UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

Lucas Rafael Oliveira Santos Eugênio

Produção e caracterização de NtRh35, uma helicase de tabaco envolvida em *splicing*

São Carlos
2024

Lucas Rafael Oliveira Santos Eugênio

Produção e caracterização de NtRh35, uma helicase de tabaco envolvida em *splicing*

Trabalho de conclusão de curso apresentado ao Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Bacharel em Ciências Físicas e Biomoleculares.
Orientadora: Profa. Dra. Ana Paula Ulian de Araújo-Instituto de Física de São Carlos (IFSC-USP).

São Carlos
2024

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

RESUMO

A RNA helicase 35 (NtRh35 ou NtDDX35) foi identificada como uma parceira de interação da proteína SCI1 (*Stigma/style Cell-cycle Inhibitor 1*), a qual é descrita estar envolvida na regulação e proliferação de células do meristema floral de *Nicotiana tabacum* (tabaco). Este trabalho teve o foco voltado a otimizar a produção da NtRh35 e também de sua parceira NtSCI1 em *E. coli*, de forma estável, a fim de viabilizar a realização de estudos biofísicos e estruturais. Entre esses, estudos para avaliar seu estado oligomérico, utilizando SEC-MALS, e sua estrutura secundária por Dicroísmo Circular foram realizados. Além disso, a técnica de termoforese em microescala (MST) e de fluorescência intrínseca foram testadas na tentativa de determinar a ligação de ATP e de RNA. Análises *in silico* foram usadas para gerar modelos e determinar domínios estruturais de ambas proteínas visando definir construções para cristalização e co-cristalização. Espera-se com o trabalho, contribuir para a caracterização desta nova RNA helicase e, assim, permitir validar putativos transcritos-alvos nos quais a NtRh35 atua.

Palavras-chave: Biofísica Molecular; RNA helicase; Biologia Estrutural.

1. INTRODUÇÃO

O estudo do desenvolvimento das flores é de suma importância para o entendimento do funcionamento de seus órgãos sexuais, tendo grande relevância econômica, já que das flores derivam os frutos, os quais fornecem uma ampla variedade de fontes de alimento. De pistilos da flor de *Nicotiana tabacum*, foi identificado o gene SCI1 (*Stigma/style Cell-cycle Inhibitor 1*) que apresenta uma alta expressão durante o processo de proliferação celular antes da diferenciação das células do meristema floral. SCI1 tem como função a regulação dessa proliferação, controlando o início e o término do processo de diferenciação.[1] Para um melhor entendimento do papel da proteína SCI1, a busca de parceiros de interação foi avaliada através de ensaios de duplo híbrido e, dentre as potenciais parceiras, foi encontrada a DEAD-box RNA helicase 35 (NtRh35).[2]

As helicases são enzimas dependentes de adenosina trifosfato (ATP) que tem como principal função separar fitas duplas de ácido ribonucleico (RNA), ácido desoxirribonucleico

(DNA) e/ou híbridos DNA-RNA e, com isso, têm funções durante todo o processo de metabolismo de ácidos nucleicos, como na transcrição, *splicing*, replicação, entre outros.[2] Estruturalmente, as helicases são compostas de dois domínios principais mais conservados que compõem o núcleo helicase onde ocorrem a ligação de ATP e do ácido nucleico a ser desenrolado (Fig. 1). Estes domínios são conhecidos como RecA1 e RecA2 (domínio helicase), pela semelhança de seu formato com a ATPase RecA, sendo ligados por um *linker* flexível. Além disso, possuem domínios N e C-terminais predominantemente desestruturados que não são conservados e que estão associados a funções específicas de cada helicase, como interações com outras proteínas parceiras ou transporte dentro das células.[3]

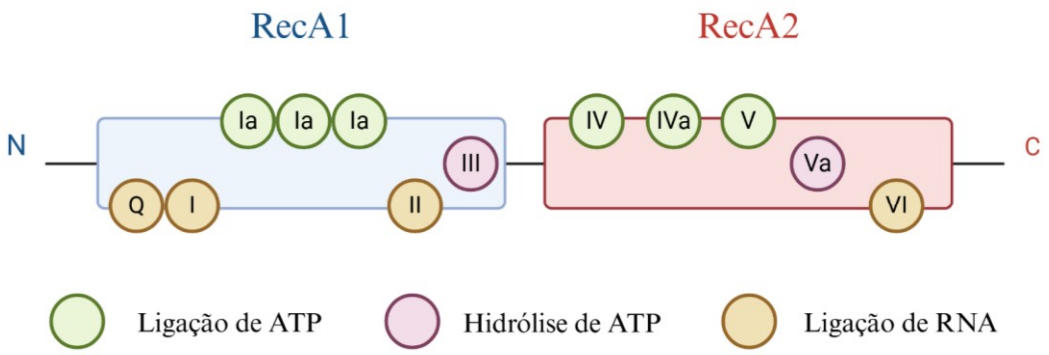


Figura 1 - Representação dos principais domínios e motivos conservados em helicases, com indicação por cor de suas principais funções. Adaptado de (LEITÃO et al., 2015).[4]

Estruturas de helicases eucarióticas já foram resolvidas por cristalografia de raios-x e criomicroscopia eletrônica, mas ainda há aspectos que não são bem entendidos.[4] Sabe-se que a ligação de ATP é necessária para que a helicase possa entrar em um estado conformacional fechado, no qual ambos domínios se aproximem e o ácido nucleico se ligue e possa ocorrer a separação das fitas (Fig. 2). Com a hidrólise do ATP ocorre o desligamento do ácido nucleico e, assim, a separação das fitas duplas não é dependente da hidrólise de ATP.[3,5] Sobre essas mudanças conformacionais, algumas helicases, como a RNA helicase DEAD-Box-19, apresentam um mecanismo de auto regulação via estruturação de seu N-terminal que, na ausência do RNA, ocupa o sítio de ligação até que a molécula de RNA se ligue.[6]

Em virtude da gama de funções, essas proteínas apresentam múltiplas classificações, como a partir do principal ligante (RNA ou DNA) ou pela sua superfamília, para a qual existem seis no total (SV 1, 2, 3, 4, 5, 6). As superfamílias 1 e 2 se diferenciam das demais por serem monoméricas ao invés de hexâmeros que se estruturam em um formato distinto de anel. Todas

as helicases eucarióticas estão contidas nessas duas superfamílias.[3] A superfamília 2 é a maior e mais diversa das superfamílias, tendo o maior grupo representado pelas RNA helicases DEAD-box. Tal denominação foi dada pela sua preferência por RNA e por apresentarem, em seu domínio RecA1, uma sequência conservada Asp-Glu-Ala-Asp (DEAD) no motivo II. Além de apresentarem a sequência DEAD, diferenciam-se das demais helicases por ter um mecanismo diferente de separação das fitas duplas: ao invés de translocar pelas fitas, o processo ocorre localmente, assim elas apresentam uma preferência por duplexes com sequências menores, em torno de 12 nucleotídeos.[3,5]

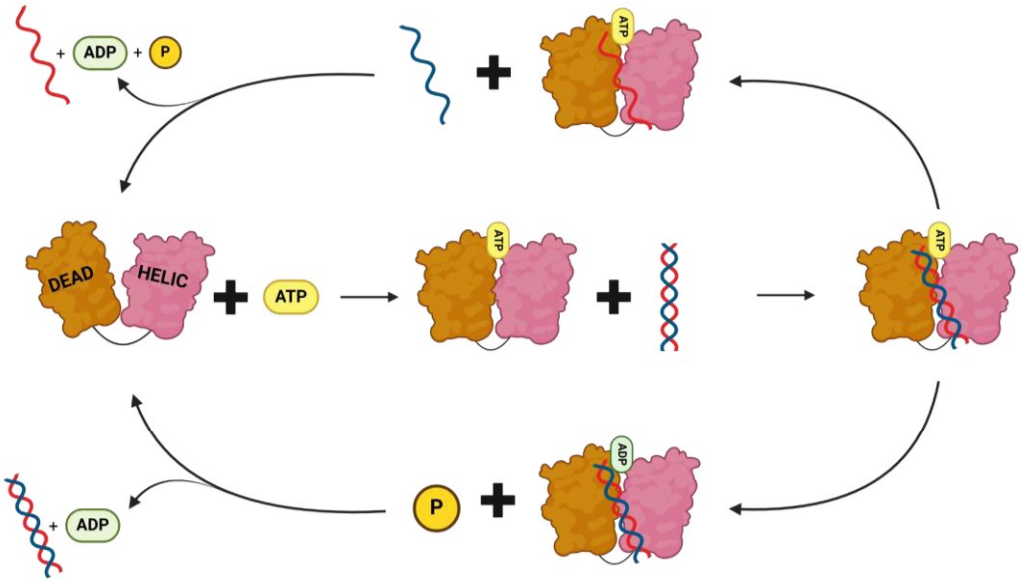


Figura 2 - Mecanismo de ligação de helicases com ATP e com ácido nucleico de dupla fita. Em um primeiro momento, é preciso que a proteína se ligue ao RNA para entrar em um estado conformacional fechado. Essa conformação é a preferencial para a ligação do ácido nucleico. Contudo, há duas possibilidades: haver a hidrólise do ATP sem o desenrolamento das fitas, e consequente liberação da fita dupla no chamado ciclo ocioso; ou haver a separação das fitas e liberação da fita separada e hidrólise do ATP e liberação da fita simples, o ciclo ativo. Elaborada pelo autor utilizando Biorender.

