FuzDrop on AlphaFold: visualizing the sequence-dependent propensity of liquid–liquid phase separation and aggregation of proteins

Andras Hatos [1], Silvio C.E. Tosatto [1], Michele Vendruscolo [2] and Monika Fuxreiter [1,*]

[1]Department of Biomedical Sciences, University of Padova, via Ugo Bassi 58/B, 35131 Padova, Italy and [2]Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

ABSTRACT

Many proteins perform their functions within membraneless organelles, where they form a liquid-like condensed state, also known as droplet state. The FuzDrop method predicts the probability of spontaneous liquid–liquid phase separation of proteins and provides a sequence-based score to identify the regions that promote this process. Furthermore, the FuzDrop method estimates the propensity of conversion of proteins to the amyloid state, and identifies aggregation hot-spots, which can drive the irreversible maturation of the liquid-like droplet state. These predictions can also identify mutations that can induce formation of amyloid aggregates, including those implicated in human diseases. To facilitate the interpretation of the predictions, the droplet-promoting and aggregation-promoting regions can be visualized on protein structures generated by AlphaFold. The FuzDrop server (https://fuzdrop.bio. unipd.it) thus offers insights into the complex behavior of proteins in their condensed states and facilitates the understanding of the functional relationships of proteins.

GRAPHICAL ABSTRACT

INTRODUCTION

Many proteins have been observed to form a dense, liquid-like state (1–4), which can be accessed from the native state in a reversible manner under cellular conditions (5), and contribute to a wide range of cellular activities (6). This state is thus emerging as a fundamental state of proteins along with the native and amyloid states (7). Among its biological functions, the droplet state can: (i) increase the local concentrations of different cellular components, which can accelerate enzymatic reactions and amplify signals, as in the case of cyclic GMP–AMP synthase (cGAS) in innate immune signaling (8), (ii) form signaling clusters for low-affinity effectors and ligands, such as in T-cell receptor (9) and Wnt (10) signaling, (iii) facilitate nucleation in polymerization reactions, such as microtubulin for centrosome formation (11,12), (iv) orchestrate components of a given cellular pathway, such in the case of p53-binding protein 1 (53BP1) (13), where droplets concentrate components for DNA repair, or of heterochromatin protein 1 (HP1), where droplets regulate gene silencing (14).

Given the widespread nature of the phenomenon of liquid–liquid phase separation, it is relevant to identify the proteins that can form the droplet state spontaneously under cellular conditions. In particular, many proteins observed in membraneless organelles do not readily form droplets in the test-tube by themselves (15,16). Thus, it is important to classify droplet-driver proteins, which can spontaneously undergo liquid–liquid phase separation, and droplet-client proteins, which require interactions with partners or specific cellular conditions to form the droplet state. These classifications distinguish proteins comprising droplet-promoting regions, and proteins with stable self-assembly of droplet-promoting regions. Recently the FuzDrop method was developed to address this problem (4).

Another important question concerns the tendency of proteins to aggregate within the droplet state. Although the liquid-like droplet state is functional in most cases, it can prompt an irreversible transition to the dysfunctional solid-like amyloid state (17,18). Familial mutations driving this
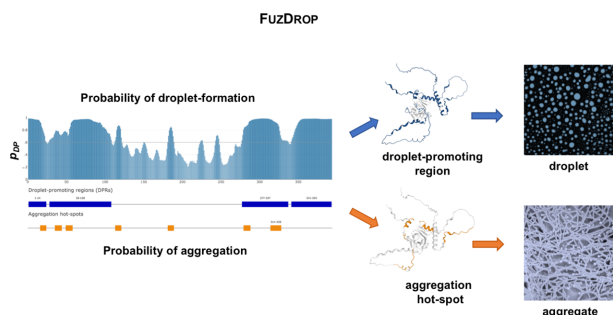
*To whom correspondence should be addressed. Tel: +39 049 827 6136; Email: monika.fuxreiter@unipd.it

process are associated with a wide variety of human diseases (19), including neurological disorders (20), cancers (21) and viral infections (22). Identifying regions where mutations can drive the conversion from liquid-like to solid-like condensates is likely to provide insights into the molecular mechanisms of disease and help develop therapeutic interventions. By analysing the context-dependence of the interactions within the droplet state, the FuzDrop method can also identify aggregation hot-spots (23), which are the regions that initiate the irreversible maturation of condensates leading to fibrillar aggregates.

