UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

RAISSA FERREIRA GUTIERREZ

Estudos funcionais de CrNIP7 de *Chlamydomonas reinhardtii*: uma
proteína envolvida na biogênese de ribossomos

São Carlos
2016

RAISSA FERREIRA GUTIERREZ

Estudos funcionais de CrNIP7 de *Chlamydomonas reinhardtii*: uma
proteína envolvida na biogênese de ribossomos

Dissertação apresentada ao Programa de Pós-
Graduação em Física do Instituto de Física de
São Carlos da Universidade de São Paulo, para
obtenção do título de Mestre em Ciências.

Área de concentração: Física Aplicada
Opção: Física Biomolecular

Orientador: Prof^a. Dr^a. Ana Paula Ulian de
Araújo

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos
2016

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.

Ficha catalográfica revisada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Gutierrez, Raissa Ferreira
Estudos funcionais de CrNIP7 de *Chlamydomonas reinhardtii*: uma proteína envolvida na biogênese de ribossomos / Raissa Ferreira Gutierrez; orientadora Ana Paula Ulian de Araújo - versão corrigida -- São Carlos, 2016.
141 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2016.

1. *Chlamydomonas reinhardtii*. 2. Biogênese de ribossomo. 3. CrNIP7. I. Araújo, Ana Paula Ulian de, orient. II. Título.

Com muito amor e gratidão, dedico este trabalho à minha mãe, Cláudia Maria, por ter me ensinado o prazer pelos estudos e pelo conhecimento.

AGRADECIMENTOS

À Profa. Dra. Ana Paula Ulian de Araújo pela orientação, amizade e confiança depositada no meu trabalho.

Ao Prof. Dr. Nilson Ivo Tonin Zanchin pelo apoio e pelos materiais enviados durante o projeto.

Ao Prof. Dr. Ricardo De Marco pela ajuda essencial na análise de dados e esclarecimento de dúvidas.

À Dra. Juliana Sposto Avaca Crusca pela co-orientação, pelos ensinamentos científicos e pela grande amizade que cultivamos.

À aluna de iniciação científica Angélica L. Carrillo Barra pela dedicação e ajuda fundamental para a realização de vários experimentos, pelas conversas, conselhos e risadas.

À Dra. Joci N. A. Macedo pelos ensinamentos sobre a técnica de duplo híbrido.

Ao Dr. José Luiz Lopes pelas considerações a respeito das análises de predição de proteínas intrinsecamente desordenadas.

À Heloisa Ciol pela amizade e por me socorrer nos experimentos no tempo que estive fora.

Às técnicas do laboratório, Andressa e Bel, pelo carinho e pela boa disposição em nos socorrer no laboratório.

À minha mãe, Cláudia Maria, pelo amor incondicional, incentivo e pelo apoio durante toda minha vida.

Ao meu pai, Sérgio e sua esposa, Simone, pelo amor e aconselhamentos.

À minha irmã querida, Gabriela, pelo amor, amizade, compreensão e aconselhamentos nos momentos mais difíceis.

Aos meus irmãos Marcelo e Tomás pelo amor, carinho e compreensão da minha ausência.

A toda minha família e aos colegas de vida, Ana Paula, Marília, Maicon e Caio que compreenderam minha ausência em várias confraternizações e momentos importantes.

Aos amigos queridos que fiz durante minha morada em São Carlos, Andrea, Bruno O., Éverton, Daiana, Débora O., Juliana, Ítalo, Ivani, Laís R., Laís B., Laura, Marina A., Naiara, Paola, Raphaela, Tahiana, Victor e Thomás.

À Dilma, Ronaldo e Ricardo pelo carinho durante todos os anos de convivência.

A todos os colegas do laboratório de Biofísica Molecular pela convivência, ajuda e troca de aprendizado científico.

A todos os docentes da graduação e pós-graduação pela qualidade das disciplinas oferecidas.

Aos funcionários e funcionárias da Seção de Pós-Graduação e Biblioteca e do Grupo de Biofísica Molecular que sempre me atenderam com carinho e simpatia.

Ao Instituto de Física de São Carlos, pela oportunidade de realização da graduação e pós-graduação e pela excelência em infraestrutura de laboratórios, salas de aula e acervo bibliográfico.

À Fundação CAPES – Coordenação de Aperfeiçoamento Pessoal de Nível Superior – pela concessão da bolsa de mestrado e pelo apoio financeiro para a realização desta pesquisa.

A todos que direta ou indiretamente contribuíram para a realização deste trabalho.

“O essencial é invisível para os olhos.”
O Pequeno Príncipe - Antoine Saint-Exupéry

RESUMO

GUTIERREZ, R.F. **Estudos funcionais de CrNIP7 de *Chlamydomonas reinhardtii***: uma proteína envolvida na biogênese de ribossomos. 2016. 141 p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

