8-25-2023

Intrinsic Disorder in PRAME and Its Role in Uveal Melanoma

Michael Antonietti

David J. Taylor Gonzalez

Mak Djulbegovic

Guy W. Dayhoff

Vladimir N. Uversky

See next page for additional authors

Authors

Michael Antonietti, David J. Taylor Gonzalez, Mak Djulbegovic, Guy W. Dayhoff, Vladimir N. Uversky, Carol L. Shields, and Carol L. Karp

Intrinsic disorder in PRAME and its role in uveal melanoma

Michael Antonietti[1], David J. Taylor Gonzalez[1], Mak Djulbegovic[1], Guy W. Dayhoff II[2], Vladimir N. Uversky[3], Carol L. Shields[4] and Carol L. Karp[1*]

Abstract

Introduction The PReferentially expressed Antigen in MElanoma (PRAME) protein has been shown to be an independent biomarker for increased risk of metastasis in Class 1 uveal melanomas (UM). Intrinsically disordered proteins and regions of proteins (IDPs/IDPRs) are proteins that do not have a well-defined three-dimensional structure and have been linked to neoplastic development. Our study aimed to evaluate the presence of intrinsic disorder in PRAME and the role these structureless regions have in PRAME(+) Class 1 UM.

Methods A bioinformatics study to characterize PRAME's propensity for the intrinsic disorder. We first used the AlphaFold tool to qualitatively assess the protein structure of PRAME. Then we used the Compositional Profiler and a set of per-residue intrinsic disorder predictors to quantify the intrinsic disorder. The Database of Disordered Protein Prediction (D^2P^2) platform, IUPred, FuzDrop, flDPnn, AUCpred, SPOT-Disorder2, and metapredict V2 allowed us to evaluate the potential functional disorder of PRAME. Additionally, we used the Search Tool for the Retrieval of Interacting Genes (STRING) to analyze PRAME's potential interactions with other proteins.

Results Our structural analysis showed that PRAME contains intrinsically disordered protein regions (IDPRs), which are structureless and flexible. We found that PRAME is significantly enriched with serine (*p*-value < 0.05), a disorder-promoting amino acid. PRAME was found to have an average disorder score of 16.49% (i.e., moderately disordered) across six per-residue intrinsic disorder predictors. Our IUPred analysis revealed the presence of disorder-to-order transition (DOT) regions in PRAME near the C-terminus of the protein (residues 475–509). The D^2P^2 platform predicted a region from approximately 140 and 175 to be highly concentrated with post-translational modifications (PTMs). FuzDrop predicted the PTM hot spot of PRAME to be a droplet-promoting region and an aggregation hotspot. Finally, our analysis using the STRING tool revealed that PRAME has significantly more interactions with other proteins than expected for randomly selected proteins of the same size, with the ability to interact with 84 different partners (STRING analysis result: *p*-value < 1.0×10^{-16}; model confidence: 0.400).

Conclusion Our study revealed that PRAME has IDPRs that are possibly linked to its functionality in the context of Class 1 UM. The regions of functionality (i.e., DOT regions, PTM sites, droplet-promoting regions, and aggregation hotspots) are localized to regions of high levels of disorder. PRAME has a complex protein–protein interaction (PPI) network that may be secondary to the structureless features of the polypeptide. Our findings contribute to our understanding of UM and suggest that IDPRs and DOT regions in PRAME may be targeted in developing new therapies for this aggressive cancer.

*Correspondence:
Carol L. Karp
ckarp@med.miami.edu
Full list of author information is available at the end of the article

Introduction

Uveal melanoma (UM) is the most common primary intraocular tumor [1]. The neoplasm most commonly arises from the choroid but also develops in the ciliary body (6%) and iris (4%) [1]. UM affects both men and women equally, and it most commonly affects middle-aged White individuals with a median age of 58 years [2]. Though there have been advances in the detection, prognostication, and treatment of UM, the risk for metastasis is high (50%) and is often associated with poor outcomes [1, 3].

Gene expression profiling (GEP) has been developed to assess the risk of metastasis for UM. GEP allows clinicians to group UM into either Class 1 (low-grade metastatic risk) or Class 2 (high-grade metastatic risk) based on their molecular profiles [4]. Class 1 UM is associated with upregulation of PReferentially expressed Antigen in MElanoma (PRAME), and class 2 UM mutations are associated with mutations or biallelic inactivation of BRCA1 (breast cancer type 1 susceptibility protein)-associated protein 1 (BAP1) [5–8]. The 5-year actuarial rate for metastasis of class 1 PRAME(-), class 1 PRAME(+), class 2 tumors is 0%, 38%, and 71% respectively [9]. These numbers suggest that PRAME may serve as an independent biomarker for the risk of metastasis in Class 1 UM.

Our group has studied ocular malignancies and characterized proteins involved in ocular surface squamous neoplasia, conjunctival melanoma, and uveal melanoma [8]. We are interested in the intrinsic disorder phenomenon (i.e., proteins or regions of proteins that have no definitive 3D structure) and how it may play a role in the pathogenesis of ocular neoplasms [8, 10]. We previously demonstrated that BAP1 has elements of intrinsic disorder [8]. Dysregulation of these structureless regions secondary to BAP1 mutations may promote metastatic behavior in Class 2 UM. This study recapitulates our previous analysis and applies it to Class 1 UM and PRAME. While there are many interactions to consider when looking at the role of PRAME in the tumor pathogenesis of UM, our group wanted to understand PRAME's intrinsic disorder propensity and its potential to provide a framework for developing biologics or immunotherapies to target PRAME.

Intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs) are proteins and the regions of proteins that do not adopt a well-defined three-dimensional structure and instead are disordered but maintain function [11]. These functional proteins or protein regions are found in all known proteomes and do not adopt a tertiary structure [12]. IDPs and IDPRs have been linked extensively to the pathogenesis of disease and malignancies [13–17]. In this study, we aim to evaluate the intrinsic disorder of PRAME and its role in pathogenesis and driving neoplastic development in Class 1 UM. Furthermore, the IDPRs of PRAME may serve as targets for treating PRAME(+) UM, thereby decreasing the risk of malignancy of the primary tumor in those most vulnerable.

Methods

Our study focused on the intrinsic disorder phenomenon and its association with PRAME expression and its association with uveal melanoma. To accomplish this, we conducted a disorder-based bioinformatics analysis that utilized publicly available databases and computational tools to probe PRAME for intrinsically disordered protein regions.

Protein sequences

The Universal Protein Resource (UniProt) is a protein database with information related to function, taxonomy, subcellular location, disease associations, post-translation modifications, expression profiles, interaction networks, structure, protein families, and sequences (available at: https://uniprot.org; date last accessed: Dec 26, 2022) [18]. The search term used was "protein: PRAME." The human variant was selected for PRAME (UniProtID: P78395). The default (canonical) amino acid sequence was selected for further analysis.

Structural assessment

The UniProt entries for PRAME also contain structural information. There are computational and experimental-based structures (i.e., nuclear magnetic resonance, cryo-electron microscopy, and x-ray crystallography) housed on the UniProt server. We selected the "AlphaFoldDB" entry. AlphaFold is the most up-to-date and accurate computational method to predict protein structure [19–21].

Quantification of intrinsic disorder
Compositional profiler

We sought to quantify the elements of intrinsic disorder in PRAME. We utilized the Compositional Profiler to measure the relative proportion of specific amino acids in

each protein (available at: http://www.cprofiler.org/; accessed on December 26, 2022) [22]. Each amino acid was compared to a background set of proteins (i.e., Protein Data Bank Select (PBD) 25 [22]) that are a standard set of highly ordered proteins (i.e., proteins that contain many elements of structure in the form of alpha helices or beta-pleated sheets) [22]. We evaluated the normalized enrichment or depletion of each amino acid as $\frac{C_x - C_{order}}{C_{order}}$, where C_x is the content of a given residue in a query protein (i.e., PRAME) and C_{order} is the content of the same residue in the PDB Select 25.

Per-residue disorder prediction

The intrinsic disorder of PRAME was quantified on a per-residue basis using a set of ten well-established disorder predictors, namely: PONDR® VLS2 [23, 24], PONDR® VL3 [25], PONDR® VLXT [26], PONDR® FIT [27], IUPred-Long, IUPred-Short [28], flDPnn [29], AUCpreD [30], SPOT-Disorder2 [31] and metapredict V2 [32]. The Rapid Intrinsic Disorder Analysis Online (RIDAO) platform (https://ridao.app) was used to yield bulk predictions for the PONDR family of predictors (i.e. VLS2, VL3, VLXT, and FIT) as well as IUPred-Long and IUPred-Short. The remaining predictions were obtained using local installations of the respective predictors—flDPnn, AUCpreD, SPOT-Disorder2, and metapredict V2. Each predictor used in this work takes a primary sequence as input and yields a measure of intrinsic disorder propensity, referred to as the disorder score, for each amino acid contained therein. Disorder scores range from 0 to 1, where 0 is the most ordered measurement and 1 is the most disordered. Taken as a whole, the complete set of disorder scores for a given sequence constitute a disorder profile.