A DEAD-box RNA helicase 35 (NtDDX35 ou NtRh35 ou Rh35), foi identificada através da interação com a proteína SCI1, tendo seu nome devido à homologia com a Rh35 de *Arabidopsis thaliana*. Até o momento, a NtRh35 não foi caracterizada, sendo a helicase humana hDDX41 sua homóloga mais estudada, especialmente por seu envolvimento em supressão tumoral, além de apresentar funções na imunidade inata e um papel no *splicing* de pré-mRNA, fazendo parte do spliceossomo humano.[7] Desta forma, é possível inferir que a relação de NtRh35 e SCI1 possa ser relacionada ao processo de *splicing*. Para corroborar esta hipótese, a interação dessas proteínas, bem como a caracterização da NtRh35 precisam ser esclarecidas.

2. MATERIAIS E MÉTODOS

2.1. Análises *in silico*

A análise de domínios das proteínas estudadas foi realizada utilizando o servidor *SMART*.[8] Para o alinhamento, foi utilizado o *BlastP,* dentro da versão *web* do *Blast*, utilizando a base de dados de organismos modelo e de estruturas depositadas no *Protein Data Bank* (PDB).[9] Para análise do alinhamento e construção de árvore filogenética foi utilizado o software *Jalview*.[10] Além de análises de identificação de classes de domínios, foi realizada a predição de estruturas não ordenadas (predição de proteína intrinsecamente desordenadas), utilizando os serviços *GlobPlotWS*, *JRonn*, *IUPRED* e *DisemblWS* do pacote *JABAWS* dentro do *Jalview*.[11] Para a predição de estrutura terciária, foi utilizado o *AlphaFold Server*, a versão web do *AlphaFold* 3.[12] Os parâmetros de massa e absorbância foram preditos pelo *ProtParam* para cada proteína e para o complexo.[13]

2.2. Clonagem, expressão e purificação

Todos os procedimentos para clonagens, subclonagens e manipulação de DNA em geral foram conduzidos de acordo com as técnicas padrão de clonagem molecular descritas por Sambrook et al., 2001 e *Gibson assembly* .[14] Modelos gerados pelo *AlphaFold* 3 foram usados para definir as construções parciais, envolvendo os domínios isolados.[12] Os plasmídeos de expressão para as proteínas integrais (pET28a-NtRH35 e pET28a-SCI1) foram produzidos e cedidos pela profa. Maria Helena Goldman, da FFCLRP-USP. Células de *E. coli* Rosetta 2 (DE3) foram transformadas com tais plasmídeos (individualmente) e as proteínas recombinantes His-SCI1 e His-NtRH35 foram expressas em meio LB seletivo, após indução com 0.5 mM IPTG, por 16 horas, a 18 ºC. As culturas foram centrifugadas a 6000 xg por 10 minutos. Em seguida, as células foram ressuspensas em tampão A (50 mM Hepes, pH 8.0, contendo 0.5 M de cloreto de sódio (NaCl), 10% glicerol e 0.5 mM de Tris (2-carboxietil)fosfina (TCEP) e submetidas à lise celular por meio de um sonicador *Fisher Scientific - Sonic Dismembrator Model 500*, operando em 20% da potência por 4 minutos, sendo 15 segundos ativo e 30 segundos em descanso. O extrato resultante foi centrifugado a 20000 xg por 30 minutos a 4 ºC. A fração solúvel foi aplicada na coluna de afinidade *Ni-NTA superflow* (*Qiagen*). Após ciclo de lavagem com tampão A contendo 10 mM de imidazol, as proteínas foram eluídas em 250 mM imidazol no mesmo tampão. As alíquotas obtidas foram

analisadas por SDS-PAGE 12%. As frações que continham as proteínas de interesse foram submetidas a cromatografia de exclusão molecular em coluna *Superdex* 200 (10/300 GL, *GE Healthcare*), em sistema *AKTA Purifier* (*Cytiva*). Para o complexo, as culturas referentes a expressão de cada proteína foram combinadas e centrifugadas a 6000 xg por 10 minutos, antes de serem purificadas em coluna de afinidade, sob as mesmas condições descritas acima.

2.3. Análise de estado oligomérico

Para análise via espalhamento de luz multiangular (SEC-MALS), as alíquotas das proteínas purificadas foram analisadas, a fim de determinar o estado oligomérico de cada uma das amostras. O equipamento utilizado para o procedimento foi um detector de espalhamento *miniDAWN TREOS* e um refratômetro *Optilab T-rEX*, ambos devidamente acoplados a um sistema de cromatografia líquida de alta eficiência (HPLC) *Waters 2695*, *Separations Module*. Utilizando uma coluna *Superdex* 200 10/300 *Increase* (*GE Healthcare*), foram injetados 50 µL de cada amostra de proteína com concentração de 2 mg/mL, a um fluxo de 0,5 mL/min.

2.4. Dicroísmo circular (CD)

A fim de analisar a estrutura secundária das proteínas, medidas de CD foram realizadas em um espectropolarímetro *Jasco J-815*, equipado com sistema de controle de temperatura (*Peltier*). Para cada amostra, os espectros foram coletados em cubeta de quartzo retangular com caminho ótico de 1 mm, no intervalo de 195-270 nm, sob as configurações de velocidade de varredura 50 nm/min, largura da banda de 2 nm e tempo de resposta de 2 s. Os experimentos de desnaturação térmica foram feitos sob uma rampa de aquecimento de 1 °C/min, no intervalo de 5 a 70 °C, com coleta de espectro a cada 5 ºC e monitoramento em 222 nm a cada 2 ºC. Para o tratamento dos espectros, utilizou-se o programa *CDToolX*.[15] Com os espectros de CD ajustados para cada proteína isolada, a média dos espectros foi calculada e as barras de erro geradas, através de fórmulas de propagação de erro padrões, para gerar uma curva teórica para o complexo. Essa curva teórica representa o espectro CD esperado da mistura das proteínas caso não houvesse alguma alteração conformacional.

2.5. Termoforese em microescala (MST)

A termoforese é o fenômeno físico no qual se induz a migração de moléculas por meio de um gradiente de temperatura. A técnica de MST tem alta sensibilidade e permite a determinação de constantes de afinidade na faixa nanomolar-micromolar, usando volumes na ordem de poucos microlitros das amostras.[16]

A marcação das proteínas desejadas para os experimentos de MST foi realizada utilizando o fluoróforo N-hidroxisuccinimida(NHS) (*RED-NHS 2nd Generation, NanoTemper*), que se liga a aminas nas cadeias laterais de lisinas ou no N-terminal da proteína, a fim de avaliar a afinidade da helicase pelos diferentes nucleotídeos: AMP-PNP, ADP e ATP.

Em relação aos ensaios da NtRh35 com RNA, foi utilizado um RNA (poli-U 15) marcado com uma sonda de carboxifluoresceína (FAM), como sugerido em CHEN et al. (2022) [17], sendo testadas várias condições tanto na presença, como na ausência de nucleotídeos, como o AMP-PNP. Para esses experimentos foi utilizado o seguinte tampão: 25 mM Hepes pH 8.0, 5% Glicerol, 125 mM NaCl, 0.15% BSA, 2.5 mM AMP-PNP, 2.5 mM MgCl2, 0.05% Tween, 1 mM TCEP.

