significant amount of functionally important intrinsic disorder. Importance of the long disordered regions in this protein is supported by finding 10 AiBSs (residues 1–6, 18–27, 47–75, 87–95, 99–106, 135–137, 142–144, 186–201, 243–256, and 277–287). In agreement with these predictions, the authors showed that several regions of Purβ were (residues 41–112, 125–210, and 125–303) were intrinsically unstable.[221]

The host restriction factor tetherin

Hotter et al. provided a comprehensive review of the antiviral role of one of the host restriction factors (i.e., specific cellular antiviral factors that inhibit retroviral replication at different steps of the viral life cycle), human tetherin.[225]

Among various antiviral mechanisms ascribed to tetherin are the inhibition of the release of diverse enveloped viruses by tethering them to the cell surface, induction of an inflammatory response, activation of NF-κB, action as an innate immune sensor of viral infections, activation of immune response through interactions with the immunoglobulin-like transcript 7 (ILT7, LILRA4).[225] As commonly seen in the virus-host arm race, effective antiviral strategies of the host are

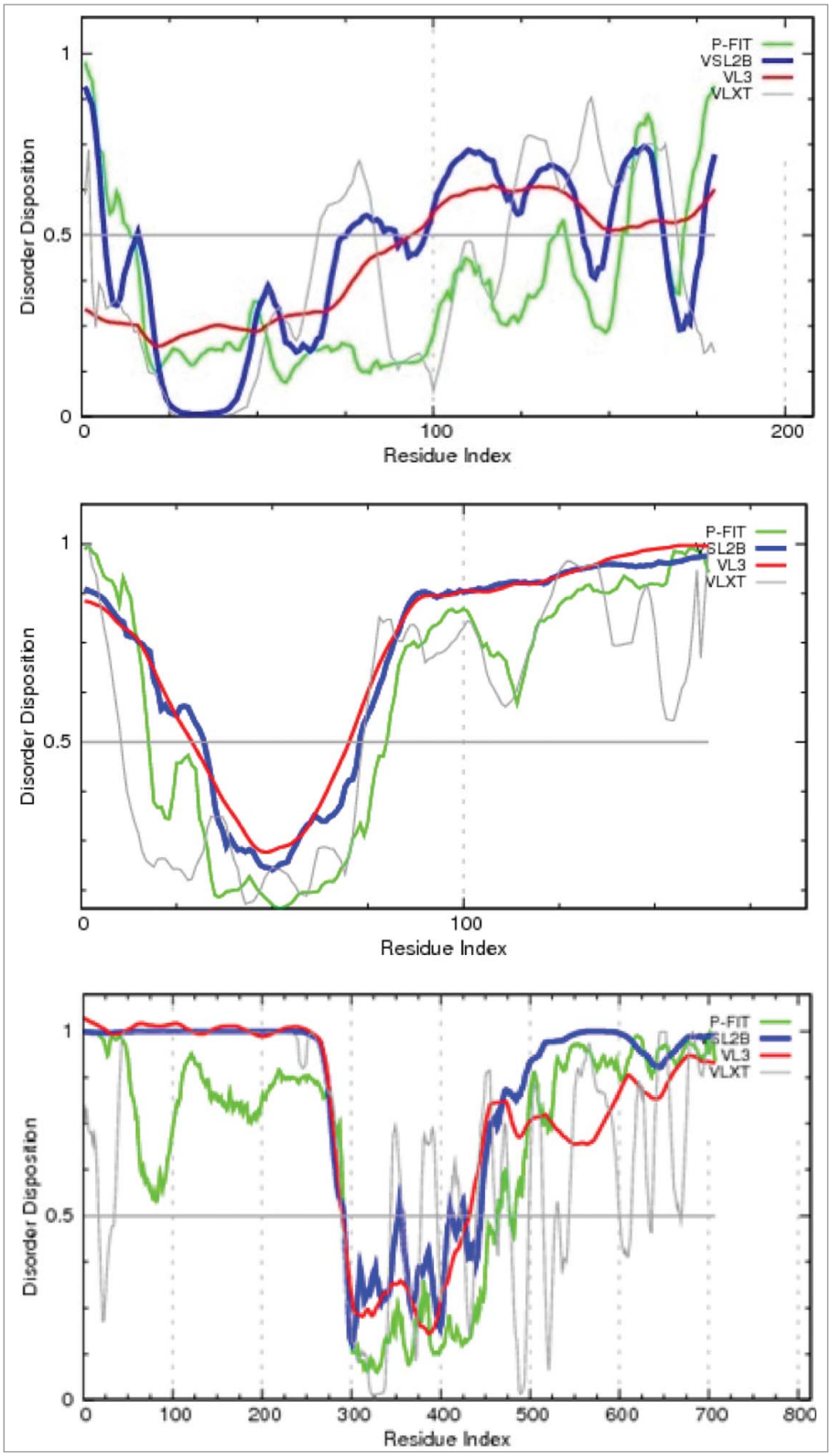

Figure 20. Intrinsic disorder propensities of human bone marrow stromal antigen 2 or tetherin (**A**, UniProt ID: Q10589); human cyclin-dependent kinase inhibitor 1 (**B**, UniProt ID: P38936); human proline- and glutamine-rich splicing factor (**C**, UniProt ID: P23246).

counterbalanced by the development of novel invasive strategies, and various viruses have evolved antagonists against this host restriction factor, such as accessory HIV-1 (human immunodeficiency virus type 1) Vpu protein that counteracts the effects of tetherin.[225]

Tetherin is a type II transmembrane protein that contains both an N-terminal transmembrane region and a C-terminal glycosyl-phosphatidylinositol (GPI) anchor.[226] Structurally, human tetherin exists as a disulfide-linked parallel homodimer, which is formed via the extracellular part that is involved in the formation of a canonical coiled-coil structure, where a long α-helix contains 3 cysteines that form disulfide bonds with the corresponding cysteines of a second tetherin molecule.[227] It is believed that tetherin-caused viral retention relies on this unusual architecture, where one transmembrane anchor may remain in the cellular plasma membrane, whereas the other is able to stick to the viral membrane.[228,229] **Figure 20A** represents the results of the multi-tool disorder predictions for human tetherin (also known as bone marrow stromal antigen, UniProt ID: Q10589) and shows that this protein is expected to be mostly disordered. This is not a very surprising observation since coiled-coil proteins are expected to be intrinsically disordered in their monomeric states and fold to a helical structure at the formation of coiled-coil structure caused by the interaction with corresponding partners.[230]

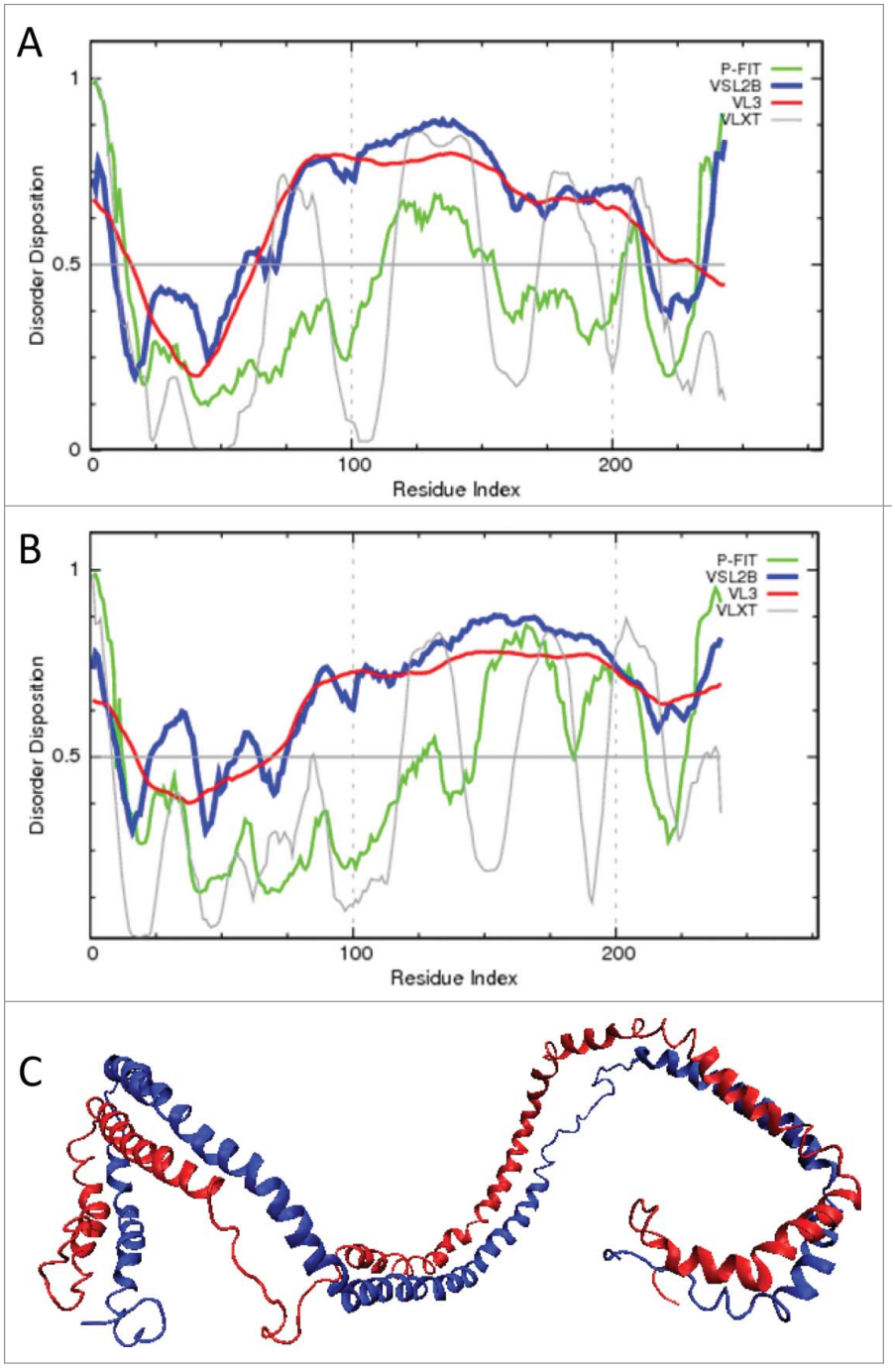

Figure 21. Intrinsic disorder propensities of human (**A**, UniProt ID: P02647) and mouse (**B**, UniProt ID: Q00623) ApoA-I proteins evaluated by by PONDR® FIT (red lines), PONDR® VLXT (blue line), PONDR® VSL2B (green line), and PONDR® VL3 (cyan line). (**C**). A model structure of the dimeric complex of human ApoA-I with lipids (PDB ID: 3K2S).

Cyclin D1/Cdk2 complex

To clarify some uncertainties in previous studies regarding formation and activities of the cyclin D1/cyclin-dependent kinase 2 (Cdk2) complexes Jahn et al. utilized a novel p21-PCNA fusion protein and p21 mutant proteins to understand the mechanisms of the cyclin D1/Cdk2 complex formation.[231] The authors show that p21 serves as an important scaffolding protein, which is required for the formation of the functional cyclin D1/Cdk2 complex, since cyclin D1 and Cdk2 failing to complex in its absence.[231] **Figure 20B** shows that the scaffolding protein p21 (which is a cyclin-dependent kinase inhibitor 1, UniProt ID: P38936) is predicted to be mostly disordered and packed with disorder-based binding sites. In fact, there are 5 AiBSs in this relatively small protein, residues 36–41, 68–79, 100–105, 109–123, and 145–164. Using this approach, the authors were able to identify a novel Cdk2 substrate, polypyrimidine tract binding protein-associated splicing factor (PSF). **Figure 20C** shows that PSF (UniProt ID: P23246) is predicted to have highly disordered tails. ANCHOR analysis[15,16] revealed that intrinsic disorder is crucial for the PSF

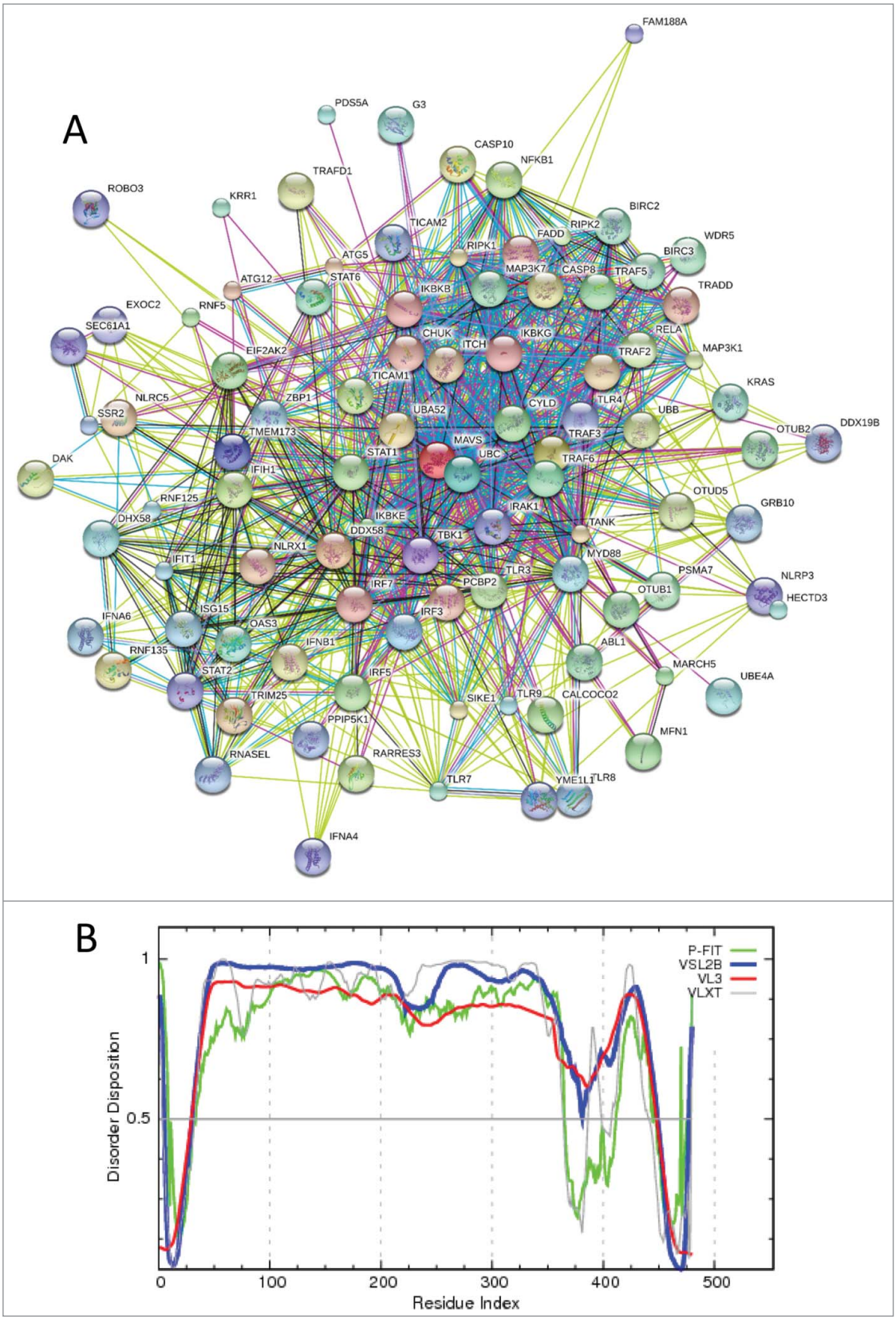

Figure 22. (**A**) Evaluating interactome of human MAVS with STRING. (**B**) Intrinsic disorder propensitof the human MAVS (UniProt ID: Q7Z434) evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