In this article, we describe the FuzDrop server (https://fuzdrop.bio.unipd.it), which provides two key sequence-based predictions concerning the condensed states of proteins: (i) the probability to form the droplet state through spontaneous liquid-liquid phase separation (4), and (ii) the likelihood of regions to aggregate within the liquid droplets (23) (Figure 1). The web server thus offer insights into the complex behavior of proteins in their condensed states and facilitates developing functional relationships based on the condensed-state behavior.

PRINCIPLES TO ENABLE THE PREDICTION OF THE CONDENSED STATE BEHAVIOR OF PROTEINS

Interactions in the droplet state of proteins

Because of its liquid-liquid nature, the droplet state is characterized by a high conformational entropy. In this view, the droplet state is a state where proteins interconvert rapidly among many different binding configurations (7). Such disordered binding mode can be achieved through contacts via multiple, alternative binding sites (24), a behaviour known as multivalency (25,26). These binding sites, which are often associated with low-complexity or disordered regions (27), however, are often difficult to identify from the amino acid sequences (4). The presence of a complex code for droplet formation is also suggested by the observation that also structured, globular proteins can undergo liquid-liquid phase separation (28,29).

The FuzDrop approach is based on a model in which the droplet state is stabilized by disordered interactions (4). Sequences sampling disordered binding modes lack the local compositional bias that drives structural ordering upon binding (24). Disordered binding modes are also observed in stoichiometric protein complexes. Structures formed by these sequences exhibit high density of frustrated contacts in both free and bound forms (30,31). The FuzPred algorithm (24) determines the degree of local sequence bias with respect to composition, hydrophobicity and structural disorder examining large number of possible sequence contexts (24). Determining these biases using different contexts enables the method to provide robust results over many different binding partners and cellular conditions.

The FuzDrop method (4) combines the prediction of disordered binding modes by the FuzPred method (24) with the estimation of the degree of protein disorder by the Espritz algorithm (32). The method consists in two main steps: (i) determination of the probability of protein regions to promote droplet formation and (ii) evaluation of the stability of their self-interactions to predict the probability of the protein to sample the droplet state. The first part of method

was trained on a set of 120 protein regions observed to drive liquid-liquid phase separation (33). As a negative set, regions from proteins not known to condensate with similar length distribution as in the positive set were used (4). The coefficients for disordered interactions and structural disorder were determined by logistic regression, resulting in AUC values of 86% and 87% on the training and test set, respectively (4).

The probability of proteins to form the droplet state was based on estimating the propensity of droplet-promoting regions as well as the stability of their self-interactions (4). The latter term was based on hydrophobic motifs embedded in disordered protein regions that stabilize self-assembly via hydrophobic effects. To predict the probability of proteins to form the droplet state, the method was trained on a set of ~400 protein sequences observed to undergo liquid-liquid separation in vitro or in vivo (4). As a negative set, all other sequences of the same organisms not known to form condensates were used. The parameters were trained using logistic regression resulting in AUC values of 88% and 91% on the training and test set, respectively (4).

Interactions driving aggregation of proteins within their droplet state

Proteins in their droplet state may in some cases convert into a more stable amyloid state (7). Since this process is associated with a wide range of human disorders (34), there is a great interest in identifying regions that can drive this process. The formation of both condensed states is driven by non-native interactions (7), which are ordered in the amyloid state and disordered in the droplet state. Thus, protein regions that drive aggregation within liquid-like condensates are capable of switching between disordered and ordered interactions (23). This interaction property, which changes between interaction modes, is denoted as context-dependence, which leads to a wide variety of cellular behaviors under different conditions (35,36).

The context-dependence of protein interactions can be estimated by the FuzPred method (37). This approach is based on the analysis of different interaction modes, which are predicted with different possible binding interfaces, corresponding to different partners and cellular conditions. The ability to sample different binding modes is quantified by the Shannon entropy of the binding mode distribution (37). Thus, the FuzPred method can estimate the interaction behavior of protein regions at individual residue resolution: (i) the most likely interaction mode, i.e. the extent to which the protein residue remains ordered or disordered upon binding (24) and (ii) the likelihood that the residue samples other binding modes (37).