Os experimentos foram realizados em diluições seriadas das moléculas parceiras não marcadas e, só então, foram adicionadas as moléculas marcadas, mantidas em concentrações constantes. Os experimentos foram realizados a 25 °C, em triplicata experimental, no equipamento *NanoTemper Monolith NT.115 Blue/Red excitation* (*NanoTemper Technologies*), equipamento multiusuário localizado no IFSC, USP. Para análise dos dados, foi utilizado o próprio *software* do equipamento, o *MO.Control* e *MO.Analysis*.

2.6. Fluorescência intrínseca

A técnica de fluorescência intrínseca do triptofano é uma ferramenta sensível para investigar a estrutura e dinâmica de proteínas em várias condições experimentais. Neste trabalho, utilizou-se um espectrofluorímetro modelo *K2 ISS*, modo estático, para as medidas de fluorescência intrínseca do triptofano. As proteínas foram excitadas em comprimento de onda de 295 nm, sendo a emissão monitorada de 310 a 450 nm. As medidas foram realizadas em cubetas de quartzo com 1 cm de caminho óptico. As medidas foram feitas na presença e ausência do ligante no tampão, a fim de ser subtraído das medidas com a proteína para se retirar as contribuições do mesmo. Para esses experimentos foi utilizado o tampão semelhante

ao utilizado nos ensaios de MST, mas sem Tween e sem BSA, ou seja: 25 mM Hepes pH 8.0, 5% Glicerol, 125 mM NaCl, 2.5 mM AMP-PNP, 2.5 mM MgCl2, 1 mM TCEP.

3. RESULTADOS

3.1. Análises *in silico*

A análise de domínios para a NtRh35 (Fig. 3) teve como principal resultado a predição dos domínios DEAD e helicase nas posições 163-374 e 409-490, respectivamente. Além disso, houve a predição de duas regiões de baixa complexidade em seu N-terminal e também de um domínio "dedo de zinco" na posição 547-563, característico de proteínas que interagem com ácidos nucleicos.[18]

Através de uma busca com o servidor *BlastP,* usando a base de dados de organismos modelos, foi revelado que a Rh35 tem maior identidade com helicases já conhecidas como a DDX41 humana e a helicase *abstrakt* de *Drosophila melanogaster*. Assim, as sequências tiveram uma identidade maior que 50%, sendo a maior parte em seus domínios conservados. A maior identidade nas regiões N-terminais foi observada somente nas sequências filogeneticamente mais próximas: *Arabidopsis thaliana* e *Glycine max*.

Figura 3 - Predições de domínio pelo *SMART*. Em rosa, estão regiões de baixa complexidade, o losango *DEXDc* representa o domínio DEAD, o quadrado *HELICc* representa o domínio helicase e o retângulo *ZnF_C2HC* representa o domínio dedo de zinco.

Em buscas no banco de dados envolvendo proteínas com estruturas depositadas no PDB, a única semelhança de *E value* nulo foi com a DDX41 dentro de um complexo spliceosomal humano, junto a demais proteínas de interação (PDB 8C6J). No entanto, outras estruturas referentes a essa helicase humana foram encontradas, mas correspondentes a domínios separados. As demais estruturas encontradas são bem mais distantes da Rh35, apresentando identidades inferiores a 50%.

Já para a SCI1, não houve predição de domínio. Porém, na base de dados de organismos modelo no *BlastP*, SCI1 apresentou identidade somente com sequências de

Arabidopsis thaliana e *Glycine max*. Com relação aos resultados do PDB, nenhuma estrutura com identidade significativa a SCI1 foi encontrada. Com base nesses resultados, sua sequência foi submetida à predição de estruturas não ordenadas (Tabela 1). De forma geral, esses resultados predizem a estrutura de SCI1 como majoritariamente não ordenada.

Tabela 1- Predições de domínios globulares/ estrutura ordenada utilizando os diversos pacotes de predição, que estão incluídos dentro do *plugin JABAWS* incluído dentro do *software Jalview*.

Programas de predição	*GlobPlotWS*	*JRonn*	*IUPRED*	*DisemblWS*
Trechos da sequência da SCI1	27-158	94-149; 153-157	95-155	97-105; 117-131

Nas análises utilizando o *AlphaFold*, realizadas para as proteínas isoladas, o modelo obtido para a NtRh35 mostrou domínios bem definidos e estruturados: o domínio DEAD e o domínio helicase, com predominância de alfa-hélices (Fig 4a). Já para a SCI1, o modelo mostrou uma estrutura majoritariamente desordenada (Fig. 4b), concordando com as predições mostradas anteriormente. Esta característica é compatível com a função de SCI1 na regulação gênica, bem como com o grande número de parceiros já observados pelo seu interactoma.[2] Finalmente, no modelo AlphaFold gerado para o complexo apresentou uma grande confiabilidade na região de interação, região que envolve o C-terminal de ambas as proteínas (Fig. 4c), corroborando o que já foi descrito por *pulldown in vitro*.[19] As análises também previram com confiabilidade uma razão de interação 1:1 para o complexo SCI1:NtRh35, mesmo quando mais de uma unidade de cada molécula foi submetida à predição.

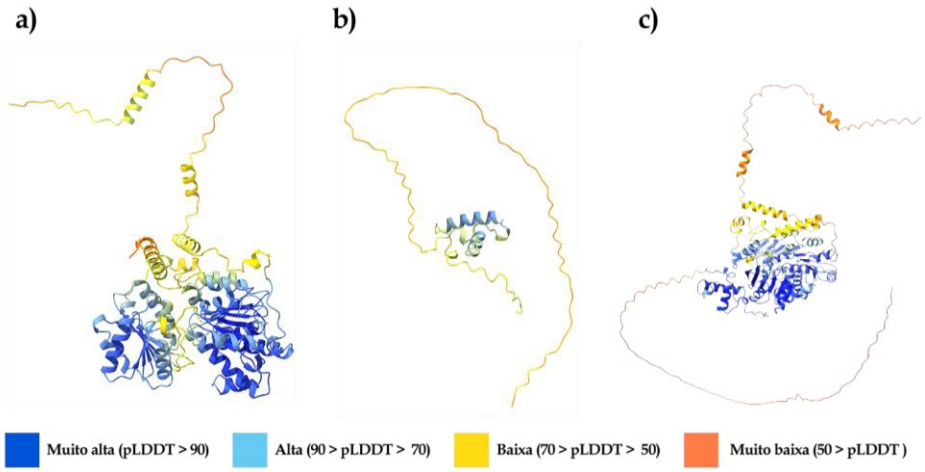


Figura 4- Estruturas preditas pelo AlphaFold, coloridas de acordo com o perfil de confiabilidade do programa para a Rh35 (a), SCI1 (b) e complexo (c).

Em relação às predições do *AlphaFold* para a NtRh35 com RNA, uma sequência de poli-U 15 (a mesma utilizada para os experimentos de MST) foi usada para realizar a análise. Primeiramente, percebeu-se que haviam poucas alterações na estrutura da helicase ao submetê-la a diferentes contextos, como na presença de nucleotídeo e de RNA (Fig.5a). Assim, constatou-se que a predição priorizava a conformação fechada da helicase. Além disso, também não houve mudanças na NtRh35 com RNA e SCI1, provavelmente porque o padrão de interação do RNA para helicases é na região entre os domínios. Ainda, a interação observada do RNA foi com os grupos fosfato e pentose, mas não com a base nitrogenada, o que já foi descrito ocorrer para outras RNA helicases, e o que, consequentemente, permite também interação com DNA.[3,4,5]