A biogênese do ribossomo é um processo complexo, altamente ordenado e regulado, no qual o transcrito primário é processado por endo e exonucleases para gerar os RNAs ribossomais maduros. Este processo foi melhor caracterizado em *Saccharomyces cerevisiae*, porém alguns fatores atuantes em humanos tiveram uma função divergente descrita. Um desses fatores é a proteína NIP7, altamente conservada em eucariotos, que atua na formação da subunidade ribossomal 60S, em levedura, e 40S, em humanos. Assim, esse trabalho propôs a caracterização funcional da proteína CrNIP7, homóloga a NIP7, presente em *Chlamydomonas reinhardtii*. *C. reinhardtii* é uma alga verde unicelular ancestral a plantas, utilizada como modelo eucarioto para estudos de fotossíntese e de flagelos. Nesse trabalho, um estudo de complementação funcional foi realizado utilizando duas linhagens de *Saccharomyces cerevisiae* diferentes e em ambas CrNIP7 complementou a função de Nip7p de leveduras, indicando uma participação na síntese da subunidade 60S do ribossomo. Uma busca por parceiros de interação de CrNIP7 foi também realizada, utilizando CrNIP7 como “isca” para rastrear uma biblioteca de cDNA de *C. reinhardtii* em sistema de duplo híbrido em leveduras, o que resultou em dois novos potenciais parceiros de interação. Esses parceiros foram identificados como proteínas preditas conceitualmente no genoma de *C. reinhardtii*, denominadas Predicted e G-patch. Adicionalmente, a interação entre CrNIP7 e CrSBDS, proteína homóloga a Sdo1 (de levedura) e HsSBDS (de humanos), foi confirmada através de um experimento de duplo híbrido dirigido. A interação entre as proteínas CrNIP7 e CrSBDS foi validada por *pull down* e um teste preliminar sugeriu que CrNIP7 e Predicted também interagem *in vitro*. Análises de bioinformática indicam que Predicted, G-patch e CrSBDS tenham regiões intrinsecamente desordenadas, as quais podem se estruturar na interação com seus parceiros. Em conjunto, os resultados desse trabalho contribuem para entendimento do papel de CrNIP7 na biogênese de ribossomos em *Chlamydomonas reinhardtii* em comparação com outros modelos eucarióticos.

Palavras-chave: *Chlamydomonas reinhardtii*. Biogênese de ribossomo. CrNIP7.

ABSTRACT

GUTIERREZ, R. F. **Functional studies of CrNIP7 from *Chlamydomonas reinhardtii***: a protein involved in ribosome biogenesis. 2016. 141 p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

Ribosome biogenesis is a complex, highly regulated and ordered process in which the primary transcript is processed by endo- and exonucleases to generate the mature ribosomal RNAs. This process was best characterized in *Saccharomyces cerevisiae*, but some factors have been described in humans with different function. One of these divergent factors is NIP7, a highly conserved protein in eukaryotes, which acts in the formation of ribosomal 60S subunit, in yeast, and 40S, in humans. Based on this, this work proposed the functional characterization of CrNIP7 protein, homologous to NIP7, from *Chlamydomonas reinhardtii*. *C. reinhardtii* is a green alga, ancestral to plants, that is used as an eukaryote model for photosynthesis and flagella studies. In this study, a functional complementation assay was performed using two different strains of *Saccharomyces cerevisiae* and, in both approaches, CrNIP7 protein complemented the function of Nip7p from yeast, indicating its participation in the synthesis of the 60S ribosomal subunit. A two-hybrid assay was carried out using CrNIP7 as bait to screen a *C. reinhardtii* cDNA library in order to find out CrNIP7 interaction partners, which resulted in two novel potentially partners. The interacting proteins were identified as conceptually predicted proteins in the genome of *C. reinhardtii* and were called Predicted and G-patch. Additionally, the interaction between CrNIP7 and CrSBDS, a protein homologous to Sdo1 (yeast) and HsSBDS (humans), was confirmed by a direct two-hybrid assay. The interaction between CrNIP7 and CrSBDS proteins was validated by pull down and a preliminary test suggested that CrNIP7 and Predicted also interact *in vitro*. Bioinformatics analyzes indicate that Predicted, G-patch and CrSBDS have intrinsically disordered regions, which can be ordered in the moment of interaction. Taken together, the results of this work contribute to understand the role played by CrNIP7 in ribosome biogenesis in *Chlamydomonas reinhardtii* compared to other eukaryotic models.

Keywords: *Chlamydomonas reinhardtii*. Ribosome biogenesis. CrNIP7.

