Published online 17 November 2021 *Nucleic Acids Research, 2022, Vol. 50, Database issue* ***D509–D517***
https://doi.org/10.1093/nar/gkab1060

FuzDB: a new phase in understanding fuzzy interactions

Andras Hatos [1], Alexander Miguel Monzon [1], Silvio C.E. Tosatto [1], Damiano Piovesan [1,*] and Monika Fuxreiter [1,2,*]

[1]Department of Biomedical Sciences, University of Padova, via Ugo Bassi 58/B, 35131 Padova, Italy and [2]Department of Biochemistry and Molecular Biology, University of Debrecen, Nagyerdei krt 98, 4010 Debrecen, Hungary

Received September 15, 2021; Revised October 13, 2021; Editorial Decision October 14, 2021; Accepted October 27, 2021

ABSTRACT

Fuzzy interactions are specific, variable contacts between proteins and other biomolecules (proteins, DNA, RNA, small molecules) formed in accord to the cellular context. Fuzzy interactions have recently been demonstrated to regulate biomolecular condensates generated by liquid-liquid phase separation. The FuzDB v4.0 database (https://fuzdb.org) assembles experimentally identified examples of fuzzy interactions, where disordered regions mediate functionally important, context-dependent contacts between the partners in stoichiometric and higher-order assemblies. The new version of FuzDB establishes cross-links with databases on structure (PDB, BMRB, PED), function (ELM, UniProt) and biomolecular condensates (PhaSepDB, PhaSePro, LLPSDB). FuzDB v4.0 is a source to decipher molecular basis of complex cellular interaction behaviors, including those in protein droplets.

INTRODUCTION

Specific molecular recognition is traditionally related to well-defined interaction patterns, using the framework of the classical structure-function paradigm. It is increasingly recognized, however, that biomolecular interactions adapt to the cellular conditions (1,2), and may vary according to cellular compartment (3), posttranslational modification (4) or stress (5). Recent advances in structure determination methods, both experimental (6) and computational (7,8) approaches focusing on the ensemble character of protein conformations indicate that conformational heterogeneity in protein assemblies is more ubiquitous than previously thought (9) (and Curr Opin Struct Biol vol 56). This is due to structural and biophysical studies carried out (i) using full proteins and not truncated constructs (ii) in the presence of functionally relevant posttranslational modification sites and/or interaction partners, (iii) on large assemblies (e.g. cryo-EM), (iv) using physiologically relevant conditions (e.g. in solution by NMR, FRET) as well as due to the improved quality of the deposited protein structures including multiple rotamers/conformers (e.g. X-ray crystallography). Accumulating experimental results thus suggest that conformational heterogeneity is general in complexes formed by disordered proteins and is frequently observed in assemblies of structured proteins. Such structural diversity is coupled to variable interactions at the interface, or dynamic, transient interactions, which significantly influence the binding properties via modulating the orientation and coordination of structured domains (10–12).

Fuzzy interactions also regulate the formation of a wide range of non-stoichiometric, higher-order assemblies (13,14), including biomolecular condensates (15,16). Recent experimental and computational results evidence conformational heterogeneity in droplets as observed in elastin (17) or Fus (18). Condensates are stabilized by disordered interactions, involving many binding configurations between redundant motifs, which can be used to predict droplet-forming propensities from sequence, as implemented by the FuzDrop method (19). Such interactions are also influenced by posttranslational modifications (20) or RNA binding (21) in accord with their cellular context-dependence to induce condensate assembly/disassembly. Interactions via redundant motifs enable coordination between many factors via weak, variable contacts, as exemplified in the eukaryotic transcriptional machinery (22). Interaction fuzziness is a major driver of the material state conversion of liquid-like condensates (23,24). Computational results indicate that interactions with many binding configurations are widespread in eukaryotic proteomes (19).

These results prompt for a relationship between conformational properties of fuzzy assemblies and their biological activities (25). The context-dependence of interactions leading to distinct biological outcomes upon minor perturbations of the heterogeneous ensemble (26) presents a

*To whom correspondence should be addressed. Tel: +39 049 827 6136; Email: monika.fuxreiter@unipd.it
Correspondence may also be addressed to Damiano Piovesan. Email: damiano.piovesan@unipd.it

challenge for the classical structure-function paradigm (2). To facilitate development of a stochastic relationship, the FuzDB database was launched in 2016 (27). The database assembled disordered protein assemblies, with experimental evidence for conformational heterogeneity upon binding to a specific partner (27). The database linked structural data to the impact on biological activity, demonstrated by changes in function (e.g. binding affinity, specificity, enzymatic or transcription activity, signaling outcome) upon modifying the fuzzy region (truncation, mutation, removal). Coupling structural and functional data therefore not only provided experimental evidence for the biological relevance of fuzzy interactions, but also gave insights into their molecular mechanisms (28).

The FuzDB database facilitated the discovery and characterization of fuzzy protein assemblies. In addition, the data was used to develop bioinformatics algorithms for predicting complex interaction behaviors from sequence (29,30). The experimental data accumulating with rapid pace motivated the update the FuzDB database. In addition, the wide interest in biomolecular condensates and interactions driving their formation require establishing connections between high-resolution structural data and experimental evidence for liquid-liquid phase separation. The FuzDB v4.0 assembles biomolecular assemblies, where fuzzy interactions mediate specific partnerships of proteins with an impact on biological activity. The FuzDB v4.0 database has been renamed (FuzDB: database of fuzzy interactions) and moved to a new site (https://fuzdb.org). It has been completely re-designed into a more user-friendly manner and it is now fully interoperable with easier access to data, cross-links to structural as well as functional databases, including those of biomolecular condensates.