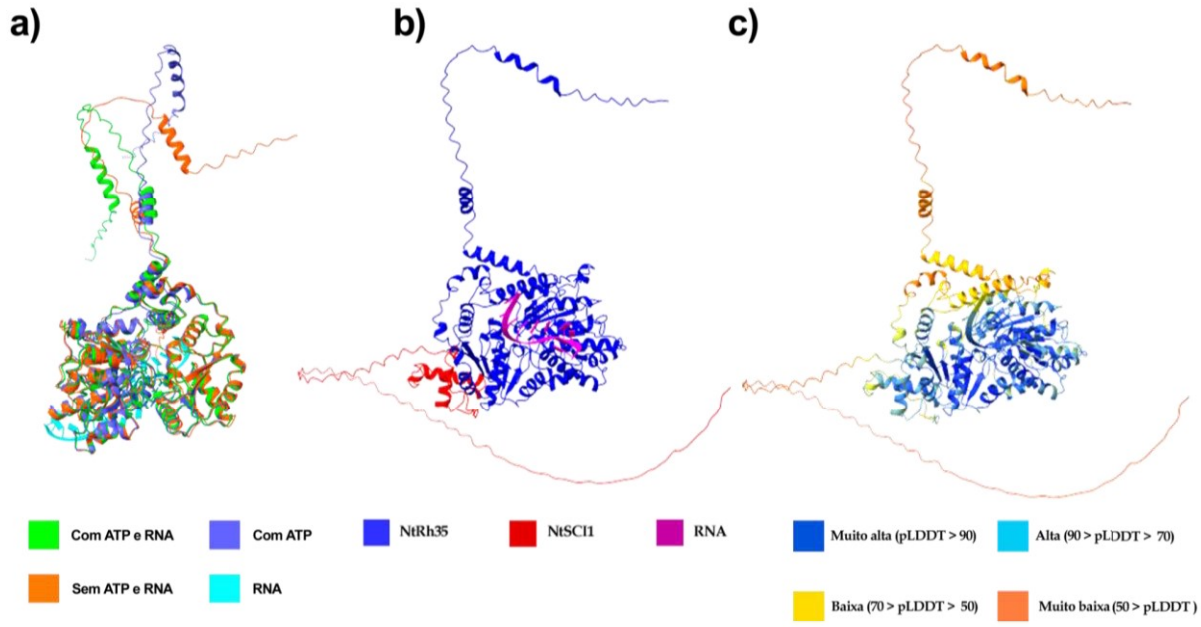


Figura 5 - Predições no *AlphaFold* para a Rh35 na presença do RNA. a) Sobreposição por alinhamento das diferentes estruturas preditas para a Rh35. b) e c) Predições para o complexo na presença de ATP e RNA, coloridas de acordo com a molécula (b) e por confiabilidade (c).

3.2. Expressão e Purificação

Após a indução da expressão em *E. coli*, o processo de lise e purificação da NtRh35 foi iniciado utilizando um tampão com baixo sal e glicerol, mas este se demonstrou inadequado. A proteína ficou em grande parte na fração insolúvel e diversas proteínas estavam sendo co-purificadas com a NtRh35. Assim, após revisão da literatura para outras helicases,

foi observado que o processo típico de purificação dessas proteínas utiliza altas concentrações de sal, de 0,5 a 1 M, a fim de eliminar as interações com outras moléculas, incluindo DNA, RNA e nucleotídeos.[17] Além disso, foi adicionado o agente redutor TCEP em uma concentração de 0,5 mM na tentativa de estabilizar a proteína, visto que a NtRh35 apresenta 11 cisteínas em sua sequência. Após as alterações, a purificação em coluna de afinidade mostrou redução de contaminantes comparado ao que se tinha inicialmente. Foi testado, também, utilizar uma terceira lavagem com 30 mM de imidazol (figura 6a) e uma segunda lavagem com 0,1% de Triton, mas não foram efetivas em melhorar a pureza. Uma alternativa a ser ensaiada futuramente será uma lavagem extra com 1 M de NaCl e substituir a coluna de cromatografia de exclusão molecular (SEC) usando a *Superdex* 75, que apresenta uma resolução melhor de separação que a *Superdex* 200, considerando a massa dos contaminantes em relação à NtRh35.

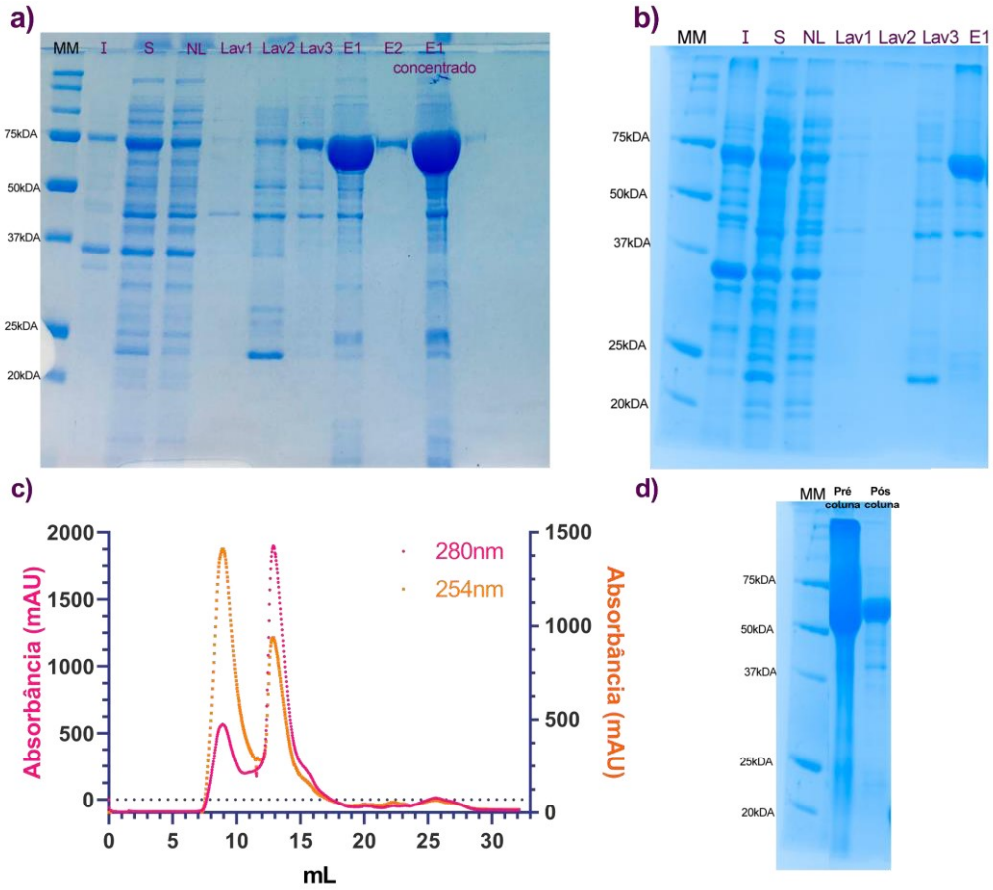


Figura 6 - Purificação da NtRh35 em coluna de níquel (a e b) e em coluna de exclusão molecular *Superdex* 200 10/300 GL (c e d). c) Cromatograma obtido para amostra apresentada em d). De a) a d): MM- Massa molecular, I - Fração insolúvel, S - Fração solúvel, NL- Extrato bruto não ligado na coluna, Lav1- Primeira lavagem com o tampão de lise, Lav2- lavagem com mesmo tampão acrescido de 10 mM de imidazol em a) e 0,01% Triton em b), Lav3- lavagem com mesmo tampão acrescido de 30 mM de imidazol em a) e 10 mM de imidazol em b), E1-

Eluição com 8 mL com tampão com 250 mM de imidazol, E2- Segunda eluição com 5mL de tampão com 250 mM de imidazol.

Em relação a purificação na SEC, sob baixas concentrações de sal, a Rh35 se mostrou instável quando concentrada. Entretanto, a Rh35 em solução mais diluída mostrou um pico mais definido na SEC quando o tampão continha menos de 250 mM de sal (como a amostra que foi utilizada para dicroísmo circular). Assim, ficou determinado que quando as soluções proteicas exigidas eram de baixa força iônica, a proteína seria purificada em alto sal e depois diluída em tampão similar para redução do sal.

Para a SCI1, o mesmo tampão e método para purificação usados para Rh35 foram utilizados, porém sem adição de TCEP. SCI1 também apresenta uma maior quantidade de contaminantes que não puderam ser retirados pela SEC na *Superdex 200* e, assim, esta coluna deverá também ser substituída pela *Superdex 75*.