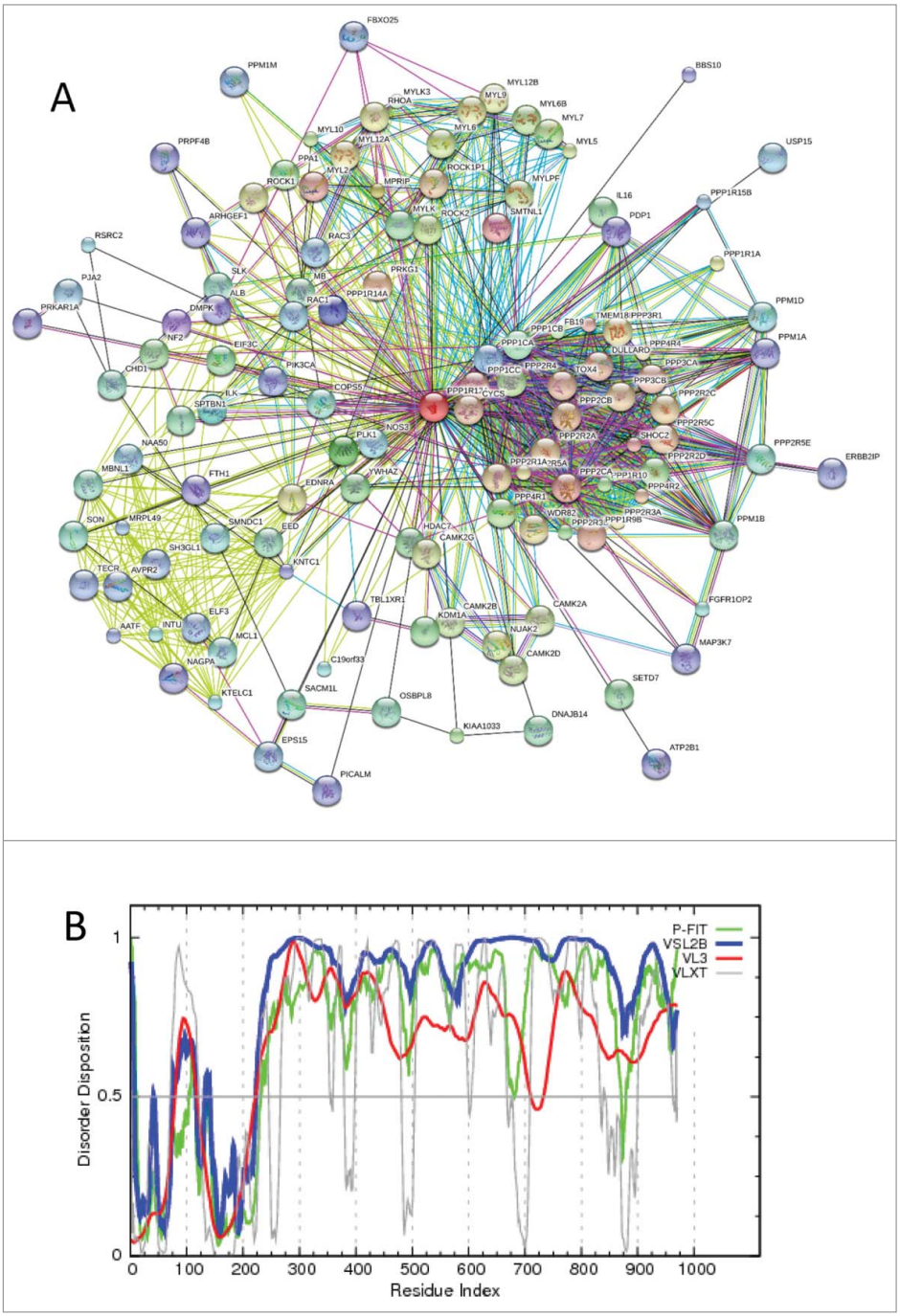

Figure 23. (**A**) Evaluating interactome of human MYPT1 with STRING. (**B**) Intrinsic disorder propensity of the human MYPT1 (UniProt ID: O14974) evaluated by PONDR® FIT (green lines), PONDR® VLXT (gray lines), PONDR® VSL2B (blue lines), and PONDR® VL3 (red lines).

functions, since this protein was predicted to have 16 AiBSs (residues 1–72, 93–97, 103–146, 154–164, 171–232, 235–257, 271–287, 293–304, 430–439, 481–496, 523–541, 579–581, 596–605, 612–678, 687–696, and 703–707).

Human and mouse apolipoproteins A-I

Nguen et al. analyzed mechanistic peculiarities of the interaction between the high density lipoprotein (HDL) particles and human or mouse apolipoprotein A-I (ApoA-I).[232]The authors established that the C-terminal domains (CTDs) of human and mouse ApoA-I proteins are the major players facilitating interactions of these proteins with HDL. Curiously, human CTD, being a bit more hydrophobic than mouse CTD, binds to the HDL particle with higher affinity. On the other hand, the isolated N-terminal helix bundle domains (residues 1–190) of human and mouse ApoA-I proteins binds HDL poorly.[232] ApoA-I stabilizes discoidal HDL particles by forming a double-belt structure[233,234] and plays a major role in the reverse cholesterol transport pathway.[235,236]

ApoA-I contains a globular N-terminal domain (residues 1–43) and a lipid-binding C-terminal domain (residues 44–243).[234] In the belt-like structure of the smallest discoidal HDL, 2 apoA-I molecules wrap around a small patch of bilayer containing 160 lipid molecules. The C-terminal domain of each monomer is ring-like, curved, planar amphipathic α-helix with the hydrophobic surface curved toward the lipids.[234] **Figure 21** shows that both human and mouse ApoA-I proteins (UniProt IDs: P02647 and Q00623) are predicted to be mostly disordered, with mouse protein possessing a bit more disordered structure. According to the ANCHOR analysis, there are 3 AiBSs in both proteins (residues 15–21, 158–164, and 221–230 in human protein, and residues 13–18, 173–204, and 207–230 in mouse ApoA-I). **Figure 21C** represents a structure of the human ApoA-I dimer in a complex with lipids (PDB ID: 3K2S). This model was built by uniting several synergistic experimental techniques, such as small angle neutron scattering (SANS) with contrast variation, isotopic deuteration of selected macromolecule components, and hydrogen/
deuterium exchange tandem mass spectrometry (HD-MS/MS).[237] In agreement with the results of disorder predictions, the ApoA-I dimer represents a highly intertwined double superhelix with a minimal number of intrachain contacts, which is stabilized mostly by the interchain interactions. In other words,

this complex is characterized by a very large interface area, and, therefore, according to Gunasekaran et al.[238] it belong to the category of 2-state dimers, where the monomers are unfolded in the unbound state and fold simultaneously with the complex formation.

Function and regulation of MAVS

Jacobs and Coyne provided an important overview of the mitochondrial antiviral signaling protein (MAVS, also known as IPS-1/VISA/Cardif).[239] MAVS is the mitochondrially-located innate immune signaling adaptor regulating and coordinating signals received from 2 independent cytosolic pathogen recognition receptors to induce antiviral genes. MAVS can be regulated by host cell factors that inhibit MAVS signaling by direct protein–protein interactions, by altering mitochondrial properties or dynamics, or by post-translational modifications.[239] The tip of the MAVS C-terminus contains a mitochondrial intermembrane region (residues 535–540) needed for the interaction with the mitochondrial membrane. According to UniProt, human MAVS (UniProt ID: Q7Z434) is involved in a wide array of protein-protein interactions. Among established MAVS binding proteins are: DDX58/RIG-I, IFIH1/MDA5, TRAF2, TRAF6, C1QBP, IRF3, FADD, RIPK1, CHUK, IKBKB, HCV and hepatitis GB virus B NS3/4A proteases, HHAV protein 3ABC, NLRX1, PSMA7, TRAFD1, PCBP2, IPS1, ITCH, CYLD, SRC, DHX58/LGP2, DDX58/RIG-I, IKBKE, TMEM173/MITA, IFIT3, TBK1, and human respiratory syncytial virus (HRSV) NS1 protein (http://www.uniprot.org/uniprot/Q7Z434). MAVS contains multiple alternatively spliced isoforms, sites of posttranslational modifications, and a proline-rich region (residues 103–153).

In other words, MAVS satisfies all the major criteria to be considered as an intrinsically disordered protein. **Figure 22A** represents the vast interactome of MAVS evaluated by STRING. **Figure 22B** shows that human MAVS is predicted to be a highly disordered protein. ANCHOR analysis[15,16] revealed that it contains 8 AiBSs located at the residues 123–148, 156–268, 277–290, 194–325, 338–377, 397–410, 416–449, and 463–479. Some of the AiBSs coincide or overlap with known functional sites of the human MAVS, e.g., regions of interaction with TRAF2 (residues 143–147), TRAF6 (residues 153–158 and 455–460). All this clearly indicates that MARV is a crucial intrinsically disordered adaptor of the innate immune signaling.

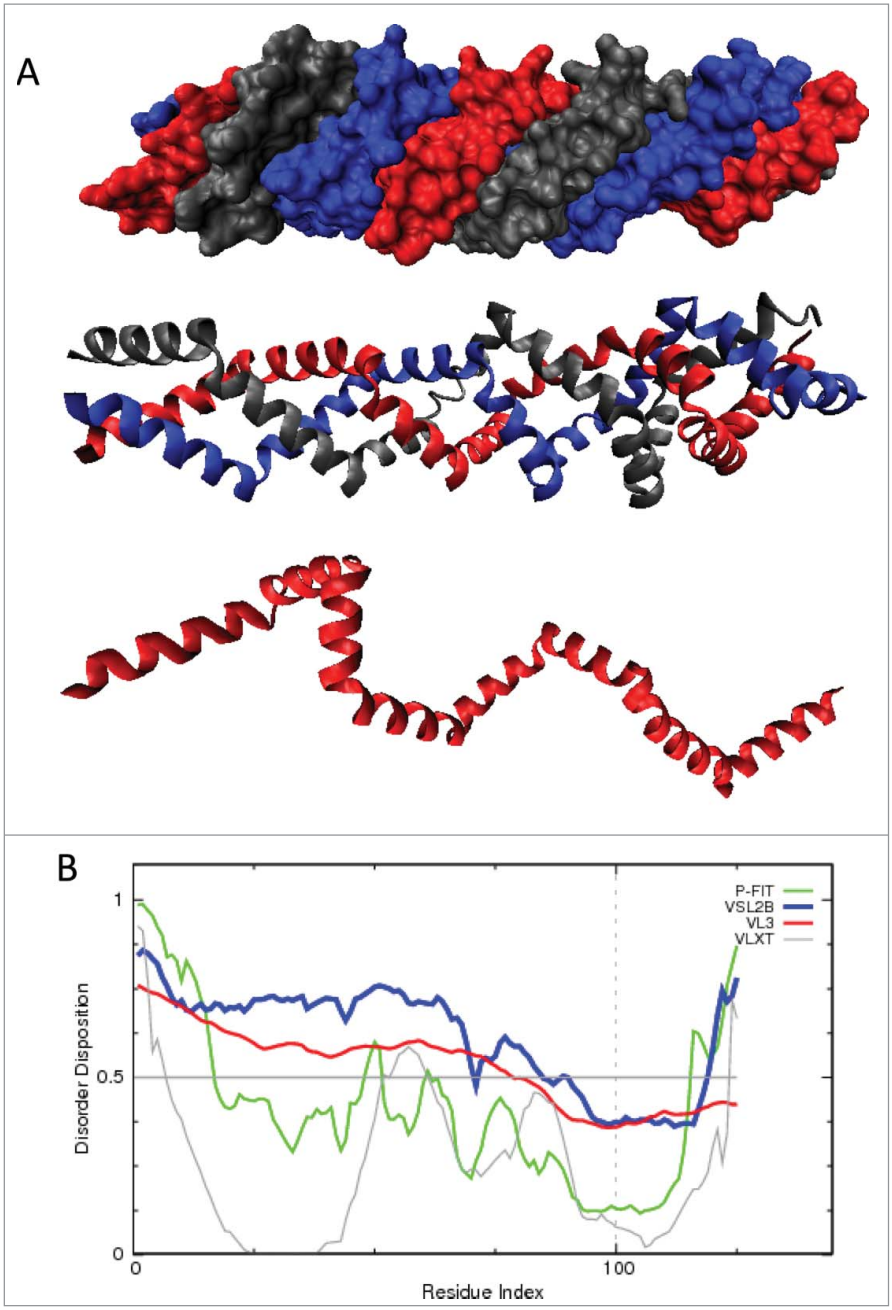

Figure 24. (**A**) Crystal structure of the *Geobacillus stearothermophilus* GerD protein (PDB ID: 4O8W). Top panel represents crystal structure of the GerD$^{60\text{-}180}$ trimer as a molecular surface, middle plot represents structure of this trimer as a ribbon diagram, and bottom panel shows the highly extended structure of one of the GerD$^{60\text{-}180}$ monomers. (**B**) Intrinsic disorder propensity of the of the *Geobacillus stearothermophilus* GerD protein (UniProt ID: Q5L3Q1) evaluated by PONDR® FIT (green line), PONDR® VLXT (gray line), PONDR® VSL2B (blue line), and PONDR® VL3 (red line).