The FuzDrop method combines the prediction of the probability to form the droplet state (4) with the estimation of interaction context-dependence from the FuzPred algorithm (37). This approach can evaluate the stability of interaction modes in the droplet state, and thus identifies protein regions that can change disordered to ordered binding modes and drive aggregation (23). The method was shown to identify mutations associated with amyotrophic lateral sclerosis (ALS) and distinguish them from mutations in the same proteins, which are not pathological (38). For exam-
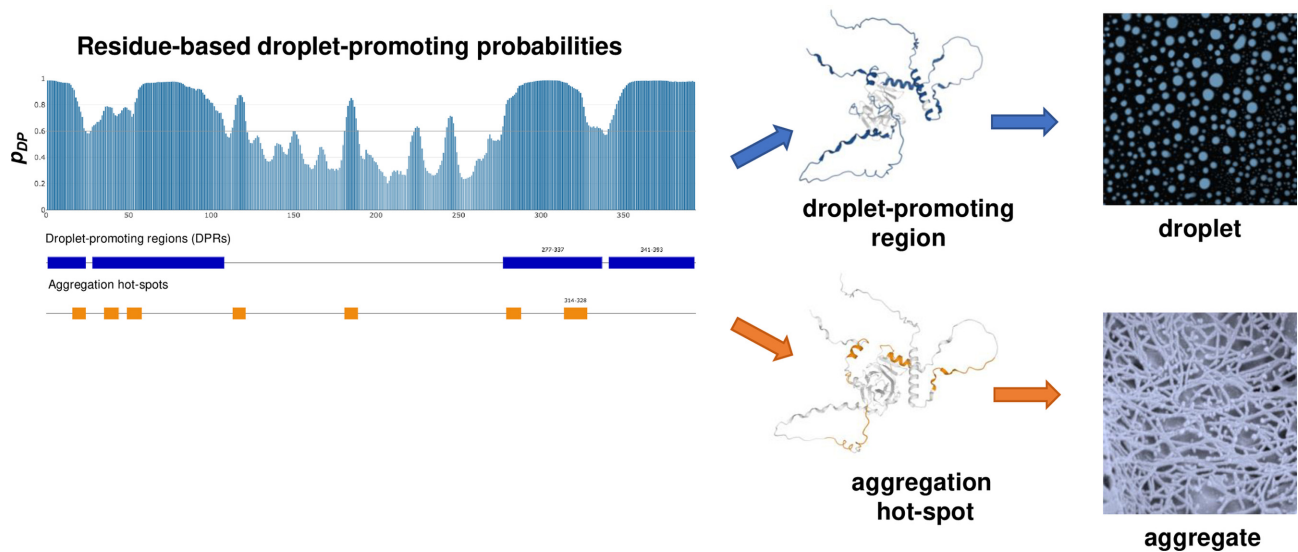
Figure 1. Functional features predicted by the FuzDrop server (https://fuzdrop.bio.unipd.it). The FuzDrop method provides two key sequence-based predictions concerning the condensed states of proteins: (i) the probability to form the droplet state through spontaneous liquid–liquid phase separation (4) and (ii) the tendency of regions to aggregate within the liquid droplets (23). Thus, the FuzDrop server gives insights into the complex behavior of proteins in their condensed states.

ple, the ALS-associated FUS G156E mutant has a much higher context-dependence than the non-pathological FUS G154 mutant, which corresponds to a higher likelihood of the conversion to the ordered amyloid state (23).

PREDICTED CONDENSED-STATE CHARACTERISTICS AVAILABLE FROM THE FUZDROP SERVER

Sequence-based probability of droplet formation

All predictions require a protein sequence, either the UniProt code (39) or the FASTA file as an input. The results are displayed on a separate page. On the top, the probability of spontaneous liquid–liquid phase separation (p_{LLPS}) is shown, which informs on the ability of protein to drive droplet formation (Figure 2A). Proteins with $p_{LLPS} \geq 0.60$ likely spontaneously phase separate and serve as droplet-drivers. The threshold was determined upon parametrization of the FuzDrop method (4).

Below the droplet-promoting probabilities of residues (p_{DP}) are displayed (Figure 2A). These values vary between 0 and 1 and inform on the ability of residues to be involved in droplet-interactions. Residues with $p_{DP} \geq 0.60$ values promote droplet formation, as indicated by a bold line (4). The graph is interactive, as one can zoom on the region of interest and then return to the original graph showing the complete protein sequence. This feature facilitates the analysis of small variations in droplet propensities, for example to guide the design of mutant forms with tailored droplet propensities.