MATERIALS AND METHODS

Assembly of structural data

Solution structures of protein complexes, which were determined by NMR, were collected from the Protein Data Bank (PDB) (31). Assemblies were defined based on the presence of multiple chains, including both homomeric or heteromeric complexes. Fuzziness was defined based on the structural disorder of the interacting residues. Residues mediating physical contacts with the partner were determined by the RING software (32) using a distance thresholds and geometrical parameters optimized separately for different types of interactions (H-bonds, ionic, pi–pi, pi–cation, Van der Waals, disulfide bridges). Protein residues forming the interface were classified as mobile and non-mobile based on changes in local conformation within the same NMR ensemble using the MOBI software (33). In case both interacting chains were classified as disordered, these were considered as fuzzy proteins and handled as separated entries. RING and MOBI outputs were available in MobiDB and candidates PDB complexes were identified by querying the database.

This selection process used for the new entries in FuzDB v4.0 can only be applied to conformational ensembles. These were based on NMR data, which was complemented by additional experimental information (SAS, SAXS, FRET), or computational data from the Protein Ensemble Database (PED) (34). Complexes of disordered proteins or regions annotated in DisProt (35), were only considered if structural evidence for disorder in the bound state was available. In case a disordered region folded into a unique structure upon interacting with the partner, it was excluded from FuzDB. Structural data was used as deposited in the databases of origin, no post-processing was carried out.

Assembly of functional data

Biological activity of the protein was automatically derived from UniProt (36) and manually curated for some of the entries based on literature. The classified functional peptide motifs, (linear motifs), located in conformationally heterogeneous regions were derived from the Eukaryotic Linear Motif (ELM) database (37). Linear motifs were only assembled in case they were experimentally defined, i.e. true positive instances in ELM. In addition, posttranslational modification sites in the conformationally heterogeneous region were collected from the UniProt database.

Assembly of condensate information

The ability of FuzDB proteins to undergo liquid-liquid phase separation was assessed based on three databases: the PhaSepDB database (http:/db.phasep.pro) (38), which contains proteins from the literature with *in vivo* and *in vitro* experimental data on liquid-liquid phase separation, proteins from UniProt associated with known membraneless organelles, and proteins identified by high-throughput experiments of liquid–liquid phase separation. The PhaSePro database (https://phasepro.elte.hu) (39), which identifies protein regions associated with liquid–liquid phase separation and the LLPSDB database (http://bio-comp.org.cn/llpsdb) (40), which assembles proteins observed to undergo *in vitro* liquid–liquid phase separation under well-defined experimental conditions. We have used all proteins reported in these datasets, including both droplet-drivers, which can spontaneously phase separate and droplet-clients, which require additional components to partition into biomolecular condensates. When the information is available, for example in PhaSepDB and PhaSePro, the boundaries of the droplet-promoting regions are also reported.

DATABASE DESCRIPTION

Definition of fuzzy interactions

Fuzzy interactions are defined as context-dependent interactions (28), where contact patterns change with the cellular conditions. Fuzzy interactions are often mediated by disordered regions (1,10), with an impact on the biological activity of the protein or the respective assembly.

Fuzzy interactions are not defined based on previously proposed molecular mechanisms (10–12), as in higher-order assemblies these scenarios are not available due to the lack of high-resolution information. Dynamic interactions were only included in the database only in case the accompanying functional information was also available.

Data sources and processing

The primary source of all new entries is the Protein Data Bank, based on structural analysis of disordered, heterogeneous interactions (Materials and Methods). The data from the PDB is analyzed by the RING and MOBI software, and the filtered data is used without modifications (Materials and Methods). The 110 entries in the earlier versions (27), have been updated using new cross-links (see below).

The structural information from the Protein Data Bank on the protein assembly is complemented by experimental information on the free form derived from the DisProt database. In addition, structural data on the conformational ensemble from the Protein Ensemble Database includes a combination of experimental and computational methods. Cross-links to PED inform on compactness and variability in secondary structure population.

Functional regions include conserved domains from Pfam (41). Functional motifs, such as annotated short linear motifs and posttranslational modification sites in the fuzzy region are derived from the UniProt database and ELM (Materials and Methods). We use only experimentally identified instances of short motifs from all functional categories. The impact of the fuzzy interaction on the biological activity is described based on the literature. This information is primarily derived from the publications associated with the structural data in PDB and complemented by additional literature for functional studies. The literature was manually curated in case of each entry.

Each FuzDB entry associates the structural information from PDB and the functional data from literature. The two data types jointly evidence interaction fuzziness and the different biological outcomes related to diverse binding modes.

The ability to form biomolecular condensates via liquid-liquid phase separation is a novel functional aspect of fuzzy assemblies. This is not criteria to qualify for a FuzDB entry, rather a useful feature for analysis (see below). Cross-links are established with PhaSepDB (http:/db.phasep.pro), PhaSePro (https://phasepro.elte.hu), and LLPSDB (http://bio-comp.org.cn/llpsdb) databases. Based on the information provided in the original sources, FuzDB defines the role of the protein in condensates. Proteins, which were observed to undergo liquid-liquid phase separation in vitro without additional partners are classified as drivers. All other protein components of membraneless organelles are defined as clients.

