



Universidade Federal do Rio de Janeiro
Centro de Ciências Matemáticas e da Natureza
Instituto de Química

Projeto Final de Curso

**Investigação da propensão à separação de fases da
proteína rica em glicina AtGRP2 de *Arabidopsis
thaliana***

Clara Malizia Leal Ferreira da Motta

Rio de Janeiro

Novembro/2019

Investigação da propensão à separação de fase da proteína rica em glicina 2 de *Arabidopsis thaliana*

Clara Malizia Leal Ferreira da Motta

PROJETO DE FINAL DE CURSO SUBMETIDO AO CORPO DOCENTE DO DEPARTAMENTO DE BIOQUÍMICA DO INSTITUTO DE QUÍMICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE QUÍMICO COM ATRIBUIÇÕES TECNOLÓGICAS.

Prof. Anderson de Sá Pinheiro (IQ - UFRJ)

Prof^a. Alexandre Braga da Rocha (IQ - UFRJ)

Prof. Marcos Dias Pereira (IQ - UFRJ)

RIO DE JANEIRO, RJ – BRASIL
NOVEMBRO/2019

Sumário

Agradecimentos	4
Índices de Figuras	6
1.Introdução	7
Organelas desprovidas de Membrana	7
Poteínas Ricas em Glicina.....	11
AtGRP2	16
2.Objetivo	20
Objetivo geral	20
Objetivos específicos.....	20
3.Material e Métodos	21
3.1.Material	21
3.1.1.Meio de Cultura.....	21
3.1.2.Vetor de Expressão.....	21
3.1.3.Cepas Bacterianas	22
3.2.Métodos.....	23
3.2.1.Transformação Bacteriana.....	23
3.2.2. Expressão de His ₆ MBP-AtGRP2	23
3.2.3. Lise Bacteriana	24
3.2.4. Purificação de AtGRP2	24
3.2.4.1. Primeira Cromatografia de Afinidade a Níquel.....	24
3.2.4.2. Diálise e Clivagem com TEV	25
3.2.4.3. Segunda Cromatografia de Afinidade a Níquel	25
3.2.4.4. Cromatografia por Exclusão Molecular	25
3.2.5. Eletroforese Desnaturante em Gel de Poliacrilamida	26
3.2.6. Dosagem de Proteína	26
3.2.7. Microscopia de Contraste de Interferência Diferencial (DIC – <i>Differential Interference Contrast</i>) e de Fluorescência	26
4. Resultados e Discussão	28
4.1. Purificação de AtGRP2	28
4.2. Propensão à Desordem e Separação de fases de AtGRP2.....	31
4.3. AtGRP2 Forma Condensados em Gel	37
4.4. Os condensados de AtGRP2 são de Natureza Proteica	39
4.5. A LLPS de AtGRP2 não é Primariamente Dirigida por Interações Eletrostáticas.....	42
5. Conclusão	43
6. Referências Bibliográficas	45

Agradecimentos

Primeiramente gostaria de agradecer aos meus pais, Luciano Leal Ferreira da Motta e Rosana de Fátima Malizia Motta, que me apoiaram na decisão de cursar química e me ajudaram sempre que precisei. Agradeço meus irmãos, Giulia Malizia Leal Ferreira da Motta e Eduardo Malizia Leal Ferreira da Motta, por servirem de base emocional, me dando todos os tipos de conselhos possíveis e imaginários, e me ajudando nos momentos difíceis pelos quais passei, e principalmente me incentivando a fazer o que eu amo e pedindo ajuda. Agradeço à minha família como um todo que, apesar de na sua maioria discordarem da minha decisão, sempre que precisei, todos estavam dispostos a me ajudar. Agradeço, também, por terem me fornecido meios que me permitiram alcançar os objetivos que eu almejei até hoje.

Agradeço aos amigos do IQ, por me suportarem em todos as aulas, e por me darem liberdade de ser como eu sou em sala, mesmo que perturbando o juízo de todo mundo. Aos meus professores, que me fizeram perceber o quanto a química é fantástica mesmo desde o primeiro período, o que me fez ser mais entusiasmada para seguir em frente até o final desse curso. Agradeço a eles, principalmente, pela nobreza da profissão, pelos conselhos e pelos conhecimentos, por decidirem ser essa parte tão importante na vida de um aluno.

Em especial, quero agradecer aos meus amigos: Mari, Sansoldo, Paulo, Rafa, Diogo, Larissa e Rodrigo por terem me aguentado todo esse tempo, por terem me dado conselhos e me guiado (nos guiamos juntos muitas vezes) e por terem feito desse ambiente acadêmico menos estressante e mais amigável. Vocês foram grande parte dessa jornada com certeza!

Agradeço ao professor Anderson Pinheiro, que é meu orientador, mas antes disso me apresentou a área da bioquímica, que me fez entender até mesmo a biologia que eu tive tanta dificuldade nos meus anos de ensino fundamental e médio. Agradeço a ele por toda paciência, broncas, pelas explicações demoradas, mas sempre muito proveitosas, por tudo que ele me deu acesso e me ensinou. Obrigado por tudo mesmo, você é uma grande inspiração para mim, todos os seus feitos e sua dedicação, espero um dia conseguir ser, no mínimo, metade do que você é.

Agradeço também à Mari, que vem me ajudando com todo esse trabalho e me ensinando muito sobre, não só sobre separação de fases, mas também de como funciona o processo científico como um todo. Obrigada ao seu carinho e atenção e a

todo esforço que você tem feito comigo e com o Anderson pra levar esse trabalho adiante.

Agradeço ao professor Alexandre Braga que me recebeu em sua sala para conversar sobre as mais diversas coisas, de carreira a gostos pessoais, que me deu conselhos importantíssimos sobre decisões que alteraram muito o curso da minha vida. Agradeço pela sua paciência e dedicação aos seus alunos

Agradeço aos meus colegas do LaBMol, Daniel, Karina, Igor, Nathalia e por terem feito os dias de trabalho não tão exaustivos quanto poderiam ser. Agradeço os momentos, as viagens, as brincadeiras, as saídas e as comemorações. Agradeço, principalmente, ao Daniel, que me ajuda frequentemente com coisas relacionadas ao laboratório e a trabalho e por ter me aguentado todo esse tempo.

Agradeço às professoras Bianca e Yraima que disponibilizaram seus equipamentos, seus materiais e, principalmente, seu tempo para garantir o sucesso desse trabalho e da minha evolução como profissional.

Agradeço também ao professor Marcos Dias e mais uma vez ao professor Alexandre Braga por terem aceitado participar da minha banca examinadora, respondendo de prontamente meu convite. Muito obrigado!

Índices de figuras

Figura 1 LLPs d: Representação esquemática das MLOs presentes nas células eucarióticas	8
Figura 2 Forças interatômicas que dirigem a LLPS	11
Figura 3 Diferentes classes de proteínas ricas em glicina.....	13
Figura 4 Esquema das possíveis conformações dos motivos de sequência ricos em glicina	15
Figura 5 Arquitetura em domínios de AtGRP2	19
Figura 6 Representação esquemática do vetor de expressão pETM30-MBP e da construção His ₆ MBP-TEV-AtGRP2	22
Figura 7 Primeira etapa de purificação de His ₆ MBP-AtGRP2.....	28
Figura 8 Clivagem enzimática da construção His ₆ MBP-AtGRP2	29
Figura 9 Purificação de AtGRP2 por cromatografia de exclusão molecular	30
Figura 10 AtGRP2 é uma proteína não globular e altamente compacta	31
Figura 11 Análise de propensão à desordem de AtGRP2	34
Figura 12 Propensão à LLPS de AtGRP2 predita por bioinformática	36
Figura 13 LLPs de AtGRP2 investigada por microscopia DIC.....	39
Figura 14 AtGRP2 forma condensados proteicos dispostos em cachos.....	40
Figura 15 Efeito de sal e 1,6-hexanodiol na LLPS de AtGRP2	41

1. Introdução

1.1. Organelas desprovidas de membrana

O planeta Terra e os seres que nele habitam vivem em constante mudança. Para uma célula, perceber essas mudanças e agir sobre elas é de fundamental importância para a sobrevivência. A capacidade de transformação, resposta, reprodução ou qualquer outra atividade celular depende de reações bioquímicas que precisam ser controladas com precisão no espaço e no tempo⁽¹⁾. Uma forma eficiente de modular tais reações e regular as atividades celulares conforme a necessidade do organismo em um dado momento é a compartimentalização⁽²⁾.

As organelas são os “pequenos órgãos” da célula. Elas reúnem funções celulares específicas por meio da compartimentalização dos componentes de uma determinada reação bioquímica, permitindo o controle espaço-temporal sobre processos metabólicos e vias de sinalização⁽³⁾. As organelas mais conhecidas, tais como, núcleo, mitocôndria, retículo endoplasmático, lisossomo e cloroplasto, são envoltas em uma membrana lipídica. Estas têm suas funções bem caracterizadas, já que são estudadas desde o início do século XIX⁽⁴⁾. Entretanto, as células também possuem as chamadas organelas desprovidas de membrana (MLOs – *Membraneless Organelles*), unidades funcionais celulares que não são delimitadas por uma bicamada lipídica⁽²⁾. As MLOs constituem complexos supramoleculares formados principalmente por proteínas e ácidos nucleicos que estão presentes tanto no núcleo (p. ex., nucléolo, corpos de Cajal e *speckles* nucleares) quanto no citoplasma (p. ex., grânulos de estresse, corpúsculos de processamento (P), corpos de Balbiani e centríolo)^(5,6). Muitas destas MLOs foram identificadas há décadas, entretanto diversas questões acerca dos mecanismos através dos quais essas organelas se formam e da relação entre suas características físicas e funções biológicas permanecem sem resposta.

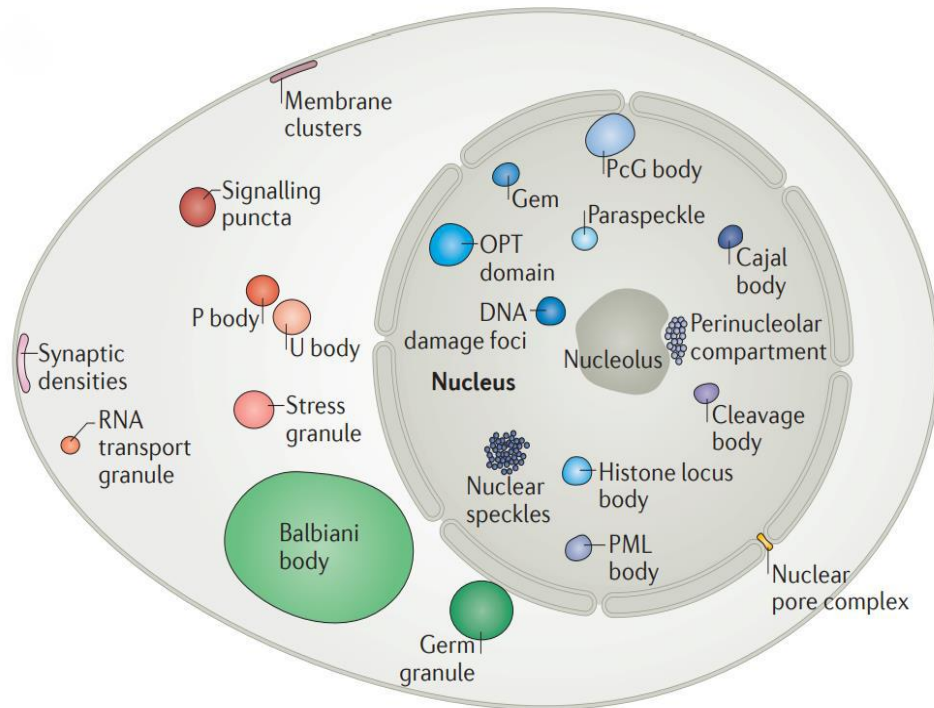


Figura 1 | Representação esquemática das MLOs presentes nas células eucarióticas. Esquema representando os possíveis condensados presentes no núcleo, citoplasma e membrana plasmática de células eucarióticas. Alguns compartimentos mostrados na imagem são típicos de células específicas e não necessariamente aparecem em qualquer célula, como por exemplo os corpos de Balbiani. (BANANI SF *et al.*, 2017).

Estudos pioneiros revelaram a natureza dinâmica das MLOs⁽⁷⁻⁹⁾. Em 2005, foi mostrado que os corpos de Cajal se comportam como “esferas semifluidas suspensas no nucleoplasma”⁽⁷⁾. Entretanto, a natureza física desses complexos ainda permanecia desconhecida. Somente em 2009, Brangwyne, Jülicher e Hyman reportaram que os grânulos P (corpos contendo proteína e RNA presentes nos embriões de *Caenorhabditis elegans*) apresentam propriedades líquidas e são formados pelo fenômeno de separação de fases líquido-líquido (LLPS – *Liquid-liquid phase separation*)⁽¹⁰⁾. Esse processo físico ocorre quando uma solução supersaturada de um determinado componente se separa espontaneamente em duas fases, uma densa e outra diluída, que coexistem de maneira estável⁽¹⁰⁾. A natureza líquida dos corpos P foi evidenciada pela sua aparência arredondada, devido à baixa tensão superficial, capacidade de sofrer eventos de fusão e fissão e intercâmbio dinâmico entre seus componentes⁽¹¹⁾. Dois anos depois, observações similares foram feitas para o nucléolo⁽¹²⁾.

A LLPS é um fenômeno bem conhecido na química de polímeros⁽¹³⁾. Todavia, o conceito de que a LLPS é o princípio formador das MLOs surgiu apenas recentemente. Li *et al.* (2012)⁽¹⁴⁾ mostraram que corpos líquidos formados por proteína e RNA, reconstituídos *in vitro* a partir de seus componentes purificados, são capazes de promover a nucleação de actina, demonstrando definitivamente o papel da LLPS na formação das MLOs. Diversas evidências experimentais sugeriram que proteínas e outras macromoléculas biológicas, como RNA, são capazes de formar condensados espacialmente organizados apresentando diferentes estados de materiais, como líquido, gel ou agregado sólido⁽¹⁵⁻¹⁷⁾.

Estudos genéticos e de proteômica identificaram os componentes proteicos de diversas MLOs⁽¹⁸⁻²¹⁾. Estes estudos mostraram que a multivalência, ou seja, a capacidade de interações múltiplas, é uma característica comum entre proteínas que sofrem LLPS. A multivalência pode estar presente tanto em proteínas enoveladas quanto desordenadas. No primeiro caso, domínios proteicos enovelados contendo superfícies de interação definidas interagem de maneira específica para formar oligômeros que podem realizar interações multivalente com outras proteínas. No segundo, motivos lineares de sequência de aminoácidos estão presentes em regiões desordenadas e servem como arcabouço para a formação de complexos proteicos multivalentes. Ainda, a característica de multivalência pode surgir da combinação de domínios enovelados e regiões intrinsecamente desordenadas em proteínas. Logo, a presença de regiões intrinsecamente desordenadas (IDRs – *Intrinsically disordered regions*) ou de sequências inteiramente desordenadas (IDPs - *Intrinsically disordered proteins*), ou seja, que não apresentam uma estrutura tridimensional definida em seu estado nativo, é uma característica frequente em proteínas que fazem LLPS⁽²²⁾. Muitas IDRs/IDPs são formadas por pequenas repetições de aminoácidos específicos, conhecidas como domínios de baixa complexidade (LCDs – *Low complexity domains*)⁽²²⁻²⁴⁾. A formação de complexos supramoleculares por proteínas enriquecidas em IDRs e LCDs leva a formação de MLOs com diferentes propriedades de materiais. Por exemplo, os corpos de Balbiani são condensados sólidos estabilizados por fortes interações entre regiões proteicas que adotam conformação em fita- β ⁽²¹⁾. Ainda, diversos grânulos de proteína e RNA são dinâmicos e de natureza líquida. As proteínas ligantes de RNA (RBPs – *RNA-binding proteins*) são conhecidas por apresentarem uma arquitetura em domínios modular e multivalente, uma característica importante para a LLPS⁽²⁵⁾.

Além de servirem como regiões de conexão entre domínios enovelados, as IDRs podem mediar interações que promovem a LLPS⁽²⁶⁾. Kato *et al.* (2012)⁽²⁷⁾ mostraram que soluções concentradas de diferentes IDRs podem formar hidrogéis espontaneamente em função do tempo. Além disso, Kim *et al.* (2013)⁽²⁸⁾ descreveram mutações associadas a doenças nas IDRs de hnRNPA1 e hnRNPA2B1 que resultam na aceleração da formação de estrutura fibrilares sólidas *in vitro*. Diversos trabalhos subsequentes mostraram que proteínas contendo IDRs, cujas sequências encontram-se associadas a determinadas patologias, como hnRNPA1 e FUS (*Fused in Sarcoma*), também são capazes de formar condensados líquidos⁽²⁹⁻³¹⁾. Estes achados reforçam a importâncias das IDRs/IDPs para a LLPS e fornecem um racional inicial para o entendimento da abundância de desordem no proteoma eucariótico.