Droplet-promoting regions and aggregation hot-spots

Droplet-promoting regions (DPRs), which are defined as sequence stretches of ≥ 10 consecutive residues with $p_{DP} \geq 0.60$ values, are displayed below the graph showing

the p_{DP} values (Figure 2A). Boundaries of DPRs are displayed above the blue boxes or by positioning the cursor into any of the blue boxes.

Aggregation hot-spots are defined as parts of droplet-promoting regions with large binding mode diversity (Table 1). These are displayed as orange boxes with the boundaries shown above. Aggregation hot-spots have a minimum length of 5 residues with a gap of maximum two residues allowed.

Cellular context-dependence (binding mode entropy)

Protein regions often change binding modes depending on the cellular conditions, partners, post-translational modifications (PTMs). We characterize the cellular context-dependence, i.e. the ability of proteins to alter their binding modes, by the binding mode entropy (S_{bind}). This quantity can be determined as a Shannon entropy computed from the frequencies of different binding modes in the presence of different, hypothetical partners (37). Below the droplet-promoting regions and aggregation hot-spots, the interactive graph displays the binding mode diversity of residues (S_{bind}) (Figure 2B). The threshold $S_{bind} \geq 2.20$ is used to identify aggregation hot-spots. Context-dependent regions, defined as ≥ 10 consecutive residues with $S_{bind} \geq 2.20$ are displayed below the graph, with residue boundaries indicated (Figure 2B). A more in-depth analysis of context-dependence is performed in the FuzPred server (https://fuzpred.bio.unipd.it).

Protein features and cross-links to other databases

Below the predicted results, sequence features, related to cellular context dependence are displayed (Figure 3). Fuzzy regions, which are regions where the structural disorder in the bound state has a functional impact, are graphically displayed. This information is derived from FuzDB (40), and

Figure 2. Predicted condensed-state characteristics (**A**) and cellular context-dependence (**B**) of the human p53 tumor suppressor (P04637). The p_{LLPS} value on the top informs on whether the protein can spontaneously phase separate and drive droplet formation ($p_{LLPS} \geq 0.60$). (A) The droplet-promoting probabilities of residues (p_{DP}) are displayed on an interactive graph. Droplet-promoting regions (DPRs), which are defined as ≥ 10 consecutive residues with $p_{DP} \geq 0.60$, are displayed below the graph. Boundaries of DPRs are displayed above the blue boxes or by positioning the cursor into any of the blue boxes. Aggregation hot-spots, which drive the conversion of the liquid-like droplet state to solid-like amyloid state, are also shown (orange boxes). These regions have high interaction mode diversity and are capable of sampling both disordered and ordered binding modes. (B) Cellular context-dependence is characterized by the binding mode diversity (S_{BIND}) computed from the frequencies of different binding modes in the presence of different, hypothetical partners (37). This quantity characterizes the ability of residues to switch between disordered and ordered interactions. Residues with a wide spectrum of interaction behaviors will be most affected by the cellular conditions.

Table 1. Main applications of the FuzDrop server. The choices for the thresholds were described in the original references (4,23)

Biological problem	Parameter, threshold
Identification of proteins undergoing liquid–liquid phase separation.	$p_{LLPS} \geq 0.60$
Identification of protein regions promoting droplet formation or partitioning in liquid droplets.	$p_{DP} \geq 0.60$, $l \geq 10$ aa
Identification of droplet-driver or droplet-client proteins.	Droplet driver: $p_{LLPS} \geq 0.60$ Droplet client: $p_{LLPS} < 0.60$; $p_{DP} \geq 0.60$ for at least 10 consecutive residues
Identification of aggregation hot-spots within the droplet state.	$p_{DP} \geq 0.60$; $S_{bind} \geq 2.2$; $l \geq 5$ aa

the overlap of droplet-promoting region with known fuzzy regions indicates different functional behaviors under different cellular conditions. In addition, a further layer of regulation can be provided by PTMs, which can interfere with the formation of protein droplets or modulate biophysical properties of the liquid-like state (41). PTMs derived from UniProt (39) are shown by red dots, and positioning the cursor above them will display the modified residue, the PTM type and the modifying enzyme. These data may inform on which enzymes regulate liquid–liquid phase separation of the given sequence. In addition, the Pfam domains (42) are displayed (Figure 3), so that their involvement in droplet formation can be assessed. Showing the domains and PTM sites together may distinguish PTMs that are likely to regulate domain functions and those that may be regulate the process of liquid–liquid phase separation.