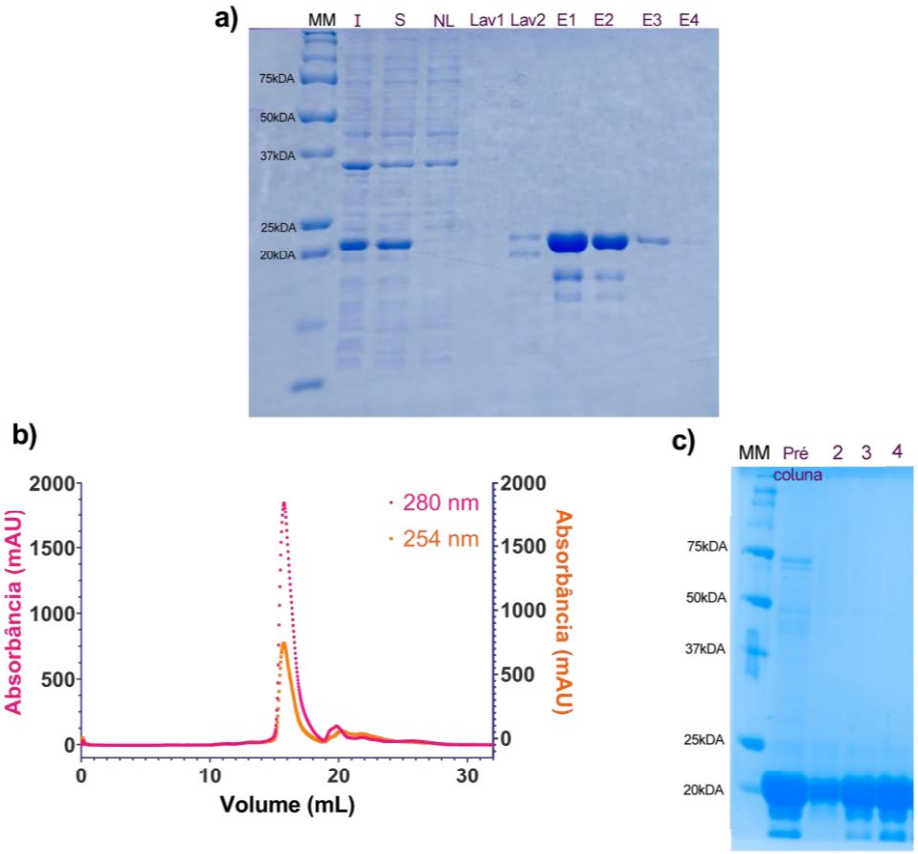


Figura 7 - Purificação da SCI1 em coluna de níquel (a) e em coluna de exclusão molecular *Superdex* 200 10/300 GL (b e c). Legenda em a) e MM- Massa molecular, I - Fração insolúvel, S - Fração solúvel, NL- Extrato bruto não ligado na coluna, Lav1- Primeira lavagem com o tampão A, Lav2- lavagem com mesmo tampão acrescido de 25 mM de imidazol, E1-E4 - Eluições com tampão com 0,5 M de imidazol; c) colunas 2 a 4 referem-se às amostras coletadas ao longo do maior pico de eluição.

Para a expressão e purificação do complexo, os mesmos procedimentos das proteínas isoladas foram utilizados, porém as células provenientes de cada expressão foram misturadas pré-centrifugação. Como ambas Rh35 e SCI1 apresentavam a *His-tag*, a co-eluição a partir da cromatografia de afinidade não permitiu avaliar se havia a formação de um complexo. Desta forma, somente após a co-eluição da cromatografia de exclusão molecular foi possível verificar que o complexo estava sendo formado. Entretanto, para confirmar este resultado, uma amostra eluída da cromatografia de exclusão molecular foi submetida à análise por SEC-MALS.

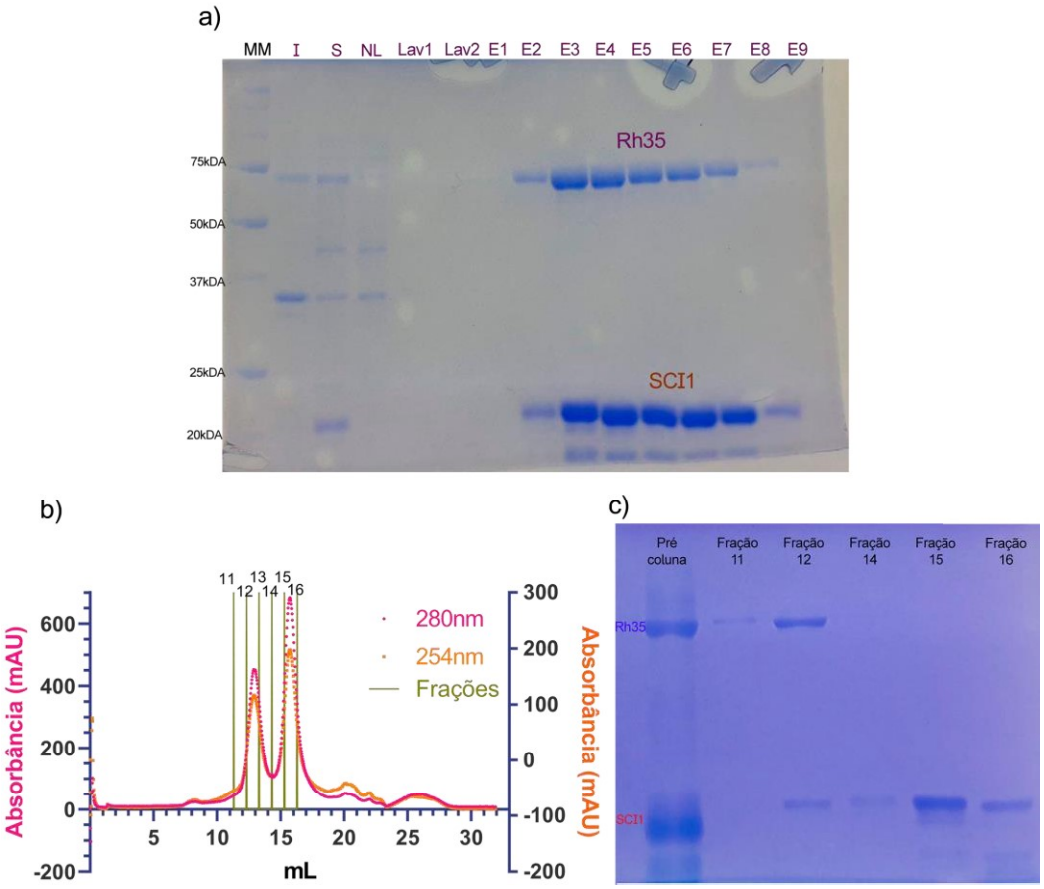
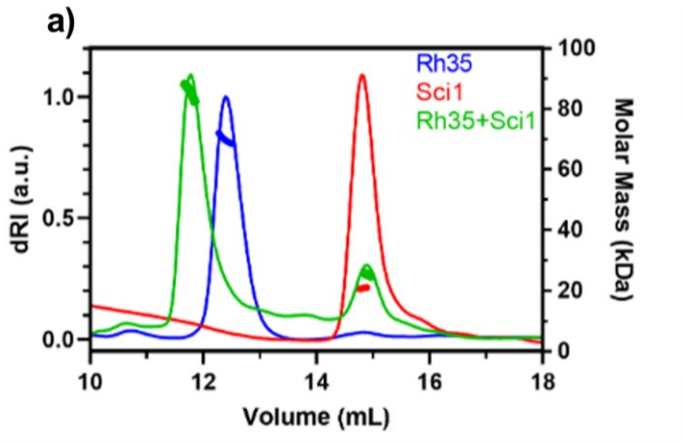


Figura 8 -Análise e purificação do complexo Rh35-SCI1. (a) Purificação do complexo por cromatografia de afinidade e (b e c) por cromatografia de exclusão molecular *Superdex* 200 10/300 GL. Cromatograma obtido em (b) para amostra apresentada em c). Legenda: MM- Massa molecular, I - Fração insolúvel, S - Fração solúvel, NL- Extrato bruto não ligante na coluna, Lav1- Primeira lavagem com o tampão de lise, Lav2- lavagem com mesmo tampão acrescido de 250 mM de imidazol, E1-E9 - Eluições com 1,5 mL com tampão com 250 mM de imidazol, Pré Coluna - Amostra antes de ser aplicada na exclusão molecular, Frações 11-16- Amostras referentes ao fracionamento dos picos apresentado em b).

3.3. SEC-MALS

As amostras purificadas foram analisadas por SEC-MALS para avaliação dos estados oligoméricos. Nesta análise (Fig. 9), ambas proteínas, quando isoladas, comportaram-se como

monômeros e com massas concordantes com as estimadas. Porém, quando foram co-purificadas, passaram a interagir mesmo em condições de alta força iônica (tampão com 500mM de NaCl), eluindo como um complexo. Através da massa obtida, foi possível determinar que o complexo formado entre NtRh35 e NtSCI1 se apresentava numa razão 1:1, assim como também foi predito pelo modelo AlphaFold. O resultado reforça os resultados da literatura obtidos via duplo-híbrido (Y2H), *pulldown* e BiFC (*Bimolecular fluorescence Complementation*).[19]



b)

Proteínas	Massa esperada (ProtParam)	Massa obtida
Rh35	68472,11	69850 ± 0,24%
SCI1	20825,63	20820 ± 8,82%
Rh35 + SCI1	88807,73	85360 ± 0,41%

Figura 9- Análise por SEC-MALS para as amostras de NtRh35, SCI1 e o complexo. Em a) as linhas finas correspondem ao índice de refração diferencial (dRi), sendo as linhas grossas referentes as massas moleculares. Em b) são mostrados os valores esperados e obtidos para a massa das proteínas em Daltons.