As IDRs/IDPs capazes de realizar LLPS são frequentemente ricas em aminoácidos polares não-carregados (glutamina, asparagina, glicina, serina e prolina), aminoácidos polares (arginina, lisina, aspartato e glutamato) e aminoácidos aromáticos (fenilalanina e tirosina). Entretanto, estes aminoácidos não estão randomicamente distribuídos na sequência primária da proteína, mas sim dispostos em pequenos motivos lineares contendo alternância de cargas e repetições degeneradas^(11,32). Esta preferência de sequência sugere a existência de um conjunto de forças interatômicas que promovem a LLPS. Estas forças provavelmente incluem interações eletrostáticas, dipolo-dipolo, pi-pi, cátion-pi, hidrofóbicas e ligações hidrogênio (Figura 2)⁽³³⁻³⁷⁾. Estudos de mutação sítio-dirigida mostraram a importância de cada tipo de resíduo para a LLPS⁽³⁸⁾. Enquanto as regiões ricas em glicina estabilizam o estado líquido dos condensados, regiões ricas em serina e glutamina promovem agregação sólida⁽³⁹⁾.

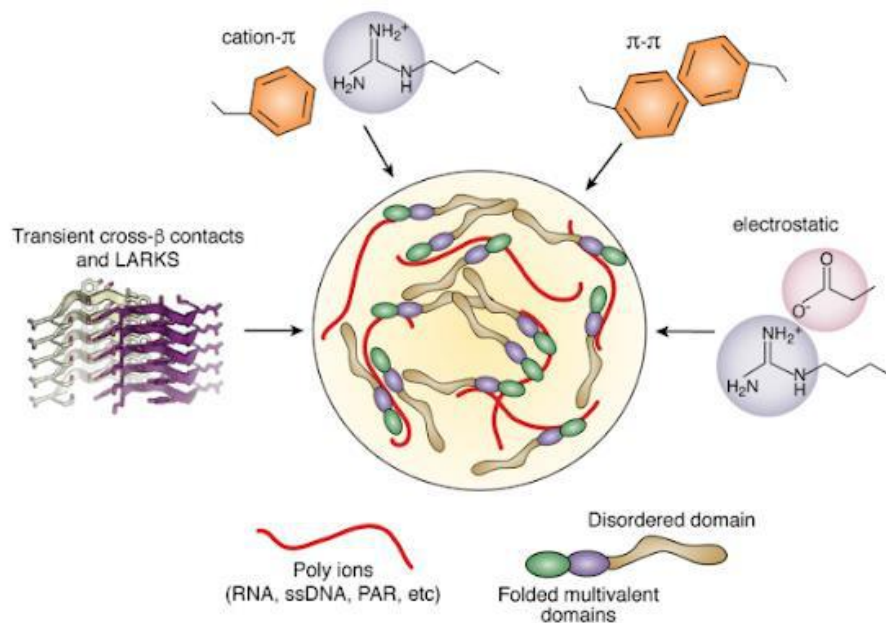


Figura 2 | Forças interatômicas que dirigem a LLPS. As interações importantes para a LLPS incluem pi-pi, cátion-pi, interações eletrostáticas e contatos entre fitas-β cruzadas. (GOMES; SHORTER, 2019).

As MLOs frequentemente contêm ácidos nucleicos, principalmente RNA. Além disso, as proteínas presentes em MLOs possuem domínios ou motivos de ligação a RNA e, de modo geral, o RNA promove a LLPS de diversas RBPs⁽⁴⁰⁻⁴²⁾. Interessantemente, a modulação da LLPS por RNA é dependente da estequiometria de interação, de forma que alta razão RNA/proteína pode inibir a LLPS^(43,44). Ainda, diversos estudos mostraram que o RNA pode afetar a propriedade de material dos condensados proteicos^(30,45). Isto sugere que estes corpos estão funcionalmente relacionados ao metabolismo de RNA e aos mecanismos pós-transcricionais de regulação da expressão gênica⁽³⁹⁾.

1.2. Proteínas ricas em glicina

As proteínas ricas em glicina (GRP – *Glycine-Rich Proteins*) são caracterizadas pelo alto conteúdo em glicina (40% a 70%) e pela presença de motivos de sequência compostos por repetições ricas em glicina, geralmente do tipo (Gly)_n-X, onde X é qualquer aminoácido^(46,47). A GRP-1 de *Petunia sp* foi a primeira proteína rica em glicina descrita na literatura. Essa proteína apresenta 384 aminoácidos em sua sequência, dos

quais 67% são glicinas⁽⁴⁸⁾. A GRP-1 é encontrada associada à parede celular da planta, sugerindo um papel estrutural para estas proteínas⁽⁴⁹⁾. Após a identificação da GRP-1, muitas outras GRPs foram descobertas em outras espécies de plantas e mamíferos. As GRPs formam um grupo de proteínas com grande heterogeneidade estrutural e funcional. A variedade de padrões de expressão e localizações subcelulares sugerem o envolvimento destas proteínas em diversos processos fisiológicos, como estrutura da parede celular, defesa contra infecção, ligantes extracelulares de proteínas cinases, resposta a múltiplos estresses abióticos, controle do tempo de floração e desenvolvimento, dentro outros⁽⁴⁷⁾.

A superfamília das GRPs é dividida em diferentes classes de acordo com a sua organização em domínios⁽⁵⁰⁾. As GRPs da classe I possuem um domínio rico em glicina formado por repetições do tipo $(\text{GlyGlyX})_n$. A maioria das proteínas desta classe encontram-se associadas à parede celular e, portanto, apresentam uma função estrutural⁽⁵¹⁾. As GRPs da classe II possuem um domínio rico em glicina formado por repetições da sequência $(\text{GlyGlyXXXGlyGly})_n$ seguido de uma região carboxi-terminal rica em cisteína. As GRPs da classe III são caracterizadas pelo baixo conteúdo em glicina, quando comparadas as demais classes, e pela presença de repetições do tipo (GlyXGlyX) . Frequentemente, estas proteínas apresentam um domínio de oleosina conservado. Essas três primeiras classes apresentam um peptídeo sinal na região amino-terminal de suas sequências. As GRPs da classe IV possuem um domínio de interação com ácido nucleico e são conhecidas por ligarem a RNA. Esta classe apresenta a maior diversidade de domínios enovelados conectados aos motivos ricos em glicina, levando a divisão em quatro subclasses (IVa a IVd). A subclasse IVa é caracterizada por apresentar um motivo de reconhecimento de RNA (RRM – *RNA Recognition Motif*). A classe IVb, além do RRM, possui um domínio *zinc finger* do tipo CCHC. A classe IVc contém um domínio *cold shock* (CSD – *Cold Shock Domain*) e dois ou mais domínios *zinc fingers*. A classe IVd apresenta dois domínios RRM. Já a classe V engloba as GRPs com padrões mistos de repetições de glicina^(47,50,52).

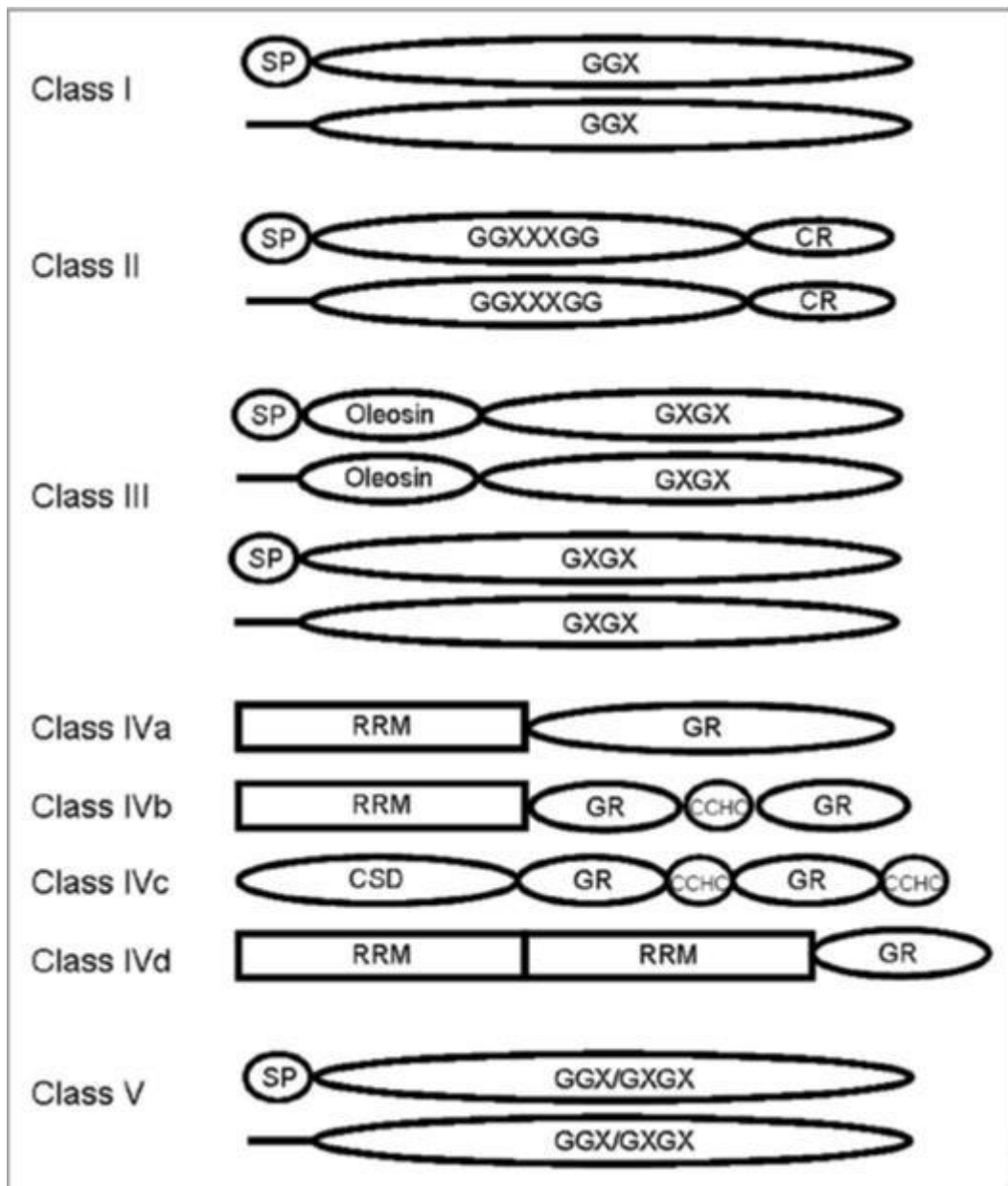


Figura 3 | Diferentes classes de proteínas ricas em glicina. A superfamília das proteínas ricas em glicina se divide em cinco classes, de I a V. A classe IV se divide em quatro subclasses de IVa a IVd. (MANGEON *et al.*, 2010).

A classe IV de GRPs se destaca por agrupar as proteínas ligantes de RNA (RBP – *RNA binding proteins*), que são conhecidas por participarem dos mecanismos pós-transcricionais de regulação da expressão gênica^(47,50,52). Os principais domínios enovelados encontrados nas GRPs de classe IV são RRM e CSD. O domínio RRM constitui o domínio ligante de RNA mais amplamente distribuído na natureza e é frequentemente encontrado em diversas RBPs, como proteínas ligantes de cauda poliA, proteínas associadas a snRNPs (*small nuclear ribonucleoproteins*), fatores de *splicing*,

etc⁽⁵³⁾. Este domínio apresenta uma assinatura molecular característica da interação com RNA, formada por dois motivos de sequência conservados: RNP1, cuja sequência consenso é (K/R)G(F/Y)(G/A)FVX(F/Y) e RNP2, cuja sequência consenso é (L/I)(F/Y)(V/I)(G/K)(G/N)L⁽⁵³⁾. Já o domínio CSD é conhecido pela habilidade de ligar-se tanto a RNA quanto a DNA. Este nome deriva da família de proteínas *cold shock* (CSPs), pequenas proteínas procarióticas cuja expressão é induzida pelo frio⁽⁵⁴⁾. As CSPs possuem um único domínio CSD e contribuem para a adaptação ao frio por atuarem como chaperonas de RNA⁽⁵⁵⁾. Em plantas, os domínios CSD estão presentes em diversas proteínas, frequentemente associados a regiões ricas em glicina⁽⁵⁶⁾.

Apesar das funções dos domínios enovelados presentes nas GRPs de classe IV serem bem entendidas, o papel das regiões ricas em glicina ainda permanece desconhecido. As estruturas dos motivos de sequência contendo repetições de glicina ainda são de difícil entendimento e, portanto, são tratadas como alças flexíveis⁽⁵⁷⁾. Algumas dessas alças interagem entre si através de interações fracas apolares dando origem a estruturas dinâmicas que se assemelham a um velcro (Figura 4A)⁽⁵⁷⁾. Outra proposta é a formação de folhas- β antiparalelas (Figura 4B) como uma potencial estrutura secundária formada pelas alças de glicina presentes nas GRPs⁽⁴⁷⁾. Nesta estrutura, a cadeia lateral de outros aminoácidos, que não glicinas, estariam voltadas para a mesma superfície, formando um domínio hidrofóbico capaz de interagir com outras proteínas ou mesmo membranas celulares, nas quais as GRPs já foram encontradas^(58,59). A grande flexibilidade das regiões ricas em glicina facilita a interação das GRPs com diversos parceiros moleculares⁽⁶⁰⁾.

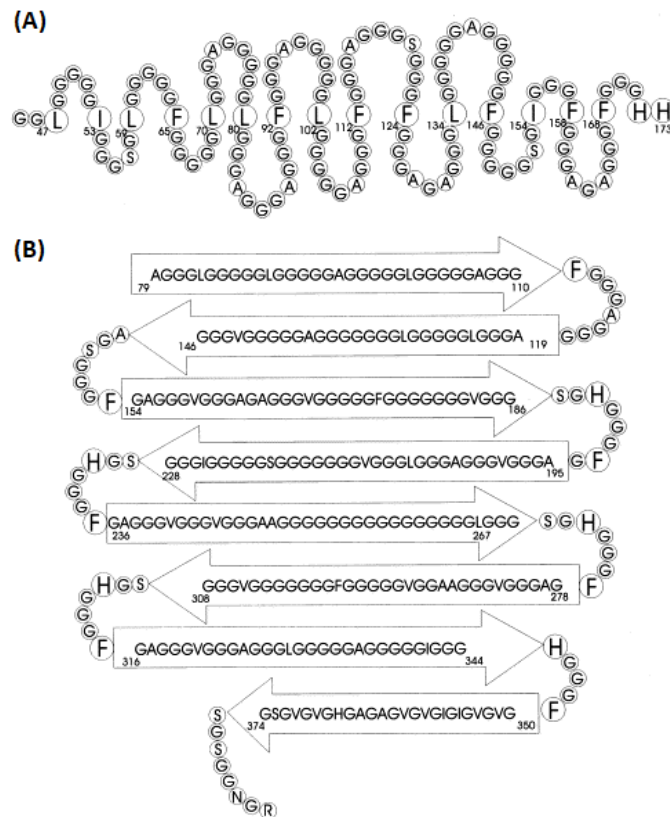


Figura 4 | Esquema das possíveis conformações dos motivos de sequência ricos em glicina. (A) Predição conformacional em velcro dos segmentos ricos em glicina da proteína AtGRP5. **(B)** Organização em folha- β das alças de glicina de PtGRP-1 proposto por Condit *et al.* (1986)⁽⁶¹⁾. (SACHETTO-MARTINS *et al.*, 2000).

Funcionalmente, as GRPs de classe IV estão envolvidas na resposta ao estresse, principalmente estresse abiótico. A expressão destas proteínas aumenta significativamente após a exposição da planta ao frio, tratamento com ácido abscísico (ABA), ácido salicílico, fermento ou estresse hídrico⁽⁶²⁾. Evidências experimentais sugerem um certo grau de conservação funcional entre as proteínas dessa família. MhGRP-RBP1 é uma GRP de classe IVa presente em *Malus hupehensis* (macieira). Esta proteína apresenta algum grau de conservação de sequência com outras GRPs, incluindo a proteína AtGRP7 (80% de identidade) de *Arabidopsis thaliana*⁽⁶³⁾. MhGRP-RBP1 é primariamente expressa em tecido foliar e sua expressão é aumentada após exposição a seca, alta salinidade, estresse oxidativo e fermento. Em contrapartida, quando a planta é tratada com ABA ou ácido jasmônico, a expressão de MhGRP-RB1 é diminuída⁽⁶³⁾. Portanto, MhGRP-RBP1, uma GRP de classe IV, parece auxiliar a macieira, uma importante planta agrícola, a adaptar-se ao estresse ambiental. De forma

análoga, o genoma de arroz codifica seis genes de GRPs de classe IV (OsGRP1 a OsGRP6)⁽⁶⁴⁾. O estresse pelo frio aumenta a expressão destes genes em mais de duas vezes. A expressão de OsGRP1 e OsGRP6 recupera a adaptação ao frio de uma cepa mutante de *Escherichia coli* sensível a temperatura⁽⁶⁴⁾. Ainda, a expressão de OsGRP1 e OsGRP4 é capaz de recuperar a resistência ao frio do mutante *atgrp7* de *Arabidopsis thaliana*, enquanto a expressão de OsGRP6 promove tolerância ao congelamento⁽⁶⁴⁾. Estes resultados confirmam a conservação funcional dos membros da classe IV de GRPs.