The FuzDrop server provides cross-links to databases of liquid–liquid phase separation (Figure 3), PhaSepDB (43), LLPSDB (44) and PhaSePro (33), where the user may refer to experimental conditions and biomolecular partners required for droplet formation. Experimental data on protein disorder can be obtained through cross-links to DisProt (45), FuzDB (40) and PED (46). DisProt provides ex-
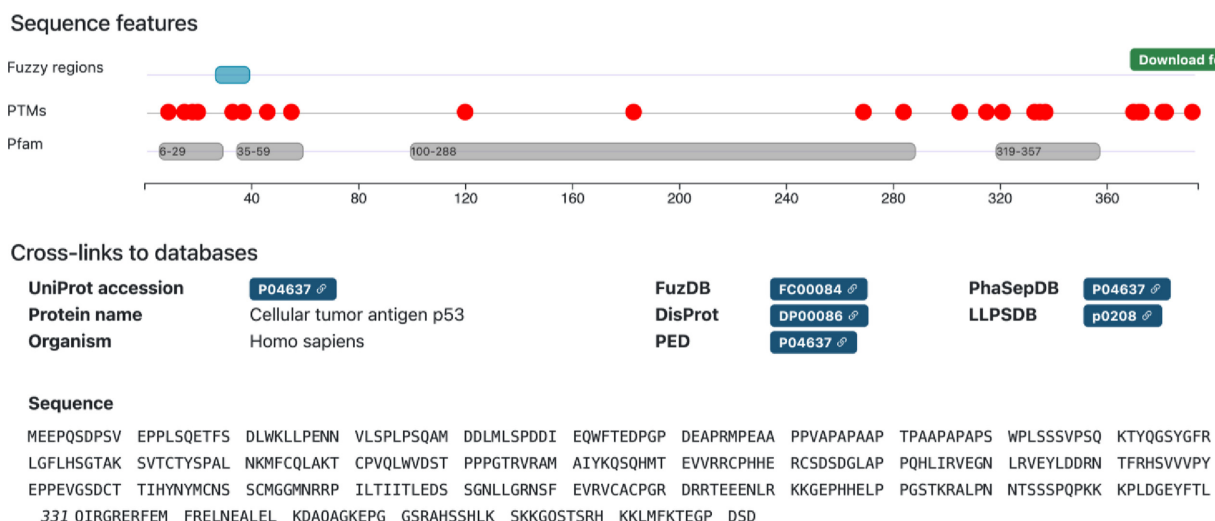
Figure 3. Protein information and cross-links to other databases. Protein features related to cellular context-dependence show fuzzy regions, derived from FuzDB (40), and post-translational modifications (PTMs), derived from UniProt (39). PTMs are shown by red dots, and positioning the cursor above them will display the modified residue, the PTM type and the modifying enzyme. In addition, the Pfam domains (42) are displayed (grey bars). Cross-links to experimental databases on liquid–liquid phase separation, PhaSepDB (43), LLPSDB (44) and PhaSePro (33) are provided. Structural information on protein disorder through cross-links to DisProt (45), FuzDB (40) and PED (46). DisProt reviews known experimental evidence on disorder in the free state of protein, whereas FuzDB assembles experimental evidence for disorder in the bound state. PED provides a detailed analysis of conformational ensembles in either free or bound states. for a limited number of systems.

perimental evidence on disorder in the free state of protein, whereas FuzDB assembles experimental evidence for disorder in the bound state. Both databases provide links to the Protein Data Bank for structural information. PED gives a detailed analysis of conformational ensembles in either free or bound states, for a limited number of systems.

Graphical representation of droplet-promoting regions and aggregation hot-spots

The predicted droplet-promoting regions and aggregation hot-spots are visualized by Mol* (47) on the structures predicted by AlphaFold (48) (AF, Figure 4A, B). The user can select the required feature; droplet-promoting regions are shown in blue and aggregation hot-spots in orange similarly to the box representation (Figure 4A, B). Orientation and size of the structures can be changed, and selected residues can be highlighted. This option depends on the availability of the predicted structure in the AlphaFold database (49).