Using the disorder profile of PRAME we computed the percent of predicted disordered residues (PPDR) for a given predictor, p, as,

$$PPDR_p = 100 * \frac{N_{disordered}}{N_{total}} \tag{1}$$

where $N_{disordered}$ is the number of residues with a disorder score greater than or equal to 0.5 (or 0.3 in the case of flDPnn) and N_{total} is the total number of residues in the sequence. A protein is classified as highly ordered when the PPDR is less than 10%, highly disordered when the PPDR is greater than 30%, and moderately disordered in all other cases [33]. Furthermore, we computed the average disorder score (ADS) over a range of residues for a given predictor, p, as,

$$ADS_p = \frac{\sum_{r=a}^{b} x_r}{b - a} \tag{2}$$

where a is the first residue to consider, b is the last residue to consider, and x_r is the disorder score of residue r, given by predictor p. A protein or protein region is characterized as ordered when the ADS is less than 0.15, disordered when the ADS is equal to or greater than 0.5, and moderately disordered, or flexible, in all other cases. Importantly, we do not compute the ADS for the flDPnn predictor for which a value greater than or equal to 0.3 is used to identify disordered residues.

Assessment of disorder-related functionality
Analyzing functionality using D²P² platform

Disorder-related functionality of PRAME was evaluated with the Database of Disordered Protein Prediction (D²P²) platform (http://d2p2.pro/; accessed on November 17, 2022) [34]. The D²P² platform integrates IUPred [35], PONDR® VLXT [26], PONDR® VSL2B [23, 24], PrDOS [36], PV2 [34], and ESpritz [37], allowing for a consolidated visualization of disorder prediction tools. The D²P² server provides the Structural Classification of Proteins (SCOP) [38, 39], SUPERFAMILY predictor [40], and InterPro prediction for the sequences of interest (https://www.ebi.ac.uk/interpro/search/sequence/, accessed on December 12, 2022) [41]. D²P²also predicts disorder-based binding sites with the ANCHOR algorithm [42]. ANCHOR-identified disordered protein-binding regions (DPBRs) that include short binding regions, which are disordered in the unbound form and expected to fold into ordered structures when interacting with specific partners (i.e., molecular recognition features (MoRFs) and other disordered protein binding domains that are longer than 25 residues [43]. Finally, the D²P²platform also predicts possible posttranslational modifications (PTMs) using the outputs of the PhosphoSitePlus [44].

Identification of redox-sensitive disorder-to-order transition (DOT) regions

The IUPred3 platform was the final step in quantifying intrinsic disorder within PRAME (available at: https://iupred3.elte.hu/; accessed on November 17, 2022) [28, 45, 46]. This tool can assess the location of disordered redox-sensitive protein regions in PRAME by determining the disorder score in both their disordered and ordered state. If the values vary significantly between the redox states, we can define a particular region as a redox-sensitive region (i.e., a disorder-to-order transition (DOT) region).

Identification of droplet-promoting regions (DPRs), aggregation hot spots, and regions with cellular context-dependent interactions

Additional functional information on human PRAME was retrieved using the FuzDrop platform (https://fuzdrop.bio.unipd.it/predictor; accessed on December 27,

Antonietti *et al. Cell Communication and Signaling*   (2023) 21:222

Page 4 of 13

2022) [47]. This tool provides information on the capability of a protein to undergo liquid–liquid phase separation, identifies the presence and localization of droplet-promoting regions (DPRs), shows the location of aggregation hot-spots (i.e., residues/regions that may promote the conversion of a protein to a liquid-like condensed state into a solid-like amyloid state) [48, 49]. FuzDrop also provides the probability of a protein having residues/regions with cellular context dependence (i.e., those characterized by the ability of residues to switch between different binding modes) [48].

Protein–protein interaction network

To further understand the protein–protein interactions of PRAME, the amino acid sequences were used as inputs for The Search Tool for the Retrieval of Interacting Genes (STRING) (URL: https://string-db.org/; accessed on November 17, 2022). STRING is a quality-controlled database that utilizes experimentally and computationally derived data to detail functional interactions between proteins [50]. The minimum required interaction score settings were set to the highest confidence (0.900), and the maximum number of interactions was assigned the highest possible value of 500.

Results

Protein sequence

Our disorder-based bioinformatics protein analysis used the canonical amino acid sequence of PRAME in FASTA format. PRAME contains 509 amino acid residues and includes nine leucine-rich repeats (LLR) at residues 116–145, 207–231, 232–258, 259–294, 295–320, 321–352, 353–371, 377–404, and 405–429. There is also a region (residues 416–509) that mediates the interaction with a retinoic acid receptor alpha (RARA).

Structural assessment

The 3D structure of PRAME modeled by AlphaFold (Fig. 1) demonstrates that PRAME has both elements of order (structure; α-helices and β-pleated sheets) and disorder (structureless, spaghetti-like regions) (Fig. 1). Structural portions of the protein are flanked by regions that lack definitive structure (residues 1–23 and 138–172). The C-terminal region of PRAME (residues 487–509) is also predicted to have low model confidence (i.e., high intrinsic disorder). Figure 1 shows that the AlphaFold model confidence is high for a large proportion of PRAME's structure, as indicated by the extensive blue color throughout the model. The terminals and disorder regions of PRAME appear to be less ordered, as indicated by low model confidence and yellow/orange coloring. This reflects the regions of PRAME with a less established structure. These regions of PRAME with less model confidence are the intrinsically disordered protein regions (IDPRs) of interest. From a qualitative point of view, it is evident that there are segments in PRAME that contain little to no structure yet likely retain biological functions as expected for most IDPRs.

Quantification of intrinsic disorder
Compositional profiler

After visualization of the IDPRs in PRAME within the AlphaFold-modeled structure, we aimed to quantify these regions. We first constructed a normalized amino acid composition profile for PRAME (Fig. 2). In the composition profile, each amino acid was arranged from the most order-promoting (i.e., cysteine (C), tryptophan (W), isoleucine (I), tyrosine (Y), phenylalanine (F), leucine (L), histidine (H), valine (V), asparagine (N), and methionine (M)) to the most disorder-promoting residues (i.e., arginine (R), threonine (T), aspartate (D), glycine (G), alanine
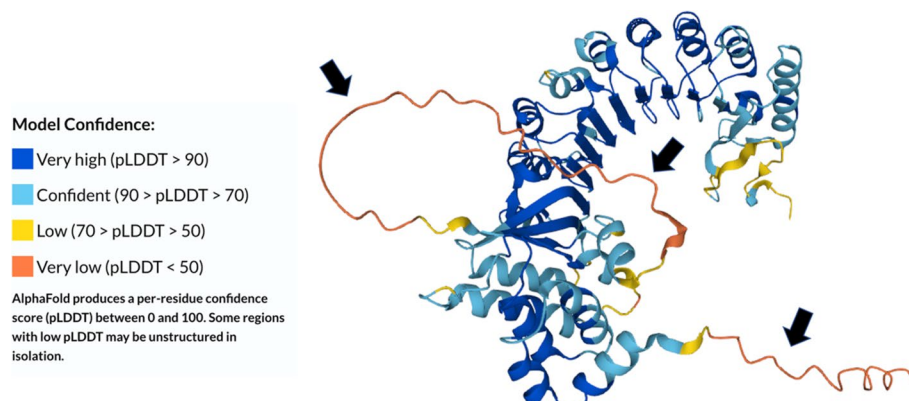
Fig. 1 Structures generated for PRAME by AlphaFold2. As measured by the predicted local distance difference test (pLDDT), regions with higher model confidence typically represent alpha-helices and beta-pleated sheets. Regions with lower levels of model confidence might represent intrinsically disordered protein regions. The black arrows indicate areas of potentially high levels of disorder that lack definitive structure
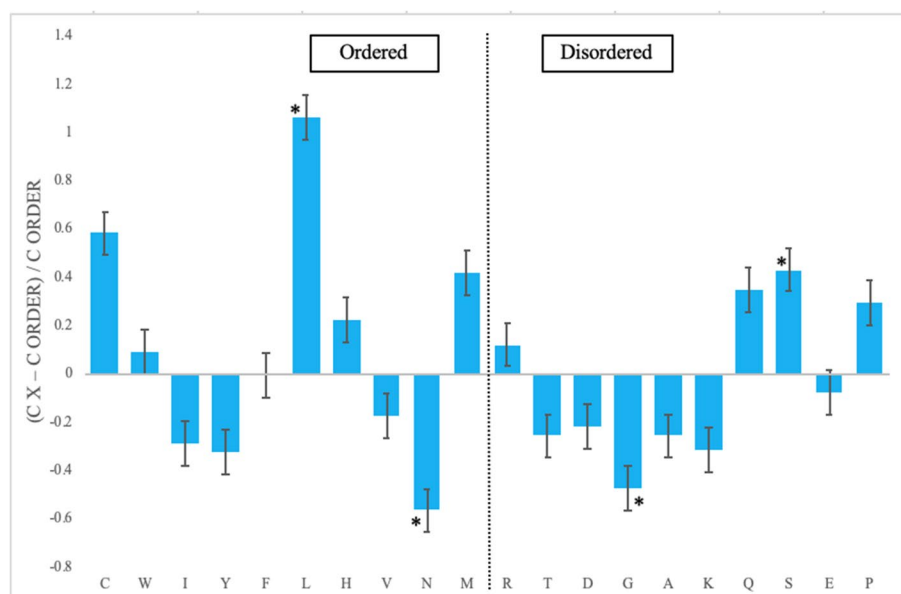
Fig. 2 Amino acid composition profile of PRAME. The fractional difference is calculated as $\frac{C_{x-}C_{order}}{C_{order}}$, where C_x is the content of a given amino acid in the query set (PRAME), and C_{order} is the content of a given amino acid in the background set (Protein Databank Select 25). Amino acids marked with (*) are statistically significant for enrichment (L and S) or depletion (G and N) (*p*-value < 0.05). The amino acid residues are ranked from most order promoting residues ((i.e., cysteine (C), tryptophan (W), isoleucine (I), tyrosine (Y), phenylalanine (F), leucine (L), histidine (H), valine (V), asparagine (N), and methionine (M)) to most disorder promoting residues ((i.e., arginine (R), threonine (T), aspartate (D), glycine (G), alanine (A), lysine (K), glutamine (Q), serine (S), glutamate (E), and proline (P)). Positive values indicate enrichment, and negative values indicate depletion of amino acids