1.3. AtGRP2

AtGRP2 (*Arabidopsis thaliana* glycine-rich protein 2), também conhecida como CSDP2 ou AtCSG2⁽⁵⁰⁾, é uma GRP encontrada em *Arabidopsis thaliana*, a primeira planta a ter seu genoma completamente sequenciado⁽⁶⁵⁾. AtGRP2 é uma RBP de 19 kDa pertencente a uma família de proteínas que contêm domínios *cold shock* em *A. thaliana* (AtGRP1 a AtGRP4)⁽⁵⁶⁾. Dentre estas, AtGRP2 é o membro mais abundantemente expresso⁽⁶⁶⁾. Esta proteína apresenta localização núcleo-citoplasmática, porém, no núcleo, encontra-se principalmente associada ao nucléolo^(67,68). AtGRP2 é preferencialmente expressa em tecidos embrionários e que apresentam alta taxa de divisão celular, como meristemas, carpelos, anteras e embriões⁽⁶⁶⁻⁶⁸⁾. A diminuição dos níveis de expressão de AtGRP2 promove floração antecipada, sugerindo que AtGRP2 funciona como um inibidor geral de floração, regulando a passagem da planta do estado vegetativo para o reprodutor⁽⁶⁶⁻⁶⁸⁾. Ainda, plantas silenciadas para AtGRP2 apresentam redução do número de estames e desenvolvimento anormal de embriões e sementes, sugerindo o envolvimento de AtGRP2 em processos de desenvolvimento⁽⁶⁶⁻⁶⁸⁾.

A expressão de AtGRP2 é fortemente induzida pelo frio⁽⁶⁷⁻⁶⁹⁾ e por alta salinidade⁽⁷⁰⁾. Sasaki *et al.* (2007) mostraram que AtGRP2 é capaz de complementar o fenótipo de sensibilidade ao frio exibido pelo mutante quádruplo de *Escherichia coli* BX04 (deletado em proteínas *cold shock* bacterianas)⁽⁶⁹⁾. Ainda, AtGRP2 é capaz de promover a desnaturação de DNA dupla-fita, sugerindo uma possível função como chaperona de RNA⁽⁶⁹⁾. Park *et al.* (2009)⁽⁷⁰⁾ mostraram que a superexpressão de AtGRP2 recupera parcialmente o fenótipo de tolerância ao congelamento de mutantes *atgrp7* de *A. thaliana*. Em contrapartida, o estudo de Sasaki *et al.* (2013)⁽⁷¹⁾ mostrou que o mutante

duplo *atgrp2* e *atgrp4* apresenta maior tolerância ao congelamento que plantas selvagens quando submetido a aclimação ao frio. De maneira análoga, a superexpressão de AtGRP2 leva a diminuição da tolerância ao congelamento após a aclimação ao frio, sugerindo que AtGRP2 funciona como um regulador negativo da tolerância ao congelamento em *A. thaliana*. Interessantemente, os mesmos resultados não foram observados em condições de não aclimação⁽⁷¹⁾. É sabido que fatores de ligação a repetições ricas em citosina (CBFs – *C-repeat binding factors*) conferem tolerância ao congelamento através da ativação da expressão de genes regulados pelo frio (COR – *Cold-regulated genes*)^(72,73). AtGRP2 regula negativamente a expressão de CBFs e genes COR durante a aclimação ao frio⁽⁷¹⁾. Conclui-se, portanto, que a AtGRP2 atua como um inibidor da via de aclimação ao frio dependente de CBFs em *A. thaliana*. Além disso, AtGRP2 atua como um inibidor da germinação de sementes. Linhagens de *Arabidopsis* que superexpressam AtGRP2 apresentam germinação retardada quando comparadas a plantas selvagens⁽⁷⁴⁾. A superexpressão de AtGRP2 reduz a expressão de genes do catabolismo de ABA (CYP707A2) e genes de biossíntese de giberelina. Os níveis de ABA encontram-se significativamente aumentados em plantas que superexpressam AtGRP2⁽⁷⁴⁾. Portanto, é proposto que AtGRP2 atua como um regulador negativo da germinação de sementes através do controle dos níveis hormonais, como ABA e giberelina. Em relação ao estresse salino, Sasaki *et al.* (2015)⁽⁷⁵⁾ mostraram que AtGRP2 possui efeitos negativos sobre a tolerância ao sal em *A. thaliana*. AtGRP2 age como um atenuador da sinalização celular induzida pelo estresse salino. A expressão de genes CBF é significativamente ativada por altas concentrações de sal^(76,77). Como AtGRP2 inibe a expressão de CBFs, é sugerido que essa proteína atenua a via de sinalização induzida por CBFs em condições de estresse salino⁽⁷⁵⁾. A ação de reguladores negativos, como AtGRP2, é essencial para a garantia de uma resposta adequada aos diversos tipos de estresses abióticos, evitando os efeitos deletérios oriundos da superativação das vias de resposta ao estresse durante o desenvolvimento vegetal.

AtGRP2 apresenta um teor de glicina de 40% e uma arquitetura modular formada por três domínios: um domínio *cold shock* (CSD) N-terminal e dois domínios *zinc fingers* C-terminais, do tipo *zinc knuckle* (ZKD) retroviral CCHC (CX₂CX₄HX₄C, onde X é qualquer aminoácido), interligados por regiões ricas em glicina⁽⁶⁷⁾ (Figura 5A). Ambos os domínios, CSD e ZKD, são conhecidos como ligantes de RNA e suas sequências primárias são bem conservadas quando comparadas a de outros domínios do mesmo

tipo. O CSD conserva resíduos hidrofóbicos típicos da interação com RNA, enquanto o ZKD apresenta o motivo CCHC bem conservado referente a interação com o zinco (Figuras 5B e 5C). Essa composição estrutural da AtGRP2 é semelhante à da proteína Lin28, com exceção das repetições de glicina. Lin28 foi primeiramente identificada como um regulador do tempo de desenvolvimento em *Caenorhabditis elegans*⁽⁷⁸⁾. Lin28 é uma RBP evolutivamente conservada formada por um domínio CSD N-terminal e dois domínios *zinc fingers* C-terminais (figura 5A)⁽⁷⁸⁾. Essa proteína é abundantemente expressa em células embrionárias e sua expressão diminui com o aumento da diferenciação celular^(79,80). Recentemente, Lin28, juntamente com outros três fatores de transcrição, foi utilizada na reprogramação de células somáticas em células tronco pluripotentes induzidas⁽⁸¹⁾. Ao nível molecular, Lin28 suprime a maturação dos microRNAs da família *let7*, que se encontram ausentes em células embrionárias, mas se acumulam durante a diferenciação celular⁽⁸²⁾. Portanto, é sugerido que Lin28 desempenhe um papel importante no bloqueio da diferenciação celular. Interessantemente, AtGRP2 apresenta um efeito semelhante, sendo majoritariamente expressa em tecidos com características embrionárias. Portanto, a semelhança de domínios entre as duas proteínas sugere que AtGRP2 possa desempenhar um papel na biogênese de microRNAs em plantas.

Embora a função de AtGRP2 no controle da floração e resposta ao estresse ambiental seja bem estabelecida, os mecanismos moleculares através dos quais essa proteína atua permanecem completamente desconhecidos. Através da sua atividade de chaperona de RNA, AtGRP2 pode regular o processamento e/ou a estabilidade de mRNAs, bem como os níveis de tradução e a biogênese de microRNAs, a fim de atenuar a sinalização celular durante o estresse abiótico. A presença de domínios de ligação a RNA juntamente com IDRs contendo motivos lineares de sequência enriquecidos em glicina fundamenta a hipótese de LLPS para AtGRP2. Portanto, a fim de caracterizar funcionalmente AtGRP2, é necessário investigar se esta proteína é capaz de formar MLOs através do fenômeno de LLPS, bem como caracterizar o estado de material e descrever as forças interatômicas que regem a formação destes biocondensados. A descrição físico-química das interações proteína-proteína e proteína-RNA responsáveis pela formação das MLOs de AtGRP2 é essencial para a elucidação das bases moleculares da tolerância ao estresse em plantas, abrindo caminhos para a produção de genótipos resistentes a múltiplos estresses com relevância agrícola para o país.

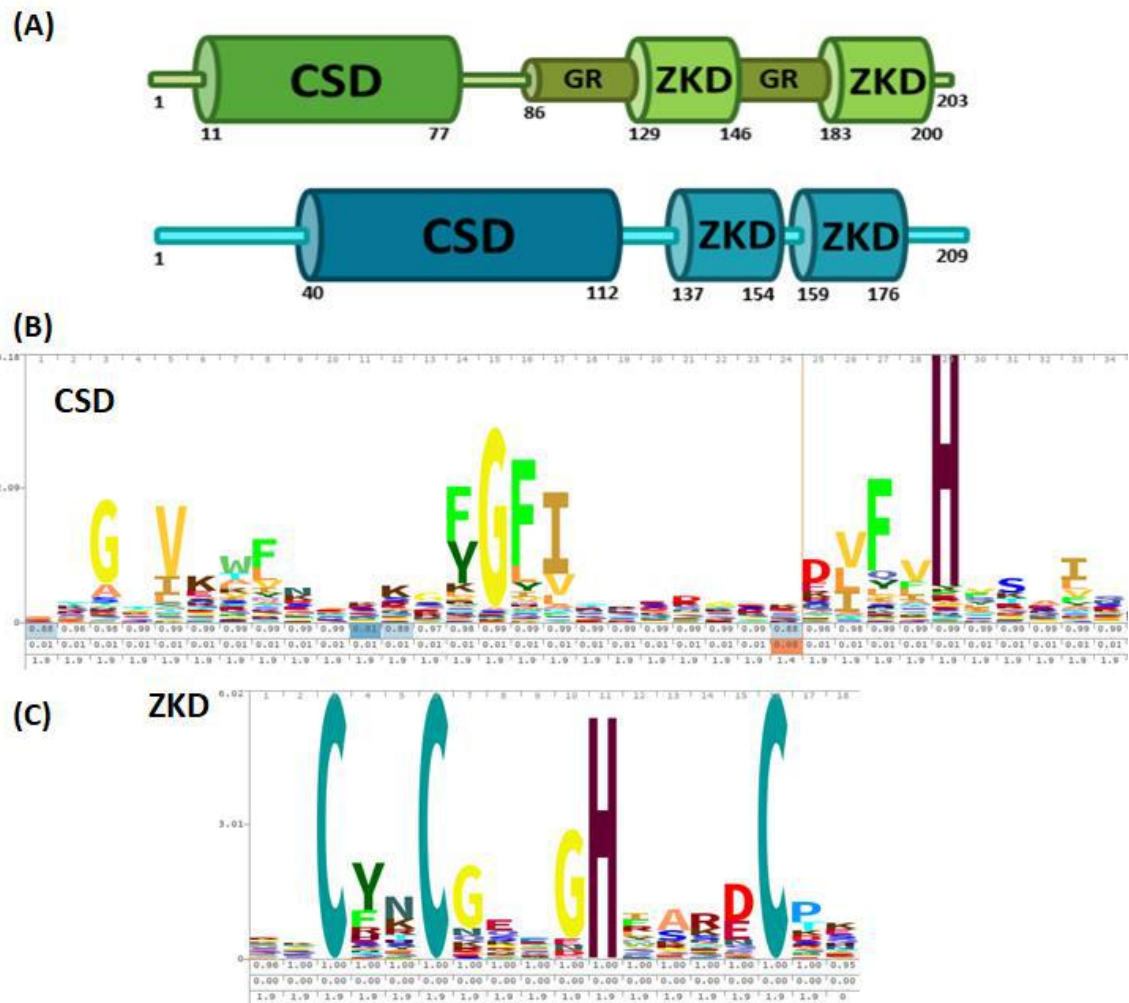


Figura 5 | Arquitetura em domínios de AtGRP2. (A) Representação da organização de domínios da AtGRP2 (verde) inclui um domínio *cold shock* (CSD) N-terminal e dois domínios *zinc-finger* do tipo *zinc-knuckle* (ZKD) C-terminais. Em azul, a semelhança da arquitetura de domínios com a proteína Lin-28. (B,C) Conservação da seqüência primária dos domínios *cold shock* (B) e *zinc finger* (C). O tamanho da letra representa o grau de conservação do resíduo. Diferentes cores representam diferentes resíduos.

2. Objetivos

2.1. Objetivo principal

O presente trabalho tem como objetivo principal estudar a propensão à separação de fases líquido-líquido da proteína AtGRP2 *in vitro*.

2.2. Objetivos específicos

- (i) Purificar e expressar AtGRP2;
- (ii) Analisar a propensão a desordem estrutural e LLPS da sequência de aminoácidos de AtGRP2;
- (iii) Caracterizar o fenômeno de LLPS de AtGRP2;
- (iv) Estudar o efeito da temperatura e da força iônica sobre a formação de biocondensados de AtGRP2;

3. Material e Métodos

3.1. Material

3.1.1. Meios de cultura

As células de bactéria foram cultivadas em meio Luria-Bertani (LB) (Sigma) [triptona 10 g/L, extrato de levedura 5 g/L, NaCl 5 g/L]. O cultivo em meio sólido foi realizado em meio LB acrescido de ágar 5 g/L (Sigma). Todos os meios de cultura foram esterilizados por autoclavação e, posteriormente, suplementados com canamicina 100 µg/mL. Todas as soluções foram preparadas com água ultrapura. As soluções de antibiótico foram esterilizadas por filtração em membrana 0,22 µm.

3.1.2. Vetor de expressão

O vetor pETM30-MBP (promotor T7, origem de replicação *f1*, gene marcador de resistência à canamicina) foi utilizado para expressão da proteína heteróloga. Este vetor fusiona a proteína de interesse a uma cauda amino-terminal composta por seis histidinas seguidas da proteína ligadora de maltose (MBP – *Maltose-binding protein*) e um sítio de clivagem para a protease TEV. O gene que codifica AtGRP2 foi subclonado no referido vetor utilizando as enzimas de restrição 5'-NcoI e 3'-XhoI (Genscript, EUA). O mapa de restrição do vetor está representado na Figura 6. A construção íntegra (contendo a cauda N-terminal de His₆MBP-TEV) possui 605 aminoácidos e massa molecular de 67,5 kDa, enquanto a proteína recombinante, obtida após a clivagem da cauda N-terminal com a protease TEV, possui 205 aminoácidos e massa molecular de 19,3 kDa (Figura 6B).