In case no predicted structure is available, for example if the sequence deviates from the canonical one, or the UniProt code is not provided, the user is prompted to initiate the structure predictions (Figure 4C). As described, the AlphaFold prediction (48) is carried out in a separate web page, and the resulted coordinate file has to be uploaded onto the FuzDrop server (Figure 4C). The condensed state properties can be displayed similarly to the structures available through the AF database (49).

We note that the AlphaFold method, by providing individual structures, may not offer an accurate representation of the disordered nature of droplet regions (Figure 4). In principle structural ensembles would enable the visualization of transient contacts that may be important to stabilize the droplet state.

Download options

The FuzDrop prediction results, the residue-based p_{DP} and S_{bind} values can be downloaded in .tsv format via the 'Download' tab on the top right of the page. The coordinates of the droplet-promoting and aggregation-promoting regions can also be downloaded in .tsv format via the 'Download' tab on the top right of the page. The graph displaying the p_{DP} values and the graphical representation of the droplet-promoting regions and aggregation hot-spots can be saved as an image by the camera icon below the 'Download' tab.

The user can save the colored AlphaFold structures representing different condensed state properties with the 'Screenshot' tab above the image. The coordinates of the predicted structure can also be downloaded as indicated by a separate tab. This feature enables the user to generate different graphical representations of the results.

FuzDrop server information

A detailed information on the background of FuzDrop predictions and a tutorial to the server are available through the Help and Tutorial menu on the top right the FuzDrop main page. The Help page describes the condensed states of proteins, and presents the view that the droplet-state is a fundamental state of proteins along with the native and amyloid states (7). The Help page introduces the principles of predicting the droplet state (4) and the aggregation hot-spots, (23) which are detailed in the references, that are also shown in the main page of the server.

The Tutorial provides a brief description of the data, which are shown in the results page.

The FuzPred link points to the FuzPred server (http://fuzpred.bio.unipd.it) for analysis of detailed interaction

Figure 4. Visualization of condensed state features on structures predicted by AlphaFold. (**A**) Droplet-promoting regions (blue) and (**B**) aggregation hot-spots (orange) can be displayed on the predicted structure of a protein from the AlphaFold Database (49). (**C**) When the predicted structure is not already available, it can be predicted separately and uploaded to obtain the visualization.

characteristics (24,26,37). FuzPred identifies regions that undergo disorder-to-order transition upon binding, as well as regions predicted to remain disordered in the bound state. Based on the degree of context-dependence, the FuzPred server helps identify regions with stable and diverse binding modes.

FuzDrop server application areas

The FuzDrop server has four main application areas, describing different condensed state characterics (Table 1).

(1) ***Identification of proteins undergoing liquid–liquid phase separation.*** The probability of the droplet state (p_{LLPS}) informs on the tendency of spontaneous liquid–liquid phase separation. Proteins with $p_{LLPS} \geq 0.60$ can act as drivers of the droplet-forming process (Table 1).

(2) ***Identification of protein regions promoting droplet formation or partitioning in liquid droplets.*** Droplet-driver proteins may have low-specificity interaction motifs embedded in disordered regions, repetitive sequence elements serving as binding motifs in low-complexity regions, multivalent signaling proteins, or structured proteins sampling multiple binding modes. Droplet-promoting regions are composed of residues with $p_{DP} \geq 0.60$ (Table 1). Based on the analysis of the available experimental data, a minimum length of 10 residues (with $p_{DP} \geq 0.60$) are required.

(3) ***Identification of droplet-driver or droplet-client proteins.*** Droplet-client proteins, which cannot spontaneously undergo liquid–liquid phase separation, but have at least one droplet-promoting region can partition into droplets via interacting with a partner (Table 1). Many proteins that were observed as components of mem-

braneless organelles cannot form droplets in the test tube (15,16). Identification of droplet-promoting regions may facilitate the design of modified constructs with tailored droplet propensities.

(4) ***Identification of aggregation hot-spots within the droplet state.*** Aggregation hot-spots, which can initiate the irreversible maturation of droplets, can be identified combining the predictions on droplet-promoting regions with the interaction diversity of FuzPred (Table 1). These regions are automatically identified by the server, and a more in-depth analysis of interaction characteristics can be performed through the FuzPred approach, which is accessible through the FuzDrop server.