Regulation of human myosin light chain phosphatase via phosphorylation at 2 sites of the regulatory subunit

Khasnis et al. analyzed the mechanisms and functional outputs of the phosphorylation/dephosphorylation events at Thr696 and Thr853 sites of the MYPT1 protein, which is a regulatory subunit of the myosin light chain phosphatase (MLCP).[240] MLCP is a cytoskeleton-associated protein phosphatase-1 (PP1) that serves as a RhoA/ROCK effector, regulating dynamic reorganization of the cytoskeleton crucial for cell motility. The authors also showed that the C-terminal domain of human MYPT1 (residues 495–1030) was responsible for the binding to the N-terminal portion of myosin light meromyosin.[240] Promiscuous interactability of MYPT1 is illustrated by **Figure 23A** representing the results of STRING analysis. **Figure 23B** shows that the very significant part of MYPT1 (UniProt ID: O14974) is expected to be intrinsically disordered, with longest disordered region covering ~72% of protein length (residues 290–1030) and containing 18 disorder-based binding sites (residues 318–338, 374–403, 406–419, 434–454, 472–477,
492–506, 548–564, 579–588, 626–649, 666–673, 697–712, 754–774, 780–815, 853–860, 887–892, 897–917–932–943, and 1021–1030). Clearly, intrinsic disorder is crucial for function of this protein.

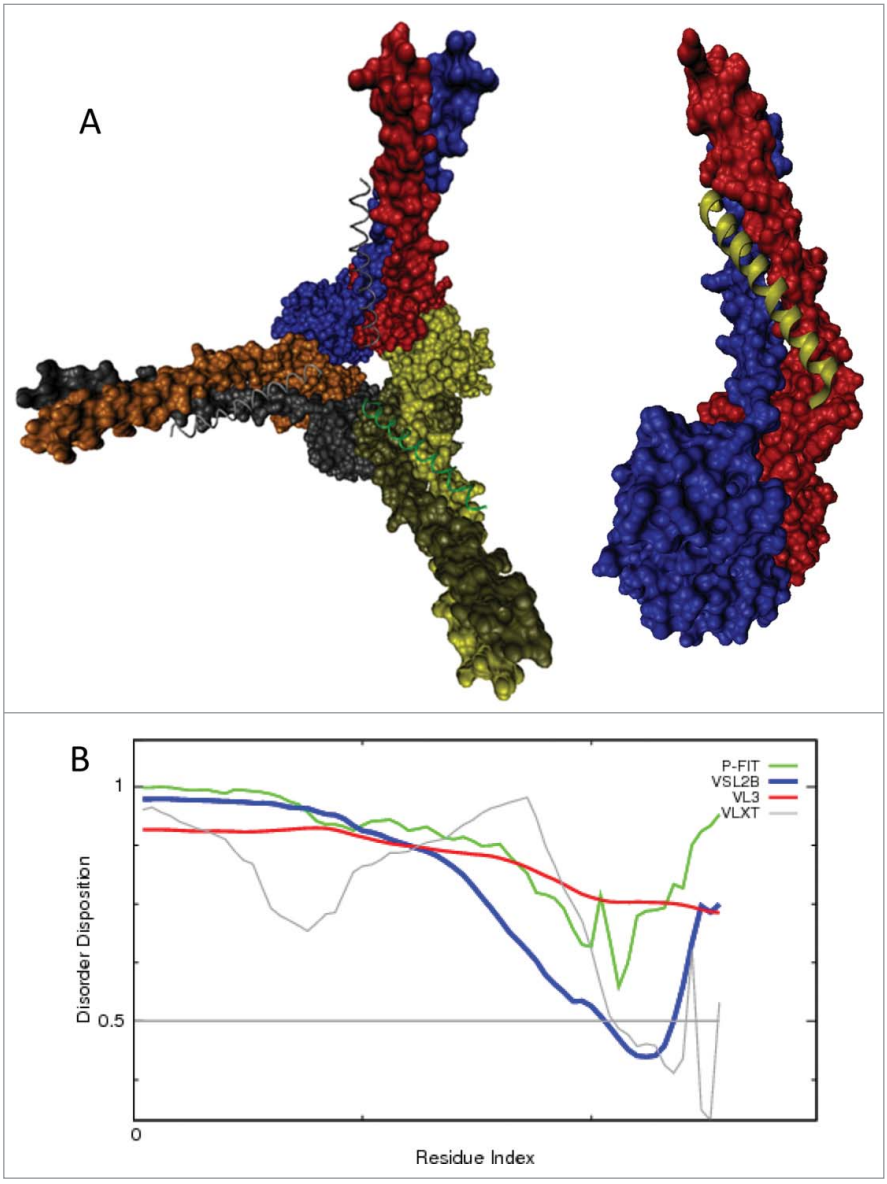

Figure 25. (**A**) Crystal structure of 3 Pup molecules bound to the Mpa hexamer (PDB ID: 3M9D) and enlarged structure of a complex between the Pup and Mpa dimer. (**B**) The results of disorder prediction for the *Mycobacterium tuberculosis* Pup (UniProt ID: P9WHN5) by PONDR® FIT (green line), PONDR® VLXT (gray line), PONDR® VSL2B (blue line), and PONDR® VL3 (red line).

GerD forms a novel trimeric superhelical rope fold

Li et al. investigated structural peculiarities of the core polypeptide of the inner membrane GerD lipoprotein from *Geobacillus stearothermophilus*.[241] Some bacteria are able to form endospores in response to adverse growth conditions.[242,243] Spores formed as a result of such sporulation process are extremely resistant to various environmental insults thereby providing the bacteria with an important means to exist in the metabolically dormant state indefinitely and remain viable for hundreds of years without water or nutrients.[244,245] Restoration of the favorable conditions triggers spore germination leading to the fast (within minutes) "awakening" of the normal metabolism in bacteria followed by outgrowth to generate growing cells.[243,244,246] This awakening is controlled by the specific nutrient sensors, cognate germinant receptors (GRs) located in the inner membrane of the spore.

GerD is one of such bacterial cognate germinant receptors that trigger spore germination in the presence of specific nutrients called germinants in the environments of the spores.[241] Li et al. have determined the crystal structure of the 121-residue core domain of the *Geobacillus stearothermophilus* GerD protein ($GerD^{60\text{-}180}$) that lacks the N-terminal signal and lipobox sequences (residues 1–28), the H01 and H02 helices (residues 29–59), and the C-terminal acidic tail (residues 181–195).[241] By a set of biophysical techniques, this $GerD^{60\text{-}180}$ was shown to form a stable, well-ordered 3-helix bundle in solution. In crystal structure, $GerD^{60\text{-}180}$ trimer consists of 3 parallel polypeptide chains twisted into a superhelical, right-handed rope (PDB ID: 4O8W).[241] In this elongated trimer, each

of the individual GerD$^{60\text{-}180}$ chains forms 8 helices (H1 to H8) linked by short turns. Structures of the individual chains can be superimposed on each other except to the loosely packed H8 helices. Helices H1–H7 in each GerD$^{60\text{-}180}$ chain twist around the central axis to form 2 complete turns of a right-handed supercoil with a pitch of ~44Å. **Figure 24A** represents crystal structure of the GerD$^{60\text{-}180}$ trimer as a molecular surface and ribbon diagram and also shows the highly extended structure of one of the GerD$^{60\text{-}180}$ monomers.[241] Analysis of the overall shapes of the highly intertwined and extended monomers within the GerD$^{60\text{-}180}$ trimer suggests that the formation of this trimer represents an example of the folding-upon-binding process. In agreement with this hypothesis, **Figure 24B** shows that the core domain of the *Geobacillus stearothermophilus* GerD protein (UniProt ID: Q5L3Q1) is mostly disordered.

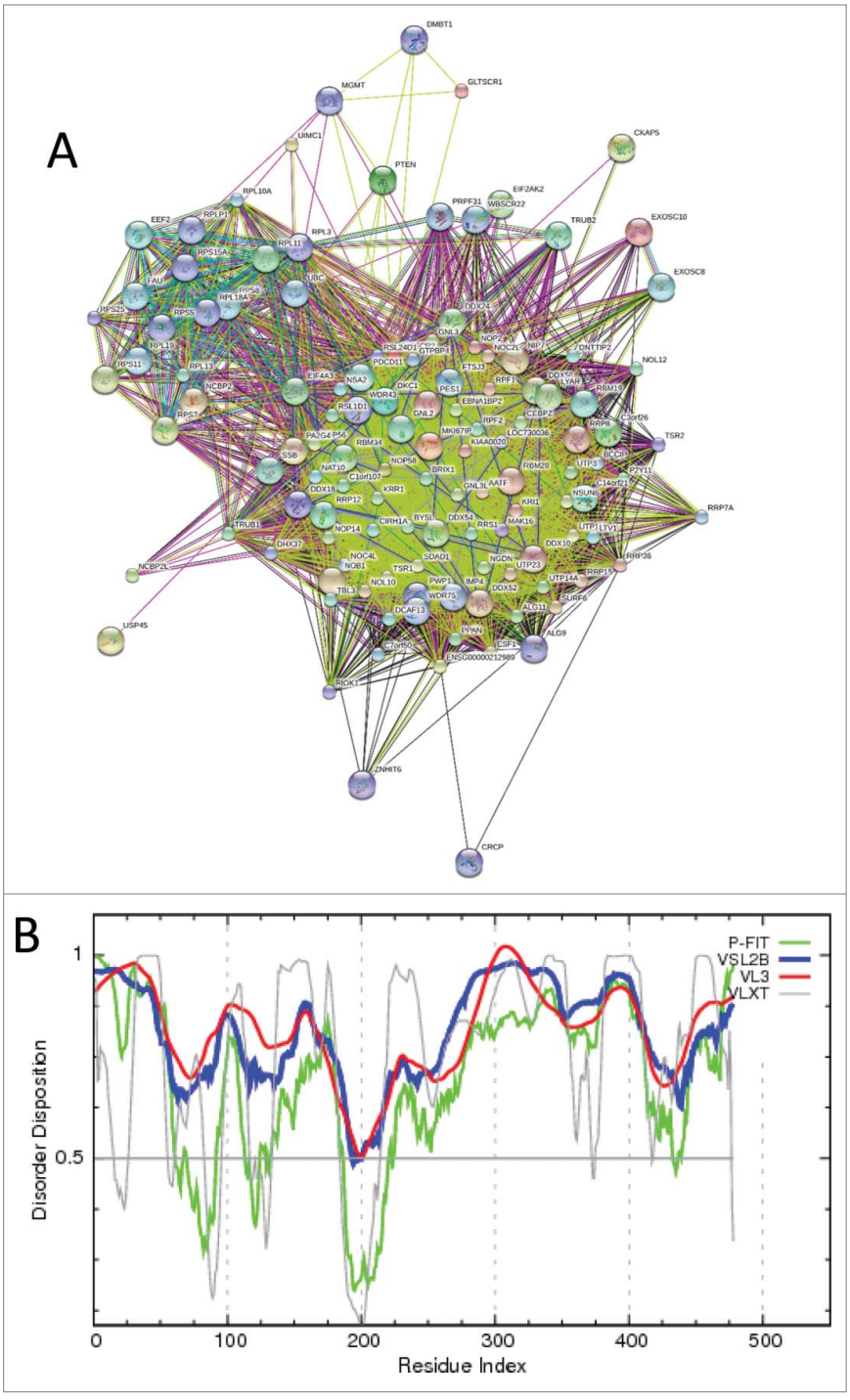

Figure 26. (**A**) Evaluation of the interactivity of the human PICT-1/GLTSCR2 (UniProt ID: Q9NZM5) by STRING. (**B**) Intrinsic disorder propensity of PICT-1/GLTSCR2 (UniProt ID: Q9NZM5) analyzed by PONDR® FIT (green line), PONDR® VLXT (gray line), PONDR® VSL2B (blue line), and PONDR® VL3 (red line).

Prokaryotic ubiquitin-like protein (Pup), the Pup ligase PafA and bacterial proteasome

Forer et al. investigated the interaction between the prokaryotic ubiquitin-like protein (Pup), the proteasome regulatory subunit Mpa (mycobacterium proteasome ATPase), and proteasome accessory factor A (PafA), an enzyme that forms an isopeptide bond between the γ-carboxylate of a glutamate at the C-terminus of Pup and the ε-amine of a substrate lysine.[247] The authors show that Pup is involved in simultaneous interaction with both Mpa and PafA, and that Mpa forms a complex with PafA, suggesting that PafA and the proteasome acts as a modular machine for the tagging and degradation of cytoplasmic proteins.[247] **Figure 25A** represents a crystal structure of 3 Pup molecules bound to the Mpa hexamer (PDB ID: 3M9D), and the results of disorder prediction for the *Mycobacterium tuberculosis* Pup (UniProt ID: P9WHN5) are shown in **Figure 25B**. Pup is obviously a highly disordered protein that partially folds at binding. In fact, of 68 residues used in the crystallization experiment, only 31 are visible in structure, whereas remaining 37 residues (1–20 and 52–68) are located within the regions of missing electron density. According to the ANCHOR analysis, Pup possesses 2 disorder-based binding regions, residues 1–15 and 34–64. Curiously, Forer et al. suggested that Pup has 2 binding sites for RafA located at residues 38–47 and 51–58, and that the first RafA binding site is overlapped with the region responsible for the Mpa binding (residues 21–51).[247]

The nucleolar PICT-1/GLTSCR2 protein

Borodianskiy-Shteinberg et al. studied self-association of the human protein interacting with carboxyl terminus-1 (PICT-1), also known as the glioma tumor suppressor candidate region 2 gene product (GLTSCR2), a nucleolar protein with unknown function.[248] PICT-1/GLTSCR2 is conserved among eukaryotes and is assumed to be essential for preimplantation embryogenesis and embryonic stem cell survival and proliferation.[249] It was suggested that PICT-1 might act as a potential tumor suppressor, being involved in interaction with the C-terminal region of the tumor suppressor phosphatase and tensin homolog (PTEN), promoting its phosphorylation and stabilization.[250] It was also suggested that PICT-1 can function as an oncogene through the inhibition of p53.[249] **Figure 26A** represents the results of the evaluation of the interactivity of the human PICT-1/GLTSCR2 (UniProt ID: Q9NZM5) by STRING and shows a dense interaction network of this intriguing protein. **Figure 26B** provides an explanation for the binding promiscuity of PICT-1 by showing that this protein is predicted to be highly disordered.