Database organization

FuzDB has a modular organization. The homepage displays a brief description of the phenomenon and provides links to the Browse, Help, Tutorial, Fuzziness and References pages, which are also displayed in the top bar.

The homepage displays a Search field for simple text or more advanced combinations of terms. The statistics on the database entries provides information on both the number of assemblies and proteins. The homepage points to the FuzPred/FuzDrop server (http://protdyn-fuzpred.org) for prediction of fuzzy interactions and the probability to form biomolecular condensates from sequence.

The Browse page displays the Table of Entries (Figure 1), which contains the core information on each entry, such as the identifier, protein name as in UniProt, the UniProt identifier, definition of the fuzzy region(s), structure identifier in PDB, LLPS identifier in specific dataset of biomolecular condensates, reference of the PDB structure, or key reference for the functional impact of fuzzy interactions. Each FuzDB identifier opens an associated Entry page. In the Browse page, it is possible to search for specific fields like protein name, cross-reference identifiers (UniProt, PDB, BMRB, ELM, PhaSepDB, LLPSDB, PhaSePro), detection method, biological function, structure properties and significance. The last three elements are manually curated text fields.

The Help page provides tools for Application Programming Interface, such as query parameters, controlled vocabularies on experimental techniques (42). The Tutorial page gives guidelines for defining fuzziness of the assembly and deriving structure-function relationships. The Fuzziness page displays some key questions related to the concept of protein fuzziness and lists topology-based (10) and mechanistic classifications (12) of fuzzy assemblies. The References page lists the citations on various aspects of fuzziness, such as the concept, the molecular mechanisms of fuzzy interactions, higher-order structures, and methods to predict interaction behaviors from sequence.

The Entry page

The Entry page provides detailed information about the protein and its assembly (Figure 2). The information available on the page is summarized in the top right.

The title displays the Entry identifier and the protein name as reported in UniProt. On the top, the sequence viewer panel shows the location and sequence of fuzzy regions as well as the embedded the linear motifs and other posttranslational modification sites.

The Function section provides information on the biological activity of the protein and the assembly, based on UniProt. The Functional sites section summarizes the conserved protein domains (43) and the functional motifs in the fuzzy region based on UniProt and the ELM database. The posttranslational modification sites (from UniProt) in the fuzzy region are listed in a separate table, which can be expanded by clicking the tab below.

Information on condensates is a novel functional feature in FuzDB v4.0. The evidence for undergoing liquid-liquid phase separation as a driver or participating in biomolecular condensates as a client is derived from the PhaSepDB, PhaSePro and LLPSDB databases and the corresponding cross-links are shown.

The Structure section lists the regions, which remain conformationally heterogeneous in the specific assembly as observed by the given structure-determination method. Structure identifiers in the PDB, DisProt and PED databases and the corresponding references are listed. These provide a comprehensive description of the conformational ensemble of the protein in its free form (DisProt) and in the bound complex (PDB, PED). The structure panel contains a structure viewer by Mol* (44), which is high-