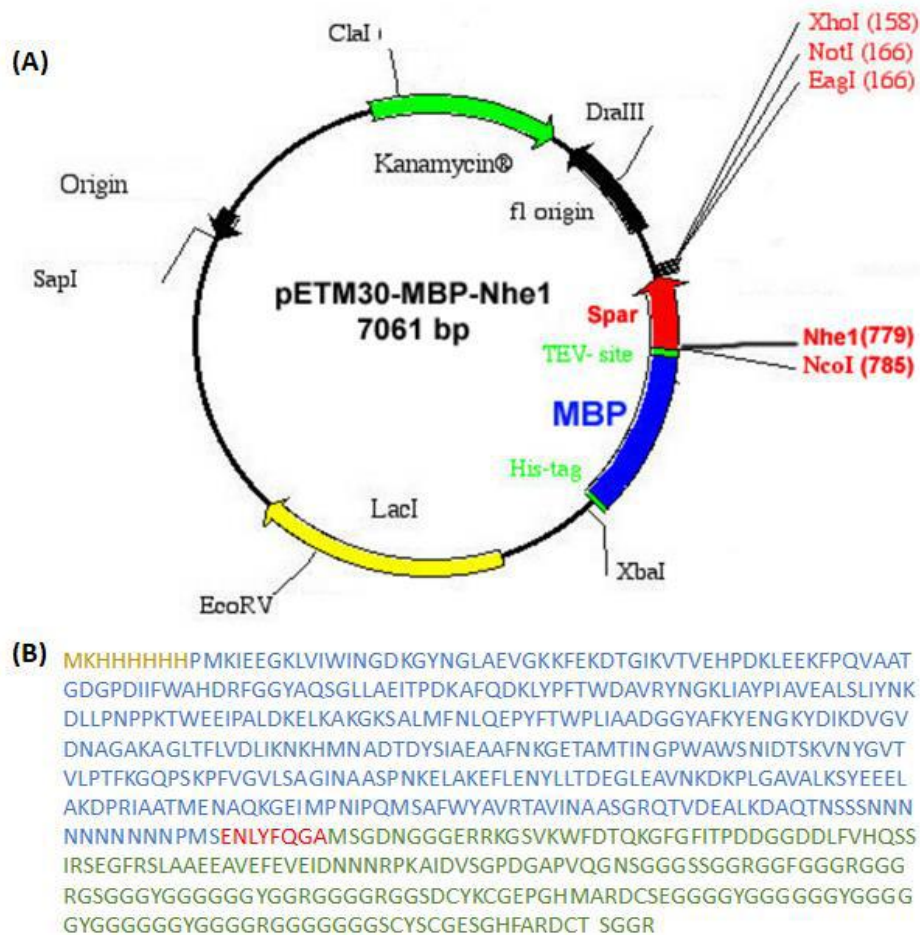


Figura 6 | Representação esquemática do vetor de expressão pETM30-MBP e da construção His₆MBP-TEV-AtGRP2. (A) Mapa de restrição do vetor pETM30-MBP utilizado na clonagem do gene *AtGRP2*, evidenciando o sítio de multiclonagem, o marcador seletivo e a cauda amino-terminal His₆MBP-TEV. **(B)** Sequência de aminoácidos de AtGRP2 fusionada a cauda His₆MBP-TEV. Em amarelo, a cauda de histidinas, em azul, a sequência da proteína MBP, em vermelho, o sítio de clivagem para TEV e, em verde, a sequência referente a AtGRP2. A clivagem enzimática ocorre entre os resíduos de glutamina (Q) e glicina (G) presentes no sítio para TEV (vermelho).

3.1.3. Cepas bacterianas

Duas cepas de bactéria foram utilizadas neste trabalho: *Escherichia coli* DH5α [genótipo: F-φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rk-,mk+) phoA supE44 λ-thi-1 gyrA96 relA1] para armazenamento e propagação dos plasmídeos e *Escherichia coli* BL21 (DE3) [genótipo: F-ompT hsdSB(rB-,mB-) gal dcm (DE3)] para a expressão heteróloga da proteína. As cepas foram gentilmente cedidas pela professora Bianca Cruz Neves do Departamento de Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro.

3.2. Métodos

3.2.1. Transformação bacteriana

Células de *E. coli* BL21 (DE3) quimicamente competentes foram preparadas através do método descrito por Hanahan *et al.* (1983)⁽⁸³⁾ e armazenadas a -80°C. A transformação foi feita adicionando 2 µL de DNA plasmidial a 200 µL de célula competente seguido de procedimento de choque térmico: 30 min em banho de gelo, 45 s a 42°C em banho maria e novamente banho de gelo por 2 min. Para a recuperação das células foram adicionados 800 µL de meio LB líquido e a cultura foi incubada a 37°C por 1 h a 200 rpm. Após esse tempo, as células foram centrifugadas a 8077 *g*. Um volume de 950 µL do sobrenadante foi descartado e o precipitado de células foi ressuscitado nos 50 µL restantes. As células foram plaqueadas em meio LB sólido contendo canamicina 100 µg/mL e incubadas a 37°C durante 16 h até formarem colônias macroscópicas.

3.2.2. Expressão de His₆MBP-AtGRP2

Uma única colônia de células transformadas com o plasmídeo recombinante pETM30-MBP-AtGRP2 foi inoculada em 10 mL de meio LB contendo canamicina 100 µg/mL e incubada a 37°C por 6 h a 200 rpm. Após esse período, 100 µL da cultura foram inoculados em 20 mL de meio LB contendo canamicina 100 µg/mL e o pré-inóculo foi incubado a 37°C por 16 h a 200 rpm. Em seguida, 10 mL dessa cultura foram inoculados em 1L de meio LB contendo canamicina 100 µg/mL e o inóculo foi incubado a 37°C e 200 rpm até que atingisse a densidade óptica a 600 nm (D.O.₆₀₀) de aproximadamente 0,6. Nesse momento, a expressão da proteína foi induzida pela adição de isopropil-β-D-tiogalactopiranosídeo (IPTG) 1 mM e a cultura foi incubada a 18°C por 30 h a 200 rpm. Ao final do cultivo, as células foram coletadas por centrifugação a 8077 *g* por 20 min a 4°C. O precipitado de células foi armazenado em ultracongelador a -80°C e a expressão da proteína de interesse foi analisada por SDS-PAGE.

3.2.3. Lise bacteriana

O precipitado bacteriano foi descongelado em banho de gelo e ressuspenso em tampão de lise [Tris-HCL 50 mM (pH 8,0), NaCl 500 mM, imidazol 5 mM]. A lise celular foi realizada por sonicação (potência de 100 W e 20% de amplitude) em 20 ciclos de 30 s com 60 s de intervalo. O lisado bacteriano foi clarificado por centrifugação a 8077 g por 40 min a 4°C.

3.2.4. Purificação de AtGRP2

A purificação de AtGRP2 se deu em três etapas distintas: (i) cromatografia de afinidade a níquel; (ii) clivagem da cauda amino-terminal pela protease TEV seguida de cromatografia de afinidade a níquel; (iii) cromatografia por exclusão molecular.

3.2.4.1. Primeira cromatografia de afinidade a níquel

A primeira etapa de purificação de AtGRP2 consistiu em uma cromatografia de afinidade a níquel. Para esse fim, foi utilizada uma coluna HisTrap HP (GE Healthcare) acoplada ao sistema de purificação FPLC Äkta Prime plus (GE Healthcare). Após a lise celular, o sobrenadante contendo a proteína de interesse foi filtrado em membrana de poro 0,22 µm e aplicado na coluna cromatográfica, previamente equilibrada em tampão A [Tris-HCl 50 mM (pH 8,0), NaCl 500 mM, imidazol 5 mM e ditrioteitol (DTT) 5 mM]. A amostra foi aplicada na coluna em um fluxo contínuo de 1,0 mL/min. A lavagem foi feita com o mesmo tampão A, sob fluxo contínuo de 5,0 mL/min, até que o traçado de absorbância a 280 nm atingisse os valores da linha base. Em seguida, foi feita uma lavagem com NaCl 2 M, a fim de remover potenciais ligantes de ácido nucleicos da amostra. A coluna foi então submetida a um gradiente linear de 0% a 100% de tampão B [Tris-HCL 50mM (pH 8,0), NaCl 500 mM, imidazol 500 mM] num fluxo contínuo de 3,0 mL/min. A coluna foi lavada com imidazol 2 M para garantir a eluição de todas as proteínas aderidas à resina. Cada fração foi recolhida em tubos cônicos de 50 mL. Alíquotas de 15 µL foram retiradas ao longo do processo de purificação para posterior análise do perfil proteico por SDS-PAGE.

3.2.4.2. Diálise e clivagem com TEV

As frações que continham a proteína His₆MBP-AtGRP2, identificadas por SDS-PAGE, foram unidas, dosadas e incubadas com a protease His₆-TEV N1a (S219V) para clivagem da cauda amino-terminal His₆MBP. A incubação se deu na proporção molar de 5:1 (AtGRP2:TEV) durante 24 h a 4°C enquanto a amostra foi dialisada para a remoção do imidazol residual da primeira etapa de purificação. A diálise foi realizada contra tampão C [Tris-HCl 50 mM (pH 8,0), NaCl 500 mM] em membrana SnakeSkin (3000 Da; Thermo Scientific). O eficiência da clivagem enzimática foi analisada por SDS-PAGE.

3.2.4.3. Segunda cromatografia de afinidade a níquel

Essa etapa se assemelha à descrita no tópico 3.2.5.1, mimetizando todas as condições experimentais e o sistema de purificação, com a diferença que a proteína de interesse foi eluída na lavagem, enquanto a cauda amino-terminal His₆MBP e a protease His₆-TEV, bem como os demais contaminantes oriundos da primeira etapa cromatográfica, permaneceram aderidas à resina de níquel. Estas proteínas foram eluídas posteriormente com tampão B. As frações recolhidas durante as diversas etapas de purificação foram submetidas a SDS-PAGE para identificação de seus perfis proteicos. As frações que continham AtGRP2 foram unidas e concentradas em sistema de ultrafiltração CENTRICON (3000 Da; Millipore) até um volume aproximado de 5 mL.

3.2.4.4. Cromatografia por exclusão molecular

A terceira e última etapa de purificação consistiu em uma cromatografia por exclusão molecular realizada em coluna Superdex 75 (HiLoad 16/60; GE Healthcare) acoplada ao sistema de purificação FPLC ÄKTAPrime Plus (GE Healthcare). O sistema foi equilibrado em tampão D [Tris-HCl 20 mM (pH 7,4), NaCl 100 mM] e a amostra foi injetada em fluxo contínuo de 1 mL/min. O perfil de eluição foi acompanhado por absorbância a 280 nm. Frações de 1 mL foram coletadas ao longo da corrida cromatográfica e posteriormente avaliadas por SDS-PAGE. As frações contendo a proteína de interesse foram unidas e concentradas em sistema de ultrafiltração CENTRICON (3000 Da; Millipore) até um volume aproximado de 3 mL, dosadas, aliquotadas, congeladas em nitrogênio líquido e armazenadas em ultracongelador a -80°C até o momento da análise.

3.2.5. Eletroforese desnaturante em gel de poliacrilamida (SDS-PAGE)

A técnica de SDS-PAGE foi realizada segundo Laemmli (1970)⁽⁸⁴⁾ utilizando o Mini-PROTEAN Tetra System 20 (BioRad). O gel concentrador foi preparado a partir de uma mistura de acrilamida 30% (acrilamida 29%, bis-acrilamida 12%), 1,0 M Tris (pH 6,8), dodecil sulfato de sódio (SDS) 0,1%, persulfato de amônio (PSA) 0,1% e TEMED 0,01%, enquanto o gel fracionador foi preparado a partir uma mistura de acrilamida 30% (acrilamida 29%, bis-acrilamida 12%), 1,5 M Tris (pH 8,8), dodecil sulfato de sódio (SDS) 0,1%, persulfato de amônio (PSA) 0,1% e TEMED 0,01%. Foram aplicados 15 µL de amostra adicionados de 5 µL de tampão de amostra [Tris-HCl 12,5 mM (pH 6,8), β-mercaptoetanol 25 mM, SDS 0,5 %, azul de bromofenol 0,25%, Glicerol 2,5%] em cada poço. A corrida foi feita a 25 mA por aproximadamente 1 h e 30 min em tampão de corrida [Tris-HCl 25 mM (pH 8,3), Glicina 250 mM, SDS 0,1 %]. Os géis foram corridos a 30 mA ao longo de 1h e 30 minutos em tampão de corrida [Tris-HCl 25 mM (pH 8,3), Glicina 250 mM e SDS 0,1% (p/v)]. Após a separação eletroforética, as proteínas foram visualizadas por coloração com Comassie Blue R.

3.2.6. Dosagem de proteínas

A concentração das amostras de proteína foi determinada através de ensaio colorimétrico utilizando o reagente *Pierce 660nm Protein Assay* (Thermo Scientific). Em uma placa de 96 poços, foram adicionados 10 µL das amostras e 150 µL do reagente de Pierce. A mistura reacional foi incubada por 5 min ao abrigo da luz. Em seguida, as absorvâncias foram determinadas a 660 nm em espectrofotômetro SpectraMax M2e (Thermo Fisher). Cada ensaio foi realizado em triplicata. As concentrações das amostras de proteína foram determinadas com base em uma curva padrão realizada com albumina bovina sérica (BSA) na faixa de 0,025 a 2 mg/mL.

3.2.7. Microscopia de contraste de interferência diferencial (DIC - *differential interference contrast*) e de fluorescência

A propensão à LLPS de AtGRP2 foi estudada por microscopia DIC e de fluorescência. As micrografias foram adquiridas em câmaras construídas a partir da adesão de uma lamínula de vidro à lâmina de microscopia utilizando fita dupla-face, de

tal forma que houvesse um espaço fixo entre a lâmina e a lamínula para a aplicação da amostra líquida. Foi utilizado um volume de 10 μL por câmara e as amostras foram incubadas por 10 min a temperatura ambiente antes da aquisição das imagens. As imagens de microscopia foram obtidas com 70 μM de AtGRP2 em tampão E [Tris-HCl 20 mM (pH 7,4), NaCl 100 mM, ZnSO_4 150 μM]. A fim de investigar o efeito de agentes de *crowding* sobre a LLPS de AtGRP2, as imagens foram adquiridas na presença de PEG-4000 na concentração final de 10% (p/v) (Thermo Scientific, EUA). A fim de investigar as propriedades físico-químicas dos condensados de AtGRP2, as imagens foram adquiridas em diferentes temperaturas (ambiente ou após incubação a 37°C) ou em alta força iônica (na presença de NaCl 1 M). Ainda, a fim de determinar a natureza proteica dos condensados de AtGRP2, as imagens foram adquiridas na presença de 1,6-hexanodiol (Merck, Alemanha) nas concentrações finais de 1% e 10% (p/v). As micrografias foram obtidas sob iluminação de Köhler num microscópio confocal invertido (Leica TCS SPE, Leica Microsystems, Alemanha) com objetiva de 63x de aumento sob imersão em óleo. Foi utilizado um *laser* de argônio a 488 nm para os ensaios de DIC. Ainda, a natureza proteica dos condensados de AtGRP2 foi confirmada por microscopia de fluorescência utilizando o fluoróforo SYPRO Orange (Sigma-Aldrich, EUA) que interage com resíduos hidrofóbicos em proteínas. O SYPRO orange foi diluído na proporção de 1:100 em amostras de 70 μM de AtGRP2. As amostras foram incubadas por 1 h e analisadas por microscopia de fluorescência, com excitação a 488 nm e emissão entre 450 e 500 nm. Todas as imagens foram coletadas a temperatura ambiente da sala do microscópio ($\sim 20^\circ\text{C}$) a não ser que seja especificada a condição do ensaio. Imagens foram processadas pelo programa Fiji 4.0. Para a quantificação de cada condição experimental, foi escolhida uma área central de 0,1 mm^2 para cada amostra. A contagem foi feita de forma manual e a estimativa de tamanho foi feita a partir da barra de escala produzida pelo programa.

4. Resultados e Discussão

4.1. Purificação de AtGRP2

A purificação de AtGRP2 foi feita a partir da combinação de duas técnicas cromatográficas. A primeira etapa constituiu em uma cromatografia de afinidade a metal imobilizado (níquel). Esta etapa se baseia na coordenação do níquel pelo anel imidazólico da cauda de polihistidinas presente na construção His₆MBP-AtGRP2. Uma vez aderida à resina, a proteína de interesse foi separada dos demais contaminantes que foram diretamente lavados. A eluição de His₆MBP-AtGRP2 foi feita com um gradiente de imidazol (5 a 500 mM), que age como um agente competidor, doando elétrons e formando ligações de coordenação com o metal de transição imobilizado⁽⁸⁵⁾. A Figura 7A mostra o cromatograma de eluição, através do qual se observa um pico na concentração de ~345 mM de imidazol. A composição proteica das amostras coletadas ao longo desta etapa foi analisada por SDS-PAGE (Figura 7B). O gel de eletroforese evidenciou a presença de uma banda intensa que corre em ~64 kDa, correspondente a massa molecular da construção His₆MBP-AtGRP2 (67,5 kDa), na etapa de eluição.

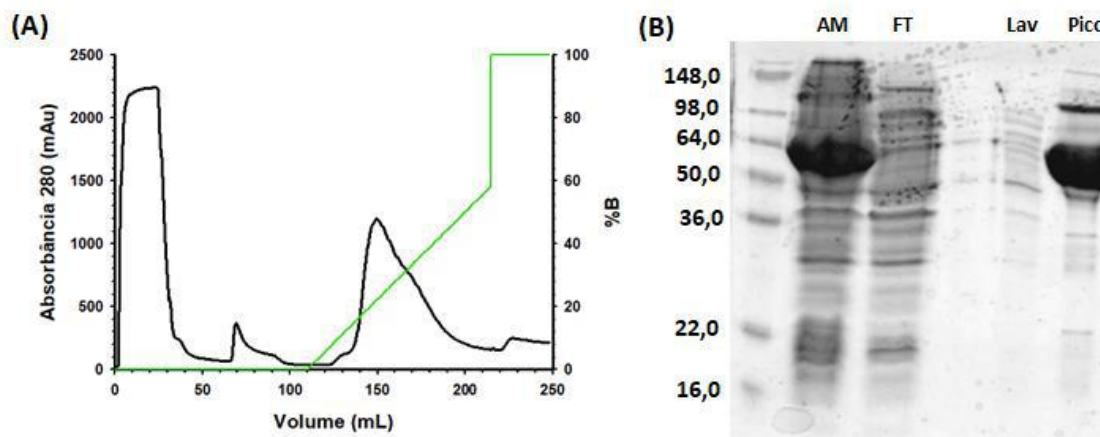


Figura 7 | Primeira etapa de purificação de His₆MBP-AtGRP2. (A) Cromatograma de eluição da construção His₆MBP-AtGRP2. O traçado de absorbância está representado em vermelho, enquanto o gradiente de imidazol está representado em verde. O pico de eluição está marcado com um traço preto (imidazol ~345 mM). (B) Análise por SDS-PAGE das amostras coletadas em (A). Am, amostra injetada na coluna; FT, flow-through; Lav, lavagem com tampão A; Pico, pico de eluição com gradiente de imidazol.