LISTA DE FIGURAS

Figura 1 –	Representação esquemática das subunidades do ribossomo e da síntese de proteínas pelo processo de tradução, mostrando os três sítios de ligação a tRNA: sítio A (A), sítio P (P) e sítio E (E).	30
Figura 2 –	Representação esquemática dos intermediários da síntese das subunidades 40S e 60S do ribossomo eucariótico de <i>S. cerevisiae</i> . Os complexos pré-ribossomais associados aos rRNAs correspondentes estão mostrados, assim como o número conhecido de fatores de montagem em cada etapa da maturação das subunidades.	31
Figura 3 –	Processamento do pré-rRNA em <i>S. cerevisiae</i> . (A) Estrutura do transcrito primário, pré-rRNA 35S, contendo os rRNAs maduros 18S, 5.8S e 25S. (B) Representação esquemática da via de processamento do RNA descrita em detalhes no texto.	33
Figura 4 –	Diagrama do processamento do pré-rRNA 47S em células humanas. O pré-rRNA 47S é convertido no pré-rRNA 45S que é processado por três vias alternativas de processamento dependendo de onde ocorre a primeira clivagem. Maiores detalhes sobre as vias estão descritos no texto.	35
Figura 5 –	Representação da arquitetura de HsNIP7 de humanos. (A) Estrutura cristalográfica de HsNIP7. (B) Diagrama da topologia de HsNIP7. As α -hélices estão representadas por cilindros verdes, as fitas- β por setas azuis e as 3_{10} -hélices por círculos verdes. As estruturas secundárias das figuras A e B estão com a mesma numeração.	38
Figura 6 –	A alga unicelular, <i>Chlamydomonas reinhardtii</i> . (A) Representação esquemática dos principais componentes de uma célula de <i>Chlamydomonas</i> : flagelos (F), corpos basais (B), núcleo (N), estigma (E) aparelho de Golgi (G), mitocôndria (M), cloroplasto (C), tilacóide (T), pirenóide (P). Adaptado de Gumen et al ²⁸ (B) Microscopia eletrônica de varredura de uma célula de <i>C. reinhardtii</i> (imagem cedida pelo Laboratório de Microscopia e Microanálise - CTMH). (C) Microscopia óptica de uma célula (12 μ m x 17 μ m) de <i>C. reinhardtii</i> (imagem cedida pelo <i>The Protist Information Server</i> , no endereço eletrônico http://protist.i.hosei.ac.jp/Protist_menuE.html).	39
Figura 7 –	Esquema do ensaio de complementação funcional. As linhagens de <i>S. cerevisiae</i> DG130, DG440 e DG442 estão especificadas na Tabela 3.	53
Figura 8 –	Ilustração sobre o funcionamento do sistema de duplo híbrido. Em A, o fator de transcrição está intacto e a transcrição dos genes repórteres é ativada. Em B, os domínios independentes, BD e AD, estão fisicamente separados não havendo transcrição. Em C e D, as proteínas híbridas, BD-Isca e AD-Presa, expressas separadamente não ativam a transcrição. Em E, as proteínas são co-expressas e se interagem, elas mediam a aproximação dos domínios BD e AD ativando a transcrição dos genes repórteres.	59
Figura 9 –	Esquema do ensaio de duplo híbrido em levedura para busca de parceiros proteicos de interação.	60

- Figura 10 – Análise da amplificação por PCR de *CrNip7* e das clonagens visando a expressão heteróloga: (A) Produto de PCR (540pb) amplificado a partir do cDNA de *C. reinhardtii* (1) Tm= 62°C; (2) Tm= 64°C. (B) Clonagem de *CrNip7* no vetor pGEMT-Easy: (3) plasmídeo não digerido, (4) análise de restrição com as enzimas *Bam*HI e *Sall*. (C) Subclonagem de *CrNip7* no vetor pETSUMO: (5) plasmídeo não digerido, (6) análise de restrição com as enzimas *Bam*HI e *Sall*. (D) Subclonagem de *CrNip7* no vetor pGEX-4T-1: (7) plasmídeo não digerido, (8) análise de restrição com as enzimas *Bam*HI e *Sall*. (E) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Eletroforese em gel de agarose 0,8% em TAE 1X..... 80
- Figura 11 – Análise da amplificação por PCR de *CrNip7* e subclonagens para os ensaios de complementação funcional e duplo híbrido: (A) Produto de PCR (540pb) amplificado a partir do molde pGEMT-*CrNip7*-BS: (1) Tm= 62°C; (2) Tm= 64°C. (B) Verificação da clonagem de *CrNip7* no vetor pGEMT-Easy: (3) clone negativo, (4 e 5) clones positivos. (C) Análise da clonagem de *CrNip7* no vetor pJN40:(6) clone negativo, (7,8) clones positivos. (D) Análise da clonagem de *CrNip7* no vetor pLEXA-N: (9 e 10) clones positivos. (E) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Todas as análises de restrição foram realizadas com as enzimas *Eco*RI e *Sall*. Eletroforese em gel de agarose 0,8% em TAE 1X..... 81
- Figura 12 – Análise da amplificação por PCR de *CrNip7* e subclonagens para os ensaios de complementação funcional: (A) Produto de PCR (540 pb) a partir da construção pGEMT-*CrNip7*-BS: (1) Tm= 62°C. (B) Análise da clonagem de *CrNip7* no vetor pGEMT-Easy: (2) plasmídeo não digerido, (3 e 4) análise de restrição com as enzimas *Bam*HI e *Xba*I. (C) Análise da clonagem de *CrNip7* no vetor YCplac33: (5,6) análise de restrição com as enzimas *Bam*HI e *Xba*I. (D) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Eletroforese em gel de agarose 0,8% em TAE 1X. 81
- Figura 13 – Análise da amplificação por PCR de *CrNip7* e subclonagens para fins de co-expressão em *E. coli*: (A) Produto de PCR (540pb) amplificado a partir da construção pGEMT-*CrNip7*-BS: (1) Tm= 62°C. (B) Análise de restrição da construção de *CrNip7* no vetor pUC19, usando as enzimas *Eco*RI e *Pst*I: (2, 3 e 4) clones positivos. (C) Análise da clonagem de *CrNip7* no vetor pETDuet-1: (5 e 7): plasmídeos não digeridos, (6 e 7) análise de restrição com as enzimas *Eco*RI e *Pst*I. (D) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Eletroforese em gel de agarose 0,8% em TAE 1X. 82
- Figura 14 – Ensaio de complementação funcional mostrando que a proteína CrNIP7, de *C. reinhardtii*, complementa a função de Nip7p, de levedura, nas duas abordagens: (A) A levedura DG130 [pJN40-*CrNip7*] expressando CrNIP7 e Nip7p-1 cresce melhor que a levedura DG130 [pJN40], controle de crescimento deficiente e termossensível, que expressa apenas a mutante Nip7p-1. (B) A levedura DG442 [YCplac33-*CrNip7*] expressando CrNIP7 tem crescimento igual ao controle positivo de fenótipo normal, DG440 [YCplac33]. Como esperado a levedura DG442 [YCplac33], utilizada como controle negativo, não cresceu no meio contendo apenas glicose..... 84
- Figura 15 – Teste da β -galactosidase em papel de filtro para verificação da ativação do gene repórter *LacZ* pela isca BD-CrNIP7 (1) e pelas potenciais interações encontradas no rastreamento da biblioteca comercial de *C. reinhardtii* (2, 3, 4 e 5). Controle negativo de interação (6) e controle positivo de interação forte (7). 87

