Can proteins be intrinsically disordered inside a membrane?

Magnus Kjaergaard*

Interdisciplinary NANO Science Center (iNANO); Aarhus University; Aarhus, Denmark

Intrinsically disorder has evolved in many soluble proteins because it confers a unique set of functional advantages. In contrast, the functions of membrane proteins are largely understood in terms of well-defined structures. This raises the question: Why would the evolutionary pressures that select for disorder leave membrane proteins untouched. In this hypothesis piece, I argue that intrinsic disorder may exist in membrane embedded proteins, but that it will take a different form due to the different environment. Disordered membrane proteins are thus likely to have fully formed secondary structure, but little tertiary structure. Furthermore, the sequence signature for disorder in membrane proteins is likely to be reversed; so disordered proteins are more hydrophobic than their folded counterparts. At present it is impossible to tell how common this type of disordered membrane protein is.

The protein structure-function paradigm has explained the function of many proteins, however not all proteins have a well-defined native structure. Instead, some proteins have an ensemble of rapidly interconverting conformations as their native state. The idea that such a protein could be functional was originally met with skepticism, but it is now established that these proteins are an integral part of the molecular machinery of the cell. This type of proteins is now called intrinsically disordered proteins (IDPs).¹ The community-wide adoption of the IDP term has driven home the point that these proteins are not anomalies, but belong to a large class of proteins, whose properties can be studied as a group.² This has led to the development of specific tools to identify

IDPs, and to study their structural properties *in vitro*³ and in live cells.⁴ As a result, we now know that IDPs are common,⁵ carry out important biological functions⁶ and are often related to disease.⁷

Protein disorder has not evolved in a third of all proteins for the amusement of scientists. The frequent occurrence of protein disorder can only be explained by a functional role. Several functions have been identified that take advantage of the structural flexibility of IDPs: Firstly, the absence of a rigid structure means that the protein can adapt to a large number of binding partners and as a result IDPs are often promiscuous in their ligand recognition. This allows them to function as signaling hubs, where they integrate different pathways.⁸ Secondly, flexible proteins are good at penetrating the active sites of enzymes and IDPs are thus more frequently modified post-translationally than folded proteins.⁹ As post-translational modifications readily change the structural ensembles of an IDPs, this can act as a molecular switch turning functions on or off. Thirdly, accessibility to enzymes is also important to IDs role in regulating degradation, where IDPs on average are degraded more quickly than folded proteins,¹⁰ probably because the proteasome initiates degradation at an unfolded chain.¹¹ Tuning the degree of folding thus provide a mechanism for tuning the half-life of a protein. Finally, ID regions act as flexible linkers connecting functional subunits of large proteins. The length and flexibility of linkers regions thus directly affect intra- and intermolecular interactions. This function of IDPs is often referred to as an entropic spacer. While this list is not exhaustive, it demonstrates that ID is common in proteins because it has important functional roles.

Keywords: binding promiscuity, hydrophobicity, Intrinsic disorder, membrane protein, sequence signature

Abbreviations: CD, Circular dichroism; ID, Intrinsic disorder; IDP, Intrinsically disordered protein; TM, trans-membrane.

*Correspondence to: Magnus Kjaergaard; Email: magnus@inano.au.dk

Submitted: 10/15/2014

Accepted: 10/15/2014

<http://dx.doi.org/10.4161/21690707.2014.984570>

About a fourth of human genes codes for integral membrane proteins. However, the recent avalanche of studies describing intrinsic disorder has focused on water-soluble proteins or the extra-membranous parts of membrane proteins, and thus ignored the parts of proteins embedded in membranes. Bioinformatic studies show that ID is common on the cytoplasmic side of α -helical membrane proteins^{12,13} and in the loops of β -barrel membrane proteins.¹⁴ Inside the membrane, however, the structure-function paradigm still reigns supreme. This leads to the central question of this hypothesis piece: Given that functional advantages of ID has been selected for in the soluble proteome, why should this selective pressure leave the membrane embedded domains of integral membrane proteins untouched?

Does a Membrane Permit Disorder?