CONCLUSIONS

The importance of the condensed states of proteins, both the liquid-like droplet state and the solid-like amyloid state, has been increasingly recognized in determining the biological activity under cellular conditions (6,7). The FuzDrop server enables users to obtain readily the probability of proteins to sample the condensed states by predicting the probability of forming the droplet state of proteins and estimating the likelihood of the conversion towards the amyloid state. Thus, the FuzDrop server can be used to identify protein regions that can drive liquid–liquid phase separation as well as to predict aggregation hot-spots that can drive the conversion of droplets to solid-like aggregates, which are visualized on protein structures predicted by AlphaFold. In summary, the FuzDrop server contributes to elaborating of state-function relationship of proteins by characterizing their complex condensed state behaviors in the cellular environment.

DATA AVAILABILITY

The authors confirm that the data in the article are publicly available.

FUNDING

REFERENCES

1. Brangwynne,C.P., Eckmann,C.R., Courson,D.S., Rybarska,A., Hoege,C., Gharakhani,J., Julicher,F. and Hyman,A.A. (2009) Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, **324**, 1729–1732.
2. 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.
3. 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.
4. 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. USA*, **117**, 33254–33262.
5. Hyman,A.A., Weber,C.A. and Julicher,F. (2014) Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.*, **30**, 39–58.
6. Lyon,A.S., Peeples,W.B. and Rosen,M.K. (2020) A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.*, **21**, 235–215.
7. Fuxreiter,M. and Vendruscolo,M. (2021) Generic nature of the condensed states of proteins. *Nat. Cell Biol.*, **23**, 587–594.
8. Du,M. and Chen,Z.J. (2018) DNA-induced liquid phase condensation of cGAS activates innate immune signaling. *Science*, **361**, 704–709.
9. Banjade,S. and Rosen,M.K. (2014) Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife*, **3**, e04123.
10. Schaefer,K.N. and Peifer,M. (2019) Wnt/Beta-Catenin signaling regulation and a role for biomolecular condensates. *Dev. Cell*, **48**, 429–444.
11. Woodruff,J.B., Ferreira Gomes,B., Widlund,P.O., Mahamid,J., Honigmann,A. and Hyman,A.A. (2017) The centrosome is a selective condensate that nucleates microtubules by concentrating tubulin. *Cell*, **169**, 1066–1077.
12. Case,L.B., Zhang,X., Ditlev,J.A. and Rosen,M.K. (2019) Stoichiometry controls activity of phase-separated clusters of actin signaling proteins. *Science*, **363**, 1093–1097.
13. Kilic,S., Lezaja,A., Gatti,M., Bianco,E., Michelena,J., Imhof,R. and Altmeyer,M. (2019) Phase separation of 53BP1 determines liquid-like behavior of DNA repair compartments. *EMBO J.*, **38**, e101379.
14. Larson,A.G., Elnatan,D., Keenen,M.M., Trnka,M.J., Johnston,J.B., Burlingame,A.L., Agard,D.A., Redding,S. and Narlikar,G.J. (2017) Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature*, **547**, 236–240.
15. Jain,S., Wheeler,J.R., Walters,R.W., Agrawal,A., Barsic,A. and Parker,R. (2016) ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell*, **164**, 487–498.
16. Youn,J.Y., Dyakov,B.J.A., Zhang,J., Knight,J.D.R., Vernon,R.M., Forman-Kay,J.D. and Gingras,A.C. (2019) Properties of stress granule and P-Body proteomes. *Mol. Cell*, **76**, 286–294.
17. 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.
18. Hardenberg,M.C., Sinnige,T., Casford,S., Dada,S., Poudel,C., Robinson,E.A., Fuxreiter,M., Kaminksi,C., Kaminski-Schierle,G.S., Nollen,E.A.A. *et al.* (2021) Observation of an alpha-synuclein liquid droplet state and its maturation into lewy body-like assemblies. *J. Mol. Cell Biol.*, **13**, 282–294.
19. Kim,H.J., Kim,N.C., Wang,Y.D., Scarborough,E.A., Moore,J., Diaz,Z., MacLea,K.S., Freibaum,B., Li,S., Molliex,A. *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*, **495**, 467–473.
20. Mathieu,C., Pappu,R.V. and Taylor,J.P. (2020) Beyond aggregation: pathological phase transitions in neurodegenerative disease. *Science*, **370**, 56–60.
21. Boija,A., Klein,I.A. and Young,R.A. (2021) Biomolecular condensates and cancer. *Cancer Cell*, **39**, 174–192.
22. Risso-Ballester,J., Galloux,M., Cao,J., Le Goffic,R., Hontonnou,F., Jobart-Malfait,A., Desquesnes,A., Sake,S.M., Haid,S., Du,M. *et al.* (2021) A condensate-hardening drug blocks RSV replication in vivo. *Nature*, **595**, 596–599.
23. Vendruscolo,M. and Fuxreiter,M. (2021) Sequence determinants of the aggregation of proteins within condensates generated by liquid–liquid phase separation. *J. Mol. Biol.*, **434**, 167201–167212.
24. Miskei,M., Horvath,A., Vendruscolo,M. and Fuxreiter,M. (2020) Sequence-Based prediction of fuzzy protein interactions. *J. Mol. Biol.*, **432**, 2289–2303.
25. Tompa,P. and Fuxreiter,M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci*, **33**, 2–8.
26. Fuxreiter,M. (2020) Classifying the binding modes of disordered proteins. *Int. J. Mol. Sci.*, **21**, 8615.
27. Lin,Y., Protter,D.S., Rosen,M.K. and Parker,R. (2015) Formation and maturation of phase-separated liquid droplets by RNA-Binding proteins. *Mol. Cell*, **60**, 208–219.
28. Thomson,J.A., Schurtenberger,P., Thurston,G.M. and Benedek,G.B. (1987) Binary liquid phase separation and critical phenomena in a protein/water solution. *Proc. Natl. Acad. Sci. USA*, **84**, 7079–7083.
29. Jang,S., Xuan,Z., Lagoy,R.C., Jawerth,L.M., Gonzalez,I.J., Singh,M., Prashad,S., Kim,H.S., Patel,A., Albrecht,D.R. *et al.* (2021)