Figura 16 –	Teste de força de interação das potenciais interações por crescimento em meio restritivo, SD-WLH, com adição gradativa de 3-AT. O experimento foi realizado com dois transformantes diferentes para cada interação.	88
Figura 17 –	Teste de expressão de 6xHis-SUMO-CrNIP7 e purificação. Expressão a partir de <i>E. coli</i> Rosetta (DE3) pETSUMO-CrNip7 e purificação por cromatografia de afinidade utilizando 2 mL de resina Ni-NTA (Qiagen) para um extrato proteico proveniente de 250 mL de meio de cultura LB. Análise por SDS-PAGE 15%.....	90
Figura 18 –	Teste de clivagem da 6xHis-SUMO-CrNIP7 imobilizada na resina de afinidade (Ni-NTA) utilizando 6xHis-SUMO protease. Análise por SDS-PAGE 15%.	90
Figura 19 –	Expressão de 6xHis-CrNIP7 pela indução da construção pETDuet-1-CrNip7 em <i>E. coli</i> Rosetta (DE3) e purificação por cromatografia de afinidade. Foram utilizadas células provenientes de 500 mL de cultura em meio LB e 2 mL de resina Ni-NTA (Qiagen). Análise por SDS-PAGE 15%.....	91
Figura 20 –	Expressão de GST-CrNIP7 pela indução da construção pGEX-CrNip7 em <i>E. coli</i> Rosetta (DE3) e purificação por cromatografia de afinidade. Foram utilizados 100 mL de cultura em LB e 0,5 mL de resina Glutathione Sepharose 4 Fast Flow (GE).	92
Figura 21 –	Alinhamento gerado pelo programa ClustalO comparando as sequências homólogas de CrNIP7 (XP_001696971.1): PaNIP7 (CAB49179.1), Nip7p (EIW07327.1), AtNIP7 (AEE83650.1), CeNIP7CCD65161.1, e HsNIP7 (NP_057185.1). Os aminoácidos conservados estão representados por asteriscos, os aminoácidos com alta similaridade com dois pontos e aminoácidos com baixa similaridade com um ponto.	93
Figura 22 –	Comparação entre a estrutura secundária resolvida por cristalografia de raio X da proteína HsNIP7 (PDB:1SQW) (linha 1, em vermelho) e às predições de estrutura secundária pelos programas GOR4 (linha 4, em azul) e PSIPRED (linha 5, em azul) de CrNIP7 (linha 3). As α -hélices estão representadas por um cilindro e as fitas- β por setas. Os aminoácidos realçados em cinza fazem parte do domínio PUA.	95
Figura 23 –	Alinhamento múltiplo realizado pelo programa ClustalO das proteínas preditas das algas verdes <i>C. reinhardtii</i> (XP_001690112.1), <i>G. pectorale</i> (KXZ55342.1) e <i>V. carteri</i> (Vocar.0001s0712.1). Os aminoácidos conservados estão representados por asteriscos, os aminoácidos com alta similaridade com dois pontos e aminoácidos com baixa similaridade com um ponto.	97
Figura 24 –	Análise da amplificação e subclonagens de <i>predicted</i> em vetores de expressão: (A) “M” = Marcador molecular GeneRuler DNA 1kb (Fermentas). (B) Amplificação da CDS <i>predicted</i> (~480 pb) a partir de pGAD-HA- <i>predicted</i> (1), utilizando Tm= 54°C. (C) Clones positivos em vetor de propagação pUC19: (2,4) análise de restrição com <i>NdeI</i> e <i>XhoI</i> , liberando o inserto correspondente à <i>predicted</i> e uma banda de ~230 pb; (3,5) plasmídeos não digeridos. (D) Clone positivo de <i>predicted</i> em pETDuet-1-CrNip7; (6) plasmídeo não digerido; (7) análise de restrição com <i>NdeI</i> e <i>XhoI</i> . (E) Clones positivos de <i>predicted</i> em pETSUMO; (8,10) plasmídeos não digeridos; (9,11) análise de restrição com <i>BamHI</i> e <i>XhoI</i> . (F) Clones positivos de <i>predicted</i> inseridos em pGEX-4T; (12,14) plasmídeos não	