Recognition of modified tRNA by the HIV-1 nucleocapsid protein p7 and related peptides

Spears et al. studied the molecular mechanisms underlying specific recognition of highly post-transcriptionally modified human $tRNA^{Lys3}_{UUU}$ (which is the primer for HIV replication) by the HIV-1 nucleocapsid protein p7.[251] It is known that p7 recruits $htRNA^{Lys3}_{UUU}$ from the host cell by binding to and remodeling the tRNA structure. An intrigue here is in the fact that $htRNA^{Lys3}_{UUU}$ is one of the most uniquely processed tRNAs that has a broad spectrum of chemically different post-transcriptional modifications playing crucial roles in regulation of the conformation and function of this tRNA during protein synthesis.[252] Using phage display approach, Spears et al. showed that p7 contains a specific $htRNA^{Lys3}_{UUU}$–binding region whose interaction with fully modified tRNA can be mimicked by the 15- and 16-amino acid peptides with the signature sequence of R-W-Q/N-H-X_2-F-Pho-X-G/A-W-R-X_2-G (where X can be most amino acids, and Pho is any hydrophobic residue).[251]

HIV-1 nucleocapsid protein p7 is a 55-residues-long protein containing 2 zinc finger domains flanked by basic amino acids required for interaction with nucleic acids.[253,254] In addition to the discussed above capability of p7 to specifically recognize $htRNA^{Lys3}_{UUU}$, the major function of p7 is to bind specifically to the packaging signal of the full-length viral RNAs and to deliver them into the assembling virion.[255] As it is a highly charged basic protein, p7 binds single-stranded nucleic acids nonspecifically. Consequently, it coats the genomic RNA protecting it from nucleases and compacting viral RNA within the core. Protein p7 also serves as an RNA chaperone that enhances several nucleic acid-dependent steps of viral life, such as melting RNA secondary structures, promoting DNA strand exchange reactions during reverse transcription,[256-258] and stimulating integration.[259]

In the NMR solution structure of p7, regions corresponding to the 2 zinc fingers (residues 15–28 and 36–49) possess well-resolved structures, whereas residues 1–13, 32–34, and 52–55 were highly dynamic and did not converge to the unique conformations.[260,261] Recent intrinsic disorder propensity analysis revealed that p7 is a highly disordered protein, with regions corresponding to the zinc fingers predicted to be more ordered than the remainder of the protein and identified as potential α-MoRFs.[262]

Concluding Remarks

This article represents a set of papers dealing with interesting and important proteins with diverse functions from various organisms. The only uniting theme for all these studies is the notion that they represent noticeable disorder overlooks. In fact, all proteins studied there contain intrinsic disorder. The amount of disorder ranges, and in some proteins only relative short regions are disordered or highly flexible, whereas some other proteins are almost entirely disordered. However, very often this disorder has functional implementations, and consideration of studied proteins in light of intrinsic disorder often provides interesting mechanical clues on the molecular mechanisms of their action.

We hope that the readers will find this new series useful, since it might increase the awareness of the scientific community about importance of intrinsic disorder for protein function. We also hope to see submissions from readers who found some important intrinsic disorder overlooks in published articles. Obviously, papers in this seris will range in their size, and we envision several types of submissions, such as comprehensive reviews of unreported disorder in papers published in particular journals (i.e., similar to this review), reviews of unreported disorder in specific areas of protein-related research, as well as brief notes about overlooked disorder in individual papers. The biggest hope though is that this series will become obsolete, and protein intrinsic disorder will find its way to be robustly present in the scientific literature.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Uversky VN. Digested disorder: quarterly intrinsic disorder digest (January/February/ March, 2013). Intrinsically Disord Proteins 2013; 1:e25496; http://dx.doi.org/10.4161/idp.25496
2. DeForte S, Reddy KD, Uversky VN. Digested disorder, issue #2: quarterly intrinsic disorder digest (April/May/June, 2013). Intrinsically Disord Prot 2013; 1: e27454; http://dx.doi.org/10.4161/idp.27454
3. Reddy KD, DeForte S, Uversky VN. Digested disorder, issue #3: quarterly intrinsic disorder digest (July-August-September, 2013). Intrinsically Disord Prot 2014; 2:e27833; http://dx.doi.org/10.4161/idp.27833
4. Peng ZL, Kurgan L. Comprehensive comparative assessment of in-silico predictors of disordered regions. Curr Protein Pept Sci 2012; 13: 6-18; PMID:22044149; http://dx.doi.org/10.2174/138920312799277938
5. Andreou AZ, Klostermeier D. eIF4B and eIF4G jointly stimulate eIF4A ATPase and unwinding activities by modulation of the eIF4A conformational cycle. J Mol Biol 2014; 426:51-61; PMID:24080224; http://dx.doi.org/10.1016/j.jmb.2013.09.027
6. Grifo JA, Tahara SM, Morgan MA, Shatkin AJ, Merrick WC. New initiation factor activity required for globin mRNA translation. J Biol Chem 1983; 258:5804-10; PMID:6853548
7. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol 2010; 11:113-27; PMID:20094052; http://dx.doi.org/10.1038/nrm2838
8. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. Microbiol Mol Biol Rev 2011; 75:434-67, first page of table of contents; PMID:21885680; http://dx.doi.org/10.1128/MMBR.00008-11

9. Parsyan A, Svitkin Y, Shahbazian D, Gkogkas C, Lasko P, Merrick WC, Sonenberg N. mRNA helicases: the tacticians of translational control. Nat Rev Mol Cell Biol 2011; 12:235-45; PMID:21427765; http://dx.doi.org/10.1038/nrm3083
10. Andreou AZ, Klostermeier D. The DEAD-box helicase eIF4A: paradigm or the odd one out? RNA Biol 2012; 10:19-32; PMID:22995829; http://dx.doi.org/10.4161/rna.21966
11. Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK. Exploiting heterogeneous sequence properties improves prediction of protein disorder. Proteins 2005; 61(Suppl 7):176-82; PMID:16187360; http://dx.doi.org/10.1002/prot.20735
12. Obradovic Z, Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK. Predicting intrinsic disorder from amino acid sequence. Proteins 2003; 53(Suppl 6):566-72; PMID:14579347; http://dx.doi.org/10.1002/prot.10532
13. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. Proteins 2001; 42:38-48; PMID:11093259
14. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim Biophys Acta 2010; 1804:996-1010; PMID:20100603; http://dx.doi.org/10.1016/j.bbapap.2010.01.011
15. Meszaros B, Simon I, Dosztanyi Z. Prediction of protein binding regions in disordered proteins. PLoS Comput Biol 2009; 5:e1000376; PMID:19412530; http://dx.doi.org/10.1371/journal.pcbi.1000376
16. Dosztanyi Z, Meszaros B, Simon I. ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics 2009; 25: 2745-6; PMID:19717576; http://dx.doi.org/10.1093/bioinformatics/btp518
17. Florova G, Karandashova S, Declerck PJ, Idell S, Komissarov AA. Remarkable stabilization of plasminogen activator inhibitor 1 in a "molecular sandwich" complex. Biochemistry 2013; 52(27):4697-709; PMID:23734661; http://dx.doi.org/10.1021/bi400470s
18. Perumal SK, Nelson SW, Benkovic SJ. Interaction of T4 UvsW helicase and single-stranded DNA binding protein gp32 through its carboxy-terminal acidic tail. J Mol Biol 2013; 425:2823-39; PMID:23732982; http://dx.doi.org/10.1016/j.jmb.2013.05.012
19. Kerr ID, Sivakolundu S, Li Z, Buchsbaum JC, Knox LA, Kriwacki R, White SW. Crystallographic and NMR analyses of UvsW and UvsW.1 from bacteriophage T4. J Biol Chem 2007; 282:34392-400; PMID:17878153; http://dx.doi.org/10.1074/jbc.M705900200
20. West SC. DNA helicases: new breeds of translocating motors and molecular pumps. Cell 1996; 86:177-80; PMID:8706121; http://dx.doi.org/10.1016/S0092-8674(00)80088-4
21. Perumal SK, Raney KD, Benkovic SJ. Analysis of the DNA translocation and unwinding activities of T4 phage helicases. Methods 2010; 51:277-88; PMID:20170733; http://dx.doi.org/10.1016/j.ymeth.2010.02.011
22. Dudas KC, Kreuzer KN. UvsW protein regulates bacteriophage T4 origin-dependent replication by unwinding R-loops. Mol Cell Biol 2001; 21:2706-15; PMID:11283250; http://dx.doi.org/10.1128/MCB.21.8.2706-2715.2001
23. Gauss P, Park K, Spencer TE, Hacker KJ. DNA helicase requirements for DNA replication during bacteriophage T4 infection. J Bacteriol 1994; 176:1667-72; PMID:8132462
24. Nelson SW, Benkovic SJ. The T4 phage UvsW protein contains both DNA unwinding and strand annealing activities. J Biol Chem 2007; 282: 407-16; PMID:17092935; http://dx.doi.org/10.1074/jbc.M608153200
25. Nelson SW, Perumal SK, Benkovic SJ. Processive and unidirectional translocation of monomeric UvsW helicase on single-stranded DNA. Biochemistry 2009; 48:1036-46; PMID:19154117; http://dx.doi.org/10.1021/bi801792q
26. Carles-Kinch K, George JW, Kreuzer KN. Bacteriophage T4 UvsW protein is a helicase involved in recombination, repair and the regulation of DNA replication origins. EMBO J 1997; 16:4142-51; PMID:9233823; http://dx.doi.org/10.1093/emboj/16.13.4142
27. Webb MR, Plank JL, Long DT, Hsieh TS, Kreuzer KN. The phage T4 protein UvsW drives Holliday junction branch migration. J Biol Chem 2007; 282:34401-11; PMID:17823128; http://dx.doi.org/10.1074/jbc.M705913200
28. Spiering MM, Nelson SW, Benkovic SJ. Repetitive lagging strand DNA synthesis by the bacteriophage T4 replisome. Mol Biosyst 2008; 4:1070-4; PMID:18931782; http://dx.doi.org/10.1039/b812163j
29. Shamoo Y, Friedman AM, Parsons MR, Konigsberg WH, Steitz TA. Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. Nature 1995; 376:362-6; PMID: 7630406; http://dx.doi.org/10.1038/376362a0
30. Morrical SW, Beernink HT, Dash A, Hempstead K. The gene 59 protein of bacteriophage T4. Characterization of protein-protein interactions with gene 32 protein, the T4 single-stranded DNA binding protein. J Biol Chem 1996; 271:20198-207; PMID:8702746; http://dx.doi.org/10.1074/jbc.271.33.20198
31. Salinas F, Benkovic SJ. Characterization of bacteriophage T4-coordinated leading- and lagging-strand synthesis on a minicircle substrate. Proc Natl Acad Sci U S A 2000; 97:7196-201; PMID:10860983; http://dx.doi.org/10.1073/pnas.97.13.7196
32. Ishmael FT, Alley SC, Benkovic SJ. Identification and mapping of protein-protein interactions between gp32 and gp59 by cross-linking. J Biol Chem 2001; 276:25236-42; PMID:11309384; http://dx.doi.org/10.1074/jbc.M100783200
33. Koo J, Tang T, Harvey H, Tammam S, Sampaleanu L, Burrows LL, Howell PL. Functional mapping of PilF and PilQ in the Pseudomonas aeruginosa type IV pilus system. Biochemistry 2013; 52:2914-23; PMID:23547883; http://dx.doi.org/10.1021/bi3015345
34. Burkhardt J, Vonck J, Averhoff B. Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium Thermus thermophilus HB27. J Biol Chem 2011; 286:9977-84; PMID: 21285351; http://dx.doi.org/10.1074/jbc.M110.212688
35. Korotkov KV, Gonen T, Hol WG. Secretins: dynamic channels for protein transport across membranes. Trends Biochem Sci 2011; 36:433-43; PMID:21565514; http://dx.doi.org/10.1016/j.tibs.2011.04.002
36. Bonet R, Vakonakis I, Campbell ID. Characterization of 14-3-3-zeta Interactions with integrin tails. J Mol Biol 2013; 425:3060-72; PMID:23763993; http://dx.doi.org/10.1016/j.jmb.2013.05.024
37. Wilker E, Yaffe MB. 14-3-3 Proteins–a focus on cancer and human disease. J Mol Cell Cardiol 2004; 37:633-42; PMID:15350836; http://dx.doi.org/10.1016/j.yjmcc.2004.04.015
38. Zhao J, Meyerkord CL, Du Y, Khuri FR, Fu H. 14-3-3 proteins as potential therapeutic targets. Semin Cell Dev Biol 2011; 22:705-12; PMID:21983031; http://dx.doi.org/10.1016/j.semcdb.2011.09.012
39. Yaffe MB, Rittinger K, Volinia S, Caron PR, Aitken A, Leffers H, Gamblin SJ, Smerdon SJ, Cantley LC. The structural basis for 14-3-3:phosphopeptide binding specificity. Cell 1997; 91:961-71; PMID:9428519; http://dx.doi.org/10.1016/S0092-8674(00)80487-0
40. Fuglsang AT, Visconti S, Drumm K, Jahn T, Stensballe A, Mattei B, Jensen ON, Aducci P, Palmgren MG. Binding of 14-3-3 protein to the plasma membrane H(+)-ATPase AHA2 involves the three C-terminal residues Tyr(946)-Thr-Val and requires phosphorylation of Thr(947). J Biol Chem 1999; 274:36774-80; PMID:10593986; http://dx.doi.org/10.1074/jbc.274.51.36774
41. Waterman MJ, Stavridi ES, Waterman JL, Halazonetis TD. ATM-dependent activation of p53 involves dephosphorylation and association with 14-3-3 proteins. Nat Genet 1998; 19:175-8; PMID:9620776; http://dx.doi.org/10.1038/542
42. Gardino AK, Smerdon SJ, Yaffe MB. Structural determinants of 14-3-3 binding specificities and regulation of subcellular localization of 14-3-3-ligand complexes: a comparison of the X-ray crystal structures of all human 14-3-3 isoforms. Semin Cancer Biol 2006; 16:173-82; PMID:16678437; http://dx.doi.org/10.1016/j.semcancer.2006.03.007
43. Obsil T, Obsilova V. Structural basis of 14-3-3 protein functions. Semin Cell Dev Biol 2011; 22:663-72; PMID:21920446; http://dx.doi.org/10.1016/j.semcdb.2011.09.001
44. Bustos DM, Iglesias AA. Intrinsic disorder is a key characteristic in partners that bind 14-3-3 proteins. Proteins 2006; 63:35-42; PMID:16444738; http://dx.doi.org/10.1002/prot.20888
45. Hynes RO. Integrins: bidirectional, allosteric signaling machines. Cell 2002; 110:673-87; PMID:12297042; http://dx.doi.org/10.1016/S0092-8674(02)00971-6
46. Harburger DS, Calderwood DA. Integrin signalling at a glance. J Cell Sci 2009; 122:159-63; PMID: 19118207; http://dx.doi.org/10.1242/jcs.018093
47. Anthis NJ, Campbell ID. The tail of integrin activation. Trends Biochem Sci 2011; 36:191-8; PMID: 21216149; http://dx.doi.org/10.1016/j.tibs.2010.11.002
48. Liu S, Calderwood DA, Ginsberg MH. Integrin cytoplasmic domain-binding proteins. J Cell Sci 2000; 113(Pt 20):3563-71; PMID:11017872
49. Legate KR, Fassler R. Mechanisms that regulate adaptor binding to beta-integrin cytoplasmic tails. J Cell Sci 2009; 122:187-98; PMID:19118211; http://dx.doi.org/10.1242/jcs.041624
50. Shattil SJ, Kim C, Ginsberg MH. The final steps of integrin activation: the end game. Nat Rev Mol Cell Biol 2011; 11:288-300; PMID:20308986; http://dx.doi.org/10.1038/nrm2871
51. Takala H, Nurminen E, Nurmi SM, Aatonen M, Strandin T, Takatalo M, Kiema T, Gahmberg CG, Ylanne J, Fagerholm SC. Beta2 integrin phosphorylation on Thr758 acts as a molecular switch to regulate 14-3-3 and filamin binding. Blood 2008; 112:1853-62; PMID:18550856; http://dx.doi.org/10.1182/blood-2007-12-127795
52. Rajsbaum R, Garcia-Sastre A, Versteeg GA. TRIMmunity: the roles of the TRIM E3-ubiquitin ligase family in innate antiviral immunity. J Mol Biol 2014; 426:1265-84; PMID:24333484; http://dx.doi.org/10.1016/j.jmb.2013.12.005
53. Reddy BA, Etkin LD. A unique bipartite cysteine-histidine motif defines a subfamily of potential zinc-finger proteins. Nucleic Acids Res 1991; 19:6330; PMID:1956795; http://dx.doi.org/10.1093/nar/19.22.6330
54. Reddy BA, Etkin LD, Freemont PS. A novel zinc finger coiled-coil domain in a family of nuclear proteins. Trends Biochem Sci 1992; 17:344-5; PMID: 1412709; http://dx.doi.org/10.1016/0968-0004(92)90308-V
55. Freemont PS. The RING finger. A novel protein sequence motif related to the zinc finger. Ann N Y Acad Sci 1993; 684:174-92; PMID:8317827; http://dx.doi.org/10.1111/j.1749-6632.1993.tb32280.x
56. Saurin AJ, Borden KL, Boddy MN, Freemont PS. Does this have a familiar RING? Trends Biochem Sci 1996; 21:208-14; PMID:8744354; http://dx.doi.org/10.1016/0968-0004(96)10036-0
57. Meroni G, Diez-Roux G. TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. Bioessays 2005; 27:1147-57; PMID:16237670; http://dx.doi.org/10.1002/bies.20304
58. Helbig KJ, Beard MR. The role of viperin in the innate antiviral response. J Mol Biol 2014; 426: 1210-9; PMID:24157441; http://dx.doi.org/10.1016/j.jmb.2013.10.019

59. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 2011; 39:D561-8; PMID:21045058; http://dx.doi.org/10.1093/nar/gkq973
60. Dibenedetto D, Rossetti G, Caliandro R, Carloni P. A molecular dynamics simulation-based interpretation of nuclear magnetic resonance multidimensional heteronuclear spectra of alpha-synuclein dopamine adducts. Biochemistry 2013; 52:6672-83; PMID:23964651; http://dx.doi.org/10.1021/bi400367r
61. Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, Walker PA, Kelly G, Haire LF, Yap MW, et al. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. Nature 2011; 480:379-82; PMID:22056990; http://dx.doi.org/10.1038/nature10623
62. Beloglazova N, Flick R, Tchigvintsev A, Brown G, Popovic A, Nocek B, Yakunin AF. Nuclease activity of the human SAMHD1 protein implicated in the Aicardi-Goutieres syndrome and HIV-1 restriction. J Biol Chem 2013; 288:8101-10; PMID:23364794; http://dx.doi.org/10.1074/jbc.M112.431148
63. White TE, Brandariz-Nunez A, Valle-Casuso JC, Amie S, Nguyen L, Kim B, Brojatsch J, Diaz-Griffero F. Contribution of SAM and HD domains to retroviral restriction mediated by human SAMHD1. Virology 2013; 436:81-90; PMID:23158101; http://dx.doi.org/10.1016/j.virol.2012.10.029
64. Ahn J, Hao C, Yan J, DeLucia M, Mehrens J, Wang C, Gronenborn AM, Skowronski J. HIV/simian immunodeficiency virus (SIV) accessory virulence factor Vpx loads the host cell restriction factor SAMHD1 onto the E3 ubiquitin ligase complex CRL4DCAF1. J Biol Chem 2012; 287:12550-8; PMID:22362772; http://dx.doi.org/10.1074/jbc.M112.340711
65. Lenz G, Ron EZ. Novel interaction between the major bacterial heat shock chaperone (GroESL) and an RNA chaperone (CspC). J Mol Biol 2013; PMID:24148697
66. Rasouly A, Ron EZ. Interplay between the heat shock response and translation in Escherichia coli. Res Microbiol 2009; 160:288-96; PMID:19379808; http://dx.doi.org/10.1016/j.resmic.2009.03.007
67. Meyer AS, Baker TA. Proteolysis in the Escherichia coli heat shock response: a player at many levels. Curr Opin Microbiol 2011; 14:194-9; PMID:21353626; http://dx.doi.org/10.1016/j.mib.2011.02.001
68. Jones PG, Inouye M. The cold-shock response–a hot topic. Mol Microbiol 1994; 11:811-8; PMID:8022259; http://dx.doi.org/10.1111/j.1365-2958.1994.tb00359.x
69. Phadtare S. Recent developments in bacterial cold-shock response. Curr Issues Mol Biol 2004; 6:125-36; PMID:15119823
70. Phadtare S, Alsina J, Inouye M. Cold-shock response and cold-shock proteins. Curr Opin Microbiol 1999; 2:175-80; PMID:10322168; http://dx.doi.org/10.1016/S1369-5274(99)80031-9
71. Phadtare S, Severinov K. RNA remodeling and gene regulation by cold shock proteins. RNA Biol 2010; 7:788-95; PMID:21045540; http://dx.doi.org/10.4161/rna.7.6.13482
72. Tompa P, Csermely P. The role of structural disorder in the function of RNA and protein chaperones. FASEB J 2004; 18:1169-75; PMID:15284216; http://dx.doi.org/10.1096/fj.04-1584rev
73. Kalli AC, Devaney I, Sansom MS. Interactions of phosphatase and tensin homologue (PTEN) proteins with phosphatidylinositol phosphates: insights from molecular dynamics simulations of PTEN and voltage sensitive phosphatase. Biochemistry 2014; 53:1724-32; PMID:24588644; http://dx.doi.org/10.1021/bi5000299
74. Maehama T, Dixon JE. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. J Biol Chem 1998; 273:13375-8; PMID:9593664; http://dx.doi.org/10.1074/jbc.273.22.13375
75. Malaney P, Pathak RR, Xue B, Uversky VN, Dave V. Intrinsic disorder in PTEN and its interactome confers structural plasticity and functional versatility. Scientific reports 2013; 3:2035; PMID:23783762; http://dx.doi.org/10.1038/srep02035
76. Malaney P, Uversky VN, Dave V. The PTEN Long N-tail is intrinsically disordered: increased viability for PTEN therapy. Mol Biosyst 2013; 9:2877-88; PMID:24056727; http://dx.doi.org/10.1039/c3mb70267g
77. Uversky VN, Dave V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, Joerger AC. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. Chem Rev 2014; 114:6844-79; PMID:24830552; http://dx.doi.org/10.1021/cr400713r
78. Singh S, Plaks JG, Homa NJ, Amrich CG, Heroux A, Hatfull GF, VanDemark AP. The structure of Xis reveals the basis for filament formation and insight into DNA bending within a mycobacteriophage intasome. J Mol Biol 2014; 426:412-22; PMID:24112940; http://dx.doi.org/10.1016/j.jmb.2013.10.002
79. Bushman W, Yin S, Thio LL, Landy A. Determinants of directionality in lambda site-specific recombination. Cell 1984; 39:699-706; PMID:6239693; http://dx.doi.org/10.1016/0092-8674(84)90477-X
80. Lewis JA, Hatfull GF. Control of directionality in L5 integrase-mediated site-specific recombination. J Mol Biol 2003; 326:805-21; PMID:12581642; http://dx.doi.org/10.1016/S0022-2836(02)01475-4
81. Lewis JA, Hatfull GF. Identification and characterization of mycobacteriophage L5 excisionase. Mol Microbiol 2000; 35:350-60; PMID:10652095; http://dx.doi.org/10.1046/j.1365-2958.2000.01695.x
82. Abbani MA, Papagiannis CV, Sam MD, Cascio D, Johnson RC, Clubb RT. Structure of the cooperative Xis-DNA complex reveals a micronucleoprotein filament that regulates phage lambda intasome assembly. Proc Natl Acad Sci U S A 2007; 104:2109-14; PMID:17287355; http://dx.doi.org/10.1073/pnas.0607820104
83. Sam MD, Papagiannis CV, Connolly KM, Corselli L, Iwahara J, Lee J, Phillips M, Wojciak JM, Johnson RC, Clubb RT. Regulation of directionality in bacteriophage lambda site-specific recombination: structure of the Xis protein. J Mol Biol 2002; 324:791-805; PMID:12460578; http://dx.doi.org/10.1016/S0022-2836(02)01150-6
84. Li J, Zoldak G, Kriehuber T, Soroka J, Schmid FX, Richter K, Buchner J. Unique proline-rich domain regulates the chaperone function of AIPL1. Biochemistry 2013; 52(12):2089-96; PMID:23418749; http://dx.doi.org/10.1021/bi301648q
85. Uversky VN. Flexible nets of malleable guardians: intrinsically disordered chaperones in neurodegenerative diseases. Chem Rev 2011; 111:1134-66; PMID:21086986; http://dx.doi.org/10.1021/cr100186d
86. Takeyama M, Wintermute JM, Manithody C, Rezaie AR, Fay PJ. Variable contributions of basic residues forming an APC exosite in the binding and inactivation of factor VIIIa. Biochemistry 2013; 52:2228-35; PMID:23480827; http://dx.doi.org/10.1021/bi301632g
87. Esmon CT. Molecular events that control the protein C anticoagulant pathway. Thromb Haemost 1993; 70:29-35; PMID:8236111
88. Walker FJ, Fay PJ. Regulation of blood coagulation by the protein C system. FASEB J 1992; 6:2561-7; PMID:1317308
89. Kalafatis M, Rand MD, Mann KG. The mechanism of inactivation of human factor V and human factor Va by activated protein C. J Biol Chem 1994; 269:31869-80; PMID:7989361
90. Mather T, Oganessyan V, Hof P, Huber R, Foundling S, Esmon C, Bode W. The 2.8 A crystal structure of Gla-domainless activated protein C. EMBO J 1996; 15:6822-31; PMID:9003757
91. Szymanski MR, Jezewska MJ, Bujalowski W. The Escherichia coli primosomal DnaT protein exists in solution as a monomer-trimer equilibrium system. Biochemistry 2013; 52:1845-57; PMID:23418648; http://dx.doi.org/10.1021/bi301568w
92. Allen GC, Jr., Kornberg A. Assembly of the primosome of DNA replication in Escherichia coli. J Biol Chem 1993; 268:19204-9; PMID:8366072
93. Sandler SJ. Requirements for replication restart proteins during constitutive stable DNA replication in Escherichia coli K-12. Genetics 2005; 169:1799-806; PMID:15716497
94. Heller RC, Marians KJ. Replication fork reactivation downstream of a blocked nascent leading strand. Nature 2006; 439:557-62; PMID:16452972; http://dx.doi.org/10.1038/nature04329
95. Uversky VN. The most important thing is the tail: multitudinous functionalities of intrinsically disordered protein termini. FEBS Lett 2013; 587:1891-901; PMID:23665034; http://dx.doi.org/10.1016/j.febslet.2013.04.042
96. Tourdot BE, Brenner MK, Keough KC, Holyst T, Newman PJ, Newman DK. Immunoreceptor tyrosine-based inhibitory motif (ITIM)-mediated inhibitory signaling is regulated by sequential phosphorylation mediated by distinct nonreceptor tyrosine kinases: a case study involving PECAM-1. Biochemistry 2013; 52:2597-608; PMID:23418871; http://dx.doi.org/10.1021/bi301461t
97. Newman PJ, Newman DK. Signal transduction pathways mediated by PECAM-1: new roles for an old molecule in platelet and vascular cell biology. Arterioscler Thromb Vasc Biol 2003; 23:953-64; PMID:12689916; http://dx.doi.org/10.1161/01.ATV.0000071347.69358.D9
98. Paddock C, Lytle BL, Peterson FC, Holyst T, Newman PJ, Volkman BF, Newman DK. Residues within a lipid-associated segment of the PECAM-1 cytoplasmic domain are susceptible to inducible, sequential phosphorylation. Blood 2011; 117:6012-23; PMID:21464369; http://dx.doi.org/10.1182/blood-2010-11-317867
99. Aubol BE, Jamros MA, McGlone ML, Adams JA. Splicing kinase SRPK1 conforms to the landscape of its SR protein substrate. Biochemistry 2013; 52 (43):7595-605; PMID:24074032; http://dx.doi.org/10.1021/bi4010864
100. Korneta I, Bujnicki JM. Intrinsic disorder in the human spliceosomal proteome. PLoS Comput Biol 2012; 8:e1002641; PMID:22912569; http://dx.doi.org/10.1371/journal.pcbi.1002641
101. Coelho Ribeiro Mde L, Espinosa J, Islam S, Martinez O, Thanki JJ, Mazariegos S, Nguyen T, Larina M, Xue B, Uversky VN. Malleable ribonucleoprotein machine: protein intrinsic disorder in the Saccharomyces cerevisiae spliceosome. PeerJ 2013; 1:e2; PMID:23638354; http://dx.doi.org/10.7717/peerj.2
102. Perreira JM, Chin CR, Feeley EM, Brass AL. IFITMs restrict the replication of multiple pathogenic viruses. J Mol Biol 2013; 425:4937-55; PMID:24076421; http://dx.doi.org/10.1016/j.jmb.2013.09.024
103. Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, Ryan BJ, Weyer JL, van der Weyden L, Fikrig E, et al. The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. Cell 2009; 139:1243-54; PMID:20064371; http://dx.doi.org/10.1016/j.cell.2009.12.017
104. Huang IC, Bailey CC, Weyer JL, Radoshitzky SR, Becker MM, Chiang JJ, Brass AL, Ahmed AA, Chi X, Dong L, et al. Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. PLoS Pathog 2011; 7:e1001258; PMID:21253575; http://dx.doi.org/10.1371/journal.ppat.1001258
105. Wilkins C, Woodward J, Lau DT, Barnes A, Joyce M, McFarlane N, McKeating JA, Tyrrell DL, Gale M, Jr. IFITM1 is a tight junction protein that inhibits