The sequence signature of soluble IDPs is tightly connected to the properties of water. Water has a high dielectric constant and thus stabilizes exposed polar groups, and forces burial of hydrophobic groups in the interior of the protein. Sequestration of hydrophobic groups from the aqueous solvent provides the main driving force in protein folding, and the low hydrophobicity of soluble IDPs thus implies a lack of thermodynamic impetus for folding. This is exacerbated by the high frequency of charged residues, which often result in charge-charge repulsion. IDPs are thus disordered because they have a combination of a low hydrophobicity and a high net charge.¹⁵

In contrast, the properties of biological membranes are radically different. The hydrophobic lipid tails of the membrane interior interact favorably with exposed hydrophobic groups, but impose a steep energetic penalty on exposed polar groups.¹⁶ Hydrophilic groups need to be bonded intra-molecularly and ideally sequestered from the lipids. This suggests that the sequence signature of putative membrane embedded IDPs will differ from that of soluble IDPs: While folded proteins in an aqueous environment sequester hydrophobic groups in the

interior, membrane proteins sequester polar groups. A putative disordered membrane embedded domain cannot do this, and a disordered membrane embedded domains should thus be more hydrophobic than their folded counterparts.

At this stage it is worthwhile to consider how disordered a protein can possibly be inside a membrane. Soluble IDPs form a continuum of structure ranging from random coils over pre-molten globules and molten globules (Fig 1).^{17,18} In the random coil state, there are few intramolecular hydrogen bonds, whereas these are common in molten globules. Unsatisfied hydrogen bonds are highly unfavorable inside the hydrophobic interior of the membrane.¹⁶ A membrane embedded random coil with dozens of unsatisfied hydrogen bonds is thus close to thermodynamically impossible. Putative membrane embedded IDPs will thus minimally have to form backbone hydrogen bonds internally, most likely in an α -helical structure, and thus be analogous to soluble pre-molten globules or molten globules (Fig. 1). In soluble proteins, these states are commonly regarded as IDPs. Since a pre-molten globule, i.e. a protein with secondary structure elements but no tertiary structure, is the maximally permitted degree of disorder for a membrane-embedded protein, membrane proteins that resemble pre-molten globules should also be regarded as IDPs.

While fully disordered membrane-embedded chains are unlikely, poly-topic membrane proteins consisting of independent trans membrane (TM) helices are entirely possible. In fact, freely diffusing TM-helices has been postulated as a general intermediate in the folding of α -helical membrane proteins,¹⁹ and it may thus be a state that most membrane proteins have to go through. A single TM-helix is the most common membrane domain, and helices that are stable as separate entities are thus commonplace in nature. The TM-helices of bitopic membrane proteins can thus reveal the sequence signature of disorder inside the membrane. The TM-helices of bitopic membrane proteins are significantly more hydrophobic than TM-helices from their poly-topic counterpart,²⁰ which strengthens the hypothesis that any putative disordered polytopic

membrane proteins will also have more hydrophobic TM-helices. This is in spite of the many bitopic membrane proteins where the TM-helices dimerize.

Functional Advantages of Disorder in The Membrane

What could the functional advantage of disorder be in a membrane protein? Many advantages have been proposed for soluble IDPs. The question is thus how many of these functions could be fulfilled by a fluctuating ensemble of TM-helices (Fig. 2):

- I. Binding promiscuity is useful to membrane proteins and soluble proteins alike. Fluctuating ensembles of helices have been found to allow binding to structurally dissimilar ligands for soluble proteins,^{21,22} and it seems plausible that a membrane receptor with a flexible transmembrane domain will be quite plastic as well.
- II. Post-translational modification often requires the peptide chain to enter a cleft to reach the active site of the enzyme. For soluble IDPs, this means that they are more readily modified than their folded counterparts.⁹ This seems likely to be the case for any putative membrane embedded IDPs as well.
- III. Many soluble IDPs are only disordered a part of the time, and thus undergo a folding transition regulated by external stimuli. A folding transition provides a general mechanism for switching between an on- and an off-state, or switching between a small or a large separation between different parts of the molecule. This is likely to be as useful for membrane proteins as for soluble proteins.
- IV. Entropic spacers tune the thermodynamics of association between 2 parts of a large protein or ligands bound to different parts of a large protein. Disorder in the membrane embedded domain could contribute to the entropic spacer effect of the soluble loops.

V. It is less clear whether disorder in a membrane embedded domain will alter its degradation kinetics. The degradation pathway of plasma membrane proteins is different from soluble proteins and initially requires targeting to a lysosome. Lysosome targeting is regulated by signals and in the intracellular domain,²³ and furthermore by ubiquitination.²⁴ There are conflicting reports on how disorder affects ubiquitination^{25,26} and it is unclear how applicable studies of ubiquitination of soluble proteins are to membrane proteins.