	digeridos; (13, 15) análise de restrição com <i>Bam</i> HI e <i>Xho</i> I. Eletroforese em gel de agarose 0,8% em TAE 1X.	98
Figura 25 –	Teste de expressão da proteína Predicted em três sistemas diferentes: (A) “M” = Marcador de massa molecular de proteínas; “MB”= Marcador de baixa massa molecular de proteínas. (B) Teste de co-expressão de 6xHis-CrNIP7 e Predicted pelo vetor pETDuet-1; (1) cultura não induzida; (2) cultura induzida. (C) Teste de expressão da proteína de fusão “GST-predicted”; (3) Cultura não induzida; (4) Cultura induzida; (5) Fração insolúvel; (6) Fração solúvel. (D) Teste de expressão da fusão “6xHis-SUMO-predicted”; (7) Cultura não induzida; (8) Fração insolúvel; (9) Fração solúvel. SDS-PAGE Tricina 10%.	99
Figura 26 –	Teste preliminar de clivagem da fusão 6xHis-SUMO-predicted com SUMO protease na resina de níquel. SDS-PAGE Tricina 10%.	100
Figura 27 –	Teste preliminar de <i>pull down</i> das proteínas CrNIP7 e Predicted. SDS-PAGE Tricina 10%.	101
Figura 28 –	Sequência de aminoácidos de Predicted (linha 1). As previsões de estrutura secundária feitas pelos programas GOR4 (linha 2) e PSIPRED (linha3) utilizam o código linha = <i>coil</i> , seta = fita β e cilindro = α -hélice. As previsões de regiões intrinsecamente desordenadas pelos programas Globplot (linha 4), DisEMBL (linha 5), tem o código: “D”= desordenado, e “–” = ordenado. Previsões de possíveis sítios de ligação pelo programa ANCHOR estão representadas em laranja (linha 6).	103
Figura 29 –	Amplificação da CDS da proteína G-patch, apresentando bandas amplificadas por PCR, usando $T_m = 54^\circ \text{C}$	105
Figura 30 –	Análise da amplificação e subclonagens de <i>G-patch</i> : (A) “M” = Marcador molecular GeneRuler DNA 1kb (Fermentas). (B) Clones positivos de <i>G-patch</i> em vetor de propagação pUC19: (1 e 4) plasmídeo não digerido; (2 e 5) análise de restrição com <i>Nde</i> I e <i>Sal</i> I; (3 e 6) análise de restrição com <i>Bam</i> HI, as manchas ao fim do gel correspondem a RNA remanescente da extração plasmidial pelo método lise alcalina. (C) Clone positivo de <i>G-patch</i> em pETDuet-1-CrNip7; (7,9) plasmídeo não digerido; (8, 9) restrição com <i>Bam</i> HI e <i>Xho</i> I. (D) Clones positivos de <i>G-patch</i> em pETSUMO; (12,14) plasmídeo não digerido; (13,15) análise de restrição com <i>Bam</i> HI e <i>Sal</i> I. Eletroforese em gel de agarose 0,8% em TAE 1X.	106
Figura 31 –	Sequência de aminoácidos de G-patch (linha1), o domínio conservado está realçado cinza. As linhas 2 e 3 representam as previsões de estrutura secundária pelos programas GOR4 e PSIPRED, respectivamente. As linhas 4 e 5 representam as previsões de IDR pelos programas Globplot e DisEMBL, respectivamente. A linha 6 mostra os possíveis sítios de ligação (em laranja) preditos pelo programa ANCHOR.	108
Figura 32 –	Análise da amplificação por PCR e subclonagens de <i>CrSBDS</i> : (A) Produto de PCR (1140 pb) amplificado a partir do <i>cDNA</i> de <i>C. reinhardtii</i> utilizando $T_m = 57^\circ \text{C}$; (1) com DMSO, (2) sem DMSO. (B) Análise da clonagem de <i>CrSBDS</i> no vetor pUC19: (3), plasmídeo não digerido (4) análise de restrição com <i>Sma</i> I e <i>Xho</i> I; (C) Análise da clonagem de <i>CrSBDS</i> em pET28a; (5) plasmídeo não digerido; (6) análise de restrição com <i>Nde</i> I e <i>Xho</i> I. (D) Análise da clonagem de <i>CrSBDS</i> em pETDuet-1-	

CrNip7; (7) plasmídeo não digerido; (8) análise de restrição com *NdeI* e *XhoI*. (E) “M” = Marcador molecular GeneRuler DNA 1kb (Fermentas); (F) Análise do produto de PCR para troca de sítios de restrição a partir de um clone positivo, Tm= 55°C. (G) Clones positivos de *CrSBDS* em pUC19; (10) plasmídeo não digerido; (11) análise de restrição com *SmaI* e *XhoI*, mostrando plasmídeo não digerido (ND), plasmídeo sem o inserto (pUC19) e inserto (CrSBDS). (H) Análise da clonagem de *CrSBDS* em pGAD-HA: (12) plasmídeo não digerido; (13) Análise de restrição com *SmaI* e *XhoI*. Eletroforese em gel de agarose 0,8% em TAE 1X. ... 112

Figura 33 – Teste de interação *in vivo* utilizando duplo híbrido dirigido para verificar a interação das proteínas BD-CrNIP7 e AD-CrSBDS. (A) Ensaio da β -galactosidase em papel de filtro mostrando a expressão do gene repórter *LacZ*. (B) Teste de força de interação por gradiente de 3AT, mostrando que o clone expressão BD-CrNIP7 e AD-CrSBDS é capaz de crescer em meio sem histina, porém é sensível à dição de 3AT no meio de cultura. 113