3.4. Dicroísmo circular (CD)

Para avaliar a estrutura secundária das proteínas em estudo, espectros de CD foram medidos a 4 ºC, em tampão de 20 mM fosfato de sódio pH 8 e 125 mM NaCl. Na figura 10a, o espectro da Rh35 mostra um perfil de estrutura secundária com predomínio de alfa-hélices, o que indica que a proteína possui estrutura secundária compatível com aquela predita pelo *Alphafold*. Com intuito de avaliar a estabilidade térmica, amostras da NtRh35 foram também submetidas à desnaturação por calor, com aquisição dos espectros a cada etapa de aquecimento. A curva obtida é apresentada na figura 10b, na qual se observa uma temperatura de *melting* de aproximadamente 57 ºC para a NtRH35.

O perfil de CD exibido pela SCI1 (Fig. 11a) é bem distinto do observado para a NtRh35, sendo característico de estruturas secundárias não ordenadas. Em relação a sua estabilidade térmica (Fig. 11b), ao longo do gradiente de temperatura não houve alteração detectável no espectro quando analisado em 222 nm. O espectro CD de uma proteína

intrinsecamente desordenada (IDP), normalmente é dominado pelas características de estruturas *random-coil*, com uma elipticidade negativa perto de 200 nm, por uma elipticidade próxima de zero em 185 nm e sendo o mínimo em 222 nm não pronunciado.[20] Portanto, os resultados obtidos aqui concordam com as predições *in silico* de que SCI1 exibe uma estrutura não ordenada.

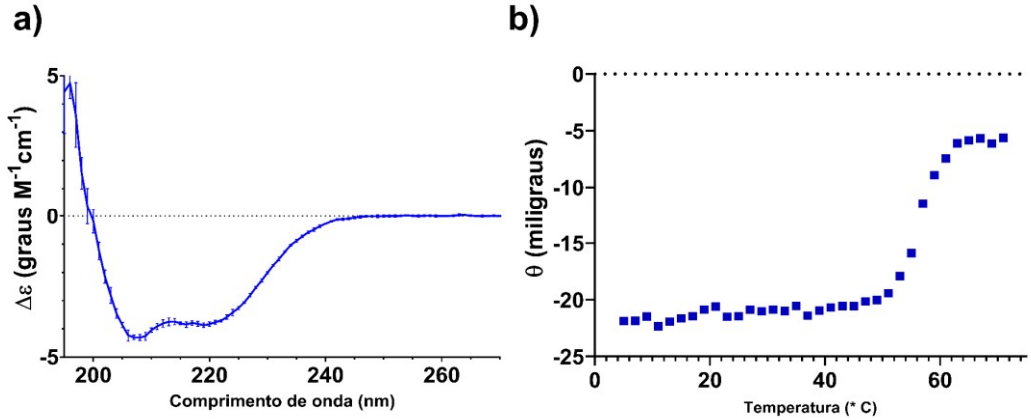


Figura 10- Análises de CD para a Rh35. a) Espectro da Rh35 em uma concentração de 0,2 mg/mL, a 4 ºC, em fosfato de sódio 20 mM contendo 125 mM de NaCl, em uma cubeta de caminho óptico de 1 mm. b) Curva de desnaturação térmica obtida após variar a temperatura, monitorando o comprimento de onda de 222 nm.

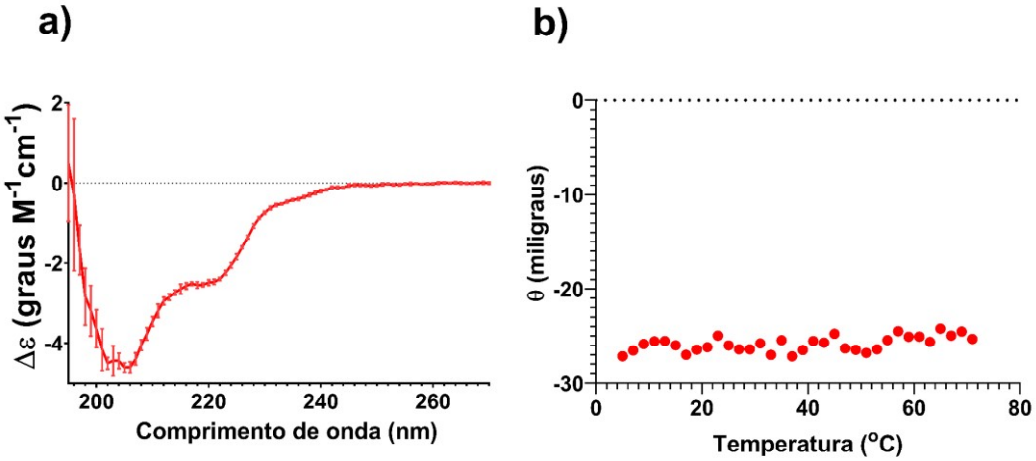


Figura 11- Análises de CD para a SCI1. a) Perfil da SCI1 a 4 oC em uma concentração de 0,23 mg/mL, em fosfato de sódio 20 mM contendo 125 mM de NaCl em uma cubeta de caminho óptico de 1 mm. b) Curva de desnaturação térmica obtida após variar a temperatura, fixando o comprimento de onda de 222 nm.

Com relação ao perfil de CD exibido pelo complexo (Fig. 12), este é bem parecido com o observado para a Rh35, possivelmente devido à predominância do sinal de alfa-hélice. O espectro experimental obtido para o complexo, quando comparado com o teórico dado pelas médias dos espectros das proteínas isoladas, apresenta um mínimo em 222 nm mais negativo,

embora em 208 nm essa diferença não seja visível. Essa medida é preliminar ainda, mas sugere que alguma alteração conformacional (detectável pelo CD) está ocorrendo com a formação do complexo.

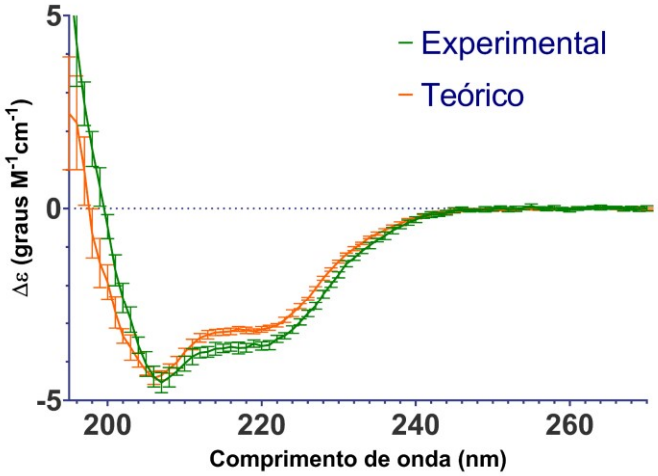


Figura 12 - Perfil de CD do complexo a 4 oC em uma concentração de 0,166 mg/mL, em fosfato de sódio 20 mM contendo 125 mM de NaCl em uma cubeta de caminho óptico de 1 mm. Em verde tem-se o resultado obtido, já em laranja, o calculado pela média entre os espectros da SCI1 e da Rh35.

3.5. Termoforese em Microescala (MST)

Para os experimentos de MST, um grande esforço foi realizado inicialmente, a fim de encontrar as condições ideais de tampão para a proteína NtRh35 e, ao mesmo tempo, viabilizar os experimentos de MST. Embora muitas helicases já tenham sido descritas na literatura, a NtRh35 em especial ainda não foi estudada, nem suas homólogas em plantas. Sua homóloga mais bem estudada é a DDX41 humana, mas essa helicase também não está completamente caracterizada em termos de afinidade e interação com RNA, dificultando determinar parâmetros de partida para os experimentos.[7,21]

Cientes da dependência de nucleotídeos para a interação entre helicases e o RNA, decidiu-se iniciar os ensaios com a avaliação da interação da NtRh35 com nucleotídeos (ATP, AMP-PNP e ADP).[3] Num primeiro experimento, a NtRh35 foi marcada via aminas primárias, sendo que a ligação dessa sonda ocorre em resíduos de Lisina. A partir da marcação, percebeu-se que a proteína apresentava bom rendimento com essa marcação, possivelmente por ter 49 lisinas em sua sequência. Contudo, durante os experimentos, percebeu-se adsorção nos capilares do tipo *Standard (Nanotemper),* mesmo em condições na presença de BSA, e não houve razão sinal-ruído suficiente para realizar as medidas de afinidade. Assim, checando as

predições das regiões de interação da NtRh35 com o nucleotídeo, viu-se que existem lisinas nos motivos responsáveis pela ligação com o nucleotídeo e na literatura recomenda-se não utilizar a marcação se houver lisinas no sítio de interação.[5,21] Desse modo, a marcação provavelmente interferiu na interação com os ligantes testados, inviabilizando a medida. Assim, é recomendado que a marcação da proteína seja alterada, sendo uma das opções a sonda que se liga à cauda de histidinas (*RED-tris-NTA 2nd Generation, Nanotemper*).