In total, many of the functional advantages of disorder could apply to soluble and membrane proteins alike, which suggest that ID could be selected for membrane proteins as well.

Dynamics vs. Disorder

While intrinsic disorder is relatively unknown in the membrane protein field, the role of structural dynamics has long been appreciated. What is the difference between a dynamic protein and a disordered protein? The IDP term has never been rigidly defined, and there is a smooth transition from highly structured, disordered proteins to folded proteins. There is a distinction between the dynamics that has been described for many membrane proteins, and the type of disorder considered here. Many membrane transport proteins work by alternating between an outwards open, an inwards open and occluded states such as e.g. the P-type ATPases or members of the major facilitator superfamily. These proteins can be accurately described by a few well-defined conformations and should thus not be considered disordered. The difference between a dynamic, but structured protein and a disordered protein is best understood in terms of energy landscapes (Fig. 3). The dynamic membrane transporters have a few local minima with steep sides. A truly disordered protein, in contrast, will have a wide energy well with many local minima and low energy barriers. In other words, membrane transporters can be

Figure 1. The disorder-to-order continuum for soluble and membrane embedded proteins. Soluble proteins can form roughly be divided into random coil-like chains, pre-molten globule states with isolated elements of secondary structure, molten globules with fully formed secondary structure but fluctuating tertiary structure, and finally fully folded states. Membrane embedded domains can in principle for all but the fully disordered chains as the large number of unsatisfied hydrogen bonds in these states are extremely unfavorable. The states are only stereotypical examples as proteins can fall anywhere on the continuum from disordered to fully folded.

Figure 2. Functional advantages of disorder in membrane embedded domains. Several of the functional advantages identified for soluble IDPs are likely to disordered membrane embedded domains: (A) Disordered proteins can fold into diverse folded states upon binding to ligands. (B) Equilibrium between a folded and a disordered state provides a general mechanism for switching between an active and an inactive site. (C) The disordered state is more accessible for post-translational modifications as a flexible chain can more easily penetrate into the active site. (D) Disordered chains can act as an entropic spacer between either ligands or protein domains. (E) It is unclear whether disorder in a membrane protein will lead to faster degradation.

Figure 3. Cartoon representations of stereotypical energy landscape of proteins with different dynamics. A rigid protein can be described by an energy funnel with a single well and steep sides. A protein can have multiple local energy minima without being disordered. If a few local minima are separated by significant free energy barriers, the protein is better described as being dynamic with a few discrete well-defined states. A truly disordered protein has many conformations with similar energy and low energy barriers between them.

described by discrete states, whereas disordered proteins occupy a continuous ensemble.

Association of TM Helices

For decades, biophysicists and structural bioinformaticians have posed the question: Why do TM-helices pack together the way that they do? This is essentially the opposite question of the one posed here: Under which circumstances would TM helices fail to pack together? The packing principles identified to answer the first question are thus useful for the emerging field of membrane IDPs, as they define what is absent in disordered TM-helices.

In model systems of TM interactions, a single polar residue is sufficient to cause helix-helix association.^{27,28} The packing of these helices is probably not sufficiently specific to form a rigid crystallizable structure, but will retain considerable flexibility. All polar residues are, however, not created equal. An important distinction is thus whether the side chain can form hydrogen bonds to the main chain of the backbone or not. Especially, serine and threonine form side chain to backbone hydrogen bonds in an α -helical context,²⁹ and is thus tolerated in TM helices as illustrated by their intermediate position in biological hydrophobicity scales.³⁰ This suggests that these residues may also be tolerated in putative disordered membrane proteins, and thus that the signature for disorder may primarily to avoid the most hydrophilic residues.

GXXXG motifs are frequently found in helix-helix packing interactions,³¹ and are sufficient to induce dimerization of model TM helices.³² The absent side chain in glycine residues allows a close approach of the TM helices and formation of favorable van der Waals interactions and CH backbone hydrogen bonds. Additionally, the variation between large and small side chains creates a rough surface that is difficult for lipids to pack against. A rough surface on a helix will thus force the surrounding lipids to either adopt specific packing interactions or leave open cavities, both of which are entropically costly.^{16,33} Removal of cavities by close packing of rough trans-membrane helices potentially favors helix-helix associations, which may be an important driving force in GXXXG packing motifs. The presence of known helix-helix association motifs such as GXXXG motifs are thus likely to be avoided in disordered membrane proteins.

