

# The MemMoRF database for recognizing disordered protein regions interacting with cellular membranes

Georgina Csizmadia<sup>1,†</sup>, Gábor Erdős<sup>2,†</sup>, Hedvig Tordai<sup>1</sup>, Rita Padányi<sup>1</sup>, Silvio Tosatto<sup>3</sup>, Zsuzsanna Dosztányi<sup>2</sup> and Tamás Hegedűs<sup>1,\*</sup>

<sup>1</sup>Department of Biophysics and Radiation Biology, Semmelweis University, Budapest 1094, Hungary, <sup>2</sup>MTA-ELTE Lendület Bioinformatics Research Group, Department of Biochemistry, Eötvös Loránd University, Budapest 1117, Hungary and <sup>3</sup>Department of Biomedical Sciences, University of Padua, Padua 35131, Italy

Received August 08, 2020; Revised September 25, 2020; Editorial Decision October 06, 2020; Accepted October 28, 2020

## ABSTRACT

Protein and lipid membrane interactions play fundamental roles in a large number of cellular processes (e.g. signalling, vesicle trafficking, or viral invasion). A growing number of examples indicate that such interactions can also rely on intrinsically disordered protein regions (IDRs), which can form specific reversible interactions not only with proteins but also with lipids. We named IDRs involved in such membrane lipid-induced disorder-to-order transition as MemMoRFs, in an analogy to IDRs exhibiting disorder-to-order transition upon interaction with protein partners termed Molecular Recognition Features (MoRFs). Currently, both the experimental detection and computational characterization of MemMoRFs are challenging, and information about these regions are scattered in the literature. To facilitate the related investigations we generated a comprehensive database of experimentally validated MemMoRFs based on manual curation of literature and structural data. To characterize the dynamics of MemMoRFs, secondary structure propensity and flexibility calculated from nuclear magnetic resonance chemical shifts were incorporated into the database. These data were supplemented by inclusion of sentences from papers, functional data and disease-related information. The MemMoRF database can be accessed via a user-friendly interface at <https://memmorf.hegelab.org>, potentially providing a central resource for the characterization of disordered regions in transmembrane and membrane-associated proteins.

## INTRODUCTION

Many proteins contain intrinsically disordered regions (IDRs) that do not fold into well-defined structures in isolation and are best represented by conformational ensembles (1,2). The flexibility enables IDRs to participate in highly specific and reversible interactions, which form a common molecular basis for a wide range of physiological processes, including signalling, gene regulation, cell cycle regulation, scaffolding, or chaperoning and pathological conditions (3–5). Such segments, which show disorder-to-order transition upon forming protein–protein interactions, are called molecular recognition features (MoRFs) (6–8). Recent evidence suggests that IDRs are also important components of membrane-associated proteins (MAPs) and transmembrane proteins (TMPs). In these proteins disordered regions of variable sizes can be located in loops between transmembrane segments or in N- and C-terminal regions (9). Many of these disordered regions possess a common feature, namely, they exhibit lipid induced alteration of ordered/disordered status. Here we suggest to term the segments involved in protein–lipid interactions as MemMoRFs, since they represent a distinct category of context-dependent behaviour when compared to MoRFs.

The fundamental role of MemMoRFs in various cellular functions is demonstrated by their ubiquitous presence in proteins associated with a wide range of cellular processes. A set of proteins with MemMoRFs participates in the regulation of cell cycle, apoptosis and phagocytosis (e.g. NOTCH1 (10) and interferon alpha-inducible protein 27-like protein 1 (11)). Other MemMoRF containing proteins play important roles in cellular trafficking and shaping the cell membrane (e.g. Myc box-dependent-interacting protein 1 (12),  $\alpha$ -synuclein (13) and pro-neuregulin-1 (14)). MemMoRFs are expected to be abundant in signalling proteins and were detected in both transmembrane receptors (e.g. integrin  $\beta$ 3 (15), PTEN tumor suppressor (16) and receptor tyrosine kinases, such as EGFR (17) and erbB-2 (18)) or membrane-associated proteins (e.g. tyrosine-protein ki-

\*To whom correspondence should be addressed. Tel: +36 1 4591500/60233; Fax: +36 1 266 6656; Email: hegedus@hegelab.org

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

nase Src (19)). MemMoRFs also play important roles in pathological processes. An important MemMoRF group is associated with neurodegenerative diseases and includes  $\alpha$ -synuclein, amyloid- $\beta$  precursor protein, and prion protein (20). In addition, MemMoRFs have been identified in several viral proteins, such as capsid proteins of HIV1 and HCV (21,22), and TMPs, including a potassium channel of the Influenza A virus (23). Most likely, these viral proteins and various toxins, such as melittin (24) and the GPCR impairing toxin of *Gila monster* (25), require MemMoRF for engaging with or penetrating through membrane bilayers.