(A), lysine (K), glutamine (Q), serine (S), glutamate (E), and proline (P)). If the amino acid is enriched in the protein, then it receives a positive value. If the amino acid is depleted in the protein, then it receives a negative value. PRAME is significantly enriched in one order-promoting amino acid (L) and significantly depleted in one disorder-promoting amino acid (N). PRAME is significantly enriched in one disorder-promoting residue (S) and significantly depleted in (G).

Per-residue disorder prediction

After determining the amino acid composition of PRAME, we aimed to quantify the intrinsic disorder on a per-residue basis. We used PONDR® VLXT [26], PONDR® VL3, PONDR® VSL2, and PONDR® FIT, as well as IUPred-Long and IUPred-Short to predict the disorder score of each amino acid in PRAME (Fig. 3A). The six disorder predictors that make up the RIDAO suite [51] allow for the classification of proteins and protein regions using the established criteria for the classification of proteins based on their percent of the predicted disordered residues (PPDR scores), where proteins are considered highly ordered (PPDR < 10%), moderately disordered (10% ≤ PPDR < 30%), and highly disordered (PPDR ≥ 30%) [22]. PRAME was predicted to be 28.36% disordered by PONDR® VLXT (i.e., moderately

disordered), 28.02% disordered by PONDR® VL3 (i.e., moderately disordered), 30.89% disordered by PONDR® VSL2 (i.e., highly disordered), 20.61 disordered by PONDR-FIT (i.e., moderately disordered), 10.31% disordered by IUPred-Long (i.e., moderately disordered), and 9.39% disordered by IUPred-Short (i.e., highly ordered). To provide further validation of the presence of intrinsic disorder in the PRAME protein, we used a set of predictors that showed good performance at a biannual blind test [52]. These predictors are fIDPnn [53], AUCpreD [54], and SPOT-Disorder2 [31]. We also utilized a deep-learning-based predictor of consensus disorder scores, metapredict V2 [32]. The corresponding data are shown in Fig. 3B illustrating that at least three regions of PRAME (residues 1–15, 140–172, and 505–509) are predicted as disordered by some of these tools as well. Notably, PONDR® VSL2, IUPred-Long and IUPred-Short also participated in the CAID competition, where they performed reasonably well [52]. Furthermore, we utilized flDPnn to predict both the intrinsic disorder status of PRAME and also predicts its disorder-based functions [53]. We utilized this tool to look for the presence of disorder-based protein-, DNA- and RNA-biding regions. This analysis revealed that the intrinsically disordered region comprising residues 154–164 can be engaged in interactions with proteins, DNA and RNA.

Antonietti *et al. Cell Communication and Signaling*   (2023) 21:222

Page 6 of 13
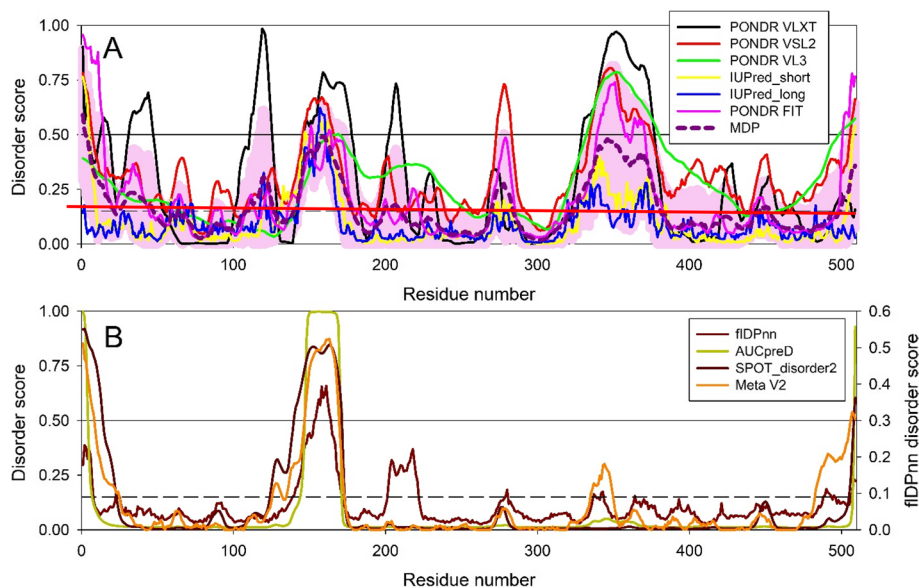
Fig. 3 Evaluation of the intrinsic disorder predisposition of PRAME. A. Per-residue intrinsic disorder profiles generated by RIDAO that assembles the outputs of PONDR® VLXT, PONDR® VL3, PONDR® VSL2, PONDR® FIT, IUPred_long and IUPred_short. The mean disorder profile (MDP) of the protein is calculated by averaging the disorder profiles of individual predictors. The light pink shade represents the MDP error distribution. The thin black line at the disorder score of 0.5 is the threshold between order and disorder, where residues/regions above 0.5 are disordered, and residues/regions below 0.5 are ordered. The solid red line at the disorder score of 0.15 is the threshold between order and flexibility, where residues/regions above 0.15 are flexible, and residues/regions below 0.15 are highly ordered. B. Per-residue intrinsic disorder profiles generated by predictors that performed better at a biannual blind test, Critical Assessment of protein Intrinsic Disorder prediction (CAID) [52]. These predictors are fIDPnn [53], AUCpreD [54], and SPOT-Disorder2 [31]. We also included the outputs of the Metapredict V2, which is a deep-learning-based predictor of consensus disorder scores [32]. Note that most predictors used in this study utilize the threshold of 0.5 to discriminate between ordered and disordered residues. The only exception is given by fIDPnn that utilizes the threshold of 0.3 for the same purpose (see right-hand Y axis in the plot B). Figure shows that PRAME is largely ordered protein but possesses elements of disorder throughout (i.e., regions comprising residues 1–15, 136–177, 330–376, and 510–509 are predicted as disordered by several predictors)

Assessment of disorder-related functionality
### *Analyzing functionality using D²P² platform*
To assess the functionality of the intrinsic disorder in PRAME, we used the D^2P^2 platform (Fig. 4). The D^2P^2 platform allows quantification of the functional intrinsic disorder of each protein and assess its relationship with post-translational modifications (PTMs) and disorder-based binding potential. PRAME's D^2P^2 output (Fig. 4) demonstrates relative agreement between the nine disorder-based predictors, where scattered regions of disorder predicted throughout the protein are present. The region of PRAME that demonstrates the most agreement among the nine per-residue disorder predictors is approximately from amino acid region 130 to 160. PRAME contains several conserved functional domains. There were two RNI like domains (residue 118–133, 210–439) identified in the SUPERFAMILY database, and the InterPro server identified domains; PB007380 (residues 2–165), PB001702 (residues 120–248), PB004892 (residues 11–505), PB014386 (residues 1–165), PB000609 (residues 326–463), PB004908 (residues 191–392), and PB016350 (residues 230–286). The InterPro server places the human PRAME into the PRAME family (InterPro ID:

IPR026271). LLR-containing domain (residues 188–469) of this protein is a member of the leucine-rich repeat domain superfamily (InterPro ID: IPR032675). InterPro also identifies this region as a ribonuclease inhibitor by CATH-Gene3D, where part of this region (residues 188–443) is included in Functional Families (InterPro ID: FF000354). Finally, according to the InterPro annotations, human PRAME is expected to be involved in the negative regulation of the apoptotic process (GO:0043066), positive regulation of cell population proliferation (GO:0008284), negative regulation of DNA-templated transcription (GO:0045892), and negative regulation of cell differentiation (GO:0045596). There are no molecular recognition features (MoRFs) in PRAME as predicated by the D^2P^2 platform. PRAME has 7 predicted post-translational modifications (PTMs), 3 ubiquitination sites, 2 acetylation sites, and 2 phosphorylation sites.

