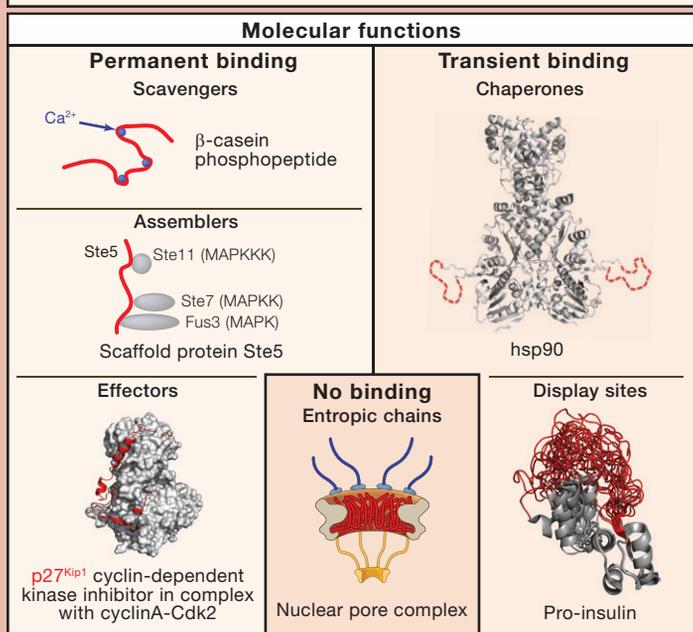
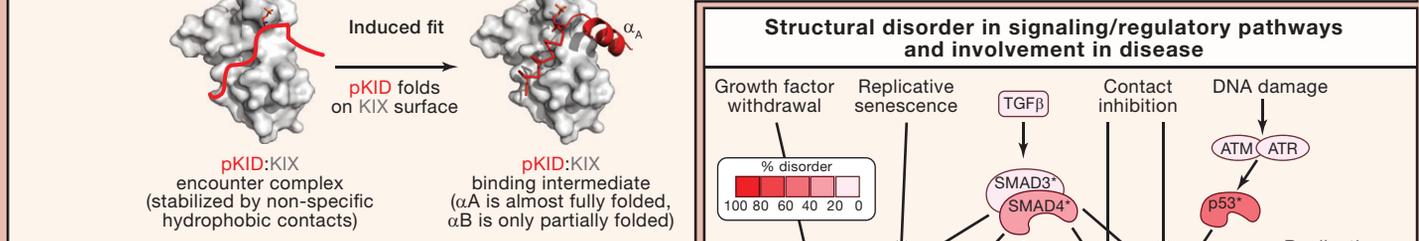
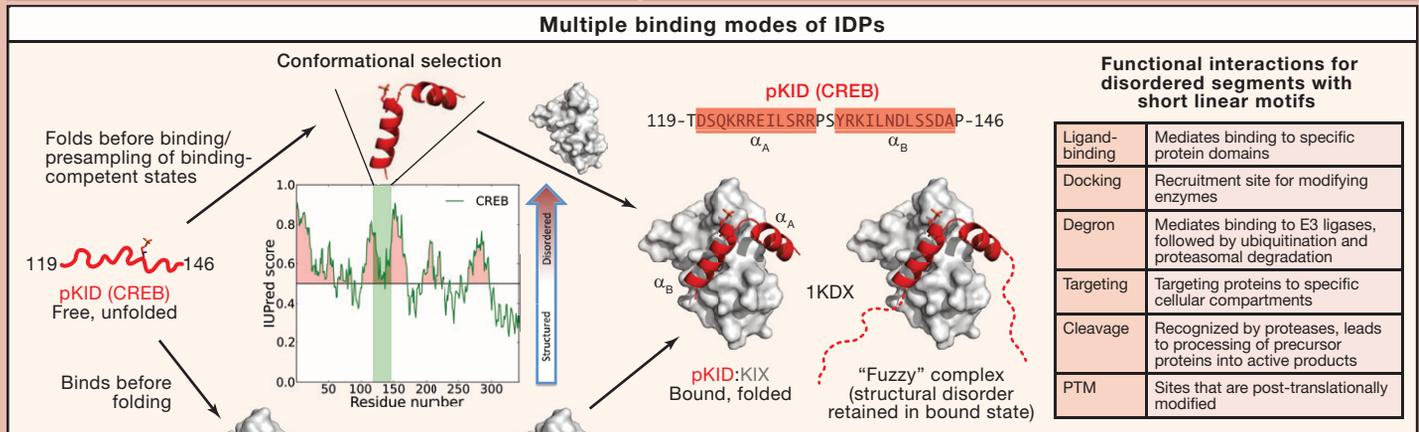
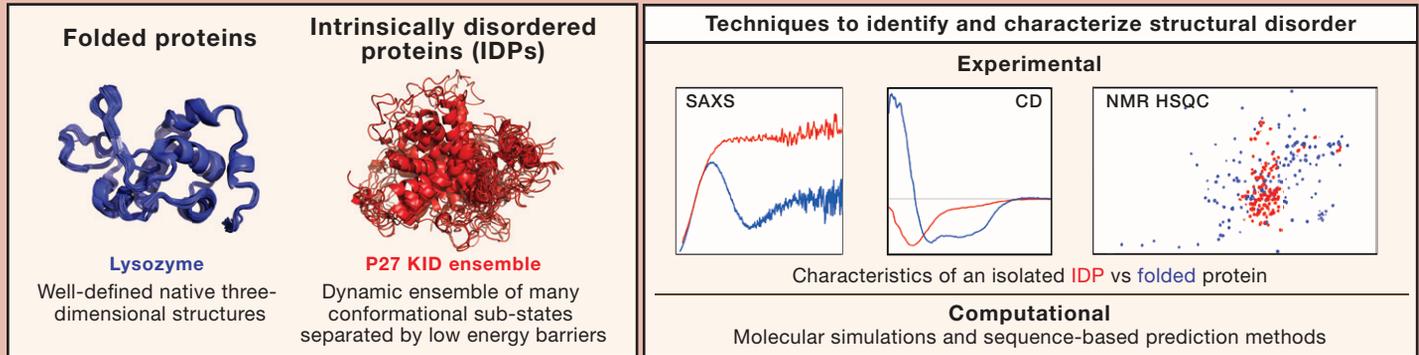