Figura 34 – Análise por SDS-PAGE 15% da expressão e purificação da proteína 6xHis-CrSBDS..... 114

Figura 35 – Análise mostrando a co-expressão e co-purificação de CrSBDS e 6xHis-CrNIP7 por indução do vetor pETDuet-1 em *E. coli* Rosetta (DE3). SDS-PAGE Tricina 16%.116

Figura 36 – Análise do experimento de *pull down* das proteínas 6xHis-CrSBDS com CrNIP7. SDS-PAGE Tricina 16%. 116

Figura 37 – Alinhamento gerado pelo programa Clustal O das sequências proteicas de SBDS de *C. reinhardtii* (Cre01.g017500), *Arabidopsis thaliana* (NP_564488.1). Os aminoácidos conservados estão representados por asteriscos, os aminoácidos com alta similaridade com dois pontos e aminoácidos com baixa similaridade com um ponto. Os aminoácidos sombreados em cinza fazem parte do domínio *zinc finger* C2H2. 117

Figura 38 – Comparação entre a estrutura secundária resolvida de HsSBDS (PDB:2KDO) (linha 1 e 2) e a predição feita pelo programa PSIPRED para CrSBDS (linha 3 e 4, os aminoácidos realçados fazem parte do domínio SBDS conservado encontrado pelo Pfam). As fitas- β estão representadas por setas, as α -hélices por cilindros e coils por linha. Estruturas em vermelho fazem parte do domínio N-terminal, em amarelo, do domínio central; em azul, do domínio C-terminal e em verde, são elementos preditos para a extensão C-terminal de CrSBDS. As linhas 5 e 6 correspondem às predições de IDRs na sequência de CrSBDS pelos programas Globplot e DisEMBL, e a linha 7 corresponde aos possíveis sítios de ligação preditos (em laranja) pelo programa ANCHOR. 119

Figura 39 – (A) Representação esquemática do experimento de sincronização de *C. reinhardtii*. O diagrama mostra a zona temporal (ZT) dos períodos de luz (ZT0 a ZT12) e escuro (ZT12 a ZT24/0) de uma cultura de células sincronizadas com a fase de crescimento (G1), divisão (S/M) e descanso (G0) demarcadas pelas setas em cinza. As amostras da cultura foram coletadas a cada uma hora com amostras adicionais a cada meia hora durante a divisão celular (linhas em marrom). Esquema reproduzido de Zones et. al.⁶² (B) Nível de expressão de CrNIP7, CrSBDS, G-patch e Predicted ao longo do ciclo celular de *C. reinhardtii*.

	RRKM, unidade de expressão gênica utilizada em Zones et. al, do inglês <i>reads per kilobase per million mapped</i>	121
Figura 40 –	Sequência de nucleotídeos da CDS CrNip7(NCBI: XM_001696919.1).	137
Figura 41 –	Sequência de aminoácidos da proteína CrNIP7 (NCBI: XP_001696971.1).	137
Figura 42 –	Sequência de nucleotídeos codificadores (CDS) da proteína Predicted (NCBI: XM_001690060.1).	137
Figura 43 –	Sequência de aminoácidos da proteína Predicted (NCBI: XP_001690112.1) de <i>C. reinhardtii</i>	137
Figura 44 –	Sequência de nucleotídeos codificadores da proteína G-patch completa. A sequência em vermelho corresponde à parte clonada no vetor pGAD-HA-isca. A base guanina sublinhada é um ponto de mutação silenciosa entre a sequência clonada no vetor de duplo híbrido e a sequência correta de G-patch. Os nucleotídeos realçados em cinza correspondem às duas partes faltam na tradução conceitual do genoma na anotação do NCBI (XM_001692395.1), mas que está correta no banco de dados Phytosome11 com a identificação de acesso Cre07.g347400.	138
Figura 45 –	Sequência de aminoácidos da proteína G-patch completa (Phytozome11: Cre07.g347400). Os aminoácidos destacados em vermelho são resultado expressão de G-patch parcial pelo vetor pGAD-HA-isca no ensaio de duplo híbrido. Os 45 aminoácidos realçados em cinza correspondem ao domínio conservado G-patch encontrado pela predição no programa Pfam.	138
Figura 46 –	Sequência 1140 pares de nucleotídeos da CDS de CrSBDS. Os nucleotídeos grifados em cinza foram encontrados por análise do banco de dados SRA em <i>C. reinhardtii</i> e correspondem à sequência anotada no banco de dados Phytozome 11 com identificação Cre01.g017500.	139
Figura 47 –	Sequência correta de aminoácidos da proteína CrSBDS (Cre01.g017500).	139
Figura 48 –	Alinhamento múltiplo (Clustal O) de SBDS dos organismos <i>A. fulgidus</i> (WP_010877998.1), <i>P. abyssi</i> (WP_010867732.1), <i>T. cruzi</i> (ACD88761.1), <i>S. cerevisiae</i> (NP_013122.1), <i>C. reinhardtii</i> (Cre01.g017500), <i>A. thaliana</i> (NP_564488.1), <i>H. sapiens</i> (NP_057122.2) e <i>C. elegans</i> (NP_497226.2.). Os aminoácidos conservados estão representados por asterisco, os aminoácidos com alta similaridade; com dois pontos e aminoácidos com baixa similaridade; com um ponto.	140