O próximo passo no processo de purificação consistiu em submeter a amostra de His₆MBP-AtGRP2 à clivagem com a protease TEV⁽⁸⁶⁾. Esta protease é muito utilizada para a remoção de caudas de purificação de proteínas recombinantes por ser extremamente específica, reconhecendo a sequência heptapeptídica (EXXYXQ*S/G) (X pode ser qualquer resíduo)⁽⁸⁷⁾, além de ser ativa em baixas temperaturas e na presença de vários inibidores normalmente adicionados durante o processo de purificação. A incubação se deu a 4°C por um período de 24 h em razão molar 1:5 (TEV: His₆MBP-AtGRP2). Após a incubação com TEV, observa-se o aparecimento de 3 outras bandas eletroforéticas: (i) uma banda em ~45 kDa, referente à cauda His₆MBP; (ii) uma banda em ~28 kDa referente à protease His₆-TEV; (iii) uma banda em ~20 kDa (corre entre os padrões de 18,4 kDa e 25 kDa) referente à AtGRP2.

Outra etapa de cromatografia de afinidade a níquel foi usada como segunda etapa de purificação, com o objetivo de separar a proteína de interesse da cauda His₆MBP e da protease His₆-TEV. Como His₆MBP e His₆-TEV possuem uma cauda de polihistidinas, estas interagem com o níquel imobilizado na resina cromatográfica. Assim, enquanto a proteína de interesse foi eluída diretamente no lavado, His₆MBP e His₆-TEV ficaram retidas e foram eluídas posteriormente pela adição de imidazol (Figura 8B).

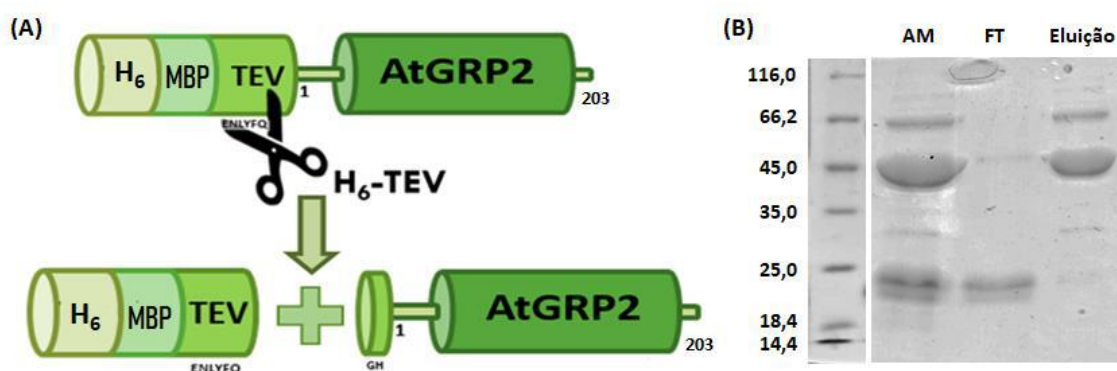


Figura 8 | Clivagem enzimática da construção His₆MBP-AtGRP2. (A) Representação esquemática da clivagem de His₆MBP-AtGRP2 com a protease TEV, evidenciando os dois resíduos de glicina e alanina adicionais que permanecem na sequência da proteína recombinante (artifícios de clonagem). (B) Análise por SDS-PAGE das amostras coletadas durante a segunda etapa de cromatografia de afinidade a níquel. Am, amostra injetada na coluna; FT, *flow-through*; Eluição, eluição com imidazol 500 mM.

A amostra de AtGRP2 foi então submetida a última etapa de purificação por cromatografia de exclusão molecular. A Figura 9A mostra o cromatograma de eluição da coluna de filtração em gel, onde se observa um pico de eluição majoritário em um volume de 84 mL. A análise por SDS-PAGE revelou a presença de uma banda proteica de ~20 kDa (corre entre os padrões de 18,4 kDa e 25 kDa), correspondente à AtGRP2 (Figura 9B). O cromatograma mostra ainda outros picos de eluição referentes a agregados proteicos de AtGRP2. A curva de calibração da coluna de filtração em gel, feita com proteínas globulares, revelou que o volume de eluição de 84 mL corresponde a uma proteína globular de massa molecular de ~10 kDa (Figura 10B). Curiosamente, essa massa é menor que a esperada para AtGRP2 (19 kDa), sugerindo que: (i) AtGRP2 é monomérica e não adota outros estados oligoméricos de maior ordem; (ii) AtGRP2 não é globular; (iii) AtGRP2 é uma proteína compacta.

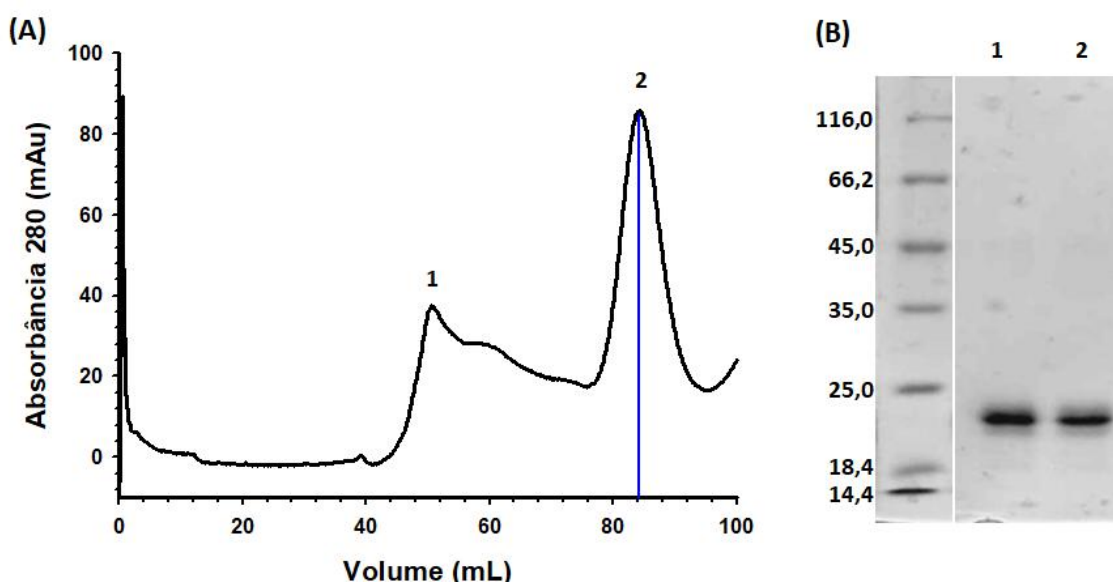


Figura 9 | Purificação de AtGRP2 por cromatografia de exclusão molecular. (A) Cromatograma de eluição de AtGRP2 da coluna de filtração em gel, evidenciando o pico em ~84 mL. (B) O gel de SDS-PAGE é referente a terceira etapa de purificação da proteína. O pico da amostra (AM) injetada na coluna de SEC coincide com a o pico 1 marcado em azul na imagem A e ambos coincidem com o peso esperado da proteína de ~20 kDa, estando entre os padrões de 18,4 e 25,0 kDa.

A fim de comparar as propriedades hidrodinâmicas de AtGRP2 íntegra com as de seu domínio CSD isolado, os cromatogramas de eluição da coluna de exclusão molecular de ambas as proteínas foram sobrepostos (Figura 10). Curiosamente, o domínio CSD isolado elui em ~81 mL, enquanto a proteína íntegra elui em ~84 mL. Considerando a

propriedade da cromatografia de exclusão molecular de separar proteínas em um gradiente decrescente de volume molecular, AtGRP2 íntegra parece ser menor que seu domínio isolado. Isso se dá por dois motivos: (i) o domínio CSD livre apresenta equilíbrio entre dois estados conformacionais, um enovelado e outro desenovelado (dados não mostrados) e, portanto, elui em volume maior que o esperado para o seu tamanho (~8 kDa); (ii) AtGRP2 é uma proteína não globular e altamente compacta.

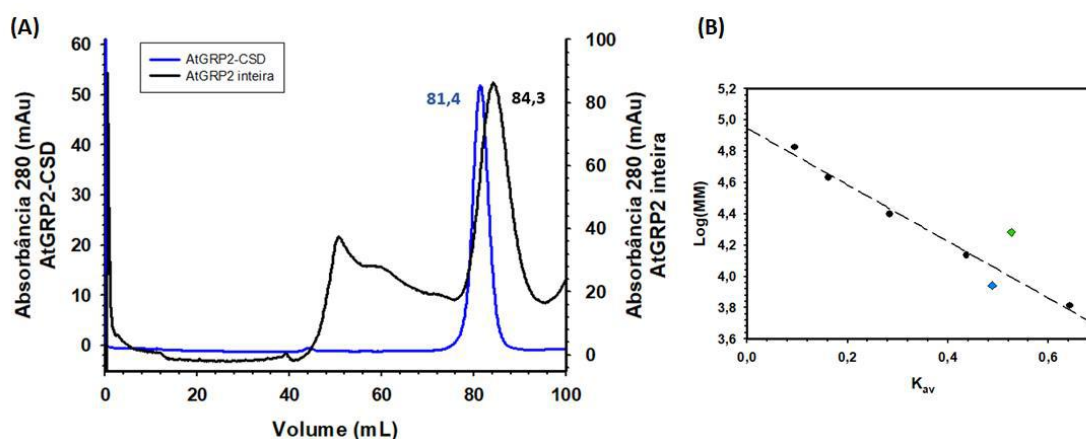


Figura 10 | AtGRP2 é uma proteína não globular e altamente compacta. (A) Cromatograma de exclusão molecular de AtGRP2 (preto) e AtGRP2-CSD (azul). AtGRP2 íntegra elui em ~84 mL, enquanto AtGRP2-CSD elui em ~81 mL. **(B)** Curva de calibração da coluna de filtração em gel (*inset*). As proteínas usadas para calibração foram: Aprotina (6,5 kDa), Ribonuclease A (13,7 kDa), Anidrase Carbônica (29,0 kDa), ovoalbumina (44,0 kDa) e Conalbumina (75,0 kDa). AtGRP2-CSD (azul) e AtGRP2 íntegra (verde) não se encontram sob a linha do ajuste linear da curva.

4.2. Propensão à desordem e separação de fases de AtGRP2

De modo geral, o fenômeno de LLPS é dirigido por proteínas intrinsecamente desordenadas (IDPs) ou que contêm regiões intrinsecamente desordenadas (IDRs), ou seja, que não possuem uma estrutura tridimensional definida⁽⁸⁸⁾. Além disso, IDPs/IDRs normalmente apresentam regiões de baixa complexidade, caracterizadas por segmentos de sequência enriquecidos em repetições de aminoácidos⁽¹¹⁾. AtGRP2 possui uma cauda C-terminal rica em glicina, portanto de baixa complexidade sequencial, o que a torna uma boa candidata para sofrer LLPS.

Assim, a sequência primária de AtGRP2 foi analisada a partir de diferentes algoritmos de bioinformática, a fim de prever o nível de desordem estrutural e de propensão à LLPS da proteína (Figura 11). A função de distribuição acumulativa (CDF - *Cumulative Distribution Function*) é um preditor de desordem binária que classifica proteínas como completamente ordenadas ou desordenadas, caso a maioria dos pontos do ajuste se localize acima ou abaixo da fronteira representada pela linha preta⁽⁸⁹⁾. Como mostra a predição, a curva de CDF para AtGRP2 se localiza completamente abaixo da fronteira, sugerindo alto grau de desordem.

Contatos pi-pi, assim como cátion-pi, são uma das principais forças interatômicas responsáveis pela promoção da LLPS. Estas interações ocorrem entre resíduos de aminoácidos que contêm orbitais pi livres e são comumente frequentes em proteínas que separam de fase e não possuem estrutura definida⁽³⁶⁾. Embora a ocorrência de contatos pi-pi seja normalmente associada a anéis aromáticos, orbitais pi também podem ser encontrados em átomos ligados que apresentam hibridização sp², como grupos amida da cadeia principal, bem como grupos amida, carboxila e guanidino de cadeias laterais. Portanto, aminoácidos como Tyr, Phe, Trp, His, Gln, Asn, Glu, Asp e Arg fazem interações pi-pi através de suas cadeias laterais. Por outro lado, aminoácidos como Gly, Ser, Thr e Pro são pequenos o bastante para terem suas ligações peptídicas expostas o suficiente para realizarem contatos pi-pi. Notavelmente, IDRs de baixa complexidade implicadas na separação de fases de FUS, EWS, hnRNPA1, TIA-1, TDP-43 e do domínio C-terminal da RNA polimerase II⁽⁹⁰⁻⁹²⁾ são enriquecidas nesses resíduos com alto potencial de formação de interações pi-pi em relação à ocorrência média no proteoma⁽³⁶⁾. A análise de PScore considera a relação entre os contatos pi-pi e a sequência primária da proteína, resultando numa pontuação referente ao maior valor da curva. AtGRP2 apresenta uma pontuação PScore de 9,07 e seus maiores valores, ou seja, os segmentos da proteína com maior propensão a formarem contatos pi-pi, coincidem com a cauda C-terminal rica em glicina.

A carga líquida por resíduo e a hidrofobicidade são outras duas formas de analisar desordem a partir da sequência primária da proteína. Quanto maior o conteúdo de cargas positivas e negativas intercaladas na sequência, maior a tendência à desordem⁽⁹³⁾. De maneira análoga, quanto menor o grau de hidrofobicidade, menor a tendência da proteína adotar uma conformação globular, portanto, maior o grau de desordem⁽⁹⁴⁾. AtGRP2 apresenta alternância de carga por resíduo, mostrando ser um polianfólio fraco, o que concorda com as características de IDP. Além disso, AtGRP2 apresenta baixa

hidropatia como um todo (hidropatia média de $\sim 0,5$), sobretudo para a cauda C-terminal. O índice de hidropatia é calculado a partir da normalização da hidrofobicidade de Kyte-Doolittle e varia de 0 a 1, onde 0 é completamente hidrofílico e 1, completamente hidrofóbico⁽³³⁾. A baixa hidropatia corrobora a tendência a desordem mostrada para AtGRP2, principalmente para a cauda C-terminal rica em glicina.

Outra possibilidade de analisar a desordem estrutural é por meio da comparação entre a carga e a hidropatia, resultando em um gráfico de duas dimensões⁽⁹⁵⁾. De modo geral, proteínas desordenadas são mais ricas em aminoácidos carregados e pobres em aminoácidos hidrofóbicos que proteínas enoveladas⁽⁹⁵⁾. Esse preditor de desordem calculado pelo PONDR® classifica o nível de desordem da proteína de acordo com a posição que ela corresponde no diagrama. Caso esteja acima da fronteira, representada pela linha preta, a proteína é predita como desordenada. Caso se localize abaixo da fronteira, a proteína é predita como ordenada⁽⁹⁶⁾. AtGRP2 se localiza sobre a fronteira, sugerindo um comportamento dúbio em relação a desordem, ou seja, não se comporta nem como uma proteína totalmente enovelada nem como uma proteína totalmente desenovelada. Logo, é possível que a proteína possua regiões estruturadas intercaladas com regiões desordenadas.

Com o auxílio dos programas PONDR® e IUPred, avaliou-se a tendência à desordem da sequência de AtGRP2 por resíduo de aminoácido. Curiosamente, somente o preditor PONDR® VLXT sugere ordenação completa para o domínio CSD de AtGRP2 (resíduos 1 a 80). Os demais preditores mostram um grau de desordem em torno de 0,5, ou seja, alguma propensão a desordem mesmo para essa região que coincide com um domínio enovelado. A predição estrutural concorda com dados do nosso grupo de pesquisa que mostra que o domínio CSD isolado apresenta equilíbrio conformacional entre dois estados, um enovelado e outro desenovelado (dados não mostrados). Em contrapartida, para a cauda C-terminal rica em glicina, os preditores da família PONDR® sugerem alto grau de desordem, com ordenação local somente ao redor dos domínios *zinc fingers* (~ 15 aminoácidos cada), enquanto os preditores da família IUPred sugerem alguma ordenação para essa região. Esses dados suportam novamente a hipótese de comportamento dúbio em relação a desordem e da existência de regiões ordenadas intercaladas com regiões desordenadas em AtGRP2.