Já para os experimentos de interação com RNA, o monitoramento do RNA se torna mais vantajoso por se resumir a ligado ou não a proteína, enquanto isso, se fosse marcada e monitorada a proteína, poderia se obter um sinal conflitante por se ter tanto a afinidade pelo nucleotídeo quanto pelo ácido nucleico. Para estes ensaios foi utilizado um análogo de ATP não hidrolisável (AMP-PNP), que se trata de uma molécula de ATP com um de seus fosfato substituído por um nitrogênio, visto que as helicases DEAD-BOX hidrolisam ATP após se ligarem ao RNA, o que provavelmente geraria um sinal ruidoso e não confiável.[17] Como se tratava de um primeiro experimento sem parâmetros de comparação com helicases homólogas, iniciou-se com altas concentrações da proteína, assim, por isso o uso de aditivos como o glicerol e o TCEP.

No experimento de MST é necessário, antes do experimento de fato, realizar a checagem de ligação. Por esta, não foi vista uma razão sinal-ruído consistente entre a medida na presença e na ausência de nucleotídeo. Nessa situação, não foi detectada adsorção nos capilares, mas o sinal de resposta foi baixo. Em outras tentativas, também não se viu diferenças entre as medições sem e com nucleotídeo. Assim, o experimento precisa ser reformulado para que as condições necessárias para detecção sejam atingidas.

3.6. Fluorescência Intrínseca

Numa análise do modelo estrutural, foi observado que um triptofano, dos dois resíduos presentes na NtRh35, situa-se próximo do sítio de ligação ao ATP, permitindo monitorá-lo como sonda intrínseca da interação. Assim, alternativamente, a técnica de emissão de fluorescência intrínseca foi usada para avaliar qualitativamente a ligação de nucleotídeos à NtRh35, já que esta técnica dispensa marcação. Os resultados mostram uma supressão da emissão de fluorescência do triptofano da NtRh35 quando na presença do nucleotídeo (Fig. 14), indicando que o ATP está interagindo nestas condições.

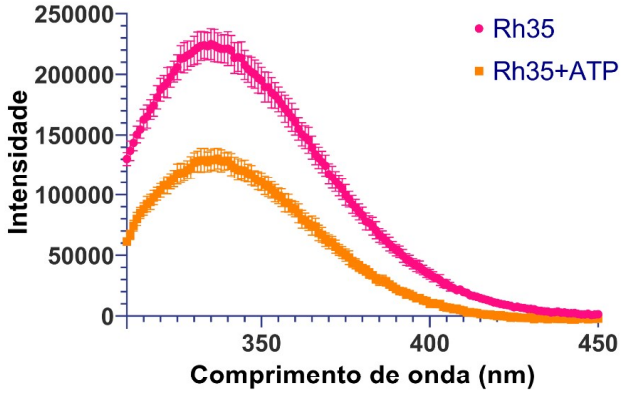


Figura 14 - Emissão de fluorescência da Rh35 na ausência e presença de 2,5 mM de ATP. Excitação a 295 nm. Amostra em tampão 25 mM de Hepes pH 8, 2,5 mM ATP, 2,5 mM MgCl2, 5 % glicerol, 1 mM TCEP.

3.7. Construções dos domínios proteicos

Visando resolver a estrutura das regiões estruturadas do complexo, iniciou-se experimentos visando produzir construções truncadas de ambas proteínas (Fig 15a), sendo a Rh35 *cut* (369-53, que corresponde ao domínio helicase em sua maior parte); e a SCI1 *cut* (82-158, região correspondente ao seu C-terminal). Os *primers* foram desenhados para realizar a clonagem via técnica de *Gibson Assembly* em vetor pSUMO. Os resultados das amplificações demonstram bandas únicas para ambos os insertos (Fig. 15b), com o inserto da Rh35 *cut* com 533 pb e da SCI1 *cut* com 267 pb. Já para os vetores, foram observadas várias amplificações inespecíficas (Fig. 15c), possivelmente devido às baixas temperaturas de anelamento dos *primers*, mas que não prejudicaram a purificação do mesmo. Até o momento, somente a nova construção de SCI1 *cut* foi obtida e confirmada por análise de restrição (Fig. 15d), restando pendente o teste de expressão, sequenciamento e também a construção da Rh35 *cut*.

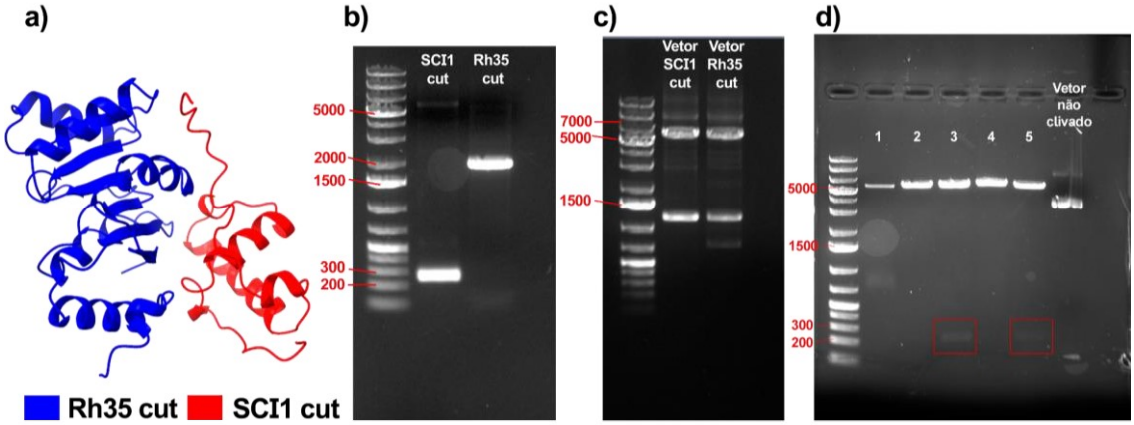


Figura 15 - a) Estrutura predita pelo *AlphaFold* para o complexo com as construções da Rh35 *cut* e a SCI1 *cut*. b) Resultado da PCR para os insertos, mostrado por eletroforese em gel de agarose. c) Resultado da PCR para os vetores. d) Teste de restrição utilizando as enzimas BamHI e XhoI para 5 colônias além do vetor original não clivado como referência. Em vermelho, estão as bandas que tiveram o inserto aparente no resultado da clivagem.

4. CONCLUSÕES E CONSIDERAÇÕES FINAIS

As condições de expressão para NtRh35 e SCI1 foram bem estabelecidas, viabilizando a purificação e obtenção das proteínas, particularmente de NtRh35, bem como do complexo. O bom rendimento tem permitido as análises biofísicas, mas a pureza da amostra final pode ainda ser melhorada. Em relação à estabilidade da NtRh35, as condições do tampão e concentração de sal foram definidas para estabilização da proteína. Porém, tem-se ainda como perspectiva realizar ensaios de Fluorimetria Diferencial de Varredura (em inglês, DSF), a fim de avaliar sua Temperatura de *melting* (Tm) em diferentes condições e avaliar via Espalhamento Dinâmico de Luz (em inglês, DLS). A partir disso, pretende-se estabelecer condições mais promissoras para os ensaios de MST com os ligantes (RNA e nucleotídeos). Um dado preliminar interessante foi dado pela técnica de fluorescência, mostrando que ATP interage com NtRH35, abrindo caminho para medidas quantitativas dessa interação. Para melhor caracterizar a NtRh35, atividades ATPásica e helicásica, na qual se monitora a atividade de separação de fitas duplas de ácidos nucleicos, também pretendem ser realizadas.

Para a SCI1, os resultados preditos pelas análises *in silico* e resultados de CD em conjunto mostram que se trata de uma proteína intrinsecamente desordenada, dado que é compatível com as funções na proliferação celular com as quais está envolvida.