The specific conformational properties of these IDRs are shaped by the chemical composition and physical properties of the membrane, and the relevant IDRs can also shape the properties of the membrane (26–28). An important mode to regulate these protein–lipid interactions is phosphorylation. For example, the intracellular domain of the T-cell receptor  $\zeta$  forms transient helices in the intracellular membrane leaflet (Supplementary Figure S1) (29). One of the transient helices (LID region) possesses an ITAM (immunoreceptor tyrosine-based activation motif), which includes a tyrosine immersed in the membrane bilayer and is a target for phosphorylation (29). When phosphorylated, the ITAM MemMoRF undocks from the membrane and loses its  $\alpha$ -helical character. A similar feature can be observed in transporter regulation, involving single pass TMPs possessing MemMoRFs, such as phospholamban (PLN) and FXYD proteins (e.g. phospholemman, FXYD domain-containing ion transport regulator 4 and Na<sup>+</sup>/K<sup>+</sup> ATPase subunit  $\gamma$ ) (Supplementary Figure S1) (30–32). The small size of these proteins makes them attractive for drug targeting (e.g. PLN in cardiac diseases) (33).

Although disorder-to-order transition was detected in numerous studies targeting protein–lipid interactions, this phenomenon has not been examined in a generalized manner. As a consequence, information about these protein segments are scattered in the literature, and the detailed structural properties of MemMoRFs are difficult to interpret in the context of additional structural and functional modules. Here, we report a novel, comprehensive database of experimentally validated MemMoRFs, currently containing 131 examples in 96 proteins as a gold standard set. The MemMoRF database complements existing databases of disordered proteins in general (34,35) and also databases of interactions of disordered proteins with ordered or other disordered proteins (36–38). A separate database was built for MoRFs of protein–protein interactions of MAPs and TMPs (38). In contrast, we collected those flexible protein regions into our MemMoRF database, which participate in reversible membrane binding, associated with conformational changes.

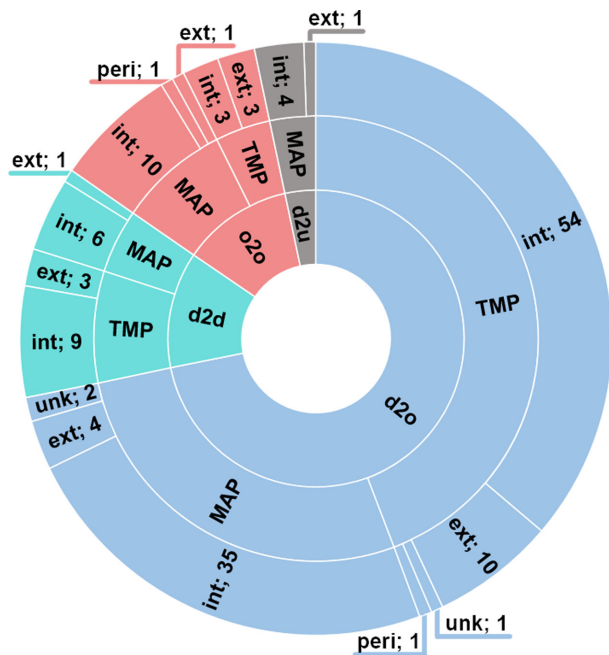
## IDENTIFICATION AND SYSTEMATIC COLLECTION OF MemMoRFs

The main aim of this database is to facilitate the investigations and targeting of proteins involving MemMoRFs. As a first step, we collected nuclear magnetic resonance (NMR) structures of TMPs and MAPs, since IDRs of these proteins can be located in the close vicinity of biological membranes. We strongly relied on NMR data, since this spectroscopy