SnapShot: Intrinsic Structural Disorder

Mainak Guharoy,^{1,2} Kris Pauwels,^{1,2} and Peter Tompa^{1,2,3}

¹VIB Structural Biology Research Center (SBRC), Vlaams Instituut voor Biotechnologie, 1050 Brussel, Belgium,

²Structural Biology Brussels (SBB), Vrije Universiteit Brussel, 1050 Brussel, Belgium, ³Institute of Enzymology, Research Centre for Natural Sciences of the Hungarian Academy of Sciences, 1117 Budapest, Hungary



Protein functions	Associated IDPs	Disease
Transcription functions	p53, c-Myc	Cancer
Chromatin-organizing proteins	High mobility group proteins	Breast, colon, gastric, and gastrointestinal cancer
Signaling adaptors and scaffolds	BRCA-1, titin, MAP2, axin	Cancer, cardiomyopathy
Cytoskeletal proteins	Tau protein, MAP2	Alzheimer's disease, tauopathies
Cell-cycle regulation	p21, p27Kip1, inhibitor-2	Breast and prostate cancer, Alzheimer's disease

SnapShot: Intrinsic Structural Disorder **Cell**

Mainak Guharoy,^{1,2} Kris Pauwels,^{1,2} and Peter Tompa^{1,2,3}

¹VIB Structural Biology Research Center (SBRC), Vlaams Instituut voor Biotechnologie, 1050 Brussel, Belgium,

²Structural Biology Brussels (SBB), Vrije Universiteit Brussel, 1050 Brussel, Belgium, ³Institute of Enzymology, Research Centre for Natural Sciences of the Hungarian Academy of Sciences, 1117 Budapest, Hungary

Many proteins (intrinsically disordered proteins, IDPs) or regions of proteins (intrinsically disordered regions, IDRs) lack a well-defined 3D structure under physiological conditions. Albeit unfolded and highly dynamic, these proteins are not denatured; rather, intrinsic structural disorder is their native, functional state. Structural disorder reaches high proportions in higher eukaryotes (35% of human proteins are predicted to possess IDRs consisting of at least 30 consecutive disordered residues). Disordered proteins play important roles in cell signaling and regulation, due to which IDPs/IDRs are also often involved in diseases.

Evidence for Structural Disorder of Proteins

There are multiple lines of evidence that structural disorder is the native state of IDPs. The primary techniques used for their characterization are NMR, SAXS, CD, and other biophysical techniques. Their results can be integrated using computational tools, which can yield structural ensembles, currently the best description of the unusual, highly dynamic structural state of IDPs. Also instrumental in studying structural disorder are bioinformatics tools, which can assess the ordered or disordered state based on an amino acid sequence. Data related to structural disorder are deposited in databases, such as PED (structural ensembles), DisProt (sequences and metadata), and MobiDB (predictions of disorder and other functional features). Experimental observations suggest that disorder encompasses a broad spectrum of distinct states, ranging from fully disordered to almost folded conformations.

Function by Induced Folding

The function of IDPs, intimately linked with their disorder, results either directly from their disordered state or from molecular recognition, when they undergo folding induced by interaction with their partner (another protein, DNA, RNA, or membrane). Induced folding (or disorder-to-order transitions), as observed for p53 binding to MDM2, E-cadherin binding to β -catenin, p27 binding to cyclin A/Cdk2, or the KID domain of CREB binding to the KIX domain of CBP, for example, may entail different advantages such as fast binding kinetics (also termed fly casting), regulation by post-translational modifications (PTMs), structural adaptability to multiple partners (functional pleiotropy, or moonlighting), and “uncoupling” specificity from binding strength. Disorder-to-order transition upon binding may proceed by two extreme mechanisms, conformational selection or induced fit (binding-induced folding), as illustrated by the example of the CREB KID binding to CBP KIX.