Where are Membrane Embedded IDPs Likely to Be Found?

Bioinformatic predictors can reliably predict which soluble proteins are disordered. These predictors are unlikely to work for membrane embedded domains as the sequence signature for disorder will be different. So while it is generally accepted that ID is common for soluble proteins, there is currently no reliable method for estimating how common it is inside the membrane. Despite a tremendous effort by the structural biology

community, we still only have atomic resolution structures for a few percent of known membrane proteins, although this number is increasing rapidly. There is thus still ample room for a significant portion of the membrane proteome to be disordered.

A good hunter knows the habits of the prey and thus optimizes the chance of success by looking where it is most likely to be found. Likewise, an experimental search for membrane IDPs needs to consider where disorder is likely to be found. But what traits are likely to be associated with disorder?

- I. Many proteins, soluble and membrane embedded alike, fall in to highly conserved families as the need to maintain a folded structure constrains the available sequence space. These constraints are relaxed for disordered proteins and they are generally less evolutionarily conserved than folded proteins.³⁴ Poorly conserved proteins are more likely to be disordered. Bioinformatic studies suggest that 9% of known polytopic membrane proteins do not fall in an identifiable family,³⁵ and many more fall in small families.
- II. Eukaryotes generally have a higher proportion of IDPs than prokaryotes.⁵ Eukaryotic membrane proteins without prokaryotic orthologs are thus have higher *a priori* probability of ID.
- III. Comparison between databases of known IDPs and functional keywords have shown that proteins involved in signaling and regulation⁶ are more likely to be disordered. This seems to be likely to be the case for putative membrane embedded IDPs as well.
- IV. Disorder is common in water exposed parts of membrane proteins, and disorder in soluble regions adjacent to the membrane will thus increase the probability of disorder in the membrane embedded domain by at least 2 mechanisms: The entropic cost of forming an interaction between 2 TM helices will be greater if the intervening linker is disordered as more degrees of freedom are lost

upon folding. Furthermore, soluble disordered regions often have a high net charge, and forcing these linkers together is thus electrostatically unfavorable.

The above criteria and the *a priori* expectations for the sequence signature of membrane protein disorder thus provide a good foundation for experimental identification of any putative membrane IDPs

Experimentally Characterized Hyperflexible Membrane Proteins

Structural characterization of integral membrane proteins is a difficult problem in its own right, and it is much more difficult for flexible systems. There are thus only few published studies demonstrating hyperflexible proteins inside membranes. There are certainly too few studies currently to suggest that it represents a general phenomenon, however, it worth emphasizing a few recent examples where proteins were found to be highly dynamic as inspiration for how such proteins can be described.

Are there any known examples of hyperflexible membrane proteins? One prominent example of what could be considered an integral membrane molten globule protein, the mitochondrial protein TSPO was recently studied by NMR. In the absence of ligands, the NMR signals were broadened and displayed poor dispersion, while retaining an α -helical CD spectrum.³⁶ Upon binding to a ligand, TSPO folds as a five-helical bundle, which confirms that the protein is functional.³⁷ Interestingly, the helices are highly hydrophobic and stable as isolated helices³⁶ apparently in agreement with the suggested characteristics of membrane embedded IDPs delineated above.

The C99 fragment of amyloid precursor protein has only a single TM helix with 2 associated membrane associated but not membrane spanning helices. These helices do not pack in a single unique structure, but populates a flexible ensemble.³⁸ Interestingly, this flexible state may facilitate cleavage of the TM helix by the intramembranous protease γ -secretase. This cleavage event releases

the amyloid β peptide, and is thus a key event in Alzheimer disease.

How Can We Characterize Membrane Embedded IDPs?

Structural characterization of disordered proteins is challenging due to their conformational heterogeneity. Crystallography has been the workhorse of structural biology and has recently started to make inroads into the membrane protein field. This technique is unlikely to play a significant role for IDPs as it requires proteins to have a well-defined structure. For soluble proteins, protein NMR is the most powerful method for characterizing IDPs at atomic resolutions,³⁹⁻⁴¹ although it is not obvious that NMR will play a similar role for membrane embedded IDPs. For soluble IDPs, the exchange between individual conformations usually falls into the fast exchange regime, which results in sharp and intense NMR signals. The more viscous environment of the membrane is likely to slow down transitions between conformations with the risk of pushing the system into the intermediate exchange regime, which has made NMR characterization of relatively folded disordered states such as molten globules difficult.⁴² Furthermore, the addition of lipids dramatically increase the tumbling time of the protein with associated problems for solutions NMR which is a difficult problem even for well-behaved membrane proteins.⁴³ However, if the conformational exchange falls in appropriate time scales, NMR can be very powerful as described in the preceding section.