Identification of redox-sensitive disorder-to-order transition (DOT) regions
We next sought to understand further PRAME's ability to under disorder-to-order transition (DOT). To assess whether this protein can undergo conformational
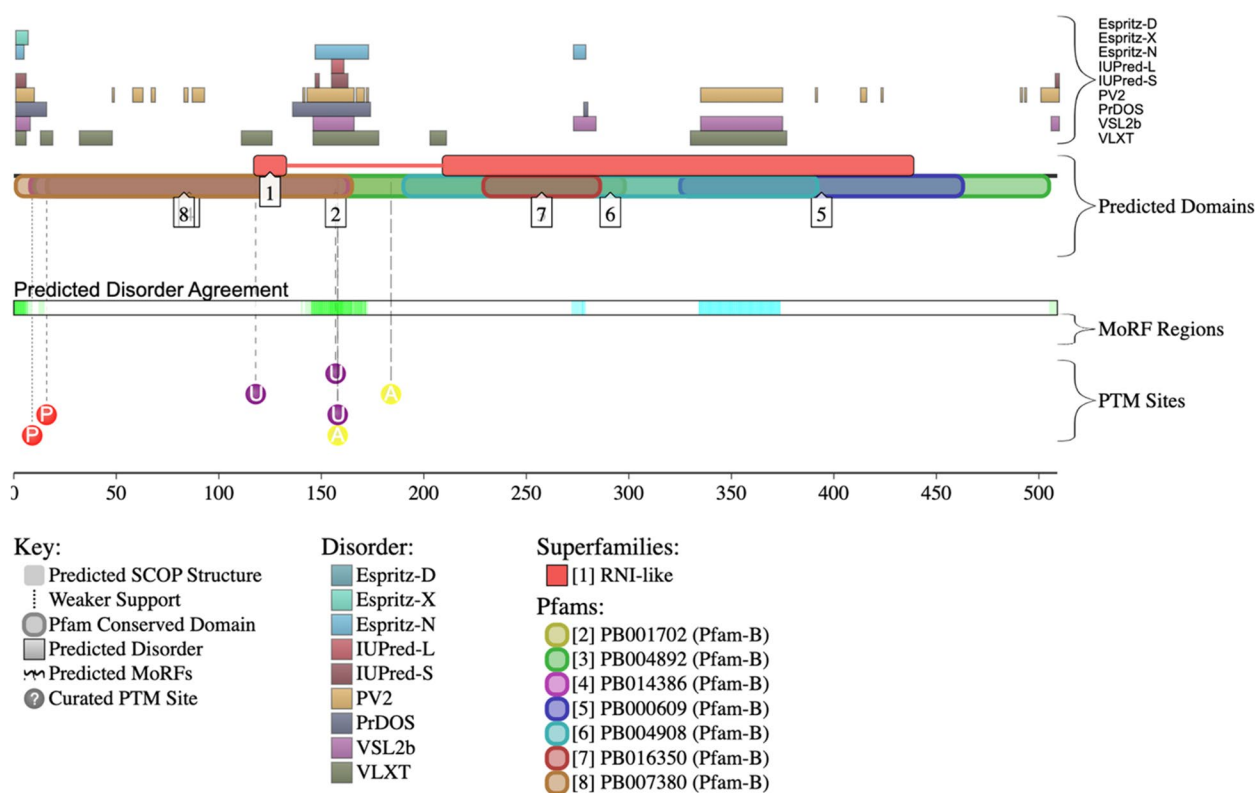
Fig. 4 The D^{2}P^2 platform output assesses the functional disorder profile for PRAME. On the figure's left are four identifiers for predictors used on the D^{2}P^2 platform. The top right of the figure shows outputs from various per-residue disorder predictors (i.e., Espritz-D, Espritz-X, Espritz-N, IUPred-L, IUPred-S, PV2, PrDOS, VSL2b, VLXT); the second bracket shows predicted protein domains, the third bracket shows where molecular recognition features are located (MoRFs; i.e., disorder regions that become ordered when binding), and the fourth bracket shows posttranslational modifications (PTM) sites. The protein domains fall within regions predicted to be ordered or disordered. The MoRF regions are localized to regions that demonstrate extensive disorder. The PTMs are also predominantly localized to areas of each protein that contain intrinsic disorder

reorganization following changes in the redox state of its environment (i.e., regions potentially capable of interaction with binding partners in a redox-sensitive manner), the IUPred's redox tool was utilized. We identified one segment of PRAME that is predicted to be moderately disordered in the reduced state and undergoes a transition to a highly ordered state in the oxidized form (i.e., DOT). This redox-related DOT of PRAME is near the C-terminus of the protein (i.e., residues 475−509) (Fig. 5). This region was marked by high levels of disorder, as evidenced by our per-residue disorder analysis (see Fig. 3).

Identification of droplet-promoting regions (DPRs), aggregation hot-spots, and regions with cellular context-dependent interactions

The propensity of a given protein to undergo liquid–liquid phase separation (LLPS) can be evaluated computationally with FuzDrop [47]. One of the advantages of this tool is its capability to group proteins based on their LLPS propensity (p_{LLPS}), where proteins capable of spontaneous LLPS are classified as droplet-driving

proteins, whereas proteins that need some additional interactions to form droplets are classified as clients. As per FuzDrop, LLPS drivers are proteins with the overall $p_{LLPS} \geq 0.60$, whereas proteins with lower overall p_{LLPS} but containing droplet-promoting regions (DPRs) defined as consecutive residues with $p_{LLPS} \geq 0.60$ will likely serve as droplet-clients [47]. Fig. 6 shows that although PRAME is characterized by a low LLPS propensity of 0.1229, it contains one DPR (residues 151–174). Therefore, although PRAME cannot phase separate by itself, it can serve as a droplet client. Importantly, the DPR of PRAME represents a part of its long internal IDPR (residues 146–177).