Phosphofructokinase relocalizes into subcellular compartments with liquid-like properties in vivo. *Biophys. J.*, **120**, 1170–1186.

30. Freiberger,M.I., Wolynes,P.G., Ferreiro,D.U. and Fuxreiter,M. (2021) Frustration in fuzzy protein complexes leads to interaction versatility. *J. Phys. Chem. B*, **125**, 2513–2520.

31. Gianni,S., Freiberger,M.I., Jemth,P., Ferreiro,D.U., Wolynes,P.G. and Fuxreiter,M. (2021) Fuzziness and frustration in the energy landscape of protein folding, function, and assembly. *Acc. Chem. Res.*, **54**, 1251–1259.

32. Walsh,I., Martin,A.J., Di Domenico,T. and Tosatto,S.C. (2012) ESpritz: accurate and fast prediction of protein disorder. *Bioinformatics*, **28**, 503–509.

33. 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.

34. Knowles,T.P., Vendruscolo,M. and Dobson,C.M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.*, **15**, 384–396.

35. 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.

36. Fuxreiter,M. (2020) Fuzzy protein theory for disordered proteins. *Biochem. Soc. Trans.*, **48**, 2557–2564.

37. 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.

38. Rhine,K., Makurath,M.A., Liu,J., Skanchy,S., Lopez,C., Catalan,K.F., Ma,Y., Fare,C.M., Shorter,J., Ha,T. *et al.* (2020) ALS/FTLD-Linked mutations in FUS glycine residues cause accelerated gelation and reduced interactions with wild-type FUS. *Mol. Cell*, **80**, 666–681.

39. UniProt,C. (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.

40. Hatos,A., Monzon,A.M., Tosatto,S.C.E., Piovesan,D. and Fuxreiter,M. (2021) FuzDB: a new phase in understanding fuzzy interactions. *Nucleic Acids Res.*, **50**, D509–D517.

41. 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.

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

47. 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.

48. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with alphafold. *Nature*, **596**, 583–589.

49. Varadi,M., Anyango,S., Deshpande,M., Nair,S., Natassia,C., Yordanova,G., Yuan,D., Stroe,O., Wood,G., Laydon,A. *et al.* (2022) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.*, **50**, D439–D444.