Ainda como resultado deste trabalho, o complexo NtRh35-SCI1 pode ser confirmado via SEC-MALS e essa interação se mostrou independente da adição de nucleotídeos ou de RNA, sugerindo tratar-se de uma interação forte, visto que alta força iônica não foi suficiente para

desfazê-la. Para obtenção de novas informações estruturais sobre o complexo, planejam-se experimentos para avaliar possíveis alterações conformacionais na presença de ligantes ou em variações de pH e sal, além de averiguar a estabilidade térmica desse complexo.

Sobre a técnica de MST, embora seja muito versátil e rápida, neste projeto a interação com RNA ainda não foi possível de ser medida por esta técnica, sendo que as condições experimentais precisam ser ajustadas.

REFERÊNCIAS

1) CRUZ, Joelma O.; MARTIN, Juca A. B. San; LUBINI, Greice; STRINI, Edward J.; SOBRAL, Rómulo; PINOTI, Vitor F.; FERREIRA, Pedro B.; THOMÉ, Vanessa; QUIAPIM, Andréa C.; DORNELAS, Marcelo C.; PRANCHEVICIUS, Maria Cristina S; MADUEÑO, Francisco; COSTA, M Manuela R; GOLDMAN, Maria Helena S. SCI1 Is a Direct Target of AGAMOUS and WUSCHEL and Is Specifically Expressed in the Floral Meristematic Cells. **Frontiers In Plant Science**, v. 12, 18 mar. 2021. Frontiers Media SA. http://dx.doi.org/10.3389/fpls.2021.642879.
2) STRINI, Edward José. A análise do interactoma de SCI1 (Stigma/Style Cell Cycle Inhibitor 1) revela possíveis mecanismos de controle da proliferação celular. 2014. Tese (Doutorado em Genética) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2014. doi:10.11606/T.17.2014.tde-21052014-103120. Acesso em: 2024-06-02
3) JANKOWSKY, Eckhard. RNA helicases at work: binding and rearranging. **Trends In Biochemical Sciences**, v. 36, n. 1, p. 19-29, jan. 2011. Elsevier BV. http://dx.doi.org/10.1016/j.tibs.2010.07.008
4) LEITÃO, Ana; COSTA, Marina; ENGUITA, Francisco. Unzippers, Resolvers and Sensors: a structural and functional biochemistry tale of RNA helicases. **International Journal Of Molecular Sciences**, v. 16, n. 2, p. 2269-2293, 22 jan. 2015. MDPI AG. http://dx.doi.org/10.3390/ijms16022269.
5) LINDER, Patrick; JANKOWSKY, Eckhard. From unwinding to clamping — the DEAD box RNA helicase family. **Nature Reviews Molecular Cell Biology**, v. 12, n. 8, p. 505-516,

22 jul. 2011. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/nrm3154.

6) COLLINS, Ruairi et al. The DEXD/H-box RNA Helicase DDX19 Is Regulated by an α-Helical Switch. **Journal Of Biological Chemistry**, v. 284, n. 16, p. 10296-10300, abr. 2009. Elsevier BV. http://dx.doi.org/10.1074/jbc.c900018200.
7) WINSTONE, Lacey; JUNG, Yohan; WU, Yuliang. DDX41: exploring the roles of a versatile helicase. **Biochemical Society Transactions**, v. 52, n. 1, p. 395-405, 13 fev. 2024. Portland Press Ltd. http://dx.doi.org/10.1042/bst20230725.
8) LETUNIC, Ivica; KHEDKAR, Supriya; BORK, Peer. SMART: recent updates, new developments and status in 2020. **Nucleic Acids Research**, v. 49, n. 1, p. 458-460, 26 out. 2020. Oxford University Press (OUP). http://dx.doi.org/10.1093/nar/gkaa937.
9) JOHNSON, M.; ZARETSKAYA, I.; RAYTSELIS, Y.; MEREZHUK, Y.; MCGINNIS, S.; MADDEN, T. L. NCBI BLAST: a better web interface. **Nucleic Acids Research**, v. 36, p. W5-W9, 19 maio 2008. Oxford University Press (OUP). http://dx.doi.org/10.1093/nar/gkn201.
10) WATERHOUSE, Andrew M.; PROCTER, James B.; MARTIN, David M. A.; CLAMP, Michèle; BARTON, Geoffrey J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, v. 25, n. 9, p. 1189-1191, 16 jan. 2009. Oxford University Press (OUP). http://dx.doi.org/10.1093/bioinformatics/btp033.
11) TROSHIN, Peter V; PROCTER, James B; SHERSTNEV, Alexander; BARTON, Daniel L; MADEIRA, Fábio; BARTON, Geoffrey J. JABAWS 2.2 distributed web services for Bioinformatics: protein disorder, conservation and RNA secondary structure. **Bioinformatics**, v. 34, n. 11, p. 1939-1940, 30 jan. 2018. Oxford University Press (OUP). http://dx.doi.org/10.1093/bioinformatics/bty045.
12) ABRAMSON, Josh et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. **Nature**, v. 630, n. 8016, p. 493-500, 8 maio 2024. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/s41586-024-07487-w.
13) PONTING, Chris P.; BIRNEY, Ewan. Protein Sequence Analysis and Domain Identification. **The Proteomics Protocols Handbook**, p. 527-541, 2005. Humana Press. http://dx.doi.org/10.1385/1-59259-890-0:527.
14) Gibson, D. G. (2011). “Enzymatic assembly of overlapping DNA fragments,” in **Methods in Enzymology** (Elsevier), 349–361. doi:10.1016/B978-0-12-385120-8.00015-2

15) MILES, Andrew J.; WALLACE, B. A. CDtoolX, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. **Protein Science**, v. 27, n. 9, p. 1717-1722, set. 2018. Wiley. http://dx.doi.org/10.1002/pro.3474.
16) SEIDEL, Susanne A.I. et al. Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. **Methods**, v. 59, n. 3, p. 301-315, mar. 2013. Elsevier BV. http://dx.doi.org/10.1016/j.ymeth.2012.12.005.
17) CHEN, Zijun; HUANG, Jinqing; LI, Jixi. Protein purification, crystallization, and structure determination of human DEAD-box RNA helicase DDX21 in different unwinding states. **Star Protocols**, v. 3, n. 3, p. 101642, set. 2022. Elsevier BV. http://dx.doi.org/10.1016/j.xpro.2022.101642.
18) WANG, Yishu; YU, Yu; PANG, Yidan; YU, Haojun; ZHANG, Wenqi; ZHAO, Xian; YU, Jianxiu. The distinct roles of zinc finger CCHC-type (ZCCHC) superfamily proteins in the regulation of RNA metabolism. **Rna Biology**, v. 18, n. 12, p. 2107-2126, 4 maio 2021. Informa UK Limited. http://dx.doi.org/10.1080/15476286.2021.1909320.
19) PINOTI, Vitor F et al. SCI1, a flower regulator of cell proliferation, and its partners NtCDKG2 and NtRH35 interact with the splicing machinery. **Journal Of Experimental Botany**,8 ago. 2024. Oxford University Press (OUP). http://dx.doi.org/10.1093/jxb/erae337.
20) SCHRAMM, Antoine; BIGNON, Christophe; BROCCA, Stefania; GRANDORI, Rita; SANTAMBROGIO, Carlo; LONGHI, Sonia. An arsenal of methods for the experimental characterization of intrinsically disordered proteins – How to choose and combine them? **Archives Of Biochemistry And Biophysics**, v. 676, nov. 2019. Elsevier BV. http://dx.doi.org/10.1016/j.abb.2019.07.020.
21) ANDREOU, Alexandra Z. DDX41: a multifunctional dead-box protein involved in pre-mrna splicing and innate immunity. **Biological Chemistry**, v. 402, n. 5, p. 645-651, 15 mar. 2021. Walter de Gruyter GmbH. http://dx.doi.org/10.1515/hsz-2020-0367.
22) RAINARD, Julie M.; PANDARAKALAM, George C.; MCELROY, Stuart P. Using Microscale Thermophoresis to Characterize Hits from High-Throughput Screening: a European lead factory perspective. **Slas Discovery**, v. 23, n. 3, p. 225-241, mar. 2018. Elsevier BV. http://dx.doi.org/10.1177/2472555217744728.
23) Nanotemper. Support: Monolith. Disponível em: <https://support.nanotempertech.com/hc/en-us/categories/17715099497873-Monolith >