Tertiary interactions are likely to be the main difference between a folded and disordered membrane domain. While random coil signal in the far-UV circular dichroism (CD) spectrum is a hallmark of a soluble IDP, this is unlikely to be the case for membrane embedded IDPs as the membrane embedded peptide chain has to form secondary structure to satisfy its hydrogen bonds. In contrast, the near UV region of the CD spectrum reports on the local environment of aromatic groups. A featureless near-UV spectrum is characteristic of a fluctuating tertiary structure and may thus be diagnostic for membrane

protein disorder. Furthermore, techniques that are intended for describing heterogeneous structural ensembles are likely to be appropriate. Examples of such techniques are single molecule FRET, time resolved FRET and electron paramagnetic resonance techniques. The soluble IDP field initially benefitted from methods for describing denatured states of folded proteins derived from the field of protein folding. The denatured states of membrane proteins are largely unexplored, however the possibility developing methods for membrane embedded IDPs may provide an additional incentive to study these denatured states as well.

Anyone working on ID will sooner or later face the question: "Are you sure that your protein is not just disordered because you have denatured it?" The field has largely answered this question by demonstrating the functionality of these states and furthermore deducing a clear sequence signature for disorder. For putative membrane embedded IDPs this problem is even more daunting. While native aqueous environment of soluble proteins is simple to mimic *in vitro*, biophysical studies of membrane proteins are often performed in an environment very different from the native membrane, and many membrane proteins denature during purification and reconstitution. It is thus even more important to validate the functionality of proposed membrane embedded IDPs by functional assays.

Conclusion

In summary, there is no *a priori* reason that membrane embedded proteins could not be disordered. Furthermore, at least some of the functional advantages of disorder in soluble proteins are likely to operate in integral membrane proteins as well. There are currently too few published examples of extreme flexibility in membrane domains to conclude that it may be a general phenomenon, however, that is probably to a large extent due to the experimental difficulties in characterizing these states. Before soluble IDP became well-known, there was a considerable reluctance to publish studies of functional but disordered proteins. This may have

held back the discovery of the important functions intrinsic disorder plays in soluble proteins. It is important not to repeat the same mistakes for putative membrane embedded IDPs. Publishing studies of functional and disordered membrane proteins will allow the construction of a database of experimentally validated ID membrane embedded domains similar to DisProt.⁴⁴ This will likely allow the deduction of the sequence signature of disorder in a membrane, which must surely be different from that of soluble proteins.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