As predições de bioinformática sugerem a existência de regiões intrinsecamente desordenadas em AtGRP2, principalmente localizadas na região C-terminal rica em glicina. Apesar disso, a cromatografia de exclusão molecular mostrou que AtGRP2

íntegra elui como uma proteína compacta, cujo volume molecular é menor do que o ocupado pelo seu domínio CSD isolado. Isso é contrário ao que se espera para IDRs/IDPs, que são entendidas como um conjunto de múltiplas conformações flexíveis que alcançam um raio hidrodinâmico maior do que o de uma proteína enovelada de mesma massa⁽⁹⁵⁾. Portanto, esse resultado sugere uma possível interação entre a cauda C-terminal rica em glicina, majoritariamente desordenada, com o domínio CSD N-terminal, mesmo que os contatos que promovem essa interação não sejam únicos e específicos, dando origem a uma conformação compacta.

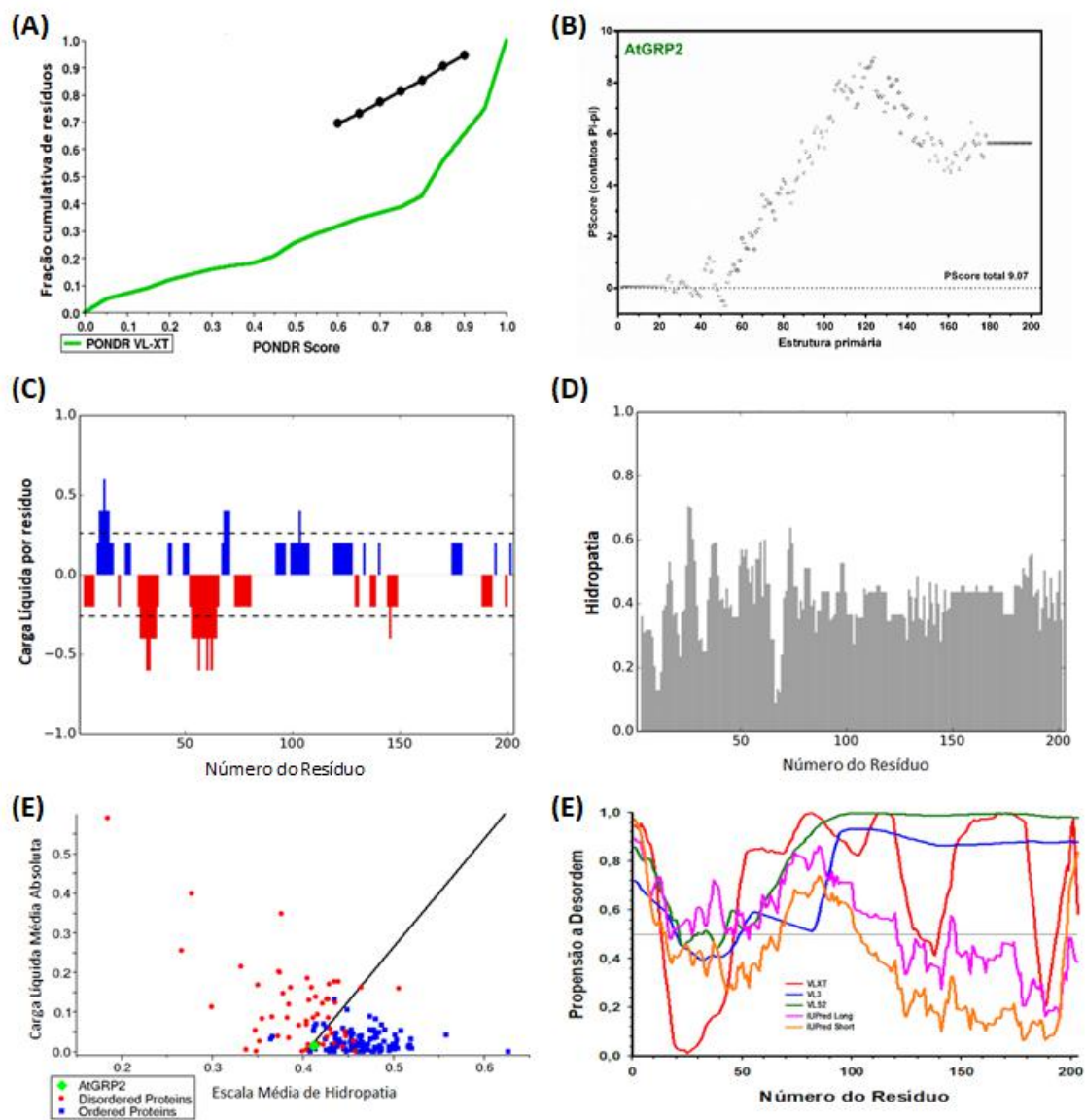


Figura 11 | Análise de propensão à desordem de AtGRP2. (A) Função de distribuição cumulativa (CDF - *Cumulative Distribution Function*) calculada para a sequência de aminoácidos de AtGRP2 (verde), localizada abaixo da linha de fronteira (preto). Espera-se, com isso, que AtGRP2 seja majoritariamente

desordenada (acesso da análise de CDF provida por <http://www.pondr.com/>). **(B)** PScore de AtGRP2 sugere propensão a desordem da cauda C-terminal rica em glicina, devido aos altos valores de contatos pi-pi. **(C)** Diagrama de carga líquida por resíduo da sequência de AtGRP2 gerado a partir do CIDER (*Classification of Intrinsically Disordered Ensemble Regions*; <http://pappulab.wustl.edu/CIDER>). Em azul, segmentos positivos e em vermelho, segmentos negativos. AtGRP2 apresenta alternância de cargas, mostrando ser um polianfólito fraco, característica de proteínas que fazem LLPS mediada por interações eletrostáticas. **(D)** Hidropatia ao longo da sequência primária de AtGRP2. O diagrama foi gerado por CIDER (<http://pappulab.wustl.edu/CIDER>). **(E)** Comparação carga-hidropatia (CH) para AtGRP2 gerado por PONDR® (*Predictor of Natural Disordered Regions*; <http://www.pondr.com/>). Pontos vermelhos representam proteínas desordenadas da literatura, pontos azuis correspondem a proteínas conhecidamente ordenadas e o ponto verde representa a proteína de interesse. **(F)** Avaliação de predisposição a desordem por resíduo da sequência de AtGRP2 por um grupo de algoritmos do programa PONDR®, PONDR VLXT (vermelho), PONDR VL3 (azul), PONDR VLS2 (verde) e pelos algoritmos da família IUPred, IUPred_short (rosa) e IUPred_long (laranja). Acesso aos algoritmos do PONDR® em <http://www.pondr.com/>, acesso ao IUPred em <https://iupred2a.elte.hu/plot>.

Considerando que os dados acima predizem a propensão à desordem, o próximo passo na avaliação bioinformática consistiu em investigar a propensão à LLPS de AtGRP2 através de dois outros algoritmos. O primeiro, PLAAC (*Prion-Like Amino Acid Composition*), procura por domínios prion-*like* dentro da sequência de aminoácidos da proteína de interesse⁽⁹⁷⁾. Príons são proteínas capazes de interconverter entre estados conformacionais distintos, sendo que ao menos um destes estados é transmissível. A transmissibilidade faz com que estas proteínas ajam como elementos não-Mendelianos de herança fenotípica, sendo capazes de se auto-propagarem através da formação de agregados amiloides⁽⁹⁸⁾. Interessantemente, elementos de sequência específicos determinam as propriedades prionóides das proteínas, de forma que podem ser preditos. Domínios prion-*like* normalmente são enriquecidos em aminoácidos polares não carregados, como asparagina e glutamina, enquanto apresentam baixos teores de aminoácidos hidrofóbicos e carregados⁽⁹⁸⁾. Como a LLPS é um fenômeno que parece preceder a agregação amiloide e proteínas que contêm domínios prion-*like* sofrem LLPS, a presença destas sequências é também um indicativo de propensão à LLPS da proteína de interesse^(89,99). O segundo algoritmo, catGRANULE, prevê a tendência de uma determinada proteína de formar grânulos de proteína-RNA através da LLPS. Para isso, o programa procura por sequências que apresentem desordem estrutural e que caracteristicamente interajam com ácido nucleico. Além disso, o programa analisa o teor

de argininas, glicinas e fenilalaninas, uma vez que estes aminoácidos se encontram enriquecidos em proteínas formadoras de grânulos⁽¹⁰⁰⁾. De acordo com a análise por PLAAC, a AtGRP2 apresenta dois domínios prion-like que coincidem com as regiões ricas em glicina C-terminais (Figura 12A). Curiosamente, estas mesmas regiões apresentam alta tendência a formação de grânulos, segundo o catGRANULE (Figura 12A). Além disso, as regiões ricas em glicina de AtGRP2 são preditas como IDRs pelos algoritmos da família PONDR®, sugerindo que AtGRP2 separa de fase e que este fenômeno é dirigido pelas caudas ricas em glicina. A pontuação catGRANULE de AtGRP2 foi comparada com a de proteínas que sabidamente separam de fase (controle positivo) e não separam de fase (controle negativo) (Figura 12B). AtGRP2 apresenta uma pontuação catGRANULE ainda maior que a da proteína FUS (*Fused in Sarcoma*), considerada protótipo para o estudo do fenômeno de LLPS⁽¹⁰¹⁻¹⁰³⁾, fundamentando mais uma vez a hipótese de LLPS para esta proteína (Figura 12B).

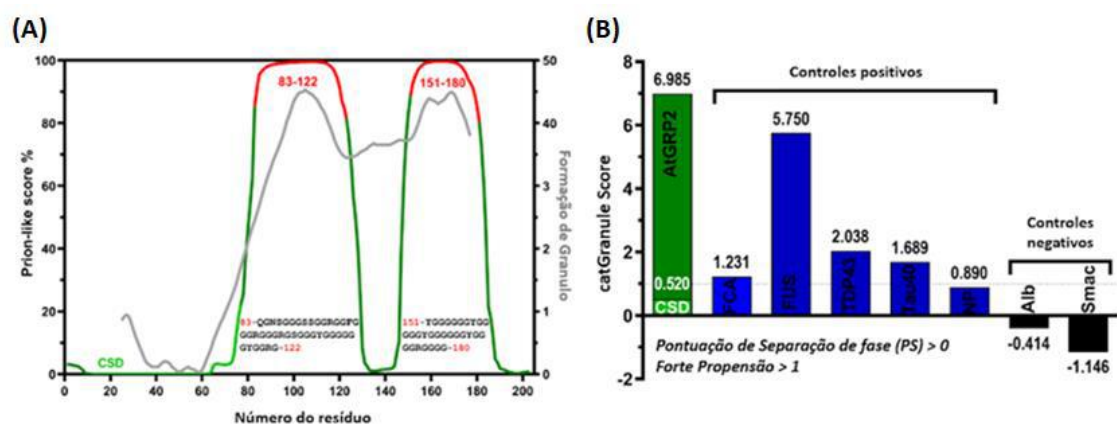


Figura 12 | Propensão à LLPS de AtGRP2 predita por bioinformática. (A) Eixo esquerdo: pontuação prion-like calculada pelo programa PLAAC (*Prion-Like Amino Acid Composition*) (Alberti et al., 2009). As duas regiões ricas em glicina presentes em AtGRP2 estão marcadas em vermelho como segmentos de baixa complexidade prion-like. Eixo direito: propensão à formação de grânulos predita pelo algoritmo catGRANULE (http://service.tartaglialab.com/new_submission/catGRANULE). **(B)** Propensão à LLPS baseada na pontuação catGranule (http://s.tartaglialab.com/page/catrapid_group), comparando AtGRP2 com proteínas controle que conhecidamente sofrem LLPS ou não. AtGRP2 em verde escuro; AtGRP2-CSD em verde claro. Barras azuis para FCA (Uniprot ID O04425), FUS (P06748), TDP-43 (Q13148), Tau40 (P10636-6) e Nucleofosmina (NP, P06748) como controle positivo. Barras pretas para as proteínas enoveladas albumina bovina sérica (Alb, P02768) e Smac (Q9NR28) como controles negativos. As pontuações para cada proteína estão localizadas acima de cada barra.

4.3. AtGRP2 forma condensados em gel

Experimentos de microscopia DIC evidenciaram a formação de condensados *in vitro* de AtGRP2 (70 μ M em Tris 20 mM (pH 7,4), NaCl 100 mM; ZnSO₄ 150 μ M), mesmo na ausência de agentes de *crowding* (Figura 13A). Os condensados de AtGRP2 são esféricos, apresentam um tamanho médio de $2,2 \pm 0,2 \mu$ m e se mantêm imóveis, sedimentados no fundo da lâmina (Figura 13 A). Além disso, não foram observados eventos de fusão ou fissão para os condensados de AtGRP2 durante todo o experimento, sugerindo que estes não apresentem características de líquidos. Portanto, com base nas características físicas aparentes dos condensados de AtGRP2, supõe-se que estes apresentem propriedades de sólidos, formando estruturas semelhantes a géis.

A formação de condensados em gel foi observada para outras proteínas que sofrem LLPS, tanto a partir da evolução de condensados líquidos⁽⁴⁰⁾ quanto sob condições de estresse^(41,105). A proteína FUS, relacionada a esclerose lateral amiotrófica (ELA), forma condensados *gel-like* quando exposta a agentes de *crowding*, uma estratégia que tenta mimetizar o ambiente intracelular superempacotado a fim de induzir a LLPS⁽⁴²⁾. As proteínas humanas prion e Tau, relacionadas a doenças neurodegenerativas, como as encefalopatias espongiformes e o mal de Alzheimer, respectivamente, formam condensados líquidos que evoluem para condensados *gel-like* e agregados sólidos fibrilares em função do tempo por mecanismos ainda pouco compreendidos^(89,104). O domínio C-terminal de TDP-43, outra RBP relacionada a ELA, forma condensados *gel-like* modulados pela alta concentração de proteína⁽⁴⁴⁾. Enquanto isso, a proteína de levedura Pub-1 forma condensados *gel-like* reversíveis quando submetida a pH ácido⁽¹⁰⁵⁾. Esses resultados sugerem que diferentes proteínas com capacidades de formação de interações interatômicas diferentes formam condensados com propriedades físicas diferentes que podem sofrer modificações em resposta a mudanças nas condições do meio. Além disso, estes estudos relacionam a propriedade de material ao estresse celular, sugerindo que a LLPS possa participar de alguma forma nos mecanismos de adaptação ao estresse celular.

Com o intuito de caracterizar as propriedades dos condensados de AtGRP2, estudou-se o efeito da adição de PEG 4000⁽¹⁰⁶⁾, um conhecido agente de *crowding*, sobre a LLPS da proteína. Normalmente, a adição de agentes de *crowding*, como PEG e FICOL, promovem a LLPS por diminuir a concentração crítica para a partição da proteína na fase condensada⁽¹⁰⁶⁾. Entretanto, o mesmo não foi observado para AtGRP2, uma vez

que não foi observado um aumento do número de condensados após a incubação com PEG 4000, de 120 do controle para 109 (Figura 13B). Todavia, na presença de PEG 4000, o tamanho médio dos condensados diminuiu para $1,10 \pm 0,08 \mu\text{m}$, sugerindo que a LLPS de AtGRP2 pode ser modulada pelas condições intracelulares (Figura 13B). Estudou-se também o efeito da temperatura sobre a LLPS de AtGRP2. Para isso, as amostras foram incubadas a 37°C por 1 hora antes da aquisição das imagens. O aumento da temperatura levou a um ligeiro aumento no número de condensados de 120 do controle para $164 \pm 12,5$ em conjunto com uma diminuição do tamanho para $0,80 \pm 0,09 \mu\text{m}$ dos condensados (Figura 13B). Interessantemente, o aumento da temperatura promoveu a formação de condensados em cacho. Esse efeito suporta a hipótese de que os condensados de AtGRP2 sejam, de fato, gelatinosos, uma vez que estes não se fundem como observado para condensados líquidos, como os formados pela proteína príon⁽⁸⁹⁾. Além disso, observou-se um aumento significativo no número, resultando em 385 condensados, e a recuperação do tamanho médio observado na situação controle de $2,3 \pm 0,1 \mu\text{m}$ quando estes foram incubados a 37°C na presença de PEG 4000 (Figura 13B), sugerindo que a LLPS de AtGRP2 é modulada pela temperatura e que a alta temperatura pode atuar como um sinal de estresse que deflagra a LLPS de AtGRP2 quando em condições de *crowding*.