Another important advantage of FuzDrop is its capability to find aggregation hot-spots, defined as residues/regions that may promote the conversion of the liquid-like condensed state into a solid-like amyloid state, capable of undergoing both ordered and disordered binding modes that can drive this process [48, 49]. Fig. 6 shows that human PRAME contains two of these hot-spots (residues 155–164 and 168–174).
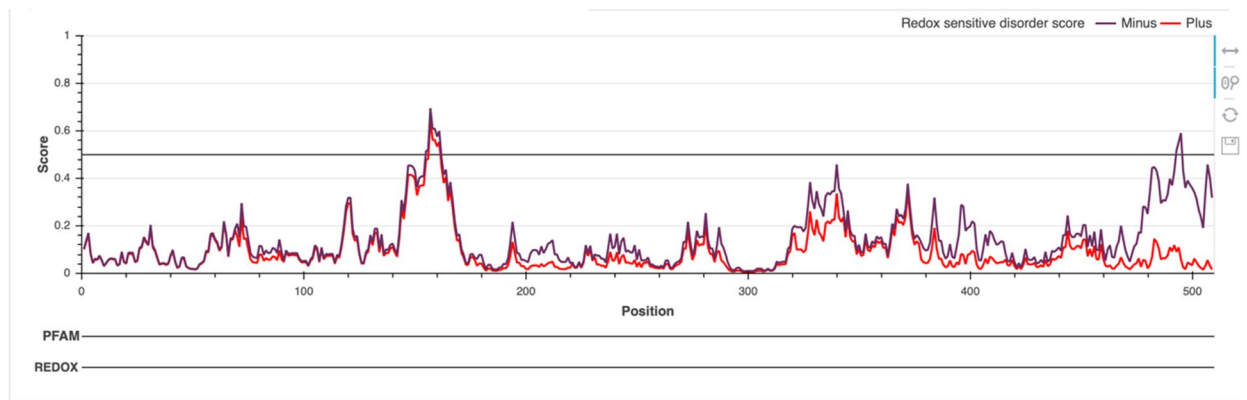
Fig. 5 Evaluation of redox-sensitive disorder predispositions of PRAME by IUPred2A_redox. The corresponding plot shows the disorder profile for the reduced (redox-minus) and oxidized (redox-plus) forms of a protein by violet and red lines, respectively. Shaded areas (not seen in PRAME) represent significant redox-sensitive regions. The C-terminus of PRAME (region is from amino acids 475 to 509) is a non-significant redox-sensitive, conditionally disordered/flexible region (i.e., a region that undergoes disorder-to-order transition (DOT) in a redox-sensitive manner). In its reduced state (redox-minus), the C-terminus is not highly disordered but flexible, characterized by an IUPred average disorder score (ADS$^{IUPred}_{redox}$) of 0.35 ± 0.11. In its reduced state (redox-plus), the C-terminus region is highly ordered (ADS.$^{IUPred}_{redox} = 0.059 \pm 0.035$)
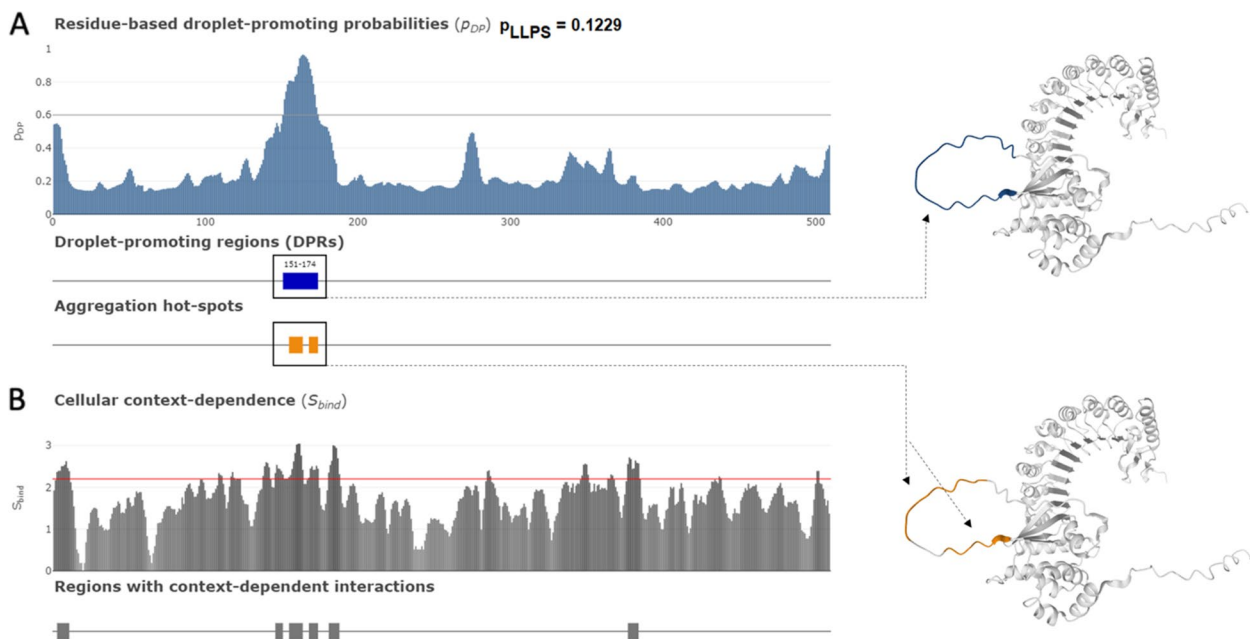
Fig. 6 Evaluation of liquid–liquid phase separation propensity of human PRAME by FuzDrop. **A** PRAME has an overall p_{LLPS} of 0.1229, indicating a low probability of promoting droplet formation or liquid–liquid phase separation. However, the protein does contain one region that promotes droplet formation, as depicted in the top AlphaFold structure, as well as two hotspots that promote aggregation, as shown in the bottom AlphaFold structure. **B** Furthermore, the behavior of PRAME is influenced by its cellular context, as indicated by the S_{BIND} value, which describe the ability of specific residues to switch between different binding modes. Regions that exhibit context-dependent interactions can alter the protein's binding behavior and binding modes in various cellular conditions. Residues or regions with S_{BIND} values of 2.25 or greater are capable of transitioning between disordered and ordered states

Finally, FuzDrop provides an outlook of the probability of a protein having residues/regions with cellular context-dependence (i.e., regions containing residues characterized by the ability to switch between disorder and order-based binding) [48]. Fig. 6 demonstrates that

PRAME contains six regions with context-dependent interactions (residues 3–11, 146–151, 155–164, 168–174, 181–188, and 377–384). Importantly, these regions with context-dependent interactions overlap with, are included in, or are located in close proximity to the

Antonietti *et al. Cell Communication and Signaling*   (2023) 21:222

Page 9 of 13

IDPRs found in PRAME (e.g., residues 1–23, 146–176, and 330–376). Furthermore, the most N-terminal disordered region with the context-dependent interactions also includes two phosphorylation sites (Ser[9] and Ser[16]), suggesting that the functionality of this region can be regulated by posttranslational modifications.

### Protein–protein interaction network

Our analysis has elucidated that intrinsic disorder is embedded into the structure of PRAME. The last step in our disorder-based bioinformatics analysis was focused on understanding how the intrinsic disorder may apply to the binding capabilities of PRAME. Our STRING analysis demonstrated that PRAME has multiple binding partners (Fig. 7). Figure 7A demonstrates PRAME's STRING network with medium model confidence (0.400), which shows 85 nodes in the network with 577 edges (interactions) and has significantly more than the 126 expected interactions in a randomly selected set of proteins the same size (p-value $< 1.0 \times 10^{-16}$). At a STRING model confidence of high (0.900, Fig. 7B), there are 4 predicted nodes in the 1st shell interaction network. There are 4

predicted edges (i.e., interactions) in the network, and the number of interactions was approximately the same as what was expected for a randomly selected set of proteins of the same size (p-value < 0.337).

Discussion

To the best of our knowledge, this work is the first to elucidate the presence of potentially functional intrinsically disordered protein regions (IDPRs) in PRAME. Our findings reveal the intrinsically disordered protein segments and regions where disorder is present within PRAME. This highlights the potential role these regions have in PRAME's function that could be a possible focus for the development of immunotherapies. We show here that PRAME has several IDPs, some of which are relatively long. Furthermore, we demonstrate that these structureless regions are associated with dynamic portions of the protein and that PRAME has a complex PPI network. Our findings must be considered when evaluating the underlying molecular processes that drive the biological behavior of PRAME(+) Class 1 uveal melanoma.
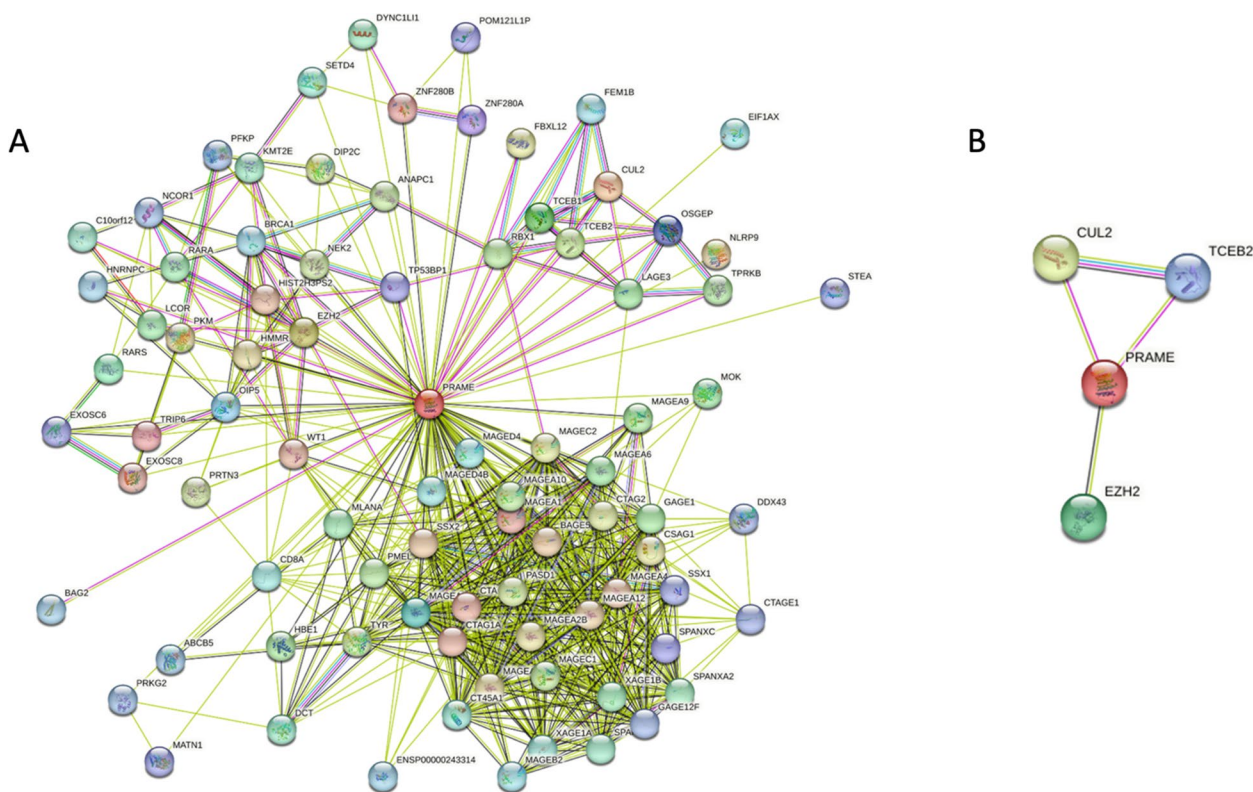
Fig. 7 First shell interaction network of PRAME as determined by The Search Tool for the Retrieval of Interacting Genes (STRING). **A** The STRING model confidence was set to medium (0.400). There are 85 nodes in the network with 577 edges (i.e., interactions). The number of interactions was significantly more than the expected 126 interactions in a randomly selected set of proteins of the same size (p-value $< 1.0 \times 10^{-16}$). **B** The STRING model confidence was set to high (0.900). There were 4 nodes in the network with 4 edges (interactions). The number of interactions was approximately the same as what was expected for a randomly selected set of proteins of the same size (p-value < 0.337)

The AlphaFold model of PRAME qualitatively demonstrated the presence of several IDPRs (i.e., spaghetti-like regions). Our quantitative analysis indicated that PRAME was most enriched with leucine (p-value < 0.05), which is partly due to the presence of nine leucine-rich repeats embedded in its sequence. Of the disorder-promoting residues, PRAME was significantly enriched with serine (p-value < 0.05). Note that at least two of those (Ser[9] and Ser[16]) were also identified as phosphorylation sites, indicating that they are likely functional hotspots for the protein. PRAME's average disordered score was 16.49%, classifying it as a moderately disordered and flexible protein.