- Dunker AK, Babu MM, Barbar E, Blackledge M, Bondos SE, Dosztányi Z, Jane H, Forman-kay J, Fuxreiter M, Gsponer J, et al. What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins* 2013; 1:1-5; <http://dx.doi.org/10.4161/idp.24157>
- Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 1999; 293:321-31; PMID:10550212; <http://dx.doi.org/10.1006/jmbi.1999.3110>
- Jensen MR, Zweckstetter M, Huang J-R, Blackledge M. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem Rev* 2014; 114:6632-60; PMID:24725176; <http://dx.doi.org/10.1021/cr400688u>
- Theillet F-X, Binolfi A, Fremberg-Kesner T, Hingorani K, Sarkar M, Kyne C, Li C, Crowley PB, Gierasch L, Pielak GJ, et al. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem Rev* 2014; 114:6661-714; PMID:24901537; <http://dx.doi.org/10.1021/cr400695p>
- Dunker A, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. *Genome Informatics* 2000; 11:161-71; PMID:11700597
- 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>
- Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys* 2008; 37:215-46; PMID:18573080; <http://dx.doi.org/10.1146/annurev.biophys.37.032807.125924>
- Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput Biol* 2006; 2:e100; PMID:16884331; <http://dx.doi.org/10.1371/journal.pcbi.0020100>
- 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>
- Tompa P, Prilusky J, Silman I, Sussman JL. Structural disorder serves as a weak signal for intracellular protein degradation. *Proteins* 2008; 71:903-9; PMID:18004785; <http://dx.doi.org/10.1002/prot.21773>
- Prakash S, Tian L, Ratliff KS, Lehotzky RE, Matoušek A. An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat Struct Mol Biol* 2004; 11:830-7; PMID:15311270; <http://dx.doi.org/10.1038/nsmb814>
- Minezaki Y, Homma K, Nishikawa K. Intrinsically disordered regions of human plasma membrane proteins preferentially occur in the cytoplasmic segment. *J Mol Biol* 2007; 368:902-13; PMID:17368479; <http://dx.doi.org/10.1016/j.jmb.2007.02.033>
- Venkatakrishnan A, Flock T, Prado DE, Oates ME, Gough J, Madan Babu M. Structured and disordered facets of the GPCR fold. *Curr Opin Struct Biol* 2014; 27C:129-37; <http://dx.doi.org/10.1016/j.sbi.2014.08.002>
- Xue B, Li L, Meroueh SO, Uversky VN, Dunker AK. Analysis of structured and intrinsically disordered regions of transmembrane proteins. *Mol Biosyst* 2009; 5:1688-702; PMID:19585006; <http://dx.doi.org/10.1039/b905913j>
- Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 2000; 41:415-27; PMID:11025552; [http://dx.doi.org/10.1002/1097-0134\(20001115\)41:3%3C415::AID-PROT130%3E3.0.CO;2-7](http://dx.doi.org/10.1002/1097-0134(20001115)41:3%3C415::AID-PROT130%3E3.0.CO;2-7)
- White SH, Wimley WC. Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* 1999; 28:319-65; PMID:10410805; <http://dx.doi.org/10.1146/annurev.biophys.28.1.319>
- Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 2005; 6:197-208; PMID:15738986; <http://dx.doi.org/10.1038/nrm1589>
- Uversky VN. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 2002; 11:739-56; PMID:11910019; <http://dx.doi.org/10.1110/ps.4210102>
- Popot JL, Engelman DM. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* 1990; 29:4031-7; PMID:1694455; <http://dx.doi.org/10.1021/bi00469a001>
- Arkin IT, Brunger AT. Statistical analysis of predicted transmembrane α -helices. *Biochim Biophys Acta - Protein Struct Mol Enzymol* 1998; 1429:113-28; [http://dx.doi.org/10.1016/S0167-4838\(98\)00225-8](http://dx.doi.org/10.1016/S0167-4838(98)00225-8)
- Kjaergaard M, Teilum K, Poulsen FM, Nørholm A-B, Hendus-Altenburger R, Pedersen SF, Krøgelund BB, Iešmantavičius V. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc Natl Acad Sci U S A* 2010; 107:12535-40; PMID:20616042; <http://dx.doi.org/10.1073/pnas.1001693107>
- Kjaergaard M, Andersen L, Nielsen LD, Teilum K. A Folded Excited State of Ligand-Free Nuclear Coactivator Binding Domain (NCBD) Underlies Plasticity in Ligand Recognition. *Biochemistry* 2013; 52:1686-93; PMID:23373423; <http://dx.doi.org/10.1021/bi4001062>
- Bonifacino JS, Traub LM. Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem* 2003; 72:395-447; PMID:12651740; <http://dx.doi.org/10.1146/annurev.biochem.72.121801.161800>
- Hicke L, Dunn R. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol* 2003; 19:141-72; PMID:14570567; <http://dx.doi.org/10.1146/annurev.cellbio.19.110701.154617>