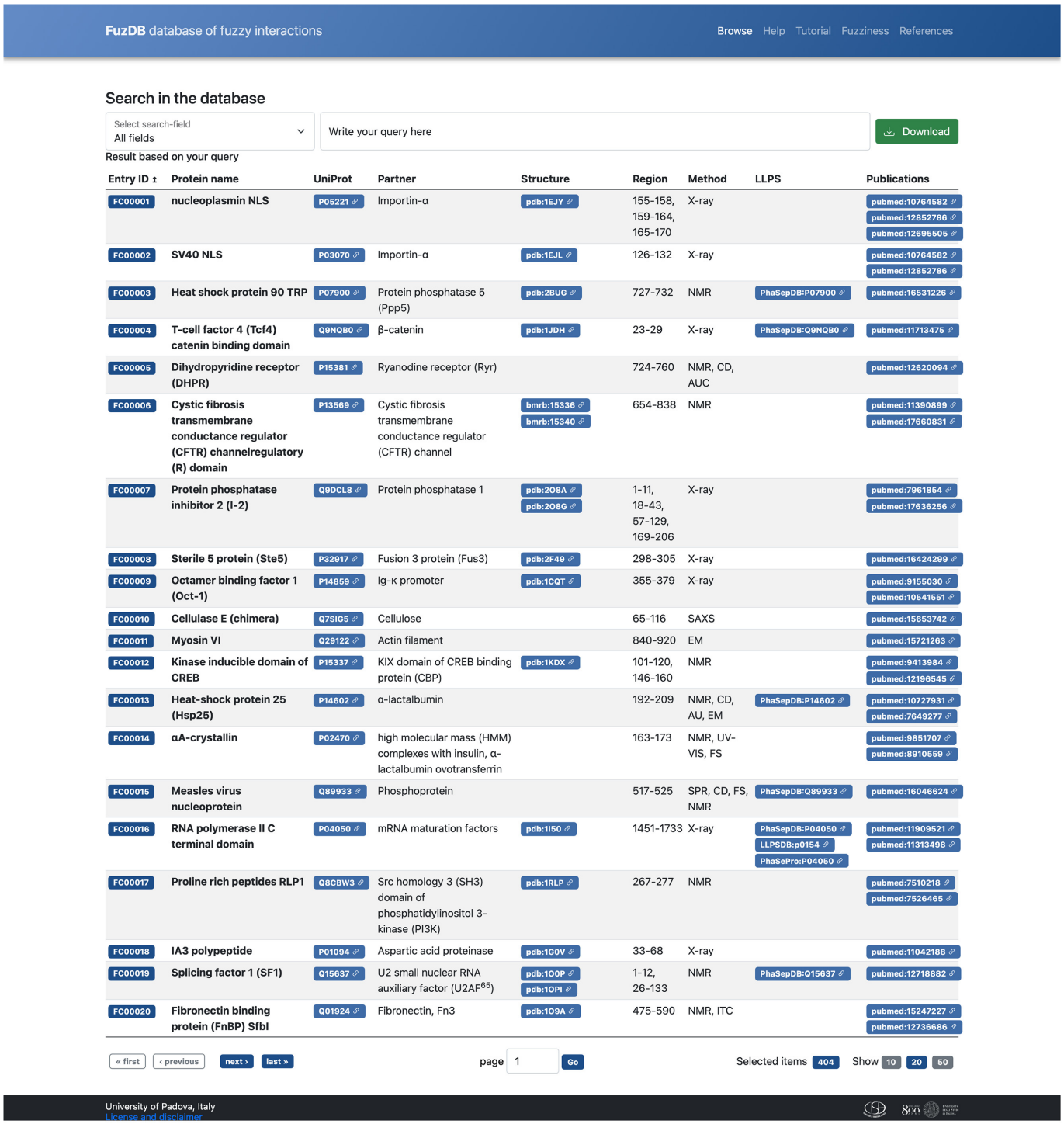

Figure 1. FuzDB Browse page. A Table of Entries (by default 20 entries/page) is shown with the following columns: FuzDB identifier, protein name as reported in the corresponding UniProt entry, the UniProt identifier, definition of the fuzzy region(s), structure identifier in the PDB, LLPS identifier in specific dataset of biomolecular condensates, and the key references for structure and function. By clicking on each FuzDB identifier the user can access the Entry page (Figure 2). The Search can be performed for free text in all fields, entry ID, protein name, UniProt identifier, structure identifier, detection method, partner protein, LLPS and ELM reference, and biological description. The Download on the top-right side is available for whole database or a selected subset of protein(s) in JSON, TSV, FASTA, XML and TXT formats.

Figure 2. FuzDB Entry page. The page is organized in different sections which are summarized in the right side (top). The top of the page shows the entry identifier and title as reported in UniProt. The feature viewer displays the fuzzy region/s, short linear motifs (SLiMs) and posttranslational modification sites (PTMs) mapped on the UniProt protein sequence. Sequence viewer highlights the amino acids involved in the fuzzy region/s. The Function section provides information on the biological activity based on the UniProt database. The Functional sites section lists the evolutionary conserved protein domains and short linear motifs (SLiMs) located in the fuzzy region/s. The boundaries of the functional regions, database cross-links and PubMed identifiers are also displayed. The Posttranslational modification site section displays the PTMs in the fuzzy region. Only the first 5 are shown, and the list can be expanded. The Condensate section lists the evidence to form biomolecular condensate either as driver or client and provide cross-links to the PhaSepDB, PhaSePro and LLPSDB databases. The Structure section lists the region/s that remain heterogeneous in the bound state based on experimental structure-determination. The table shows cross-links to PDB, PED and DisProt databases and the corresponding PubMed identifier/s. The Significance section summarizes the biological impact of fuzzy interactions on the biological activity.

Table 1. Application areas of FuzDB

Area	Specific problems
Structure-Function Analysis (*structural biology*)	Fuzzy complex discovery, experimental methods
	Structural characterization of fuzzy regions
	Molecular mechanisms of fuzzy interactions
	Structure-function relationships
Sequences Codes (*evolutionary biology*)	Sequence motifs in context-dependent interactions
	Protein domains associated with fuzzy regions
	Evolutionary conservation in fuzzy regions
	Co-evolutionary signatures in fuzzy partnerships
Biomolecular Condensates (*cell biology*)	Contact amino acid types
	Contact patterns, frequencies
	Client interaction motifs
Regulatory Motifs (*systems biology*)	PTMs associated with fuzzy interactions
	Structural features related to context-dependence
	PTMs driving higher-order organization of proteins
Method Development (*bioinformatics*)	Prediction of context-dependent interactions
	Prediction of condensates, drivers and clients
	Prediction of droplet hot-spots

lighted above the image. The structure can be rotated, moved, zoomed and the individual conformers can also be visualized.

The Significance section is a brief description of the biological impact of fuzzy interactions.

Database updates

An automatic update of the Uniprot-based information is performed regularly. Updating the database will be performed in each 6 months, using the automated structure analysis routine (Materials and Methods). These updates will follow the growing number of entries in the biomolecular condensate databases. An independent source for updates will include lower resolution structures, which are not deposited in PDB, but the experimental data is available in Biological Magnetic Resonance Data Bank (BMRB (45)). The third source for new database entries will be literature search, based on reported fuzzy interactions in specific assemblies. All entries in the updates will be manually curated.