method is the most widely used approach to determine lipid interaction of proteins at residue level. NMR structures of TMPs were gathered using the Membrane Protein Browser of Protein Data Bank (PDB) (39). MAPs were looked up in the UniProt database (40) using the ‘peripheral membrane protein’ subcellular location keyword (UniProt 2019.01.14). To characterize the transient secondary structures of IDRs, we did not rely solely on the deposited NMR structures, but also incorporated secondary structure propensity (SSPop; using  $\delta 2D$ ) (41) and flexibility ( $1-S^2$ ; using Random Coil Index, RCI) (42). We considered a region intrinsically disordered if the residues within the region exhibited ‘coil’ secondary structure population ( $>0.5$ ) and high flexibility ( $>0.15$ ). A region was also considered disordered in the case of corresponding disorder annotation in DisProt, DIBS, MFIB or PFAM databases (34,36–37,43). As a next stage, we also explored invisible regions of X-ray and cryo-EM structures, which belong to TMPs or MAPs possessing annotated disordered regions. We subjected the whole protein set to extensive literature analysis to identify IDRs and MemMoRFs or further validate MemMoRFs detected using NMR or other structural data. Thus, proofs of disorder could be derived from calculations, databases, structural information and literature evidence.

A total number of 538 proteins including 206 TMPs and 332 MAPs were screened for segments located in an IDR and exhibiting altered dynamics in a membrane mimetic. A total of 149 membrane interacting regions were identified in 107 proteins. Only 11 regions were derived exclusively from unresolved residues in X-ray or cryo-EM structures. 131 out of the membrane interacting regions are disordered in aqueous solution and become ordered ( $n = 107$ ) or retain flexibility ( $n = 19$ ) upon binding to membrane mimetics. Eighteen regions with stable secondary structure both in solution and in membrane bound state were found. Eighty-four out of 149 membrane interacting regions are in TMPs and the remaining 65 are in MAPs. Intracellular, extracellular and periplasmic localizations were observed in 121, 23 and 2 cases, respectively (Figure 1). SSPop and flexibility values were calculated from published chemical shift data with sufficient quality, available for 41% of our protein dataset. Using data from NMR experiments, X-ray and cryo-EM structures and literature, we found 92 regions among identified IDRs, which were not annotated in the DisProt database (Supplementary Tables S1 and S2).

We categorized the entries depending how their dynamical properties changed upon interaction with the lipid bilayer. In many cases, the corresponding NMR data indicates that the observed transitions resulted in increased  $\alpha$ -helical propensity values. In addition, some disordered segments retained their conformational freedom when interacted with membrane mimetics or specific membrane lipids. Although these segments still likely sample a reduced conformational space in the presence of a membrane environment, we labeled their transition as ‘disorder-to-disorder’ to emphasize the dynamics of their membrane bound state. In the case of some of the MemMoRFs derived from X-ray or cryo-EM structures, there is no information on the lipid-bound structure, thus we labeled these entries as ‘disorder-to-unknown’. We also included cases, where the binding helices were stable both in aqueous solution and interact-



**Figure 1.** Distribution of MemMoRF types in the database. A total of 149 membrane interacting regions were identified in 107 proteins. One hundred and thirty-one out of these regions are disordered in aqueous solution and become ordered ( $n = 107$ ) or retain flexibility ( $n = 19$ ) upon binding to membrane mimetics. Eighteen regions with stable secondary structure both in solution and in membrane bound state were found. d2o: disorder-to-order, d2d: disorder-to-disorder, o2o: order-to-order, d2u: disorder-to-unknown, TMP: transmembrane protein, MAP: membrane associated protein, int: intracellular side, ext: extracellular side, peri: peri-plasmic side, unk: unknown location.

ing with the membrane (e.g. helices from kinase suppressor of Ras 1 and BH3-interacting domain death agonist). We labeled these entries as ‘bistable helix’. While they are not classical MemMoRFs, these segments still exhibit a dynamic equilibrium between lipid and water phases and participate in regulatory interactions.

To demonstrate the analysis of various data for IDR and MemMoRF validation and phosphorylation-dependent regulatory disorder-to-order transition, we selected integrin  $\beta 3$ , which plays a role in angiogenesis and tumor growth (44,45). Targeting integrin  $\beta 3$  associated signalling was shown to induce apoptosis of endothelial tumor cells (44) and cell permeable peptides derived from integrin  $\beta$  cytoplasmic tails (CT) were developed for angiogenesis inhibition (46). Since these peptides overlap with a MemMoRF, a detailed description of phosphorylation dependent protein and membrane interactions of this region will help to improve the potential therapeutic use of integrin  $\beta$  CT peptides. The complexity of various conformations of the C-terminal integrin  $\beta 3$  MemMoRF under different conditions is shown in Figure 2 and accessible at <https://memmorf.hegelab.org/entry/P05106>. The C-terminal MemMoRF region (a.a. 770–784) exhibits low helix propensity ( $<0.5$ ) and high flexibility values ( $>0.15$ ) in the absence of a membrane mimetic (PDB ID: 2KNC and BMRB ID: 16496) that indicates a disordered state for these residues (47). Although the helical propensity for a.a. 771–776 is lower than the commonly used SSPop threshold (0.5), a