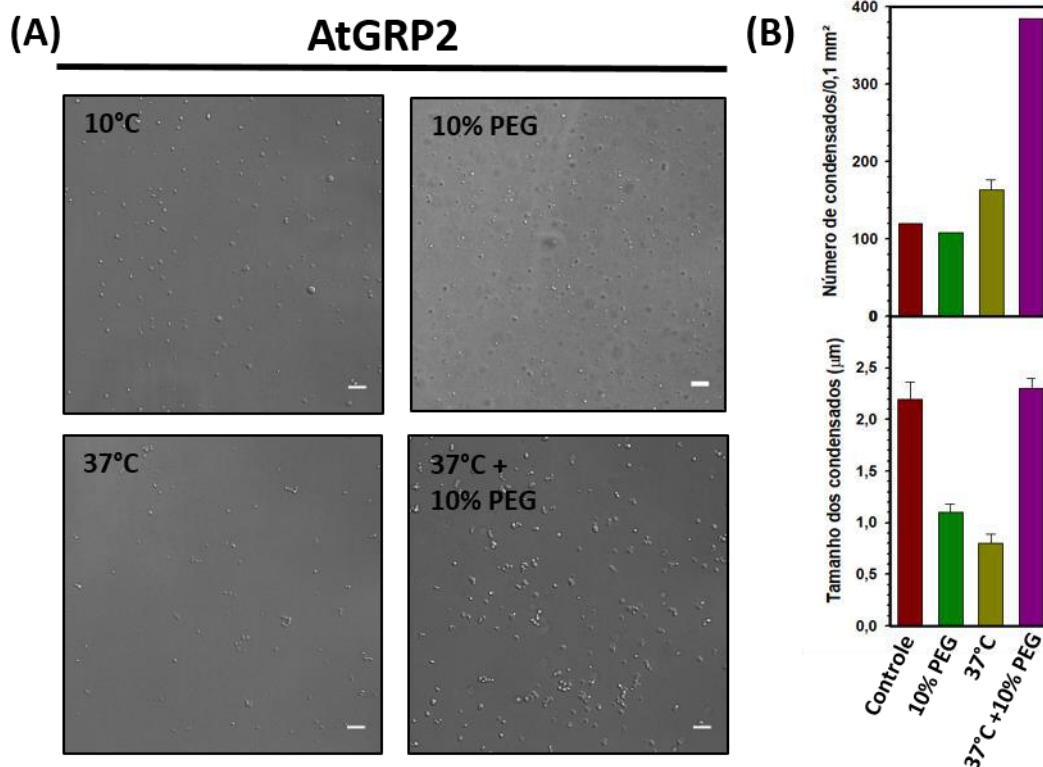


Figura 13 | LLPs de AtGRP2 investigada por microscopia DIC. Imagens representativas dos experimentos de microscopia realizados para AtGRP2 purificada a 70 μM em Tris 20 mM (pH 7,4), NaCl 100 mM, ZnSO₄ 150 μM]. Barra de escala, 10 μm. **(A) Esquerda (acima):** AtGRP2 controle a 10°C. **Direita (acima):** amostra adicionada de 10% (p/v) de PEG 4000. **Esquerda (abaixo):** amostra após 1 h de incubação a 37°C. **Direita (abaixo):** amostra adicionada de 10% (p/v) de PEG 4000 após 1 h de incubação a 37°C. **(B)** Quantificação do tamanho médio (acima) e do número (abaixo) de condensados de AtGRP2 a partir das imagens de microscopia DIC para: controle a 10°C (barra vermelha), amostra adicionada de 10% (p/v) de PEG 4000 (barra verde), amostra após 1 h de incubação a 37°C (barra amarela), amostra adicionada de 10% (p/v) de PEG 4000 após 1 h de incubação a 37°C (barra roxa).

4.4. Os condensados de AtGRP2 são de natureza proteica

O estudo da LLPS de proteínas depende ainda da comprovação de que os condensados observados são de natureza proteica. Para isso, utilizou-se duas estratégias: *(i)* adição do fluoróforo SYPRO orange e *(ii)* incubação com 10% (p/v) de 1,6-hexanodiol. O SYPRO orange é um corante fluorescente capaz de se ligar a resíduos hidrofóbicos expostos em proteínas⁽¹⁰⁷⁾. Os condensados vistos nas imagens de DIC coincidem com os observados nas imagens de microscopia de fluorescência adquiridas na presença de SYPRO Orange, o que fica melhor evidenciado na sobreposição das

duas imagens (Figura 14). Logo, estes dados sugerem que os condensados observados ao microscópio são, de fato, de natureza proteica. É interessante notar que os condensados de AtGRP2 não foram bem corados com SYPRO Orange, o que indica que a interação entre a proteína e o fluoróforo não se deu por completo. Esse resultado sugere que os condensados de AtGRP2 sejam formados por uma porção significativa de interações hidrofóbicas.

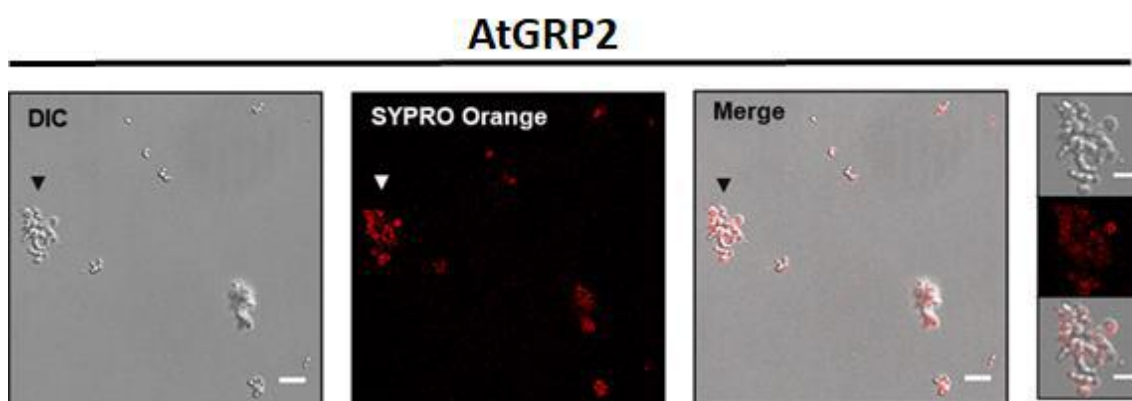


Figura 14 | AtGRP2 forma condensados proteicos dispostos em cachos. Imagens representativas dos experimentos de microscopia DIC e de fluorescência adquiridos para AtGRP2 purificada a 70 μM em Tris 20 mM (pH 7,4), NaCl 100 mM, ZnSO_4 150 μM . Barra de escala, 10 μm . A amostra foi incubada por 10 min a temperatura ambiente após a adição de SYPRO orange. Da direita para a esquerda: Imagem DIC de condensados de AtGRP2. Imagem de fluorescência do SYPRO orange ligado a resíduos hidrofóbicos. Sobreposição das duas imagens. Aproximação do aglomerado de condensados marcado com a cabeça de seta.

A fim de comprovar a natureza proteica dos condensados de AtGRP2, a amostra foi incubada com 10% de 1,6-hexanodiol. O 1,6-hexanodiol é um álcool alifático conhecido por desfazer interações hidrofóbicas fracas envolvidas em LLPS, desfazendo o condensado proteico⁽¹⁰⁸⁾. Entretanto, a adição da mesma concentração de 1,6-hexanodiol não é capaz de desfazer vesículas lipídicas. Assim, esse álcool é frequentemente utilizado para diferenciar entre condensados proteicos e vesículas lipídicas em imagens de microscopia. A Figura 15C mostra que a adição de 10% de 1,6-hexanodiol desfez por completo os condensados de AtGRP2, evidenciando sua natureza proteica. Em seguida, AtGRP2 foi incubada com 1,6-hexanodiol a 1%, no intuito de se estudar ainda a natureza das interações que regem a LLPS. Considerando as

propriedades sólidas dos condensados de AtGRP2 e o modo de atuação do 1,6-hexanodiol, a adição de 10x menos hexanodiol resultou na mudança das propriedades físicas aparentes dos condensados. Os condensados passaram a ficar suspensos e não mais sedimentados no fundo da lâmina, além de deixarem de ficar imóveis, se comportando como condensados líquidos, da mesma forma que a maioria das proteínas que sofrem LLPS, como LAF-1⁽¹⁰⁹⁾, FUS⁽¹¹⁰⁾ e TDP-43⁽¹¹¹⁾. Ainda, na presença de 1% de 1,6-hexanodiol, observou-se a formação de corpos multivesiculados, sugerindo a modificação das propriedades de material do interior dos condensados de AtGRP2 (Figura 15C).

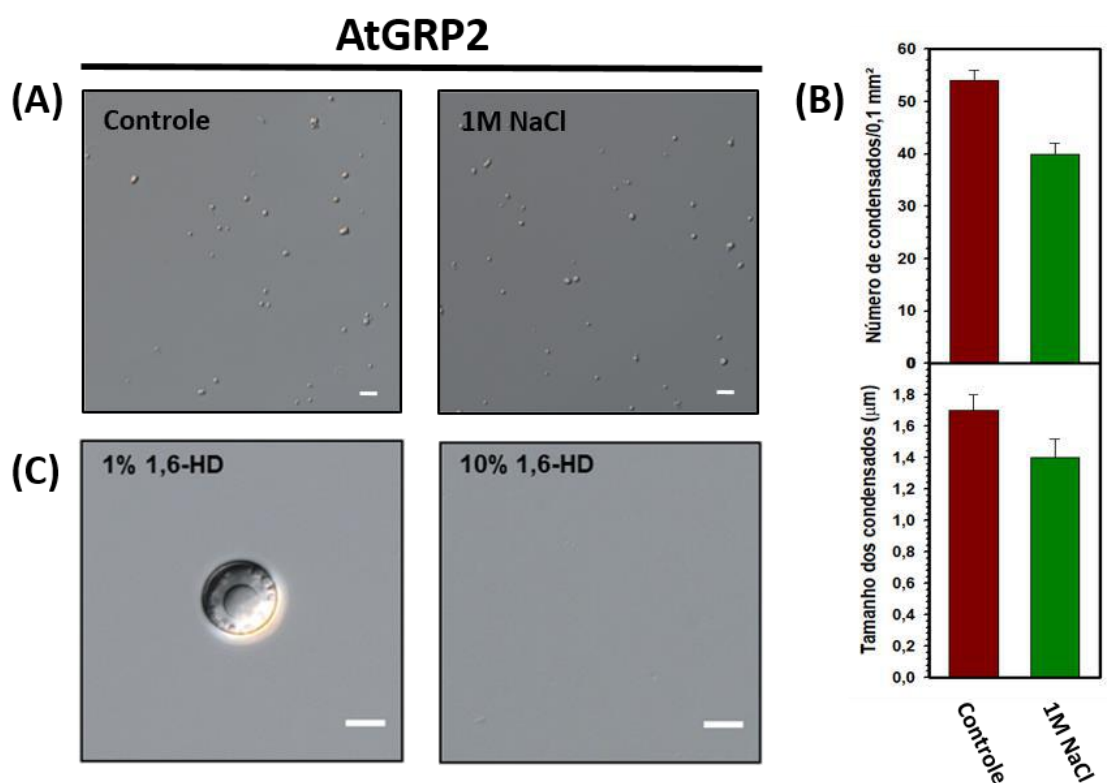


Figura 15 | Efeito de sal e 1,6-hexanodiol na LLPS de AtGRP2. Imagens representativas dos experimentos de microscopia DIC obtidos para AtGRP2 purificada a 20 μM em Tris 20 mM (pH 7,4), NaCl 100 mM, ZnSO₄ 150 μM. Barra de escala, 10 μm. **(A)** Direita: AtGRP2 controle. Esquerda: AtGRP2 na presença de NaCl 1 M. A amostra foi incubada por 10 min a temperatura ambiente após a adição de NaCl. **(B)** Quantificação do tamanho médio (acima) e do número (abaixo) de condensados de AtGRP2 a partir das imagens de microscopia DIC para: controle (barra vermelha) e amostra adicionada de NaCl 1 M (barra verde). **(C)** Incubação da amostra de AtGRP2 com 1,6-hexanodiol por 1 h. Direita: AtGRP2 na presença de 1,6-hexanodiol 1% (p/v). Esquerda: AtGRP2 na presença de 1,6-hexanodiol 10% (p/v).

4.5. A LLPS de AtGRP2 não é primariamente dirigida por interações eletrostáticas

A fim de se estudar a contribuição das interações eletrostáticas para a LLPS de AtGRP2, a amostra foi incubada com 1M de NaCl. Em comparação com o controle, contendo 54 ± 2 condensados por imagem de $1,7 \pm 0,1 \mu\text{m}$, a adição de sal promoveu uma ligeira diminuição do número (40 ± 2) e do tamanho médio ($1,4 \pm 0,1$) μm dos condensados de AtGRP2 (figura 15B). Essa resistência ao sal indica que a LLPS de AtGRP2 não é primariamente regida por interações eletrostáticas. Isso corrobora os dados obtidos na presença de SYPRO orange, sugerindo que a interação hidrofóbica parece ter uma contribuição mais significativa que a interação eletrostática para a LLPS de AtGRP2.

A AtGRP2 é capaz de formar condensados em concentrações tão baixas quanto $20 \mu\text{M}$, próxima a concentração intracelular da proteína⁽¹¹²⁾. Isso implica na possibilidade de AtGRP2 formar condensados *in vivo*, o que pode ser chave para o entendimento do mecanismo de atuação dessa proteína. AtGRP2 regula negativamente uma série de processos celulares de resposta a estresse⁽⁷⁰⁻⁷⁵⁾. Além disso, pode atuar no processamento ou estabilização de RNA, devido a sua função de chaperona⁽⁶⁹⁾. Uma possibilidade seria a de sequestro de RNAs específicos para dentro dos condensados quando a planta for submetida a algum estresse abiótico, formando “estresseosomas”. Para discutir melhor sobre os possíveis mecanismos de atuação da AtGRP2 é necessário conhecer as concentrações limite para a formação de condensados, bem como a modulação da LLPS por RNA. Como o ligante é capaz de modular a separação de fases? Qual seria o papel dos domínios isolados nesse processo? A cauda rica em glicina parece interagir com o domínio CSD? Será que essa interação ocorre no mesmo sítio de interação com RNA? Em caso positivo, o ligante é capaz de quebrar a interação com a cauda? Além disso, é possível ainda investigar a propriedade de material dos condensados por meio de FRAP (*Fluorescence Recovery After Photobleaching*), a fim de confirmar a hipótese de que a separação se dá para uma fase em gel.

5. Conclusões

- AtGRP2 foi purificada por três etapas de cromatografia: afinidade a níquel, clivagem da cauda de polihistidinas e nova etapa de cromatografia de afinidade a níquel e exclusão molecular.
- AtGRP2 é uma proteína não globular compacta, como mostrado pelos resultados de cromatografia por exclusão molecular em comparação com seu domínio CSD isolado.
- De acordo com as predições de bioinformática, AtGRP2 apresenta alta tendência a desordem estrutural, principalmente na região da cauda C-terminal rica em glicina.
- A predição à LLPS sugere que AtGRP2 apresenta alta tendência à transição de fase (maior ainda que a proteína FUS) e que esse processo é provavelmente dirigido pela cauda C-terminal rica em glicina.
- Experimentos de microscopia DIC confirmaram a habilidade AtGRP2 de separar de fase em concentrações tão baixas quanto 20 μM .
- O condensados de AtGRP2 são esféricos, de tamanho micrométrico, depositados no fundo da lâmina e imóveis, sugerindo propriedade de material sólido e formação de estruturas *gel-like*.
- Os condensados mantêm suas propriedades quando submetidos a adição do agente de *crowding* PEG-4000, não mudando significativamente em tamanho ou em número.
- A quantidade de condensados aumenta quando a amostra é incubada a 37°C por 1 h tanto em tampão quanto presença de PEG-4000, sugerindo que a temperatura possa ser um sinal de estresse que deflagra a LLPS de AtGRP2.
- O aumento da temperatura alterou o perfil dos condensados que aparecem na forma de cachos, grudados uns aos outros, mas sem se fundir, o que corrobora as suas características *gel-like*.
- A adição de SYPRO Orange indica que os condensados são proteicos. O fato de os condensados não corarem completamente pelo fluoróforo sugere que as interações hidrofóbicas possam ter participação na LLPS.
- A adição de NaCl 1 M não foi capaz de desfazer os condensados, o que sugere que contatos eletrostáticos não regem a LLPS de AtGRP2.