Using FuzDB

The FuzDB v4.0 database can be used for (Table 1) (i) the discovery of fuzzy complexes and their experimental analysis; (ii) the comparative analysis of fuzzy assemblies; (iii) the analysis of interactions in biomolecular condensates; (iv) the analysis of motifs regulating fuzzy interactions; (v) the development of sequence-based methods to predict cellular interaction behaviors and formation of biomolecular condensates.

Identification fuzzy complexes and methods for experimental analysis

Fuzzy interactions are difficult to detect using traditional structure determination methods (46,47). FuzDB v4.0 database provides detailed insights into the structure determination process of fuzzy assemblies via cross-links to spectral data in BMRB, PDB and PED as well as the corresponding references. All these data may guide the experimental characterization of the conformational ensembles of potential fuzzy assemblies in particular, in case of considerable sequence similarity or evolutionary relationship, common partner or functionality with entries already available in the database. The recognition of the phenomenon of fuzziness facilitates the analysis of segments, which exhibit considerable spatial and temporal variations in specific contacts with the partner. Based on the analyzed examples, structure-function relationships can be derived (Figure 3).

Comparative analysis of fuzzy assemblies

Our knowledge on the biophysical forces driving towards well-defined or heterogeneous interactions is incomplete. Furthermore, many proteins exhibit different binding modes depending on the partner and cellular conditions (28,29). FuzDB also contains different assemblies of the same protein, which can be analyzed for the underlying sequence codes. Comparative analysis informs on the structure and binding characteristics related to distinct recognition modes. The impact of the interacting partner on the different binding modes can be studied, dissecting intrinsic and extrinsic factors for distinct behaviors. In addition, comparative analysis of fuzzy regions can be used to derive evolutionary relationships. In case of low sequence similarity, evolutionary relationships can be indicated based on similar dynamical characteristics (48). The association of protein domains with fuzzy regions can also be systematically investigated.

Analysis of interactions in biomolecular condensates

Understanding the molecular organization of biomolecular condensates requires high resolution structure analysis, which is problematic due to high concentration, variable stoichiometry and conformational heterogeneity (17,49). Biophysicists and cell biologists working on liquid-liquid phase separation aim at identifying sequence motifs driving formation of biomolecular condensates under healthy and pathological conditions. FuzDB v4.0 highlights those assemblies, which were also identified as components of biomolecular condensates. These fuzzy, stoichiometric assemblies are likely stabilized by similar interactions, which also operate in higher-order assemblies and thus can help deciphering the biophysical driving forces of condensate interactions. In this manner, FuzDB can be used to obtain detailed information on variable, dynamic contacts of protein regions, which also promote formation or stabilize condensates. Thus, fuzzy assemblies inform on geometrical features, stability and variability of contacts, amino acid types and networks in condensates. Increasing evidence for condensate formation will underscore the potential of FuzDB in this direction.