LISTA DE TABELAS

Tabela 1 – Experimentos relacionados com cada proteína.	48
Tabela 2 – Oligonucleotídeos utilizados para amplificação das sequências codificadoras das proteínas estudadas neste trabalho.	50
Tabela 3 – Linhagens de <i>S. cerevisiae</i> utilizadas do ensaio de complementação funcional.	54
Tabela 4 – Plasmídeos utilizados nos experimentos de complementação funcional.	54
Tabela 5 – Meios de cultura utilizados no ensaio de complementação funcional.	56
Tabela 6 – Resumo das transformações realizadas e meios de cultura para seleção plasmidial.	56
Tabela 7 – Plasmídeos utilizados no ensaio de duplo híbrido.	61
Tabela 8 – Resumo das transformações realizadas e meios de cultura para seleção plasmidial.	61
Tabela 9 – Composição dos meios de cultura para leveduras.	62
Tabela 10 – Soluções utilizadas na transformação da biblioteca de cDNA.	64
Tabela 11 – Construções em plasmídeos de expressão realizadas neste trabalho.	70
Tabela 12 – Loci gênico correspondente às proteínas de interesse.	76
Tabela 13 – Descrição dos oligonucleotídeos utilizados para amplificação de <i>CrNip7</i> e construções em vetores de clonagem e de expressão.	80
Tabela 14 – Matriz de identidade entre as sequências homólogas a CrNIP7 (programa Clustal O). ...	93
Tabela 15 – Matriz de identidade entre as proteínas preditas de <i>C. reinhardtii</i> , <i>G. pectorale</i> e <i>V. carteri</i> pelo programa ClustalO.	97
Tabela 16 - Matriz de identidade entre as sequências de SBDS (programa Clustal O).	141

LISTA DE ABREVIATURAS E SIGLAS

3-AT	3-amino-1, 2, 4-triazol
3'-ETS	Sequência espaçadora externa 3' (<i>External Transcribed Spacer 3'</i>)
5'-ETS	Sequência espaçadora externa 5' (<i>External Transcribed Spacer 5'</i>)
AD	Domínio de ativação da transcrição (<i>Activation Domain</i>)
BD	Domínio de ligação ao DNA (<i>Binding Domain</i>)
cDNA	Ácido desoxirribonucleico complementar
CDS	Sequência de DNA codificante
CrNIP7	Proteína de <i>Chlamydomonas reinhardtii</i> homóloga a Nip7p
DNA	Ácido desoxirribonucleico
dNTP	Desoxirribonucleotídeo trifosfato (mistura de dATP, dCTP, dGTP e dTTP)
DTT	Ditiotreitol
EDTA	Ácido etilenodiaminotetracético
Glu	Glicose
Gal	Galactose
HsNIP7	Proteína de humanos homóloga a Nip7p
IPTG	Isopropil- β -D-tiogalactopiranosídeo
ITS1	Sequência espaçadora interna 1 (<i>Internal Transcribed Spacer 1</i>)
ITS2	Sequência espaçadora interna 2 (<i>Internal Transcribed Spacer 2</i>)
kb	Quilobase
kDa	Quilodalton
LB	Meio de cultura (<i>Lysogenic broth</i>)
LiAc	Acetato de lítio
mRNA	RNA mensageiro
Nip7p	Proteína de importação nuclear 7 de <i>Saccharomyces cerevisiae</i> (<i>Nuclear import protein 7</i>)
pb	Pares de base
PCR	Reação em cadeia da polimerase (<i>Polymerase Chain Reaction</i>)
PEG	Polietilenoglicol
pH	Potencial hidrogeniônico
poli-A	Sequência de poli-adenina
poli-U	Sequência de poli-uracila
poli-AU	Sequência de poli-adenina e uracila
PTC	Centro de peptidil transferase (<i>Peptidyl Transferase Center</i>)

PUA	Domínio conservado encontrado em pseudo-uridina sintases e archaeosina transglicosilases (<i>PseudoUridine synthases and Archaeosine tranglycosilases</i>)
RNA	Ácido ribonucleico
rRNA	RNA ribossomal
RP	Proteína ribossomal (<i>Ribossomal Protein</i>)
RT-PCR	Reação da transcriptase reversa, seguida de reação em cadeia da polimerase
SBDS	Proteína associada a síndrome de <i>Shwachman-Bodian-Diamond</i>
SD	Meio de cultura mínimo
SDS-PAGE	Eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS)
snoRNA	Pequeno RNA nucleolar (<i>small nucleolar RNA</i>)
snoRNP	Pequena partícula ribonucleoproteica nucleolar (<i>small nucleolar ribonucleoprotein</i>)
TAE	Tampão de Tris, ácido acético e EDTA
TE	Tampão Tris-EDTA
Tris	Tris-hidroximetilaminometano
TRITON X-100	Polietilenoglicol-terc-octilfenil éter
tRNA	RNA transportador
ura	Uracila
X-Gal	5-bromo-4-cloro-3-indolil-beta-D-galactopiranosídeo
YNB	<i>Difco yeast nitrogen base without amino acids</i>
YPD	Meio de cultura contendo extrato de levedura, peptona e dextrose (<i>Yeast extract, Peptone, Dextrose media</i>)