Peptide Motifs—Recognition Elements of IDPs

The interactions of IDPs/IDRs are most often mediated by short interaction modules, referred to as peptide motifs (also called short linear motifs, SLiMs, or eukaryotic linear motifs, ELMs). Peptide motifs are typically <10 residues in length, they prefer to be located within protein IDRs, and based on the presence of a few highly conserved, specificity-determining residues, they are recognized and/or modified by structured domains of their interacting partners. There are six different functional outcomes of motif recognition, including targeting, regulated degradation, or post-translational modifications. Due to their short length, peptide motifs usually engage in weak and transient interactions.

Molecular Mechanisms of IDPs

Intrinsically disordered proteins cannot have functions typical for folded proteins, such as enzymatic activity, and their very existence defies the classical structure-function paradigm. Sometimes their function stems directly from the disordered state, without partner binding, termed the entropic chain mechanism (e.g., in the nuclear pore complex). An entropic chain can generate force, allow functional inter-domain motions, reach out for remote binding partners, or simply keep proteins apart using a spacer mechanism. Very often, however, IDPs function by molecular recognition, when their motifs (SLiMs) or longer regions (disordered domains) undergo a folding transition induced by the partner. When binding is transient, it can lead to modification (display site, e.g., pro-insulin) or chaperone actions (e.g., Hsp90). In case of permanent binding, disorder results in the assembly of multi-protein complexes (e.g., scaffold proteins like Ste5), modification of the activity of the partner (e.g., p27 inhibiting Cdk2/cyclin A), or scavenging of small molecules (e.g., casein binding to calcium phosphate).

Structural Disorder in Signaling, Regulation, and Disease

The functional advantages of structural disorder, such as fast binding kinetics, facile regulation by PTMs, high functional density, and specificity without strong binding (reversibility), are all potentially beneficial in signaling and regulatory tasks, when fast responses to changes in the external or internal environment are critical for cellular fitness and survival. Accordingly, structural disorder abounds in proteins with signaling and regulatory functions (as illustrated by the G1-S restriction point pathway, proteins with at least one long disordered region marked by asterisk). For example, disorder can reach very high levels in transcription factors, chromatin-organizing proteins, signaling adaptors and scaffolds, cytoskeletal proteins, and cell-cycle regulatory proteins. For apparently the very same reasons, higher eukaryotes that rely on complex regulatory decisions to harmonize the operation of their multiple cell types possess high levels of disorder (about 30% of their residues fall into disordered regions). Due to its prominent roles in signaling, structural disorder also reaches high levels in disease-associated proteins, and some of the most-studied proteins, such as α -synuclein, tau protein, c-Myc, and p53, are noted for their high disorder content, due to which IDPs are also attractive targets in drug development efforts.

ABBREVIATIONS

Cdk, cyclin-dependent kinase; CREB, cyclic-AMP response element-binding protein; ELM, eukaryotic linear motif; MDM2, murine-double minute 2; PTM, post-translational modification; SAXS, small-angle X-ray scattering; SLiM, short linear motif.

ACKNOWLEDGMENTS

This work was supported by the Odysseus grant G.0029.12 from Research Foundation Flanders (FWO.). M.G. is the recipient of an Omics@VIB Marie Curie COFUND fellowship; K.P. is the recipient of a FWO Pegasus long-term postdoctoral fellowship.

REFERENCES

- Krishnan, N., Koveal, D., Miller, D.H., Xue, B., Akshinthala, S.D., Kragelj, J., Jensen, M.R., Gauss, C.M., Page, R., Blackledge, M., et al. (2014). *Nat. Chem. Biol.* 10, 558–566.
- Kriwacki, R.W., Hengst, L., Tennant, L., Reed, S.I., and Wright, P.E. (1996). *Proc. Natl. Acad. Sci. USA* 93, 11504–11509.
- Schwalbe, M., Ozenne, V., Bibow, S., Jaremko, M., Jaremko, L., Gajda, M., Jensen, M.R., Biernat, J., Becker, S., Mandelkow, E., et al. (2014). *Structure* 22, 238–249.
- Sugase, K., Dyson, H.J., and Wright, P.E. (2007). *Nature* 447, 1021–1025.
- Tompa, P. (2011). *Curr. Opin. Struct. Biol.* 21, 419–425.
- Tompa, P., Davey, N.E., Gibson, T.J., and Babu, M.M. (2014). *Mol. Cell* 55, 161–169.
- Uversky, V.N., Oldfield, C.J., and Dunker, A.K. (2008). *Annu. Rev. Biophys.* 37, 215–246.
- van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., et al. (2014). *Chem. Rev.* 114, 6589–6631.
- Wright, P.E., and Dyson, H.J. (2009). *Curr. Opin. Struct. Biol.* 19, 31–38.
- Wright, P.E., and Dyson, H.J. (2015). *Nat. Rev. Mol. Cell Biol.* 16, 18–29.