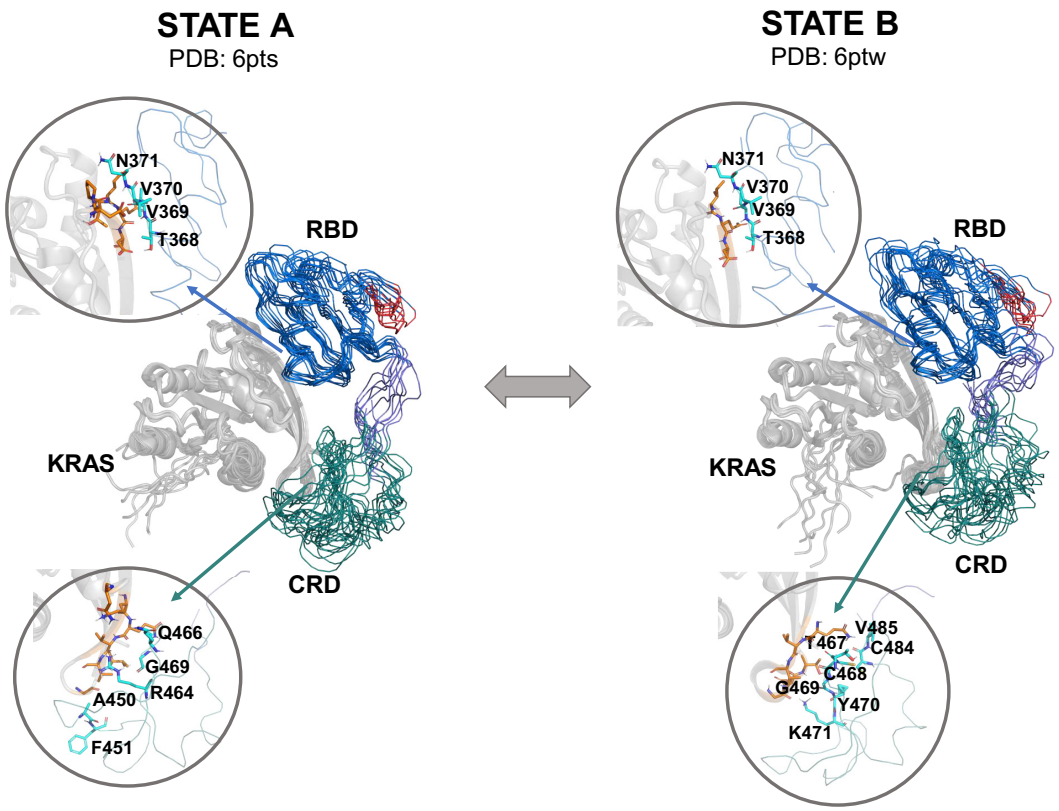

Figure 3. Fuzzy interactions by KRAS control the signaling outcome of the MAPK pathway (FC00179). Activation of RAF by KRAS is achieved via shifting its conformation equilibrium in the membrane-bound form (50). In state A (PDB:6pts), α4-α5 of KRAS and the RAF cystein-rich domain (CRD, green) interacts with the membrane surface, while the RAF RAS-binding domain (RBD, blue) is distant from the membrane. In state B (PDB:6ptw), the RAF RBD approaches the membrane surface, enabling KRAS α4-α5 separation. The exposed KRAS α4-α5 induces dimerization leading to enhanced MAPK signaling activity. Cancer mutations stabilize state B resulting in more active MAPK signaling. KRAS residues involved in fuzzy interactions are shown by orange, RAF residues are cyan. Loop residues ^{106}KGKKAR^{111} distinguished in the conformational transition are shown in red.

Analysis of motifs regulating fuzzy regions

FuzDB lists the posttranslational modification sites (from UniProt), which are located within fuzzy regions. PTMs (i) can modulate the conformational properties of the ensemble for partner recognition (ii) can mediate specific contacts or (iii) enable selectivity for given sequence contexts. Large-scale analysis of the database can identify those posttranslational modifications, which are frequently associated with fuzzy interactions. This information can also be projected to large-scale protein interaction networks or phosphoproteomic data (50).

Development of sequence-based methods to predict interaction behaviors and biomolecular condensates

FuzDB facilitated the development of the FuzPred method to identify protein regions, which undergo disorder-to-order transition or disorder-to-disorder transition upon binding (30). Such prediction can be performed using the sequence of the protein, without specifying the binding partner. In addition, one can estimate the likelihood to change the preferred binding mode, i.e. protein regions, which can sample both ordered and disordered interactions (29). These principles have been exploited for developing the FuzDrop method to predict droplet-promoting regions driving protein liquid–liquid phase separation (19) and hotspots for conversion of droplets to fibrillar aggregates (24). The growing size of the database facilitates optimization of the current predictors, as well as the development of novel ones, in particular for interactions in biomolecular condensates. A negative dataset for training prediction methods can be assembled for example based on DisProt entries, which were not included in FuzDB database. In addition, the computational algorithms used to define FuzDB entries (Materials and Methods) can be applied to automatically generate both positive and negative datasets for prediction method training.

Example of a FuzDB entry and significance of the fuzzy interaction

FC00179 illustrates how fuzzy interactions in different conformational sub-states lead to distinct biological activities. KRAS is a small GTPase involved in MAPK signalling, and activates RAF kinases via dynamic, multivalent interactions. In particular, KRAS is anchored to the membrane via a canonical lipid-binding site, and an alternative interface, the variations of which result in two functional states for the KRAS/RAF complex (50) (Figure 3). In state A (PDB:6pts), α4–α5 of KRAS interacts with the

membrane surface, while the β2–β3 loop interacts with the RAF cystein-rich domain (CRD). In contrast, in state B (PDB:6ptw), membrane interactions are dominantly mediated by the RAF RAS-binding domain (RBD, β3–β4 loop), which was distant from the membrane in state A. Consequently, the KRAS α4–α5 separates from the membrane in state B and its exposure induces KRAS dimerization leading to enhanced signaling activity. Cancer mutations stabilize state B resulting in more active MAPK signaling. Overall, the fine-tuned fuzzy interactions control the conformational equilibrium of the ternary complex and the signaling outcome (50).

CONCLUSIONS AND PERSPECTIVES

The advances in structure determination techniques provide deeper insights into the complexity of conformational and interaction behaviors of proteins, in particular, within the context of functional partners or under cellular conditions. These rapidly accumulating results highlight the importance to reformulate the classical structure-function paradigm as a relationship between ensembles of conformations and functions (1,2). Such links should reveal how structural heterogeneity with the functional partner leads to distinct biological outcomes. These efforts require the structural analysis of the underlying conformational ensembles as well as functional data on protein variants, under different cellular conditions.

FuzDB facilitates the development of such context-dependent structure-function relationships. The database establishes direct experimental evidence between conformational heterogeneity of protein assemblies and their biological impact. FuzDB v4.0 provides cross-references for individual or comparative analysis of proteins and protein regions involved in fuzzy interactions. Furthermore, the cross-links to databases on biomolecular condensates are available, highlighting the importance of fuzzy interactions in liquid-liquid phase separation. In addition, FuzDB provides high-resolution structural information on droplet-forming regions. The database will be expanded using examples, which were determined by cryo-EM or based on literature, without structure deposited in PDB. FuzDB v4.0 will be linked to the new version of the FuzPred/FuzDrop server, which is in the process of being moved to University of Padova.

Increasing recognition of conformational heterogeneity and multifunctionality of proteins will make FuzDB as a useful resource for elucidating the biophysical principles and sequence codes of complex cellular interaction behaviors.