Additionally, there are seven predicted post-translational modifications (PTMs) in PRAME. These amino acid sites are vital to the protein's function and are concentrated in areas of high levels of disorder (i.e., the N-terminus and amino acid region 140 to 160). The concentration of PTMs in intrinsically disordered regions are known to produce a range of regional structural effects [55, 56]. The PTM-rich regions are predicted to be highly dynamic as they have one droplet-promoting region and two aggregation hot spots. The functions and role of droplet-promoting regions and aggregation hot spots are not fully understood, but they may play a role in regulating the concentration of proteins in the cell [47]. In addition, there are six regions with cellular state-dependent (i.e., context-dependent) interactions, which likely have highly dynamic behavior associated with PRAME's function. This is consistent with the established notion that the propensity for the disorder of a protein can vary with cellular conditions [48].

PRAME may promote tumor progression through the inhibition of differentiation, growth arrest, and apoptosis induced by retinoic acid [57]. Through this study, we demonstrate that IDPRs, DOT regions, droplet-promoting regions, and aggregation hot spots are present in PRAME and provide it the ability to undergo complex PPI. These entities may be considered when developing novel immunotherapies or biologics targeting PRAME as they may play a role in the increased metastasis and pathogenesis of Class 1 UM. The presence of highly dynamic intrinsically disordered protein domains and a potentially small interaction network make PRAME an even more attractive target than has previously been described [57, 58].

Our study brings new considerations in the molecular behavior of PRAME. Our group's and other studies have shown that IDPs and IDPRs are associated with different ocular malignancies, including UM [10, 14, 22, 23]. IDPs, IDPRs, and DOT regions have been considered attractive drug targets for many years [12, 22–24]. There is tremendous potential to apply these same principles to developing immunotherapies directed at PRAME to treat uveal melanoma. PRAME is also expressed in other non-UM malignancies, including cutaneous melanoma, breast carcinoma, non-small cell lung cancer, and leukemia [59]. Therefore, the findings in this study may be applied to those pathologies. Additionally, our group has characterized the association of IDPRs found in a protein associated with an ocular neoplasm to the recently appreciated phenomenon of membrane-less organelles (MLOs) that is driven by liquid–liquid phase separation (LLPS). We demonstrate that PRAME can potentially promote the formation of these MLOs via LLPS. The utility of these findings will need to be researched further.

Among the most intriguing recent developments in the field of the molecular biology of the cell is the recognition of the abundant presence of MLOs. MLOs are subcellular structures formed by proteins and other macromolecules not enclosed by a lipid bilayer. These liquid droplet-like entities are formed through the process LLPS, where parts of the homogeneous solution of a cell separate to form two distinct phases. LLPS represent an ideal stress-response mechanism that allows for the spatio-temporal organization of macromolecular components in living cells and are extremely sensitive to even minute variations of the environment (e.g., ionic strength, pH, temperature, and solute concentration) [60–66]. Our group has become interested in these entities because the fluidity of MLOs is rooted in LLPS, which is thought to be driven by the interaction of intrinsically disordered protein regions of various proteins [60, 61, 67–69]. Although MLOs have multiple physiological functions, aberrant LLPS can be deleterious and are linked to various maladies, including neurodegenerative diseases and cancer [70–72]. In the context of our study, we know that PRAME is upregulated in a subset of Class 1 UM associated with an increased risk of metastasis. It is possible that PRAME's predicted ability to undergo LLPS and propensity to form droplets may alter immune-associated signaling to dampen tumor regulation and enhance cell turnover.

Our study does not come without its limitations. We used databases and bioinformatics tools, and this approach does not represent a comprehensive nor an exhaustive investigation into the molecular behavior of PRAME. We aimed to analyze the intrinsic disorder qualitatively and quantitatively in PRAME and make a presumption on how intrinsic protein disorder may play a role in the pathogenesis of UM. Further work modeling these disordered segments through molecular dynamic simulations or docking analysis would be the next steps to target these regions further effectively. If these efforts were focused on in-vitro disorder-based drug discovery efforts, it may lead to the most success. Consequently,

Antonietti *et al. Cell Communication and Signaling*   (2023) 21:222

Page 11 of 13

we could develop a novel immunotherapy targeting the intrinsic disorder in PRAME, which may be identified and eventually translated to direct care of PRAME(+) UM patients. These findings are strictly theoretical; however, our findings establish a necessary biological framework for targeting PRAME.

Through this study, we have demonstrated that intrinsically disordered protein regions are likely involved in functionality of PRAME and could be related to its pathogenesis in UM. Further investigation will be necessary to translate these findings into patient care. Modern drug discovery and therapeutic target development are grounded in structure and order-based assumptions. We are taking a different aim than that paradigm and believe these drug discovery campaigns must also be built on the assumption that proteins have inherent flexibility, thus altering the therapeutic development process. This study represents an exciting take on immunotherapy development. Successful efforts to target the intrinsically disordered-based proteins of PRAME may reduce rates of metastasis and poor outcomes in Class 1 uveal melanomas.

Conclusion

PRAME plays an important role in tumor biology and has been identified as a biomarker for Class 1 UM. PRAME(+) Class 1 tumors are associated with higher rates of metastasis and, thus, worse prognosis. Our study demonstrates that intrinsically disordered protein regions are present in PRAME. IDPRs in PRAME likely have a role in its functionality and protein–protein interactions. We show that intrinsically disordered protein regions or disorder-to-order transition regions in PRAME are likely important for the functionality of the protein and could be targeted when developing immunotherapies for this protein. Future studies should further validate our findings to better elucidate the role of IDPRs in PRAME for the treatment of Uveal Melanoma.

Abbreviations

UM	Uveal melanoma
GEP	Gene expression profiling
PRAME	Preferentially expressed antigen in melanoma
BRCA	Breast cancer type 1 susceptibility protein
BAP1	BRCA-associated protein 1
IDPs	Intrinsically disordered proteins
IDPRs	Intrinsically disordered protein regions
UniProt	Universal Protein Resource
PBD	Protein Data Bank Select
RIDAO	Rapid Intrinsic Disorder Analysis Online
PONDR	Predictor of Natural Disorder Regions
PPDR	Percent of predicted disorder residues
ADS	Average disorder score
D²P²	Database of Disordered Protein Prediction
SCOP	Structural Classification of Proteins
MoRFs	Molecular recognition features
PTMs	Posttranslational modifications
DOT	Disorder-to-order transition
DPRs	Droplet-promoting regions
STRING	Search Tool for the Retrieval of Interacting Genes
LLR	Leucine-rich repeats
RARA	Retinoic acid receptor alpha
MLOs	LLPS propensity (p_{LLPS}), membrane-less organelles
LLPS	Liquid–liquid phase separation
CAID	Critical Assessment of protein Intrinsic Disorder prediction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12964-023-01197-y.

> **Additional file 1: Figure S1.** PRAME amino acid sequence. Shown in FASTA format, the protein sequence represents PRAME canonical form. The protein contains 509 residues.

Availability of data and materials
Data and materials available upon request.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
Dr. Karp has a pending PCT/US2022/029842 with the University of Miami, Dr. Karp and Dr. Shields are on the medical advisory board for Interfeen Biologics.

Author details
[1]Bascom Palmer Eye Institute, University of Miami, 900 NW 17th Street, Miami, FL 33136, USA. [2]Department of Chemistry, College of Art and Sciences, University of South Florida, FL 33612 Tampa, USA. [3]Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, FL 33612 Tampa, USA. [4]Ocular Oncology Service, Wills Eye Hospital, Thomas Jefferson University, PA, Philadelphia, USA.

References

1. Kaliki S, Shields CL. Uveal melanoma: relatively rare but deadly cancer. Eye (Lond). 2017;31(2):241–57. https://doi.org/10.1038/eye.2016.275. (From NLM Medline).
2. Chattopadhyay C, Kim DW, Gombos DS, Oba J, Qin Y, Williams MD, Esmaeli B, Grimm EA, Wargo JA, Woodman SE, et al. Uveal melanoma: From diagnosis to treatment and the science in between. Cancer. 2016;122(15):2299–312. https://doi.org/10.1002/cncr.29727. (From NLM Medline).