small stable helix is present in all of the 20 structures deposited in the 2KNC PDB structure. However, this part of the protein is marked disordered in the DisProt and Pfam databases. Importantly, in the presence of a lipid environment (PDB ID: 2KV9 and BMRB ID: 16771) (48) the decreased flexibility, which approaches the disorder threshold (0.15), confirms the presence of a MemMoRF. At the same time, the helical propensity increases and indicates a stable helix in a smaller part of the sequence (a.a. 776–781). This is in contrast with the long  $\alpha$ -helix present in the structural ensemble deposited in the PDB. Moreover, in the paper accompanying these NMR structures (48) the authors noted the presence of transient helical contents and increasing disorder towards the C-terminus based on hydrogen-deuterium exchange experiments. Interestingly, phosphorylation of this MemMoRF promotes disorder (helical propensity  $<0.5$  and flexibility  $>0.15$ ) in the presence of a membrane mimetic as well (PDB ID: 2LJD and BMRB ID: 17930), thus serves as a conformational switch (49). Again, this NMR ensemble also exhibits a stable small helix between residues 775–780. The potential over-representation of helical content urged us to supplement the entries in the MemMoRF database with supporting statements from papers, and it also cautions database curators that making a decision about disorder should not rely solely on atomistic NMR structures. Phosphorylation-dependent regulation via MemMoRFs is further demonstrated by the T cell receptor  $\zeta$  and phospholamban (Supplementary Figure S1).

## WEB SERVER AND INTERFACE IMPLEMENTATION

In order to make MemMoRF data accessible and to facilitate research on proteins possessing MemMoRF regions, we developed a web application available at <https://memmorf.hegelab.org>.

### Browsing and searching data

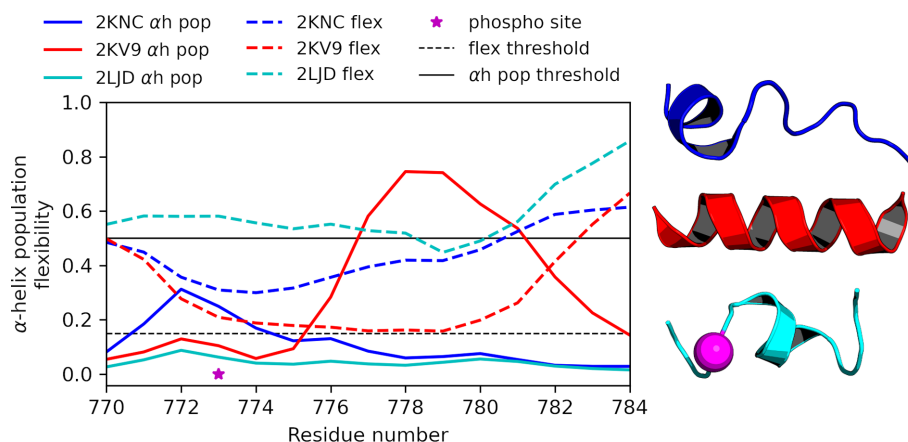
The database provides both browsing and searching functionalities in the ‘Browse’ page, where all entries are listed in a table format. Table columns contain the protein and the gene name, the source organism, the UniProt accession number type, the corresponding PDB and BMRB identifiers and MIM entry, if available, for each protein in our database. The table is sortable and can be filtered by simple queries or using dropdown lists. Searches independent from columns can be performed by typing a query in a text box located at the top of the page.

### Entry pages

Each entry starts with the most important, basic information about the protein from UniProt (protein name, accession code, gene name, organism and keywords). Then, MemMoRFs are listed along with their boundaries, type of transition upon membrane binding (e.g. disorder-to-order), localization (e.g. intracellular) and statements supporting disordered and membrane interacting nature of the region.