hepatitis C virus entry. Hepatology 2012; 57:461-9; PMID:22996292; http://dx.doi.org/10.1002/hep.26066
106. Harrison SC. Mechanism of membrane fusion by viral envelope proteins. Adv Virus Res 2005; 64:231-61; PMID:16139596
107. Feeley EM, Sims JS, John SP, Chin CR, Pertel T, Chen LM, Gaiha GD, Ryan BJ, Donis RO, Elledge SJ, et al. IFITM3 inhibits influenza A virus infection by preventing cytosolic entry. PLoS Pathog 2011; 7: e1002337; PMID:22046135; http://dx.doi.org/10.1371/journal.ppat.1002337
108. Garner E, Cannon P, Romero P, Obradovic Z, Dunker AK. Predicting disordered regions from amino acid sequence: common themes despite differing structural characterization. Genome Inform Ser Workshop Genome Inform 1998; 9:201-13; PMID:11072336
109. Garner E, Romero P, Dunker AK, Brown C, Obradovic Z. Predicting binding regions within disordered proteins. Genome Inform Ser Workshop Genome Inform 1999; 10:41-50; PMID:11072341
110. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. Coupled folding and binding with alpha-helix-forming molecular recognition elements. Biochemistry 2005; 44:12454-70; PMID: 16156658; http://dx.doi.org/10.1021/bi050736e
111. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK, Uversky VN. Analysis of molecular recognition features (MoRFs). J Mol Biol 2006; 362:1043-59; PMID:16935303; http://dx.doi.org/10.1016/j.jmb.2006.07.087
112. Cheng Y, Oldfield CJ, Meng J, Romero P, Uversky VN, Dunker AK. Mining alpha-helix-forming molecular recognition features with cross species sequence alignments. Biochemistry 2007; 46: 13468-77; PMID:17973494; http://dx.doi.org/10.1021/bi7012273
113. Amador FJ, Kimlicka L, Stathopulos PB, Gasmi-Seabrook GM, Maclennan DH, Van Petegem F, Ikura M. Type 2 ryanodine receptor domain A contains a unique and dynamic alpha-helix that transitions to a beta-strand in a mutant linked with a heritable cardiomyopathy. J Mol Biol 2013; 425:4034-46; PMID: 23978697; http://dx.doi.org/10.1016/j.jmb.2013.08.015
114. Rios E, Brum G. Involvement of dihydropyridine receptors in excitation-contraction coupling in skeletal muscle. Nature 1987; 325:717-20; PMID:2434854; http://dx.doi.org/10.1038/325717a0
115. Dulhunty AF, Haarmann CS, Green D, Laver DR, Board PG, Casarotto MG. Interactions between dihydropyridine receptors and ryanodine receptors in striated muscle. Prog Biophys Mol Biol 2002; 79:45-75; PMID:12225776
116. Lobo PA, Van Petegem F. Crystal structures of the N-terminal domains of cardiac and skeletal muscle ryanodine receptors: insights into disease mutations. Structure 2009; 17:1505-14; PMID:19913485; http://dx.doi.org/10.1016/j.str.2009.08.016
117. Lobo PA, Kimlicka L, Tung CC, Van Petegem F. The deletion of exon 3 in the cardiac ryanodine receptor is rescued by beta strand switching. Structure 2011; 19:790-8; PMID:21645850; http://dx.doi.org/10.1016/j.str.2011.03.016
118. Rodriguez AD, Dunn SD, Konermann L. ATP-induced dimerization of the FF epsilon subunit from Bacillus PS3: a hydrogen exchange-mass spectrometry study. Biochemistry 2014; 53(24):4072-80; PMID: 24870150
119. Uhlin U, Cox GB, Guss JM. Crystal structure of the epsilon subunit of the proton-translocating ATP synthase from Escherichia coli. Structure 1997; 5:1219-30; PMID:9331422; http://dx.doi.org/10.1016/S0969-2126(97)00272-4
120. Simon E, Gildor T, Kornitzer D. Phosphorylation of the cyclin CaPcl5 modulates both cyclin stability and specific recognition of the substrate. J Mol Biol 2013; 425:3151-65; PMID:23763991; http://dx.doi.org/10.1016/j.jmb.2013.06.004
121. Morgan DO. Principles of CDK regulation. Nature 1995; 374:131-4; PMID:7877684
122. Miller ME, Cross FR. Cyclin specificity: how many wheels do you need on a unicycle? J Cell Sci 2001; 114:1811-20; PMID:11329367
123. Measday V, Moore L, Retnakaran R, Lee J, Donoviel M, Neiman AM, Andrews B. A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. Mol Cell Biol 1997; 17:1212-23; PMID:9032248
124. Shemer R, Meimoun A, Holtzman T, Kornitzer D. Regulation of the transcription factor Gcn4 by Pho85 cyclin PCL5. Mol Cell Biol 2002; 22:5395-404; PMID:12101234; http://dx.doi.org/10.1128/MCB.22.15.5395-5404.2002
125. Gildor T, Shemer R, Atir-Lande A, Kornitzer D. Coevolution of cyclin Pcl5 and its substrate Gcn4. Eukaryot Cell 2005; 4:310-8; PMID:15701793; http://dx.doi.org/10.1128/EC.4.2.310-318.2005
126. Uversky VN, Dunker AK. Understanding protein non-folding. Biochim Biophys Acta 2010; 1804:1231-64; PMID:20117254; http://dx.doi.org/10.1016/j.bbapap.2010.01.017
127. Rack JG, Vanlinden MR, Lutter T, Aasland R, Ziegler M. Constitutive nuclear localization of an alternatively spliced sirtuin-2 isoform. J Mol Biol 2014; 426 (8):1677-91; PMID:24177535; http://dx.doi.org/10.1016/j.jmb.2013.10.027
128. Asher G, Schibler U. Crosstalk between components of circadian and metabolic cycles in mammals. Cell Metab 2011; 13:125-37; PMID:21284980; http://dx.doi.org/10.1016/j.cmet.2011.01.006
129. Houtkooper RH, Pirinen E, Auwerx J. Sirtuins as regulators of metabolism and healthspan. Nat Rev Mol Cell Biol 2012; 13:225-38; PMID:22395773
130. Kim HS, Vassilopoulos A, Wang RH, Lahusen T, Xiao Z, Xu X, Li C, Veenstra TD, Li B, Yu H, et al. SIRT2 maintains genome integrity and suppresses tumorigenesis through regulating APC/C activity. Cancer Cell 2011; 20:487-99; PMID:22014574; http://dx.doi.org/10.1016/j.ccr.2011.09.004
131. McBurney MW, Yang X, Jardine K, Hixon M, Boekelheide K, Webb JR, Lansdorp PM, Lemieux M. The mammalian SIR2alpha protein has a role in embryogenesis and gametogenesis. Mol Cell Biol 2003; 23:38-54; PMID:12482959; http://dx.doi.org/10.1128/MCB.23.1.38-54.2003
132. Vaquero A. The conserved role of sirtuins in chromatin regulation. Int J Dev Biol 2009; 53:303-22; PMID:19378253; http://dx.doi.org/10.1387/ijdb.082675av
133. Wang RH, Sengupta K, Li C, Kim HS, Cao L, Xiao C, Kim S, Xu X, Zheng Y, Chilton B, et al. Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. Cancer Cell 2008; 14:312-23; PMID:18835033; http://dx.doi.org/10.1016/j.ccr.2008.09.001
134. Flick F, Luscher B. Regulation of sirtuin function by posttranslational modifications. Front Pharmacol 2012; 3:29; PMID:22403547; http://dx.doi.org/10.3389/fphar.2012.00029
135. Bao J, Lu Z, Joseph JJ, Carabenciov D, Dimond CC, Pang L, Samsel L, McCoy JP, Jr., Leclerc J, Nguyen P, et al. Characterization of the murine SIRT3 mitochondrial localization sequence and comparison of mitochondrial enrichment and deacetylase activity of long and short SIRT3 isoforms. J Cell Biochem 2010; 110:238-47; PMID:20235147
136. Matsushita N, Yonashiro R, Ogata Y, Sugiura A, Nagashima S, Fukuda T, Inatome R, Yanagi S. Distinct regulation of mitochondrial localization and stability of two human Sirt5 isoforms. Genes Cells 2011; 16:190-202; PMID:21143562; http://dx.doi.org/10.1111/j.1365-2443.2010.01475.x
137. Wang HF, Li Q, Feng RL, Wen TQ. Transcription levels of sirtuin family in neural stem cells and brain tissues of adult mice. Cell Mol Biol (Noisy-le-grand) 2012; Suppl.58:OL1737-43; PMID:22992439
138. Michishita E, Park JY, Burneskis JM, Barrett JC, Horikawa I. Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins. Mol Biol Cell 2005; 16:4623-35; PMID:16079181; http://dx.doi.org/10.1091/mbc.E05-01-0033
139. North BJ, Verdin E. Interphase nucleo-cytoplasmic shuttling and localization of SIRT2 during mitosis. PLoS One 2007; 2:e784; PMID:17726514; http://dx.doi.org/10.1371/journal.pone.0000784
140. Romero PR, Zaidi S, Fang YY, Uversky VN, Radivojac P, Oldfield CJ, Cortese MS, Sickmeier M, LeGall T, Obradovic Z, et al. Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. Proc Natl Acad Sci U S A 2006; 103:8390-5; PMID:16717195; http://dx.doi.org/10.1073/pnas.0507916103
141. Buljan M, Chalancon G, Eustermann S, Wagner GP, Fuxreiter M, Bateman A, Babu MM. Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. Mol Cell 2012; 46:871-83; PMID:22749400; http://dx.doi.org/10.1016/j.molcel.2012.05.039
142. Buljan M, Chalancon G, Dunker AK, Bateman A, Balaji S, Fuxreiter M, Babu MM. Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. Curr Opin Struct Biol 2013; 23:443-50; PMID:23706950; http://dx.doi.org/10.1016/j.sbi.2013.03.006
143. Parker MW, Linkugel AD, Vander Kooi CW. Effect of C-terminal sequence on competitive semaphorin binding to neuropilin-1. J Mol Biol 2013; 425:4405-14; PMID:23871893; http://dx.doi.org/10.1016/j.jmb.2013.07.017
144. Parker MW, Hellman LM, Xu P, Fried MG, Vander Kooi CW. Furin processing of semaphorin 3F determines its anti-angiogenic activity by regulating direct binding and competition for neuropilin. Biochemistry 2010; 49:4068-75; PMID:20387901; http://dx.doi.org/10.1021/bi100327r
145. Vander Kooi CW, Jusino MA, Perman B, Neau DB, Bellamy HD, Leahy DJ. Structural basis for ligand and heparin binding to neuropilin B domains. Proc Natl Acad Sci U S A 2007; 104:6152-7; PMID: 17405859; http://dx.doi.org/10.1073/pnas.0700043104
146. Parker MW, Xu P, Li X, Vander Kooi CW. Structural basis for selective vascular endothelial growth factor-A (VEGF-A) binding to neuropilin-1. J Biol Chem 2012; 287:11082-9; PMID:22318724; http://dx.doi.org/10.1074/jbc.M111.331140
147. Delcombel R, Janssen L, Vassy R, Gammons M, Haddad O, Richard B, Letourneur D, Bates D, Hendricks C, Waltenberger J, et al. New prospects in the roles of the C-terminal domains of VEGF-A and their cooperation for ligand binding, cellular signaling and vessels formation. Angiogenesis 2013; 16:353-71; PMID: 23254820; http://dx.doi.org/10.1007/s10456-012-9320-y
148. Malashkevich VN, Almo SC, Dowd TL. X-ray crystal structure of bovine 3 Glu-osteocalcin. Biochemistry 2013; 52:8387-92; PMID:24138653; http://dx.doi.org/10.1021/bi4010254
149. Hauschka PV, Lian JB, Gallop PM. Direct identification of the calcium-binding amino acid, gamma-carboxyglutamate, in mineralized tissue. Proc Natl Acad Sci U S A 1975; 72:3925-9; PMID:1060074; http://dx.doi.org/10.1073/pnas.72.10.3925
150. Price PA, Otsuka AA, Poser JW, Kristaponis J, Raman N. Characterization of a gamma-carboxyglutamic acid-containing protein from bone. Proc Natl Acad Sci U S A 1976; 73:1447-51; PMID:1064018; http://dx.doi.org/10.1073/pnas.73.5.1447
151. Berkner KL. The vitamin K-dependent carboxylase. Annu Rev Nutr 2005; 25:127-49; PMID:16011462; http://dx.doi.org/10.1146/annurev.nutr.25.050304.092713
152. Hauschka PV, Carr SA. Calcium-dependent alpha-helical structure in osteocalcin. Biochemistry 1982;