Antonietti *et al. Cell Communication and Signaling*   (2023) 21:222

Page 12 of 13

3.   Diener-West M, Reynolds SM, Agugliaro DJ, Caldwell R, Cumming K, Earle JD, Hawkins BS, Hayman JA, Jaiyesimi I, Jampol LM, et al. Development of metastatic disease after enrollment in the COMS trials for treatment of choroidal melanoma: Collaborative Ocular Melanoma Study Group Report No. 26. Arch Ophthalmol. 2005;123(12):1639–43. https://doi.org/10.1001/archopht.123.12.1639. (From NLM Medline).

4.   Onken MD, Worley LA, Ehlers JP, Harbour JW. Gene expression profiling in uveal melanoma reveals two molecular classes and predicts metastatic death. Cancer Res. 2004;64(20):7205–9. https://doi.org/10.1158/0008-5472.CAN-04-1750. (From NLM Medline).

5.   Onken MD, Worley LA, Char DH, Augsburger JJ, Correa ZM, Nudleman E, Aaberg TM Jr, Altaweel MM, Bardenstein DS, Finger PT. Collaborative Ocular Oncology Group report number 1: prospective validation of a multi-gene prognostic assay in uveal melanoma. Ophthalmology. 2012;119(8):1596–603.

6.   Kaler CJ, Dollar JJ, Cruz AM, Kuznetsoff JN, Sanchez MI, Decatur CL, Licht JD, Smalley KS, Correa ZM, Kurtenbach S. BAP1 loss promotes suppressive tumor immune microenvironment via upregulation of PROS1 in class 2 uveal melanomas. Cancers. 2022;14(15):3678.

7.   Harbour JW, Onken MD, Roberson ED, Duan S, Cao L, Worley LA, Council ML, Matatall KA, Helms C, Bowcock AM. Frequent mutation of BAP1 in metastasizing uveal melanomas. Science. 2010;330(6009):1410–3.

8.   Djulbegovic MB, Taylor DJ, Uversky VN, Galor A, Shields CL, Karp CL. Intrinsic Disorder in BAP1 and Its Association with Uveal Melanoma. Genes. 2022;13(10):1703.

9.   Field MG, Decatur CL, Kurtenbach S, Gezgin G, van der Velden PA, Jager MJ, Kozak KN, Harbour JW. PRAME as an Independent Biomarker for Metastasis in Uveal Melanoma. Clin Cancer Res. 2016;22(5):1234–42. https://doi.org/10.1158/1078-0432.CCR-15-2071. (From NLM Medline).

10.   Djulbegovic, M. B.; Uversky, V. N.; Harbour, J. W.; Galor, A.; Karp, C. L. Structural Protein Analysis of Driver Gene Mutations in Conjunctival Melanoma. Genes (Basel) 2021, 12 (10). https://doi.org/10.3390/genes12101625 From NLM Medline.

11.   van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, et al. Classification of intrinsically disordered regions and proteins. Chem Rev. 2014;114(13):6589–631. https://doi.org/10.1021/cr400525m. (From NLM Medline).

12.   Peng Z, Yan J, Fan X, Mizianty MJ, Xue B, Wang K, Hu G, Uversky VN, Kurgan L. Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. Cell Mol Life Sci. 2015;72(1):137–51. https://doi.org/10.1007/s00018-014-1661-9. (From NLM Medline).

13.   Uversky VN, Dave V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, Joerger AC. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. Chem Rev. 2014;114(13):6844–79. https://doi.org/10.1021/cr400713r. (From NLM Medline).

14.   Uversky VN, Oldfield CJ, Midic U, Xie H, Xue B, Vucetic S, Iakoucheva LM, Obradovic Z, Dunker AK. Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. BMC Genomics. 2009;10(Suppl 1):S7. https://doi.org/10.1186/1471-2164-10-S1-S7. (From NLM Medline).

15.   Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys. 2008;37:215–46. https://doi.org/10.1146/annurev.biophys.37.032807.125924. (From NLM Medline).

16.   Trivedi, R.; Nagarajaram, H. A. Intrinsically Disordered Proteins: An Overview. Int J Mol Sci 2022, 23 (22). https://doi.org/10.3390/ijms232214050 From NLM Medline.

17.   Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: regulation and disease. Curr Opin Struct Biol. 2011;21(3):432–40. https://doi.org/10.1016/j.sbi.2011.03.011. (From NLM).

18.   UniProt C. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47(D1):D506–15. https://doi.org/10.1093/nar/gky1049. (From NLM Medline).

19.   Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022;50(D1):D439–44. https://doi.org/10.1093/nar/gkab1061. (From NLM Medline).

20.   Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Zidek A, Bridgland A, Cowie A, Meyer C, Laydon A, et al. Highly accurate protein structure prediction for the human proteome. Nature. 2021;596(7873):590–6. https://doi.org/10.1038/s41586-021-03828-1. (From NLM Medline).

21.   Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2. (From NLM Medline).

22.   Vacic V, Uversky VN, Dunker AK, Lonardi S. Composition Profiler: a tool for discovery and visualization of amino acid composition differences. BMC Bioinformatics. 2007;8:211. https://doi.org/10.1186/1471-2105-8-211. (From NLM Medline).

23.   Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. BMC Bioinformatics. 2006;7:208. https://doi.org/10.1186/1471-2105-7-208. (From NLM Medline).

24.   Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK. Exploiting heterogeneous sequence properties improves prediction of protein disorder. Proteins. 2005;61(Suppl 7):176–82. https://doi.org/10.1002/prot.20735. (From NLM Medline).

25.   Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK, Obradovic Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. J Bioinform Comput Biol. 2005;3(1):35–60. https://doi.org/10.1142/s0219720005000886. (From NLM Medline).

26.   Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. Proteins. 2001;42(1):38–48 (10.1002/1097-0134(20010101)42:1<38::aid-prot50>3.0.co;2-3 From NLM Medline).

27.   Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim Biophys Acta. 2010;1804(4):996–1010. https://doi.org/10.1016/j.bbapap.2010.01.011. (From NLM Medline).

28.   Meszaros B, Erdos G, Dosztanyi Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 2018;46(W1):W329–37. https://doi.org/10.1093/nar/gky384. (From NLM Medline).

29.   Hu XS, Nascimento TD, Bender MC, Hall T, Petty S, O'Malley S, Ellwood RP, Kaciroti N, Maslowski E, DaSilva AF. Feasibility of a Real-Time Clinical Augmented Reality and Artificial Intelligence Framework for Pain Detection and Localization From the Brain. J Med Internet Res. 2019;21(6):e13594. https://doi.org/10.2196/13594. (From NLM Medline).

30.   Wang H, Li S, Wang C, Wang Y, Fang J, Liu K. Plasma and Vitreous Metabolomics Profiling of Proliferative Diabetic Retinopathy. Invest Ophthalmol Vis Sci. 2022;63(2):17. https://doi.org/10.1167/iovs.63.2.17. (From NLM Medline).

31.   Hanson J, Paliwal KK, Litfin T, Zhou Y. SPOT-Disorder 2: Improved Protein Intrinsic Disorder Prediction by Ensembled Deep Learning. Genomics Proteomics Bioinformatics. 2019;17(6):645–56. https://doi.org/10.1016/j.gpb.2019.01.004.

32.   Emenecker RJ, Griffith D, Holehouse AS. Metapredict: a fast, accurate, and easy-to-use predictor of consensus disorder and structure. Biophys J. 2021;120(20):4312–9. https://doi.org/10.1016/j.bpj.2021.08.039.

33.   Rajagopalan K, Mooney SM, Parekh N, Getzenberg RH, Kulkarni P. A majority of the cancer/testis antigens are intrinsically disordered proteins. J Cell Biochem. 2011;112(11):3256–67. https://doi.org/10.1002/jcb.23252. (From NLM Medline).

34.   Oates ME, Romero P, Ishida T, Ghalwash M, Mizianty MJ, Xue B, Dosztanyi Z, Uversky VN, Obradovic Z, Kurgan L, et al. D(2)P(2): database of disordered protein predictions. Nucleic Acids Res. 2013;41(Database issue):D508-516. https://doi.org/10.1093/nar/gks1226. (From NLM Medline).

35.   Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics. 2005;21(16):3433–4. https://doi.org/10.1093/bioinformatics/bti541. (From NLM Medline).

36.   Ishida T, Kinoshita K. PrDOS: prediction of disordered protein regions from amino acid sequence. Nucleic Acids Res. 2007;35(Web Server issue):W460-464. https://doi.org/10.1093/nar/gkm363. (From NLM Medline).

37.   Walsh I, Martin AJ, Di Domenico T, Tosatto SC. ESpritz: accurate and fast prediction of protein disorder. Bioinformatics. 2012;28(4):503–9. https://doi.org/10.1093/bioinformatics/btr682. (From NLM Medline).