FUNDING

Hungarian Academy of Sciences [HAS-11015 to M.F.]; European Union's Horizon 2020 research and innovation program [778247, 952334]; ELIXIR, the European infrastructure for biological data. Funding for open access charge: European Union's Horizon [2020 778247, 952334]
Conflict of interest statement. None declared.

REFERENCES

1. Fuxreiter,M. (2018) Fuzziness in protein interactions - a historical perspective. *J. Mol. Biol.*, **430**, 2278–2287.
2. Fuxreiter,M. (2018) Towards a stochastic paradigm: from fuzzy ensembles to cellular functions. *Molecules*, **23**, 3008.
3. Buljan,M., Chalancon,G., Eustermann,S., Wagner,G.P., Fuxreiter,M., Bateman,A. and Babu,M.M. (2012) Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell*, **46**, 871–883.
4. Stamos,J.L., Chu,M.L., Enos,M.D., Shah,N. and Weis,W.I. (2014) Structural basis of GSK-3 inhibition by N-terminal phosphorylation and by the Wnt receptor LRP6. *eLife*, **3**, e01998.
5. Miskei,M., Gregus,A., Sharma,R., Duro,N., Zsolyomi,F. and Fuxreiter,M. (2017) Fuzziness enables context dependence of protein interactions. *FEBS Lett.*, **591**, 2682–2695.
6. Kragelj,J., Ozenne,V., Blackledge,M. and Jensen,M.R. (2013) Conformational propensities of intrinsically disordered proteins from NMR chemical shifts. *ChemPhysChem*, **14**, 3034–3045.
7. Marsh,J.A. and Forman-Kay,J.D. (2011) Ensemble modeling of protein disordered states: experimental restraint contributions and validation. *Proteins*, **80**, 556–572.
8. Bonomi,M. and Vendruscolo,M. (2019) Determination of protein structural ensembles using cryo-electron microscopy. *Curr. Opin. Struct. Biol.*, **56**, 37–45.
9. Panchenko,A.R. and Fuxreiter,M. (2019) Dynamic protein interactions - from complexes to molecular machines. *Curr. Opin. Struct. Biol.*, **56**, vi–viii.
10. Tompa,P. and Fuxreiter,M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.*, **33**, 2–8.
11. Fuxreiter,M. and Tompa,P. (2012) Fuzzy complexes: a more stochastic view of protein function. *Adv. Exp. Med. Biol.*, **725**, 1–14.
12. Fuxreiter,M. (2012) Fuzziness: linking regulation to protein dynamics. *Mol. Biosyst.*, **8**, 168–177.
13. Wu,H. and Fuxreiter,M. (2016) The structure and dynamics of higher-order assemblies: amyloids, signalosomes, and granules. *Cell*, **165**, 1055–1066.
14. Fuxreiter,M. and Vendruscolo,M. (2021) Generic nature of the condensed states of proteins. *Nat. Cell Biol.*, **23**, 587–594.
15. Banani,S.F., Lee,H.O., Hyman,A.A. and Rosen,M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.*, **18**, 285–298.
16. Boeynaems,S., Alberti,S., Fawzi,N.L., Mittag,T., Polymenidou,M., Rousseau,F., Schymkowitz,J., Shorter,J., Wolozin,B., Van Den Bosch,L. *et al.* (2018) Protein phase separation: a new phase in cell biology. *Trends Cell Biol.*, **28**, 420–435.
17. Reichheld,S.E., Muiznieks,L.D., Keeley,F.W. and Sharpe,S. (2017) Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E4408–E4415.
18. Niaki,A.G., Sarkar,J., Cai,X., Rhine,K., Vidaurre,V., Guy,B., Hurst,M., Lee,J.C., Koh,H.R., Guo,L. *et al.* (2020) Loss of dynamic RNA interaction and aberrant phase separation induced by two distinct types of ALS/FTD-Linked FUS mutations. *Mol. Cell*, **77**, 82–94.
19. Hardenberg,M., Horvath,A., Ambrus,V., Fuxreiter,M. and Vendruscolo,M. (2020) Widespread occurrence of the droplet state of proteins in the human proteome. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 33254–33262.
20. Wippich,F., Bodenmiller,B., Trajkovska,M.G., Wanka,S., Aebersold,R. and Pelkmans,L. (2013) Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell*, **152**, 791–805.
21. Roden,C. and Gladfelter,A.S. (2021) RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.*, **22**, 183–195.
22. Boija,A., Klein,I.A., Sabari,B.R., Dall'Agnese,A., Coffey,E.L., Zamudio,A.V., Li,C.H., Shrinivas,K., Manteiga,J.C., Hannett,N.M. *et al.* (2018) Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell*, **175**, 1842–1855.