- Hagai T, Azia A, Tóth-Petróczy Á, Levy Y. Intrinsic disorder in ubiquitination substrates. *J Mol Biol* 2011; 412:319-24; PMID:21802429; <http://dx.doi.org/10.1016/j.jmb.2011.07.024>
- Ng AHM, Fang NN, Comyn SA, Gsponer J, Mayor T. System-wide analysis reveals intrinsically disordered proteins are prone to ubiquitylation after misfolding stress. *Mol Cell Proteomics* 2013; 12:2456-67; PMID:23716602; <http://dx.doi.org/10.1074/mcp.M112.023416>
- Zhou FX, Merianos HJ, Brunger AT, Engelman DM. Polar residues drive association of poly-leucine transmembrane helices. *Proc Natl Acad Sci U S A* 2001; 98:2250-5; PMID:11226225; <http://dx.doi.org/10.1073/pnas.041593698>
- Gratkowski H, Lear JD, DeGrado WF. Polar side chains drive the association of model transmembrane peptides. *Proc Natl Acad Sci U S A* 2001; 98:880-5; PMID:11158564; <http://dx.doi.org/10.1073/pnas.98.3.880>
- Gray TM, Matthews BW. Intrahelical hydrogen bonding of serine, threonine and cysteine residues within α -helices and its relevance to membrane-bound proteins. *J Mol Biol* 1984; 175:75-81; PMID:6427470; [http://dx.doi.org/10.1016/0022-2836\(84\)90446-7](http://dx.doi.org/10.1016/0022-2836(84)90446-7)
- Hessa T, Kim H, Bihlmaier K, Lundin C, Boeckl J, Andersson H, Nilsson I, White SH, von Heijne G. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 2005; 433:377-81; PMID:15674282; <http://dx.doi.org/10.1038/nature03216>
- Senes A, Gerstein M, Engelman DM. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol* 2000; 296:921-36; PMID:10677292; <http://dx.doi.org/10.1006/jmbi.1999.3488>
- Brosig B, Langosch D. The dimerization motif of the glycophorin A transmembrane segment in membranes: importance of glycine residues. *Protein Sci* 1998; 7:1052-6; PMID:9568912; <http://dx.doi.org/10.1002/pro.5560070423>
- Eriksson A, Baase W, Zhang X, Heinz D, Blaber M, Baldwin E, Matthews B. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* (80-) 1992; 255:178-83; <http://dx.doi.org/10.1126/science.1553543>
- Brown CJ, Takayama S, Campen AM, Vise P, Marshall TW, Oldfield CJ, Williams CJ, Dunker AK. Evolutionary rate heterogeneity in proteins with long disordered regions. *J Mol Evol* 2002; 55:104-10; PMID:12165847; <http://dx.doi.org/10.1007/s00239-001-2309-6>
- Oberai A, Ihm Y, Kim S, Bowie JU. A limited universe of membrane protein families and folds. *Protein Sci* 2006; 15:1723-34; PMID:16815920; <http://dx.doi.org/10.1110/ps.062109706>
- Murali S, Robert J-C, Coïc Y-M, Neumann J-M, Ostuni MA, Yao Z-X, Papadopoulos V, Jamin N, Lacapère J-J. Secondary and tertiary structures of the transmembrane domains of the translocator protein TSPO determined by NMR. Stabilization of the TSPO tertiary fold upon ligand binding. *Biochim Biophys Acta* 2008; 1778:1375-81; PMID:18420025; <http://dx.doi.org/10.1016/j.bbame.2008.03.012>
- Jaremko L, Jaremko M, Giller K, Becker S, Zweckstetter M. Structure of the mitochondrial translocator protein in complex with a diagnostic ligand. *Science* 2014; 343:1363-6; PMID:24653034; <http://dx.doi.org/10.1126/science.1248725>
- Barrett PJ, Song Y, Van Horn WD, Husted EJ, Schaefer JM, Hadziselimovic A, Beel AJ, Sanders CR. The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science* 2012; 336:1168-71; PMID:22654059; <http://dx.doi.org/10.1126/science.1219988>
- Felli IC, Pierattelli R. Recent progress in NMR spectroscopy: toward the study of intrinsically disordered proteins of increasing size and complexity. *IUBMB Life* 2012; 64:473-81; PMID:22556167; <http://dx.doi.org/10.1002/iub.1045>
- Jensen MR, Ruigrok RWH, Blackledge M. Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr Opin Struct Biol* 2013; 23:426-35; PMID:23545493; <http://dx.doi.org/10.1016/j.sbi.2013.02.007>

41. Kjaergaard M, Poulsen FM. Disordered proteins studied by chemical shifts. *Prog Nucl Magn Reson Spectrosc* 2012; 60:42-51; PMID:22293398; <http://dx.doi.org/10.1016/j.pnmrs.2011.10.001>
42. Redfield C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods* 2004; 34:121-32; PMID:15283921; <http://dx.doi.org/10.1016/j.ymeth.2004.03.009>
43. Rösner HI, Kragelund BB. Structure and dynamic properties of membrane proteins using NMR. *Compr Physiol* 2012; 2:1491-539; PMID:23798308
44. Vucetic S, Obradovic Z, Vacic V, Radivojac P, Peng K, Iakoucheva LM, Cortese MS, Lawson JD, Brown CJ, Sikes JG, et al. DisProt: a database of protein disorder. *Bioinformatics* 2005; 21:137-40; PMID:15310560; <http://dx.doi.org/10.1093/bioinformatics/bth476>