The SSPop and flexibility calculations, if available, are shown in a graph, which can be saved using the controls next to it. Visibility of corresponding lines can





**Figure 2.** Experimental NMR data provide information on disorder level thus input for MemMoRF identification. Secondary structure population (e.g. helix and coil) and flexibility ( $1-S^2$ ) values were utilized to characterize the per residue disorder-order propensity in NMR ensembles, as exemplified by a MemMoRF from integrin  $\beta 3$ . Blue: in organic solvent; red: in DPC; cyan: phosphorylated in DPC; magenta: phosphorylation site;  $\alpha$ h pop:  $\alpha$ -helix population calculated by  $\delta 2D$ ; flex:  $1-S^2$  calculated by RCI,  $\alpha$ h pop threshold: 0.5, flex threshold: 0.15.

be toggled by clicking on their legend entry. NMR experiments associated with the entry can be selected in a box, right from the graph. This box lists the membrane mimetic used in the NMR experiment, the PDB ID of the calculated NMR ensemble, the BMRB ID and the ‘\_Assigned\_chem\_shift\_list.ID’ of chemical shift data used for the calculation. The ‘\_Assigned\_chem\_shift\_list.ID’ identifies a specific set of chemical shift values from the BMRB NMR-STAR files, which may contain multiple sets acquired under different experimental conditions.

Sequence specific data, including MemMoRF regions, transmembrane (TM) helices, PFAM domains (43), short linear motifs from ELM (50), post-translational modifications from PhosphoSitePlus (51), IDRs from DisProt, DIBS and MFIB (34,36–37), and segments covered in the PDB are also shown. To further characterize the sequences, Wimley-White hydrophobicity plots and IUPred short predictions were also included into the main graph (52,53). Structures strongly associated with the MemMoRF can be selected for display and manipulation using LiteMol (54). MemMoRFs and TM helices are labelled by default on the structure. In the bottom region of the entry page, additional information is listed to assess the role of MemMoRFs in pathophysiological states. These data include disease phenotypes from MIM (55), disease causing and polymorphic variations from dbSNP (56), and DrugBank records (57). Protein-protein interactions with a link to the IntAct (58) and STRING (59) entries of the protein are followed by an interactive, embedded STRING network graph.

### Help and feedback pages

Although the web application is simple and self-explanatory, we provide a comprehensive help page with sections about the following items: i) MemMoRF definition, ii) information resources linked to the memMoRF database, iii) details of our data collection pipeline, iv) search possibilities of the browse page, v) the content of the entry page and vi) links to the statistics and download pages. Users are encouraged to submit comments or ques-

tions via a contact form or email located in the feedback page.

### Server implementation and database structure

The web application is served via a DJANGO (version 2.1.1) based web interface, fuelled by an SQL database providing fast access to data even in the case of parallel queries. The SQL database contains all the information collected from the literature and various databases (e.g. UniProt, dbSNP, MIM and Drug Bank), organized into multiple tables. Each record in the UniProt table represents a single protein and has the UniProt accession number as a unique key. Other data, including records in the MemMoRF table, are linked to this table by UniProt accession number. In order to provide the best possible user experience on various devices and browsing options for users, the front-end compatibility of MemMoRF is supported by a combination of bootstrap (version 4.3.1) and JQuery (version 2.1.4).

In addition to data access through the web application interface, data can also be downloaded in JSON, XML or TSV formats, or by a RESTful API serving standard JSON format (e.g. <https://memmorf.hegelab.org/rest/Q9NR61.json>).

### CONCLUSION

MemMoRF database delivers data on membrane interacting disordered regions of TMPs and MAPs. These regions exhibit dynamic structural alterations during transition from solvent to a membrane bound state. The collected information has a high potential to contribute to understanding crucial cellular functions and pathological states associated with membrane proteins. Our database provides significant benefits for the broad scientific community because of the following: i) it is a freely accessible, easy-to-use, and organized resource; ii) it includes a gold standard set for determining the sequence requirements for lipid interaction of disordered regions; iii) it represents a high-quality set for developing novel *in silico* pipelines for MemMoRF

identification; and iv) it promotes further experimental investigation and development of drug targeting approaches for MemMoRF containing proteins.

We also provide a new set of disordered protein regions, which are not currently present in other manually maintained databases, based on careful literature curation and NMR-based calculation of secondary structure propensities and flexibility. Our aim is to establish and maintain the MemMoRF database in the long term as a central resource for membrane proteins with lipid bilayer interacting disordered segments.

## DATA AVAILABILITY

MemMoRF database and web application are available at <http://memmorf.hegelab.org>. Since the server is located in the EU, it fully adheres to the General Data Protection Regulation (GDPR).

## SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Pietro Sormanni and Michele Vendruscolo (Department of Chemistry, University of Cambridge, Cambridge, UK) for providing their scripts for preparing and running  $\delta$ 2D calculations.

## FUNDING

National Research, Development and Innovation Office [NKFIH 127961 to T.H.]; Higher Education Institutional Excellence Programme of the Ministry for Innovation and Technology in Hungary. Funding for open access charge: Semmelweis University, EFOP-3.6.3-VEKOP-16-2017-00009.

*Conflict of interest statement.* None declared.

## REFERENCES

- Sormanni,P., Piovesan,D., Heller,G.T., Bonomi,M., Kukic,P., Camilloni,C., Fuxreiter,M., Dosztanyi,Z., Pappu,R.V., Babu,M.M. *et al.* (2017) Simultaneous quantification of protein order and disorder. *Nat. Chem. Biol.*, **13**, 339–342.
- Uversky,V.N. (2016) Dancing protein clouds: the strange biology and chaotic physics of intrinsically disordered proteins. *J. Biol. Chem.*, **291**, 6681–6688.
- Dunker,A.K., Brown,C.J., Lawson,J.D., Iakoucheva,L.M. and Obradović,Z. (2002) Intrinsic disorder and protein function. *Biochemistry*, **41**, 6573–6582.
- Dyson,H.J. and Wright,P.E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, **6**, 197–208.
- 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) Classification of intrinsically disordered regions and proteins. *Chem. Rev.*, **114**, 6589–6631.
- Oldfield,C.J., Cheng,Y., Cortese,M.S., Romero,P., Uversky,V.N. and Dunker,A.K. (2005) Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry*, **44**, 12454–12470.
- Mohan,A., Oldfield,C.J., Radivojac,P., Vacic,V., Cortese,M.S., Dunker,A.K. and Uversky,V.N. (2006) Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.*, **362**, 1043–1059.
- Yang,J., Gao,M., Xiong,J., Su,Z. and Huang,Y. (2019) Features of molecular recognition of intrinsically disordered proteins via coupled folding and binding. *Protein Sci.*, **28**, 1952–1965.
- Tusnády,G.E., Dobson,L. and Tompa,P. (2015) Disordered regions in transmembrane proteins. *Biochim. Biophys. Acta*, **1848**, 2839–2848.
- Deatherage,C.L., Lu,Z., Kroncke,B.M., Ma,S., Smith,J.A., Voehler,M.W., McFeeters,R.L. and Sanders,C.R. (2017) Structural and biochemical differences between the Notch and the amyloid precursor protein transmembrane domains. *Sci. Adv.*, **3**, e1602794.
- Klammt,C., Maslennikov,I., Bayrhuber,M., Eichmann,C., Vajpai,N., Chiu,E.J.C., Blain,K.Y., Esquivies,L., Kwon,J.H.J., Balana,B. *et al.* (2012) Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nat. Methods*, **9**, 834–839.
- Lów,C., Weininger,U., Lee,H., Schweimer,K., Neundorff,I., Beck-Sickinge,A.G., Pastor,R.W. and Balbach,J. (2008) Structure and dynamics of helix-0 of the N-BAR domain in lipid micelles and bilayers. *Biophys. J.*, **95**, 4315–4323.
- Ulmer,T.S., Bax,A., Cole,N.B. and Nussbaum,R.L. (2005) Structure and dynamics of micelle-bound human alpha-synuclein. *J. Biol. Chem.*, **280**, 9595–9603.