21:2538-47; PMID:6807342; http://dx.doi.org/10.1021/bi00539a038
153. Poser JW, Price PA. A method for decarboxylation of gamma-carboxyglutamic acid in proteins. Properties of the decarboxylated gamma-carboxyglutamic acid protein from calf bone. J Biol Chem 1979; 254:431-6; PMID:762070
154. Sasi BK, Sonawane PJ, Gupta V, Sahu BS, Mahapatra NR. Coordinated transcriptional regulation of Hspa1a gene by multiple transcription factors: crucial roles for HSF-1, NF-Y, NF-kappaB, and CREB. J Mol Biol 2014; 426:116-35; PMID:24041570; http://dx.doi.org/10.1016/j.jmb.2013.09.008
155. Zhang X, Xu Z, Zhou L, Chen Y, He M, Cheng L, Hu FB, Tanguay RM, Wu T. Plasma levels of Hsp70 and anti-Hsp70 antibody predict risk of acute coronary syndrome. Cell Stress Chaperones 2010; 15:675-86; PMID:20300983; http://dx.doi.org/10.1007/s12192-010-0180-3
156. Kiang JG, Tsokos GC. Heat shock protein 70 kDa: molecular biology, biochemistry, and physiology. Pharmacol Therap 1998; 80:183-201; PMID:9839771
157. Calderwood SK. Heat shock proteins in breast cancer progression–a suitable case for treatment? Int J Hyperthermia 2010; 26:681-5; PMID:20653417; http://dx.doi.org/10.3109/02656736.2010.490254
158. Chafekar SM, Duennwald ML. Impaired heat shock response in cells expressing full-length polyglutamine-expanded huntingtin. PLoS One 2012; 7:e37929; PMID:22649566; http://dx.doi.org/10.1371/journal.pone.0037929
159. Fujimoto M, Nakai A. The heat shock factor family and adaptation to proteotoxic stress. FEBS J 2010; 277:4112-25; PMID:20945528; http://dx.doi.org/10.1111/j.1742-4658.2010.07827.x
160. Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, Goodman RH. Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature 1993; 365:855-9; PMID:8413673; http://dx.doi.org/10.1038/365855a0
161. Arias J, Alberts AS, Brindle P, Claret FX, Smeal T, Karin M, Feramisco J, Montminy M. Activation of cAMP and mitogen responsive genes relies on a common nuclear factor. Nature 1994; 370:226-9; PMID:8028671; http://dx.doi.org/10.1038/370226a0
162. Kwok RP, Lundblad JR, Chrivia JC, Richards JP, Bachinger HP, Brennan RG, Roberts SG, Green MR, Goodman RH. Nuclear protein CBP is a coactivator for the transcription factor CREB. Nature 1994; 370:223-6; PMID:7913207; http://dx.doi.org/10.1038/370223a0
163. Dorn A, Bollekens J, Staub A, Benoist C, Mathis D. A multiplicity of CCAAT box-binding proteins. Cell 1987; 50:863-72; PMID:3476205; http://dx.doi.org/10.1016/0092-8674(87)90513-7
164. Mantovani R. A survey of 178 NF-Y binding CCAAT boxes. Nucleic Acids Res 1998; 26:1135-43; PMID:9469818; http://dx.doi.org/10.1093/nar/26.5.1135
165. Bienz M. A CCAAT box confers cell-type-specific regulation on the Xenopus hsp70 gene in oocytes. Cell 1986; 46:1037-42; PMID:3757032; http://dx.doi.org/10.1016/0092-8674(86)90703-8
166. Imbriano C, Bolognese F, Gurtner A, Piaggio G, Mantovani R. HSP-CBF is an NF-Y-dependent coactivator of the heat shock promoters CCAAT boxes. J Biol Chem 2001; 276:26332-9; PMID:11306579; http://dx.doi.org/10.1074/jbc.M101553200
167. Nardini M, Gnesutta N, Donati G, Gatta R, Forni C, Fossati A, Vonrhein C, Moras D, Romier C, Bolognesi M, et al. Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell 2013; 152:132-43; PMID:23332751; http://dx.doi.org/10.1016/j.cell.2012.11.047
168. Morris KR, Lutz RD, Choi HS, Kamitani T, Chmura K, Chan ED. Role of the NF-kappaB signaling pathway and kappaB cis-regulatory elements on the IRF-1 and iNOS promoter regions in mycobacterial lipoarabinomannan induction of nitric oxide. Infection and immunity 2003; 71:1442-52; PMID:12595462; http://dx.doi.org/10.1128/IAI.71.3.1442-1452.2003
169. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. Biochemistry 2006; 45:6873-88; PMID:16734424; http://dx.doi.org/10.1021/bi0602718
170. Westerheide SD, Raynes R, Powell C, Xue B, Uversky VN. HSF transcription factor family, heat shock response, and protein intrinsic disorder. Curr Protein Pept Sci 2012; 13:86-103; PMID:22044151; http://dx.doi.org/10.2174/138920312799277956
171. Xue B, Uversky VN. Intrinsic disorder in proteins involved in the innate antiviral immunity: another flexible side of a molecular arms race. J Mol Biol 2014; 426:1322-50; PMID:24184279; http://dx.doi.org/10.1016/j.jmb.2013.10.030
172. Blasic JR, Jr., Matos-Cruz V, Ujla D, Cameron EG, Hattar S, Halpern ME, Robinson PR. Identification of critical phosphorylation sites on the carboxy tail of melanopsin. Biochemistry 2014; 53:2644-9; PMID:24678795; http://dx.doi.org/10.1021/bi401724r
173. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 2004; 32:1037-49; PMID:14960716; http://dx.doi.org/10.1093/nar/gkh253
174. Uversky VN. Intrinsic disorder-based protein interactions and their modulators. Curr Pharm Des 2013; 19:4191-213; PMID:23170892; http://dx.doi.org/10.2174/1381612811319230005
175. Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, Radivojac P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci 2014; 23:1077-93; PMID:24888500; http://dx.doi.org/10.1002/pro.2494
176. Wilson SS, Wiens ME, Smith JG. Antiviral mechanisms of human defensins. J Mol Biol 2013; 425:4965-80; PMID:24095897; http://dx.doi.org/10.1016/j.jmb.2013.09.038
177. Iakoucheva LM, Kimzey AL, Masselon CD, Bruce JE, Garner EC, Brown CJ, Dunker AK, Smith RD, Ackerman EJ. Identification of intrinsic order and disorder in the DNA repair protein XPA. Protein Sci 2001; 10:560-71; PMID:11344324; http://dx.doi.org/10.1110/ps.29401
178. Santner AA, Croy CH, Vasanwala FH, Uversky VN, Van YY, Dunker AK. Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression. Biochemistry 2012; 51:7250-62; PMID:22924672; http://dx.doi.org/10.1021/bi300653m
179. Burke JR, Liban TJ, Restrepo T, Lee HW, Rubin SM. Multiple mechanisms for E2F binding inhibition by phosphorylation of the retinoblastoma protein C-terminal domain. J Mol Biol 2013; PMID:24103329
180. Burkhart DL, Sage J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. Nat Rev Cancer 2008; 8:671-82; PMID:18650841; http://dx.doi.org/10.1038/nrc2399
181. Dick FA, Rubin SM. Molecular mechanisms underlying RB protein function. Nat Rev Mol Cell Biol 2013; 14:297-306; PMID:23594950; http://dx.doi.org/10.1038/nrm3567
182. Mihara K, Cao XR, Yen A, Chandler S, Driscoll B, Murphree AL, T'Ang A, Fung YK. Cell cycle-dependent regulation of phosphorylation of the human retinoblastoma gene product. Science 1989; 246:1300-3; PMID:2588006; http://dx.doi.org/10.1126/science.2588006
183. Bagchi S, Weinmann R, Raychaudhuri P. The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. Cell 1991; 65:1063-72; PMID:1828393; http://dx.doi.org/10.1016/0092-8674(91)90558-G
184. Radivojac P, Vacic V, Haynes C, Cocklin RR, Mohan A, Heyen JW, Goebl MG, Iakoucheva LM. Identification, analysis, and prediction of protein ubiquitination sites. Proteins 2010; 78:365-80; PMID:19722269; http://dx.doi.org/10.1002/prot.22555
185. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, Obradovic Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res 2007; 6:1882-98; PMID:17391014; http://dx.doi.org/10.1021/pr060392u
186. Vucetic S, Xie H, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. J Proteome Res 2007; 6:1899-916; PMID:17391015; http://dx.doi.org/10.1021/pr060393m
187. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. J Proteome Res 2007; 6:1917-32; PMID:17391016; http://dx.doi.org/10.1021/pr060394e
188. Pan Z, Shang Y, Jia M, Zhang L, Xia C, Zhang M, Wang W, Wen W. Structural and biochemical characterization of the interaction between LGN and Frmpd1. J Mol Biol 2013; 425:1039-49; PMID:23318951; http://dx.doi.org/10.1016/j.jmb.2013.01.003
189. Yu F, Morin X, Cai Y, Yang X, Chia W. Analysis of partner of inscuteable, a novel player of Drosophila asymmetric divisions, reveals two distinct steps in inscuteable apical localization. Cell 2000; 100:399-409; PMID:10693757; http://dx.doi.org/10.1016/S0092-8674(00)80676-5
190. Siller KH, Doe CQ. Spindle orientation during asymmetric cell division. Nat Cell Biol 2009; 11:365-74; PMID:19337318; http://dx.doi.org/10.1038/ncb0409-365
191. Knoblich JA. Mechanisms of asymmetric stem cell division. Cell 2008; 132:583-97; PMID:18295577; http://dx.doi.org/10.1016/j.cell.2008.02.007
192. Zhu J, Wen W, Zheng Z, Shang Y, Wei Z, Xiao Z, Pan Z, Du Q, Wang W, Zhang M. LGN/mInsc and LGN/NuMA complex structures suggest distinct functions in asymmetric cell division for the Par3/mInsc/LGN and Galphai/LGN/NuMA pathways. Mol Cell 2011; 43:418-31; PMID:21816348; http://dx.doi.org/10.1016/j.molcel.2011.07.011
193. Culurgioni S, Alfieri A, Pendolino V, Laddomada F, Mapelli M. Inscuteable and NuMA proteins bind competitively to Leu-Gly-Asn repeat-enriched protein (LGN) during asymmetric cell divisions. Proc Natl Acad Sci U S A 2011; 108:20998-1003; PMID:22171003; http://dx.doi.org/10.1073/pnas.1113077108
194. Yuzawa S, Kamakura S, Iwakiri Y, Hayase J, Sumimoto H. Structural basis for interaction between the conserved cell polarity proteins Inscuteable and Leu-Gly-Asn repeat-enriched protein (LGN). Proc Natl Acad Sci U S A 2011; 108:19210-5; PMID:22074847; http://dx.doi.org/10.1073/pnas.1110951108
195. Blatch GL, Lassle M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays 1999; 21:932-9; PMID:10517866
196. Willard FS, Kimple RJ, Siderovski DP. Return of the GDI: the GoLoco motif in cell division. Annu Rev Biochem 2004; 73:925-51; PMID:15189163; http://dx.doi.org/10.1146/annurev.biochem.73.011303.073756
197. Du Q, Macara IG. Mammalian Pins is a conformational switch that links NuMA to heterotrimeric G proteins. Cell 2004; 119:503-16; PMID:15537540; http://dx.doi.org/10.1016/j.cell.2004.10.028
198. Yu F, Cai Y, Kaushik R, Yang X, Chia W. Distinct roles of Galphai and Gbeta13F subunits of the heterotrimeric G protein complex in the mediation of Drosophila neuroblast asymmetric divisions. J Cell Biol 2003; 162:623-33; PMID:12925708; http://dx.doi.org/10.1083/jcb.200303174