- A adição de 1,6-hexanodiol 10% desfez os condensados por completo, confirmando a natureza proteica.
- A adição de 1,6-hexanodiol 1% modificou a propriedade de material aparente dos condensados, que passaram a se mover e ficaram suspensos em solução, sugerindo liquefação.

6. Referências bibliográficas

1. Banani SF *et al.* (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell. Biol.* **18**: 285-298.
2. Darling AL *et al.* (2018) Intrinsically Disordered Proteome of Human Membrane-Less Organelles. *Proteomics* **18**: e1700193.
3. Balzeau J, Menezes MR, Cao S, Hagan JP (2017) The LIN28/let-7 Pathway in Cancer. *Front Genet.* **8**:31.
4. Lidforss, B. (1915). In: Paul Hinneberg. Allgemeine Biologie. Leipzig, Berlin: Verlag von B. G. Teubner. Eine Neubildung dieser Organe oder Organellen findet wenigstens bei höheren Pflanzen nicht statt, *Protoplasma*.pp. **227**: 218–264.
5. Mitrea, D.M. and Kriwacki, R.W. (2016) Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.* **14**: 1.
6. Uversky, V.N. (2017) Intrinsically disordered proteins in over-crowded milieu: membrane-less organelles, phase separation, and intrinsic disorder. *Curr. Opin. Struct. Biol.* **44**: 18–30.
7. Handwerger, K.E. *et al.* (2005) Cajal bodies, nucleoli, and speckles in the *Xenopus* oocyte nucleus have a low-density, sponge-like structure. *Mol. Biol. Cell.* **16**: 202–211 18.
8. Kedersha, N. *et al.* (2000) Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J. Cell Biol.* **151**: 1257–1268 19.
9. Andrei, M.A. *et al.* (2005) A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA.* **11**: 717–727.
10. Brangwynne, C.P. *et al.* (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science.* **324**: 1729–1732.
11. Brangwynne CP, Tompa P, Pappu RV (2015) Polymer physics of intracellular phase transitions. *Nature.* **11**: 899-904.
12. Brangwynne, C.P. *et al.* (2011) Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 4334–4339.
13. Flory, P.J. (1942) Thermodynamics of high polymer solutions. *J. Chem. Phys.* **10**: 10.
14. Li, P. *et al.* (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature.* **483**: 336–340
15. Molliex, A. *et al.* (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell.* **163**: 123–133.
16. Patel, A. *et al.* (2015) A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell.* **162**: 1066–1077.
17. Kim, H.J. *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature.* **495**: 467–473.

18. Jain, S. et al. (2016) ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell*. **164**: 487–498.
19. Andersen, J.S. et al. (2005) Nucleolar proteome dynamics. *Nature*. **433**: 77–83 35.
20. Kong, K.W. et al. (2013) Whole-genome screening identifies proteins localized to distinct nuclear bodies. *J. Cell Biol.* **203**: 149–164 36.
21. Boke, E. et al. (2016) Amyloid-like self-assembly of a cellular compartment. *Cell*. **166**: 637–650.
22. Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **83**: 553–584.
23. Varadi, M. et al. (2015) Functional advantages of conserved intrinsic disorder in RNA-binding proteins. *PLoS One*. **10**: e0139731 39.
24. Kato, M. et al. (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**: 753–767.
25. Lunde, B.M. et al. (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* **8**: 479–490.
26. Banjade, S. et al. (2015) Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc. Natl. Acad. Sci. U. S. A.* **112**: E6426–E6435.
27. Kato, M. et al. (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**: 753–767.
28. Kim, H.J. et al. (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*. **495**: 467–473.
29. Lin, Y. et al. (2015) Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell*. **60**: 208–219 48.
30. Elbaum-Garfinkle, S. et al. (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **112**: 7189–7194 49.
31. Nott, T.J. et al. (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell*. **57**: 936–947.
32. Tompa, P. et al. (2014) A million peptide motifs for the molecular biologist. *Mol. Cell* **55**, 161–169
33. Monahan, Z. et al. (2017) Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* **36**: 2951–2967.
34. Pak, C.W. et al. (2016) Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell*. **63**: 72–85 54.
35. Riback, J.A. et al. (2017) Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell*. **168**: 1028–1040 55.

36. Vernon, R.M. et al. (2018) Pi-pi contacts are an overlooked protein feature relevant to phase separation. *Elife*. **7**: e31486 56.
37. Brady, J.P. et al. (2017) Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc. Natl. Acad. Sci. U. S. A.* **114**: E8194–E8203.
38. Lin, Y. et al. (2016) Toxic PR poly-dipeptides encoded by the C9orf72 repeat expansion target LC domain polymers. *Cell* **167**, 789–802
39. Loughlin FE, Wilce JA (2019) TDP-43 and FUS-structural insights into RNA recognition and self-association. *Curr Opin Struct Biol.* **59**:134-142.
40. Schwartz, J.C. et al. (2013) RNA seeds higher-order assembly of FUS protein. *Cell Rep.* **5**: 918–925 59.
41. Zhang, X. et al. (2017) RNA stores tau reversibly in complex coacervates. *PLoS Biol.* **15**: e2002183 60.
42. Zhang, H. et al. (2015) RNA controls polyQ protein phase transitions. *Mol. Cell.* **60**: 220–230.
43. Burke, K.A. et al. (2015) Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Mol. Cell.* **60**: 231–241
44. Banerjee, P.R. et al. (2017) Reentrant phase transition drives dynamic substructure formation in ribonucleoprotein droplets. *Angew. Chem. Int. Ed. Engl.* **56**: 11354–11359
45. Zhang, H. et al. (2015) RNA controls polyQ protein phase transitions. *Mol. Cell* **60**, 220–230
46. Sachetto-Martins G, Fusaro AF (2007) Blooming Time for Plant Glycine-Rich Proteins. *Plant Signaling & Behavior.* **2(5)**: 386-387.
47. Sachetto-Martins G, Franco LO, de Oliveira DE (2000) Plant glycine-rich proteins: a family or just proteins with a common motif? *Biochim. Biophys. Acta* **1492**: 1-14.
48. Condit CM, Meagher RB (1986) A gene encoding a novel glycine-rich structural protein of *Petunia*. *Nature.* **323**:178–181.
49. Condit CM, McLean BG, Meagher RB (1990) Characterization of the expression of the petuniaglycine-rich protein-1 gene product. *Plant Physiol.* **93**:596–602.
50. Mangeon A, Junqueira RM, Sachetto-Martins G (2010) Functional diversity of the plant glycine-rich proteins superfamily. *Plant Signaling & Behavior.* **5(2)**: 99-104.
51. Cassab GI (1998) Plant cell wall proteins. *Annu Rev Plant Physiol Plant Mol. Biol.* **49**: 281–30.
52. Ciuzan O et al. (2015) The evolutionarily conserved multifunctional glycine-rich RNA-binding proteins play key roles in development and stress adaptation. *Physiol. Plant.* **153**: 1-11.

53. Cléry A, Blatter M, Allain FH. (2008) RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol.* **18(3)**: 290-8.
54. Lindquist JA, Mertens PR. (2018) **Cold shock** proteins: from cellular mechanisms to pathophysiology and disease. *Cell Commun Signal.* **16(1)**: 63.
55. Sachs R, Max KEA, Heinemann U, Balbach J (2012) RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA.* **18**: 65–76
56. Sasaki K, Imai R (2012) Pleiotropic roles of cold shock domain proteins in plants. *Front Plant Sci.* **2**: 116.
57. Steinert PM, Mack JW, Korge BP, Gan SQ, Haynes SR, Steven AC (1991) Glycine loops in proteins: their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. *International Journal of Biological Macromolecules.* **13**: 130-139.
58. Keller B, Templeton MD, Lamb CJ (1989) Specific localization of a plant cell wall glycine-rich protein in protoxylem cells of the vascular system. *PNAS.* **86 (5)**: 1529-1533.
59. Varner JE, Cassab GI (1986) Molecular biology: A new protein in petunia. *Nature.* **323**: 110.
60. Steinert PM, Mack JW, Korge BP, Gan SQ, Haynes SR, Steven AC (1991) Glycine loops in proteins: their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. *Int J Biol Macromol.* **13**: 130–13
61. Condit CM, Meagher RB (1986) A gene encoding a novel glycine-rich structural protein of petunia. *Nature.* **323(6084)**: 178–181.
62. Czolpinska M, Rurek M. (2018) Plant **Glycine-Rich Proteins** in **Stress** Response: An Emerging, Still Prospective Story. *Front Plant Sci.* **9**: 302.
63. Wang S, Wang R, Liang D, Ma F, Shu H (2012) Molecular characterization and expression analysis of a glycine-rich RNA-binding protein gene from *Malushupehensis* Rehd. *Mol Biol Rep.* **39**: 4145–415.
64. Kim JY, Kim WY, Kwak KJ, Oh SH, Han YS, Kang H (2010) Glycine-rich RNA-binding proteins are functionally conserved in *Arabidopsis thaliana* and *Oryza sativa* during cold adaptation process. *J Exp Bot.* **61**: 2317–2325
65. Delatorre CA, Adriano Silva AA (2008) *Arabidopsis thaliana*: Uma pequena planta um grande papel. *Revista de Ciências Agrárias.* **31**
66. Nakaminami K *et al.* (2009) *Arabidopsis* cold shock domain proteins: relationships to floral and silique development. *J. Exp. Bot.* **60**: 1047-1062.
67. Fusaro AF, Bocca SN, Ramos RL, Barroco RM, Magioli C, Jorge VC, *et al.* (2007) AtGRP2, a cold-induced nucleo-cytoplasmic RNA-binding protein, has a role in flower and seed development. *Planta.* **225**: 1339–1351.

68. Sasaki K, Kim MH, Imai R (2007) Arabidopsis cold shock domain protein 2 is a RNA chaperone that is regulated by cold and development signals. *Biochem. Biophys. Res. Commun.* **364**: 633-638.
69. Kim JS, Park SJ, Kwak KJ, Kim YO, Kim JY, Song J, Jang B, Jung CH, Kang H (2007) Cold shock domain proteins and glycine-rich RNA-binding proteins from *Arabidopsis thaliana* can promote the cold adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **35(2)**: 506–516.
70. Park SJ *et al.* (2009) Cold shock domain proteins affect seed germination and growth of *Arabidopsis thaliana* under abiotic stress conditions. *Plant Cell Physiol.* **50**: 869-878.
71. Sasaki K, Kim MH, Imai R. (2013) Arabidopsis COLD SHOCK DOMAIN PROTEIN 2 is a negative regulator of cold acclimation. *New Phytol.* **198(1)**: 95-102.
72. Jaglo-Ottosen KR. (1998) Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science.* **280**: 104-6.
73. Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF. (2000) Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol.* **124**: 1854-65.
74. Sasaki K, Kim MH, Kanno Y, Seo M, Kamiya Y, Imai R. (2015) Arabidopsis cold shock domain protein 2 influences ABA accumulation in seed and negatively regulates germination. *Biochem Biophys Res Commun.* **456(1)**: 380-4.
75. Sasaki K, Liu Y, Kim MH, Imai R. (2015) An RNA chaperone, AtCSP2, negatively regulates salt stress tolerance. *Plant Signal Behav.* **10(8)**: e1042637.
76. Kant P, Kant S, Gordon M, Shaked R, Barak S. (2007) STRESS RESPONSE SUPPRESSOR1 and STRESS RESPONSE SUPPRESSOR2, two DEAD-box RNA helicases that attenuate Arabidopsis responses to multiple abiotic stresses. *Plant Physiol.* **145**: 814-30.
77. He Y, Li W, Lv J, Jia Y, Wang M, Xia G. (2012) Ectopic expression of a wheat MYB transcription factor gene, TaMYB73, improves salinity stress tolerance in *Arabidopsis thaliana*. *J Exp Bot.* **63**: 1511-22.
78. Moss EG, Lee RC, Ambros V. (1997) The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA. *Cell.* **88**: 637-646.
79. Richards M, Tan S-P, Tan J-H, Chan W-K, Bongso A. (2004) The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells.* **22**: 51-64.
80. Darr H, Benvenisty N. (2009) Genetic analysis of the role of the reprogramming gene LIN-28 in human embryonic stem cells. *Stem Cells.* **27**: 352-62.
81. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, *et al.* (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science.* **318**: 1917-20.

82. Viswanathan SR, Daley GQ, Gregory RI. (2008) Selective blockade of microRNA processing by Lin28. *Science*. **320**: 97-100.
83. Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*. **166**: 557-580
84. Laemmli UK (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*. **227**: 680-685.
85. François Baneux (1999) Recombinant protein expression in *Escherichia coli*. *Elsevier*. **10**: 411-421.
86. Blommel PG, Fox BG (2007) A combined approach to improving large-scale production of Tobacco Etch Virus protease. *Protein Expression and Purification* **55(1)**: 53-68.
87. Carrington JC, Dougherty WG (1998) A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proceedings of the National Academy of Sciences of the United States of America*. **85(10)**: 3391–3395.
88. Drino A, Schaefer MR (2018) RNAs, phase separation, and membrane-less organelles: are post-transcriptional modifications modulating organelle dynamics? *Bioessays* **40**: e1800085
89. Matos C *et al.* (2019) Liquid-liquid phase separation and aggregation of the prion protein globular domain modulated by a high-affinity DNA aptamer. *Biorxiv*.
90. Mitrea DM, Kriwacki RW (2016) Phase separation in biology; functional organization of a higher order. *Cell Communication and Signaling* 14:1.
91. Taylor JP, Brown RH, Cleveland DW (2016) Decoding ALS: from genes to mechanism. *Nature*. **539**:197–206.
92. Kato M *et al.* (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**:753–767.
93. Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *PNAS*. **18**: 8183 - 8188
94. Weathers EA, Paulaitis ME, Woolf TB, Hoh JH (2004) Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS*. **576(3)**: 348 - 352
95. Oldfield CJ and Dunker AK (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **83**: 553–584.
96. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK (2001) Sequence complexity of disordered protein. *Proteins*. **42**: 38 – 48.
97. Lancaster AK, Nutter-Upham A, Lindquist S, King OD. (2014) PLAAC: a web and command-line application to identify proteins with Prion-Like Amino Acid Composition. *Bioinformatics*.
98. Alberti S, Halfmann R, King O, Kapila A, Lindquist S. (2009) A systematic survey identifies

- prions and illuminates sequence features of prionogenic proteins. *Cell*. **137**: 146-58.
99. Maharana S *et al.* (2018) RNA buffers the phase separation behavior of prion-like RNA-binding proteins. *Science*. **360(6391)**: 918–921.
 100. Bolognesi B, Lorenzo Gotor N, Dhar R, Cirillo D, Baldrighi M, Tartaglia GG, Lehner B (2016) A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. *Cell Rep*. 16(1): 222-231.
 101. Qamar S *et al.* (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- π Interactions. *Cell*. **173(3)**: 720–734.e15.
 102. Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, Tycko R (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*. **171(3)**:615-627
 103. Murthy AC, Dignon GL, Kan Y, Zerze GH, Parekh SH, Mittal J, Fawzi NL (2019) Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat Struct Mol Biol*. **26(7)**:637-648.
 104. Wegmann S, et al. (2018) Tau protein liquid-liquid phase separation can initiate tau aggregation. *EMBO J*. **37(7)**.
 105. Kroschwald S, Munder MC, Maharana S, Franzmann TM, Richter D, Ruer M, Hyman AA, Alberti S (2018) Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. *Cell Rep*. **23(11)**: 3327-3339.
 106. Kaur T, Alshareedah I, Wang W, Ngo J, Moosa MM, Banerjee PR (2019) Molecular Crowding Tunes Material States of Ribonucleoprotein Condensates. *Biomolecules*. **9(2)**.
 107. Steinberg TH, Jones LJ, Haugland RP, Singer VL (1996) SYPRO orange and SYPRO red protein gel stains: one-step fluorescent staining of denaturing gels for detection of nanogram levels of protein. *Analytical biochemistry*. **239**: 223–237.
 108. Kroschwald S, Alberti S (2017) Gel or Die: Phase Separation as a Survival Strategy *Cell*. **168**: 947–948.
 109. Elbaum-Garfinkle S, Kim Y, Szczepaniak K, Chih-Hsiung C, Eckmann CR, Myong S, Brangwynne CP (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *PNAS*. **112(23)**: 7189–7194.
 110. Shin Y, Brangwynne CP (2017) Liquid phase condensation in cell physiology and disease. *Science*. **357(6357)**.
 111. McGurk L, Gomes E, Guo L, Mojsilovic-Petrovic J, Tran V, Kalb RG, Shorter J, Bonini NM (2018) Poly(ADP-Ribose) Prevents Pathological Phase Separation of TDP-43 by Promoting Liquid Demixing and Stress Granule Localization. *Mol Cell*. **71(5)**: 703-717.
 112. Milo R, Jorgensen P, Moran U, Weber G, Springer M (2010) BioNumbers-the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* (38(Database issue): D750-D753.