23. Alberti,S. and Hyman,A.A. (2021) Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.*, **22**, 196–213.
24. Vendruscolo,M. and Fuxreiter,M. (2021) Sequence determinants of the aggregation of proteins within condensates generated by liquid-liquid phase separation. *J. Mol. Biol.*, 167201.
25. Fuxreiter,M. (2018) Fold or not to fold upon binding - does it really matter? *Curr. Opin. Struct. Biol.*, **54**, 19–25.
26. Bah,A., Vernon,R.M., Siddiqui,Z., Krzeminski,M., Muhandiram,R., Zhao,C., Sonenberg,N., Kay,L.E. and Forman-Kay,J.D. (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature*, **519**, 106–109.
27. 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.
28. Fuxreiter,M. (2020) Classifying the binding modes of disordered proteins. *Int. J. Mol. Sci.*, **21**, 8615.
29. Horvath,A., Miskei,M., Ambrus,V., Vendruscolo,M. and Fuxreiter,M. (2020) Sequence-based prediction of protein binding mode landscapes. *PLoS Comp Biol*, **16**, e1007864.
30. Miskei,M., Horvath,A., Vendruscolo,M. and Fuxreiter,M. (2020) Sequence-based prediction of fuzzy protein interactions. *J. Mol. Biol.*, **432**, 2289–2303.
31. 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.
32. Piovesan,D., Minervini,G. and Tosatto,S.C. (2016) The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.*, **44**, W367–W374.
33. Piovesan,D. and Tosatto,S.C.E. (2018) Mobi 2.0: an improved method to define intrinsic disorder, mobility and linear binding regions in protein structures. *Bioinformatics*, **34**, 122–123.
34. Lazar,T., Martinez-Perez,E., Quaglia,F., Hatos,A., Chemes,L.B., Iserte,J.A., Mendez,N.A., Garrone,N.A., Saldano,T.E., Marchetti,J. *et al.* (2021) PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic Acids Res.*, **49**, D404–D411.
35. Hatos,A., Hajdu-Soltesz,B., Monzon,A.M., Palopoli,N., Alvarez,L., Aykac-Fas,B., Bassot,C., Benitez,G.I., Bevilacqua,M., Chasapi,A. *et al.* (2020) DisProt: intrinsic protein disorder annotation in 2020. *Nucleic Acids Res.*, **48**, D269–D276.
36. UniProt,C. (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic. Acids. Res.*, **49**, D480–D489.
37. Kumar,M., Gouw,M., Michael,S., Samano-Sanchez,H., Pancsa,R., Glavina,J., Diakogianni,A., Valverde,J.A., Bukirova,D., Calyseva,J. *et al.* (2020) ELM-the eukaryotic linear motif resource in 2020. *Nucleic Acids Res.*, **48**, D296–D306.
38. You,K., Huang,Q., Yu,C., Shen,B., Sevilla,C., Shi,M., Hermjakob,H., Chen,Y. and Li,T. (2020) PhaSepDB: a database of liquid-liquid phase separation related proteins. *Nucleic Acids Res.*, **48**, D354–D359.
39. Meszaros,B., Erdos,G., Szabo,B., Schad,E., Tantos,A., Abukhairan,R., Horvath,T., Murvai,N., Kovacs,O.P., Kovacs,M. *et al.* (2020) PhaSePro: the database of proteins driving liquid-liquid phase separation. *Nucleic Acids Res.*, **48**, D360–D367.
40. Li,Q., Peng,X., Li,Y., Tang,W., Zhu,J., Huang,J., Qi,Y. and Zhang,Z. (2020) LLPSDB: a database of proteins undergoing liquid-liquid phase separation in vitro. *Nucleic Acids Res.*, **48**, D320–D327.
41. Mistry,J., Chuguransky,S., Williams,L., Qureshi,M., Salazar,G.A., Sonnhammer,E.L.L., Tosatto,S.C.E., Paladin,L., Raj,S., Richardson,L.J. *et al.* (2021) Pfam: the protein families database in 2021. *Nucleic Acids Res.*, **49**, D412–D419.
42. Troilo,F., Bignon,C., Gianni,S., Fuxreiter,M. and Longhi,S. (2018) Experimental characterization of fuzzy protein assemblies: interactions of paramyxoviral NTAIL domains with their functional partners. *Methods Enzymol.*, **611**, 137–192.
43. Pal,A. and Levy,Y. (2020) Balance between asymmetry and abundance in multi-domain DNA-binding proteins may regulate the kinetics of their binding to DNA. *PLoS Comput. Biol.*, **16**, e1007867.
44. Sehnal,D., Bittrich,S., Deshpande,M., Svobodova,R., Berka,K., Bazgier,V., Velankar,S., Burley,S.K., Koca,J. and Rose,A.S. (2021) Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.*, **49**, W431–W437.
45. Ulrich,E.L., Akutsu,H., Doreleijers,J.F., Harano,Y., Ioannidis,Y.E., Lin,J., Livny,M., Mading,S., Maziuk,D., Miller,Z. *et al.* (2008) BioMagResBank. *Nucleic Acids Res.*, **36**, D402–D408.
46. Jensen,M.R., Zweckstetter,M., Huang,J.R. and Blackledge,M. (2014) Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem. Rev.*, **114**, 6632–6660.
47. Rezaei-Ghaleh,N., Blackledge,M. and Zweckstetter,M. (2012) Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *ChemBioChem*, **13**, 930–950.
48. Zsolyomi,F., Ambrus,V. and Fuxreiter,M. (2020) Patterns of dynamics comprise a conserved evolutionary trait. *J. Mol. Biol.*, **432**, 497–507.
49. Burke,K.A., Janke,A.M., Rhine,C.L. and Fawzi,N.L. (2015) Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Mol. Cell*, **60**, 231–241.
50. Woodsmith,J. and Stelzl,U. (2014) Studying post-translational modifications with protein interaction networks. *Curr. Opin. Struct. Biol.*, **24**, 34–44.