- Chukhlieb,M., Raasakka,A., Ruskamo,S. and Kursula,P. (2015) The N-terminal cytoplasmic domain of neuregulin 1 type III is intrinsically disordered. *Amino Acids*, **47**, 1567–1577.
- Vinogradova,O., Vaynberg,J., Kong,X., Haas,T.A., Plov,E.F. and Qin,J. (2004) Membrane-mediated structural transitions at the cytoplasmic face during integrin activation. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 4094–4099.
- Nanda,H., Heinrich,F. and Lösche,M. (2015) Membrane association of the PTEN tumor suppressor: neutron scattering and MD simulations reveal the structure of protein-membrane complexes. *Methods*, **77–78**, 136–146.
- Choowongkamon,K., Carlin,C.R. and Sönnichsen,F.D. (2005) A structural model for the membrane-bound form of the juxtamembrane domain of the epidermal growth factor receptor. *J. Biol. Chem.*, **280**, 24043–24052.
- Bragin,P.E., Mineev,K.S., Bocharova,O.V., Volynsky,P.E., Bocharov,E.V. and Arseniev,A.S. (2016) HER2 transmembrane domain dimerization coupled with self-association of membrane-embedded cytoplasmic juxtamembrane regions. *J. Mol. Biol.*, **428**, 52–61.
- Pérez,Y., Maffei,M., Igea,A., Amata,I., Gairi,M., Nebreda,A.R., Bernadó,P. and Pons,M. (2013) Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.*, **3**, 1295.
- Riek,R. and Eisenberg,D.S. (2016) The activities of amyloids from a structural perspective. *Nature*, **539**, 227–235.
- Serrano,S., Araujo,A., Apellániz,B., Bryson,S., Carravilla,P., de la Arada,I., Huarte,N., Rujas,E., Pai,E.F., Arrondo,J.L.R. *et al.* (2014) Structure and immunogenicity of a peptide vaccine, including the complete HIV-1 gp41 2F5 epitope: implications for antibody recognition mechanism and immunogen design. *J. Biol. Chem.*, **289**, 6565–6580.
- Penin,F., Brass,V., Appel,N., Ramboarina,S., Montserret,R., Ficheux,D., Blum,H.E., Bartenschlager,R. and Moradpour,D. (2004) Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. *J. Biol. Chem.*, **279**, 40835–40843.
- Schnell,J.R. and Chou,J.J. (2008) Structure and mechanism of the M2 proton channel of influenza A virus. *Nature*, **451**, 591–595.
- Appadu,A., Jelokhani-Niaraki,M. and DeBruin,L. (2015) Conformational changes and association of Membrane-Interacting peptides in myelin membrane models: a case of the C-Terminal peptide of proteolipid protein and the antimicrobial peptide melittin. *J. Phys. Chem. B*, **119**, 14821–14830.
- Neidigh,J.W., Fesinmeyer,R.M., Prickett,K.S. and Andersen,N.H. (2001) Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in the solution and micelle-associated states. *Biochemistry*, **40**, 13188–13200.
- Gustavsson,M., Traaseth,N.J. and Veglia,G. (2012) Probing ground and excited states of phospholamban in model and native lipid membranes by magic angle spinning NMR spectroscopy. *Biochim. Biophys. Acta*, **1818**, 146–153.
- Fakhree,M.A.A., Blum,C. and Claessens,M.M.A.E. (2019) Shaping membranes with disordered proteins. *Arch. Biochem. Biophys.*, **677**, 108163.