199. D'Andrea LD, Regan L. TPR proteins: the versatile helix. Trends Biochem Sci 2003; 28:655-62; PMID:14659697; http://dx.doi.org/10.1016/j.tibs.2003.10.007
200. An N, Blumer JB, Bernard ML, Lanier SM. The PDZ and band 4.1 containing protein Frmpd1 regulates the subcellular location of activator of G-protein signaling 3 and its interaction with G-proteins. J Biol Chem 2008; 283:24718-28; PMID:18566450; http://dx.doi.org/10.1074/jbc.M803497200
201. Dosztanyi Z, Csizmok V, Tompa P, Simon I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. J Mol Biol 2005; 347:827-39; PMID:15769473; http://dx.doi.org/10.1016/j.jmb.2005.01.071
202. Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 2005; 21:3433-4; PMID:15955779; http://dx.doi.org/10.1093/bioinformatics/bti541
203. Davenport AM, Huber FM, Hoelz A. Structural and Functional Analysis of Human SIRT1. J Mol Biol 2013; PMID:24120939
204. Donmez G, Guarente L. Aging and disease: connections to sirtuins. Aging Cell 2010; 9:285-90; PMID:20409078; http://dx.doi.org/10.1111/j.1474-9726.2010.00548.x
205. Guarente L. Sirtuins in aging and disease. Cold Spring Harb Symp Quant Biol 2007; 72:483-8; PMID:18419308; http://dx.doi.org/10.1101/sqb.2007.72.024
206. Longo VD, Kennedy BK. Sirtuins in aging and age-related disease. Cell 2006; 126:257-68; PMID:16873059; http://dx.doi.org/10.1016/j.cell.2006.07.002
207. Kang H, Suh JY, Jung YS, Jung JW, Kim MK, Chung JH. Peptide switch is essential for Sirt1 deacetylase activity. Mol Cell 2011; 44:203-13; PMID:22017869; http://dx.doi.org/10.1016/j.molcel.2011.07.038
208. Pan M, Yuan H, Brent M, Ding EC, Marmorstein R. SIRT1 contains N- and C-terminal regions that potentiate deacetylase activity. J Biol Chem 2012; 287:2468-76; PMID:22157016; http://dx.doi.org/10.1074/jbc.M111.285031
209. Costantini S, Sharma A, Raucci R, Costantini M, Autiero I, Colonna G. Genealogy of an ancient protein family: the Sirtuins, a family of disordered members. BMC Evol Biol 2013; 13:60; PMID:23497088; http://dx.doi.org/10.1186/1471-2148-13-60
210. Davies CW, Paul LN, Das C. Mechanism of recruitment and activation of the endosome-associated deubiquitinase AMSH. Biochemistry 2013; 52:7818-29; PMID:24151880; http://dx.doi.org/10.1021/bi401106b
211. McDonell LM, Mirzaa GM, Alcantara D, Schwartzentruber J, Carter MT, Lee LJ, Clericuzio CL, Graham JM, Jr., Morris-Rosendahl DJ, Polster T, et al. Mutations in STAMBP, encoding a deubiquitinating enzyme, cause microcephaly-capillary malformation syndrome. Nat Genet 2013; 45:556-62; PMID:23542699; http://dx.doi.org/10.1038/ng.2602
212. Maytal-Kivity V, Reis N, Hofmann K, Glickman MH. MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. BMC Biochem 2002; 3:28; PMID:12370088; http://dx.doi.org/10.1186/1471-2091-3-28
213. McCullough J, Row PE, Lorenzo O, Doherty M, Beynon R, Clague MJ, Urbe S. Activation of the endosome-associated ubiquitin isopeptidase AMSH by STAM, a component of the multivesicular body-sorting machinery. Current biology : CB 2006; 16:160-5; PMID:16431367; http://dx.doi.org/10.1016/j.cub.2005.11.073
214. Dinkel H, Van Roey K, Michael S, Davey NE, Weatheritt RJ, Born D, Speck T, Kruger D, Grebnev G, Kuban M, et al. The eukaryotic linear motif resource ELM: 10 years and counting. Nucleic Acids Res 2014; 42:D259-66; PMID:24214962; http://dx.doi.org/10.1093/nar/gkt1047
215. Freed DM, Lukasik SM, Sikora A, Mokdad A, Cafiso DS. Monomeric TonB and the Ton box are required for the formation of a high-affinity transporter-TonB complex. Biochemistry 2013; 52:2638-48; PMID:23517233; http://dx.doi.org/10.1021/bi3016108
216. Kohler SD, Weber A, Howard SP, Welte W, Drescher M. The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria. Protein Sci 2010; 19:625-30; PMID:20095050; http://dx.doi.org/10.1002/pro.345
217. Shultis DD, Purdy MD, Banchs CN, Wiener MC. Outer membrane active transport: structure of the BtuB:TonB complex. Science 2006; 312:1396-9; PMID:16741124; http://dx.doi.org/10.1126/science.1127694
218. Pawelek PD, Croteau N, Ng-Thow-Hing C, Khursigara CM, Moiseeva N, Allaire M, Coulton JW. Structure of TonB in complex with FhuA, E. coli outer membrane receptor. Science 2006; 312:1399-402; PMID:16741125; http://dx.doi.org/10.1126/science.1128057
219. Stark JL, Mehla K, Chaika N, Acton TB, Xiao R, Singh PK, Montelione GT, Powers R. Structure and function of human DnaJ homologue subfamily a member 1 (DNAJA1) and its relationship to pancreatic cancer. Biochemistry 2014; 53:1360-72; PMID:24512202; http://dx.doi.org/10.1021/bi401329a
220. Crnogorac-Jurcevic T, Gangeswaran R, Bhakta V, Capurso G, Lattimore S, Akada M, Sunamura M, Prime W, Campbell F, Brentnall TA, et al. Proteomic analysis of chronic pancreatitis and pancreatic adenocarcinoma. Gastroenterology 2005; 129:1454-63; PMID:16285947; http://dx.doi.org/10.1053/j.gastro.2005.08.012
221. Rumora AE, Wang SX, Ferris LA, Everse SJ, Kelm RJ, Jr. Structural basis of multisite single-stranded DNA recognition and ACTA2 repression by purine-rich element binding protein B (Purbeta). Biochemistry 2013; 52:4439-50; PMID:23724822; http://dx.doi.org/10.1021/bi400283r
222. Bergemann AD, Johnson EM. The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. Mol Cell Biol 1992; 12:1257-65; PMID:1545807
223. Bergemann AD, Ma ZW, Johnson EM. Sequence of cDNA comprising the human pur gene and sequence-specific single-stranded-DNA-binding properties of the encoded protein. Mol Cell Biol 1992; 12:5673-82; PMID:1448097
224. Johnson EM, Daniel DC, Gordon J. The pur protein family: genetic and structural features in development and disease. J Cell Physiol 2013; 228:930-7; PMID:23018800; http://dx.doi.org/10.1002/jcp.24237
225. Hotter D, Sauter D, Kirchhoff F. Emerging role of the host restriction factor tetherin in viral immune sensing. J Mol Biol 2013; 425:4956-64; PMID:24075872; http://dx.doi.org/10.1016/j.jmb.2013.09.029
226. Kupzig S, Korolchuk V, Rollason R, Sugden A, Wilde A, Banting G. Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. Traffic 2003; 4:694-709; PMID:12956872; http://dx.doi.org/10.1034/j.1600-0854.2003.00129.x
227. Swiecki M, Scheaffer SM, Allaire M, Fremont DH, Colonna M, Brett TJ. Structural and biophysical analysis of BST-2/tetherin ectodomains reveals an evolutionary conserved design to inhibit virus release. J Biol Chem 2010; 286:2987-97; PMID:21084286; http://dx.doi.org/10.1074/jbc.M110.190538
228. Thomsen ND, Berger JM. Running in reverse: the structural basis for translocation polarity in hexameric helicases. Cell 2009; 139:523-34; PMID:19879839; http://dx.doi.org/10.1016/j.cell.2009.08.043
229. Hammonds J, Wang JJ, Yi H, Spearman P. Immunoelectron microscopic evidence for Tetherin/BST2 as the physical bridge between HIV-1 virions and the plasma membrane. PLoS Pathog 2010; 6:e1000749; PMID:20140192
230. Peysselon F, Xue B, Uversky VN, Ricard-Blum S. Intrinsic disorder of the extracellular matrix. Mol Biosyst 2011; 7:3353-65; PMID:22009114; http://dx.doi.org/10.1039/c1mb05316g
231. Jahn SC, Law ME, Corsino PE, Rowe TC, Davis BJ, Law BK. Assembly, activation, and substrate specificity of cyclin D1/Cdk2 complexes. Biochemistry 2013; PMID:23627734
232. Nguyen D, Nickel M, Mizuguchi C, Saito H, Lund-Katz S, Phillips MC. Interactions of apolipoprotein A-I with high-density lipoprotein particles. Biochemistry 2013; 52:1963-72; PMID:23425306; http://dx.doi.org/10.1021/bi400032y
233. Koppaka V, Silvestro L, Engler JA, Brouillette CG, Axelsen PH. The structure of human lipoprotein A-I. Evidence for the "belt" model. J Biol Chem 1999; 274:14541-4; PMID:10329643
234. Segrest JP, Jones MK, Klon AE, Sheldahl CJ, Hellinger M, De Loof H, Harvey SC. A detailed molecular belt model for apolipoprotein A-I in discoidal high density lipoprotein. J Biol Chem 1999; 274:31755-8; PMID:10542194; http://dx.doi.org/10.1074/jbc.274.45.31755
235. Lund-Katz S, Phillips MC. High density lipoprotein structure-function and role in reverse cholesterol transport. Subcell Biochem 2010; 51:183-227; PMID:20213545; http://dx.doi.org/10.1007/978-90-481-8622-8_7
236. Rothblat GH, Phillips MC. High-density lipoprotein heterogeneity and function in reverse cholesterol transport. Curr Opin Lipidol 2010; 21:229-38; PMID:20480549; http://dx.doi.org/10.1097/MOL.0b013e328338472d
237. Wu Z, Gogonea V, Lee X, Wagner MA, Li XM, Huang Y, Undurti A, May RP, Haertlein M, Moulin M, et al. Double superhelix model of high density lipoprotein. J Biol Chem 2009; 284:36605-19; PMID:19812036; http://dx.doi.org/10.1074/jbc.M109.039537
238. Gunasekaran K, Tsai CJ, Nussinov R. Analysis of ordered and disordered protein complexes reveals structural features discriminating between stable and unstable monomers. J Mol Biol 2004; 341:1327-41; PMID:15321724; http://dx.doi.org/10.1016/j.jmb.2004.07.002
239. Jacobs JL, Coyne CB. Mechanisms of MAVS regulation at the mitochondrial membrane. J Mol Biol 2013; 425:5009-19; PMID:24120683; http://dx.doi.org/10.1016/j.jmb.2013.10.007
240. Khasnis M, Nakatomi A, Gumpper K, Eto M. Reconstituted human myosin light chain phosphatase reveals distinct roles of two inhibitory phosphorylation sites of the regulatory subunit, MYPT1. Biochemistry 2014; 53:2701-9; PMID:24712327; http://dx.doi.org/10.1021/bi5001728
241. Li Y, Jin K, Ghosh S, Devarakonda P, Carlson K, Davis A, Stewart KA, Cammett E, Pelczar Rossi P, Setlow B, et al. Structural and functional analysis of the GerD spore germination protein of Bacillus species. J Mol Biol 2014; 426:1995-2008; PMID:24530795; http://dx.doi.org/10.1016/j.jmb.2014.02.004
242. Setlow P. Spores of Bacillus subtilis: their resistance to and killing by radiation, heat and chemicals. J Appl Microbiol 2006; 101:514-25; PMID:16907802; http://dx.doi.org/10.1111/j.1365-2672.2005.02736.x
243. Paredes-Sabja D, Setlow P, Sarker MR. Germination of spores of Bacillales and Clostridiales species: mechanisms and proteins involved. Trends Microbiol 2011; 19:85-94; PMID:21112786; http://dx.doi.org/10.1016/j.tim.2010.10.004
244. Setlow P. Spore germination. Curr Opin Microbiol 2003; 6:550-6; PMID:14662349; http://dx.doi.org/10.1016/j.mib.2003.10.001

245. Moir A. How do spores germinate? J Appl Microbiol 2006; 101:526-30; PMID:16907803; http://dx.doi.org/10.1111/j.1365-2672.2006.02885.x
246. Ross C, Abel-Santos E. The Ger receptor family from sporulating bacteria. Curr Issues Mol Biol 2010; 12:147-58; PMID:20472940
247. Forer N, Korman M, Elharar Y, Vishkautzan M, Gur E. Bacterial proteasome and PafA, the pup ligase, interact to form a modular protein tagging and degradation machine. Biochemistry 2013; 52:9029-35; PMID:24228735; http://dx.doi.org/10.1021/bi401017b
248. Borodianskiy-Shteinberg T, Kalt I, Kipper S, Nachum N, Katz S, Pauker MH, Barda-Saad M, Gerber D, Sarid R. The nucleolar PICT-1/GLTSCR2 protein forms homo-oligomers. J Mol Biol 2014; 426:2363-78; PMID:24735870; http://dx.doi.org/10.1016/j.jmb.2014.04.006
249. Sasaki M, Kawahara K, Nishio M, Mimori K, Kogo R, Hamada K, Itoh B, Wang J, Komatsu Y, Yang YR, et al. Regulation of the MDM2-P53 pathway and tumor growth by PICT1 via nucleolar RPL11. Nat Med 2011; 17:944-51; PMID:21804542; http://dx.doi.org/10.1038/nm.2392
250. Okahara F, Ikawa H, Kanaho Y, Maehama T. Regulation of PTEN phosphorylation and stability by a tumor suppressor candidate protein. J Biol Chem 2004; 279:45300-3; PMID:15355975; http://dx.doi.org/10.1074/jbc.C400377200
251. Spears JL, Xiao X, Hall CK, Agris PF. Amino acid signature enables proteins to recognize modified tRNA. Biochemistry 2014; 53:1125-33; PMID:24483944; http://dx.doi.org/10.1021/bi401174h
252. Vendeix FA, Murphy FVt, Cantara WA, Leszczynska G, Gustilo EM, Sproat B, Malkiewicz A, Agris PF. Human tRNA(Lys3)(UUU) is pre-structured by natural modifications for cognate and wobble codon binding through keto-enol tautomerism. J Mol Biol 2011; 416:467-85; PMID:22227389; http://dx.doi.org/10.1016/j.jmb.2011.12.048
253. Schmalzbauer E, Strack B, Dannull J, Guehmann S, Moelling K. Mutations of basic amino acids of NCp7 of human immunodeficiency virus type 1 affect RNA binding in vitro. J Virol 1996; 70:771-7; PMID:8551614
254. Poon DT, Wu J, Aldovini A. Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. J Virol 1996; 70:6607-16; PMID:8794295
255. Frankel AD, Young JA. HIV-1: fifteen proteins and an RNA. Annu Rev Biochem 1998; 67:1-25; PMID:9759480; http://dx.doi.org/10.1146/annurev.biochem.67.1.1
256. Huang Y, Khorchid A, Wang J, Parniak MA, Darlix JL, Wainberg MA, Kleiman L. Effect of mutations in the nucleocapsid protein (NCp7) upon Pr160(gag-pol) and tRNA(Lys) incorporation into human immunodeficiency virus type 1. J Virol 1997; 71:4378-84; PMID:9151827
257. Guo J, Henderson LE, Bess J, Kane B, Levin JG. Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. J Virol 1997; 71:5178-88; PMID:9188585
258. Cameron CE, Ghosh M, Le Grice SF, Benkovic SJ. Mutations in HIV reverse transcriptase which alter RNase H activity and decrease strand transfer efficiency are suppressed by HIV nucleocapsid protein. Proc Natl Acad Sci U S A 1997; 94:6700-5; PMID:9192628; http://dx.doi.org/10.1073/pnas.94.13.6700
259. Carteau S, Batson SC, Poljak L, Mouscadet JF, de Rocquigny H, Darlix JL, Roques BP, Kas E, Auclair C. Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg2+-dependent DNA integration in vitro. J Virol 1997; 71:6225-9; PMID:9223522
260. Summers MF, Henderson LE, Chance MR, Bess JW, Jr., South TL, Blake PR, Sagi I, Perez-Alvarado G, Sowder RC, 3rd, Hare DR, et al. Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. Protein Sci 1992; 1:563-74; PMID:1304355; http://dx.doi.org/10.1002/pro.5560010502
261. Morellet N, Jullian N, De Rocquigny H, Maigret B, Darlix JL, Roques BP. Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR. EMBO J 1992; 11:3059-65; PMID:1639074
262. Xue B, Mizianty MJ, Kurgan L, Uversky VN. Protein intrinsic disorder as a flexible armor and a weapon of HIV-1. Cell Mol Life Sci 2012; 69:1211-59; PMID:22033837; http://dx.doi.org/10.1007/s00018-011-0859-3