38.   Andreeva A, Howorth D, Brenner SE, Hubbard TJ, Chothia C, Murzin AG. SCOP database in 2004: refinements integrate structure and sequence family data. Nucleic Acids Res. 2004;32(Database issue):D226-229. https://doi.org/10.1093/nar/gkh039. (From NLM Medline).

39. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol. 1995;247(4):536–40. https://doi.org/10.1006/jmbi.1995.0159. (From NLM Medline).

40. de Lima Morais DA, Fang H, Rackham OJ, Wilson D, Pethica R, Chothia C, Gough J. SUPERFAMILY 175 including a domain-centric gene ontology method. Nucleic Acids Res. 2011;39(Database issue):D427-434. https://doi.org/10.1093/nar/gkq1130. (From NLM Medline).

41. Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L, et al. InterPro in 2022. Nucleic Acids Res. 2022. https://doi.org/10.1093/nar/gkac993. (From NLM).

42. Meszaros B, Simon I, Dosztanyi Z. Prediction of protein binding regions in disordered proteins. PLoS Comput Biol. 2009;5(5):e1000376. https://doi.org/10.1371/journal.pcbi.1000376. (From NLM Medline).

43. Katuwawala A, Peng Z, Yang J, Kurgan L. Computational Prediction of MoRFs, Short Disorder-to-order Transitioning Protein Binding Regions. Comput Struct Biotechnol J. 2019;17:454–62. https://doi.org/10.1016/j.csbj.2019.03.013. (FromN LM PubMed-not-MEDLINE).

44. Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, Latham V, Sullivan M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res. 2012;40(Database issue):D261-270. https://doi.org/10.1093/nar/gkr1122. (From NLM Medline).

45. Erdos G, Pajkos M, Dosztanyi Z. IUPred3: prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. Nucleic Acids Res. 2021;49(W1):W297–303. https://doi.org/10.1093/nar/gkab408. (From NLM Medline).

46. Erdos G, Dosztanyi Z. Analyzing Protein Disorder with IUPred2A. Curr Protoc Bioinformatics. 2020;70(1):e99. https://doi.org/10.1002/cpbi.99. (From NLM Medline).

47. Hardenberg M, Horvath A, Ambrus V, Fuxreiter M, Vendruscolo M. Widespread occurrence of the droplet state of proteins in the human proteome. Proc Natl Acad Sci U S A. 2020;117(52):33254–62. https://doi.org/10.1073/pnas.2007670117. (From NLM).

48. Horvath A, Vendruscolo M, Fuxreiter M. Sequence-based Prediction of the Cellular Toxicity Associated with Amyloid Aggregation within Protein Condensates. Biochemistry. 2022;61(22):2461–9. https://doi.org/10.1021/acs.biochem.2c00499. (From NLM).

49. Vendruscolo M, Fuxreiter M. Sequence Determinants of the Aggregation of Proteins Within Condensates Generated by Liquid-liquid Phase Separation. J Mol Biol. 2022;434(1):167201. https://doi.org/10.1016/j.jmb.2021.167201. (From NLM).

50. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–13. https://doi.org/10.1093/nar/gky1131. (From NLM Medline).

51. Dayhoff GW, Uversky VN 2nd. Rapid prediction and analysis of protein intrinsic disorder. Protein Sci. 2022;31(12):e4496. https://doi.org/10.1002/pro.4496. (From NLM).

52. Necci M, Piovesan D, Predictors C, DisProt C, Tosatto SCE. Critical assessment of protein intrinsic disorder prediction. Nat Methods. 2021;18(5):472–81. https://doi.org/10.1038/s41592-021-01117-3.

53. Hu G, Katuwawala A, Wang K, Wu Z, Ghadermarzi S, Gao J, Kurgan L. flDPnn: Accurate intrinsic disorder prediction with putative propensities of disorder functions. Nat Commun. 2021;12(1):4438. https://doi.org/10.1038/s41467-021-24773-7.

54. Wang S, Ma J, Xu J. AUCpreD: proteome-level protein disorder prediction by AUC-maximized deep convolutional neural fields. Bioinformatics. 2016;32(17):i672–9. https://doi.org/10.1093/bioinformatics/btw446.

55. Bah A, Forman-Kay JD. Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. J Biol Chem. 2016;291(13):6696–705. https://doi.org/10.1074/jbc.R115.695056. (From NLM).

56. Uversky VN. Protein intrinsic disorder and structure-function continuum. Prog Mol Biol Transl Sci. 2019;166:1–17. https://doi.org/10.1016/bs.pmbts.2019.05.003. (From NLM).

57. Epping MT, Wang L, Edel MJ, Carlee L, Hernandez M, Bernards R. The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. Cell. 2005;122(6):835–47. https://doi.org/10.1016/j.cell.2005.07.003. (From NLM Medline).

58. Amir AL, van der Steen DM, van Loenen MM, Hagedoorn RS, de Boer R, Kester MD, de Ru AH, Lugthart GJ, van Kooten C, Hiemstra PS, et al. PRAME-specific Allo-HLA-restricted T cells with potent antitumor reactivity useful for therapeutic T-cell receptor gene transfer. Clin Cancer Res. 2011;17(17):5615–25. https://doi.org/10.1158/1078-0432.CCR-11-1066. (From NLM Medline).

59. Ikeda H, Lethe B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T, et al. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. Immunity. 1997;6(2):199–208. https://doi.org/10.1016/s1074-7613(00)80424-4. (From NLM Medline).

60. Fonin AV, Antifeeva IA, Kuznetsova IM, Turoverov KK, Zaslavsky BY, Kulkarni P, Uversky VN. Biological soft matter: intrinsically disordered proteins in liquid-liquid phase separation and biomolecular condensates. Essays Biochem. 2022;66(7):831–47. https://doi.org/10.1042/ebc20220052. (From NLM).

61. Antifeeva IA, Fonin AV, Fefilova AS, Stepanenko OV, Povarova OI, Silonov SA, Kuznetsova IM, Uversky VN, Turoverov KK. Liquid-liquid phase separation as an organizing principle of intracellular space: overview of the evolution of the cell compartmentalization concept. Cell Mol Life Sci. 2022;79(5):251. https://doi.org/10.1007/s00018-022-04276-4. (From NLM).

62. Nesterov SV, Ilyinsky NS, Uversky VN. Liquid-liquid phase separation as a common organizing principle of intracellular space and biomembranes providing dynamic adaptive responses. Biochim Biophys Acta Mol Cell Res. 2021;1868(11):119102. https://doi.org/10.1016/j.bbamcr.2021.119102. (From NLM).

63. Alberti S, Hyman AA. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. Nat Rev Mol Cell Biol. 2021;22(3):196–213. https://doi.org/10.1038/s41580-020-00326-6. (From NLM).

64. Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, et al. Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol. 2018;28(6):420–35. https://doi.org/10.1016/j.tcb.2018.02.004. (From NLM).

65. Shin, Y.; Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. Science 2017, 357 (6357). https://doi.org/10.1126/science.aaf4382 From NLM.

66. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol. 2017;18(5):285–98. https://doi.org/10.1038/nrm.2017.7. (From NLM).

67. Uversky VN. Recent Developments in the Field of Intrinsically Disordered Proteins: Intrinsic Disorder-Based Emergence in Cellular Biology in Light of the Physiological and Pathological Liquid-Liquid Phase Transitions. Annu Rev Biophys. 2021;50:135–56. https://doi.org/10.1146/annurev-biophys-062920-063704. (From NLM).

68. Brocca, S.; Grandori, R.; Longhi, S.; Uversky, V. Liquid-Liquid Phase Separation by Intrinsically Disordered Protein Regions of Viruses: Roles in Viral Life Cycle and Control of Virus-Host Interactions. Int J Mol Sci 2020, 21 (23). https://doi.org/10.3390/ijms21239045 From NLM.

69. Uversky, V. N.; Finkelstein, A. V. Life in Phases: Intra- and Inter- Molecular Phase Transitions in Protein Solutions. Biomolecules 2019, 9 (12). https://doi.org/10.3390/biom9120842 From NLM.

70. Ahmad A, Uversky VN, Khan RH. Aberrant liquid-liquid phase separation and amyloid aggregation of proteins related to neurodegenerative diseases. Int J Biol Macromol. 2022;220:703–20. https://doi.org/10.1016/j.ijbiomac.2022.08.132. (From NLM).

71. Darling AL, Shorter J. Combating deleterious phase transitions in neurodegenerative disease. Biochim Biophys Acta Mol Cell Res. 2021;1868(5):118984. https://doi.org/10.1016/j.bbamcr.2021.118984. (From NLM).

72. Lu J, Qian J, Xu Z, Yin S, Zhou L, Zheng S, Zhang W. Emerging Roles of Liquid-Liquid Phase Separation in Cancer: From Protein Aggregation to Immune-Associated Signaling. Front Cell Dev Biol. 2021;9:631486. https://doi.org/10.3389/fcell.2021.631486. (From NLM).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.