28. Zeno, W.F., Baul, U., Snead, W.T., DeGroot, A.C.M., Wang, L., Lafer, E.M., Thirumalai, D. and Stachowiak, J.C. (2018) Synergy between intrinsically disordered domains and structured proteins amplifies membrane curvature sensing. *Nat. Commun.*, **9**, 4152.
29. Aivazian, D. and Stern, L.J. (2000) Phosphorylation of T cell receptor zeta is regulated by a lipid dependent folding transition. *Nat. Struct. Biol.*, **7**, 1023–1026.
30. Geering, K. (2005) Function of FXFD proteins, regulators of Na, K-ATPase. *J. Bioenerg. Biomembr.*, **37**, 387–392.
31. MacLennan, D.H. and Kranias, E.G. (2003) Phospholamban: a crucial regulator of cardiac contractility. *Nat. Rev. Mol. Cell Biol.*, **4**, 566–577.
32. Akin, B.L., Hurley, T.D., Chen, Z. and Jones, L.R. (2013) The structural basis for phospholamban inhibition of the calcium pump in sarcoplasmic reticulum. *J. Biol. Chem.*, **288**, 30181–30191.
33. Ablorh, N.-A.D. and Thomas, D.D. (2015) Phospholamban phosphorylation, mutation, and structural dynamics: a biophysical approach to understanding and treating cardiomyopathy. *Biophys. Rev.*, **7**, 63–76.
34. Hatos, A., Hajdu-Soltész, B., Monzon, A.M., Palopoli, N., Álvarez, L., Aykac-Fas, B., Bassot, C., Benítez, G.I., Bevilacqua, M., Chasapi, A. *et al.* (2020) DisProt: intrinsic protein disorder annotation in 2020. *Nucleic Acids Res.*, **48**, D269–D276.
35. Piovesan, D., Tabaro, F., Paladin, L., Necci, M., Micetic, I., Camilloni, C., Davey, N., Dosztányi, Z., Mészáros, B., Monzon, A.M. *et al.* (2018) MobiDB 3.0: more annotations for intrinsic disorder, conformational diversity and interactions in proteins. *Nucleic Acids Res.*, **46**, D471–D476.
36. Schad, E., Fichó, E., Pancsa, R., Simon, I., Dosztányi, Z. and Mészáros, B. (2018) DIBS: a repository of disordered binding sites mediating interactions with ordered proteins. *Bioinformatics*, **34**, 535–537.
37. Fichó, E., Reményi, I., Simon, I. and Mészáros, B. (2017) MFIB: a repository of protein complexes with mutual folding induced by binding. *Bioinformatics*, **33**, 3682–3684.
38. Miskei, M., Antal, C. and Fuxreiter, M. (2017) FuzDB: database of fuzzy complexes, a tool to develop stochastic structure-function relationships for protein complexes and higher-order assemblies. *Nucleic Acids Res.*, **45**, D228–D235.
39. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
40. UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.*, **47**, D506–D515.
41. Camilloni, C., De Simone, A., Vranken, W.F. and Vendruscolo, M. (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry*, **51**, 2224–2231.
42. Berjanskii, M.V. and Wishart, D.S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.*, **127**, 14970–14971.
43. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A. *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.*, **47**, D427–D432.
44. Brooks, P.C., Montgomery, A.M., Rosenfeld, M., Reisfeld, R.A., Hu, T., Klier, G. and Cheresh, D.A. (1994) Integrin alpha v beta 3 antagonists promote tumor regression by inducing apoptosis of angiogenic blood vessels. *Cell*, **79**, 1157–1164.
45. Brooks, P.C., Strömblad, S., Klemke, R., Visscher, D., Sarkar, F.H. and Cheresh, D.A. (1995) Antiintegrin alpha v beta 3 blocks human breast cancer growth and angiogenesis in human skin. *J. Clin. Invest.*, **96**, 1815–1822.
46. Cao, Z., Suo, X., Chu, Y., Xu, Z., Bao, Y., Miao, C., Deng, W., Mao, K., Gao, J., Xu, Z. *et al.* (2018) Peptides derived from the integrin  $\beta$  cytoplasmic tails inhibit angiogenesis. *Cell Commun. Signal.*, **16**, 38.
47. Yang, J., Ma, Y.-Q., Page, R.C., Misra, S., Plow, E.F. and Qin, J. (2009) Structure of an integrin alphaIIb beta3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 17729–17734.
48. Metcalf, D.G., Moore, D.T., Wu, Y., Kielec, J.M., Molnar, K., Valentine, K.G., Wand, A.J., Bennett, J.S. and DeGrado, W.F. (2010) NMR analysis of the alphaIIb beta3 cytoplasmic interaction suggests a mechanism for integrin regulation. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 22481–22486.
49. Deshmukh, L., Meller, N., Alder, N., Byzova, T. and Vinogradova, O. (2011) Tyrosine phosphorylation as a conformational switch: a case study of integrin  $\beta$ 3 cytoplasmic tail. *J. Biol. Chem.*, **286**, 40943–40953.
50. Kumar, M., Gouw, M., Michael, S., Sámano-Sánchez, H., Pancsa, R., Glavina, J., Diakogianni, A., Valverde, J.A., Bukirova, D., Čalyševa, J. *et al.* (2020) ELM—the eukaryotic linear motif resource in 2020. *Nucleic Acids Res.*, **48**, D296–D306.
51. Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V. and Skrzypek, E. (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.*, **43**, D512–D520.
52. Wimley, W.C. and White, S.H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.*, **3**, 842–848.
53. Mészáros, B., Erdos, G. and Dosztányi, Z. (2018) IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.*, **46**, W329–W337.
54. Sehnal, D., Deshpande, M., Vařeková, R.S., Mir, S., Berka, K., Midlik, A., Pravda, L., Velankar, S. and Koča, J. (2017) LiteMol suite: interactive web-based visualization of large-scale macromolecular structure data. *Nat. Methods*, **14**, 1121–1122.
55. Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A. and McKusick, V.A. (2005) Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.*, **33**, D514–D517.
56. Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M. and Sirotkin, K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
57. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z. *et al.* (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.
58. Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N.H., Chavali, G., Chen, C., del-Toro, N. *et al.* (2014) The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.*, **42**, D358–D363.
59. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P. *et al.* (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*, **47**, D607–D613.