SUMÁRIO

1	INTRODUÇÃO	29
1.1	O ribossomo.....	29
1.2	Biogênese de ribossomos em eucariotos	30
1.2.1	A via de processamento do RNA	32
1.2.2	Fatores não ribossomais divergentes.....	35
1.3	NIP7	36
1.4	<i>Chlamydomonas reinhardtii</i>	39
1.4.1	Estrutura geral do genoma.....	40
1.4.2	Ribossomos	41
1.4.3	Importância e aplicações.....	43
1.5	Motivação para os estudos funcionais de CrNIP7	44
1.6	Objetivos.....	45
2	MATERIAIS E MÉTODOS.....	48
2.1	Isolamento dos Sequências de DNA codificantes (CDS)	48
2.1.1	Obtenção do DNA complementar (cDNA) total de <i>Chlamydomonas reinhardtii</i>	48
2.1.1.1	Extração de RNA mensageiro (mRNA) total	49
2.1.1.2	Transcrição Reversa e Reação em Cadeia da Polimerase (RT-PCR)	49
2.1.2	Clonagens.....	51
2.1.2.1	Transformação de bactéria por choque térmico	51
2.1.2.2	Extração plasmidial de bactéria	52
2.2	Ensaio de complementação funcional	53
2.2.1	Linhagens de levedura e plasmídeos.....	53
2.2.2	Clonagem da CDS de CrNip7 nos plasmídeos de levedura	54
2.2.3	Transformação de levedura pelo método “Quick and Easy Trafo”	55
2.2.4	Ensaio de complementação em levedura utilizando a linhagem DG130	56
2.2.5	Ensaio de complementação em levedura utilizando as linhagens DG440 e DG442.....	57
2.3	Ensaio de Duplo Híbrido em levedura	57
2.3.1	Linhagem de levedura, plasmídeos e biblioteca de <i>C. reinhardtii</i>	59
2.3.2	Transformação de levedura em pequena escala.....	61
2.3.3	Ensaio da β -galactosidade em papel de filtro.....	62
2.3.4	Transformação de levedura em grande escala para varredura da biblioteca de cDNA.....	63
2.3.5	Extração plasmidial de levedura.....	65

2.3.6	Transformação de bactérias, extração plasmidial de bactérias, sequenciamento e análise das sequências “AD-Presa”	66
2.3.7	Confirmação das interações das proteínas encontradas e teste de força de interação	67
2.4	Análises de bioinformática.....	67
2.4.1	Anotação das sequências codificadoras das proteínas CrSBDS e G-patch.....	68
2.5	Verificação da interação entre CrNIP7 e CrSBDS por duplo híbrido dirigido.....	68
2.6	Validação das interações proteicas encontradas nos ensaios de duplo híbrido	69
2.6.1	Expressão heteróloga de proteínas e purificação por cromatografia de afinidade a níquel .	71
2.6.2	Teste de clivagem com SUMO protease na coluna.....	72
2.6.3	Purificação da GST-CrNIP7.....	72
2.6.4	Co-expressão e Co-purificação	73
2.6.5	<i>Pull down</i>	73
2.7	Espectrometria de massas	74
2.8	Análise do perfil de expressão de CrNIP7, G-patch, Predicted e CrSBDS	76
3	RESULTADOS E DISCUSSÃO	79
3.1	CrNIP7	79
3.1.1	Amplificação e clonagens da CDS <i>CrNip7</i>	79
3.1.2	Ensaio de complementação funcional.....	82
3.1.3	Ensaio de Duplo Híbrido	84
3.1.4	Testes de expressão de CrNIP7 em sistema heterólogo.....	89
3.1.4.1	Expressão, purificação e teste de clivagem de 6xHis-SUMO-CrNIP7	89
3.1.4.2	Expressão e purificação de 6xHis-CrNIP7	91
3.1.4.3	Expressão e purificação de GST-CrNIP7.....	91
3.1.5	Análise de bioinformática	92
3.1.6	Conclusão parcial.....	94
3.2	Predicted	96
3.2.1	Amplificação da CDS <i>Predicted</i> e clonagem em vetores de expressão.....	97
3.2.2	Teste de expressão	98
3.2.3	Teste de clivagem de 6xHis-SUMO-predicted e <i>pull down</i> com CrNIP7	100
3.2.4	Análise de bioinformática	101
3.2.5	Conclusão parcial.....	104
3.3	G-patch.....	104
3.3.1	Amplificação da sequência codificadora de G-patch	105
3.3.2	Clonagens de G-patch e teste preliminar de expressão	106

3.3.3	Análise de bioinformática.....	107
3.3.4	O domínio G-patch	109
3.3.5	Conclusão parcial	110
3.4	CrSBDS	111
3.4.1	Amplificações e clonagens da CDS <i>CrSBDS</i>	112
3.4.2	Duplo Híbrido direcionado para análise da interação CrNIP7 e CrSBDS.....	113
3.4.3	Expressão heteróloga e purificação de 6xHis-CrSBDS	114
3.4.4	Validação da interação por <i>Pull Down</i>	115
3.4.5	Análise de bioinformática.....	116
3.4.6	Conclusão parcial	118
3.5	Discussão final.....	120
4	CONCLUSÕES E PERSPECTIVAS	125
	REFERÊNCIAS.....	129
	APÊNDICE A – Informações adicionais sobre as proteínas tratadas no trabalho	137

INTRODUÇÃO

1 INTRODUÇÃO

1.1 O ribossomo

O ribossomo é uma macromolécula, constituída por RNAs ribossomais (rRNA) e proteínas ribossomais (RPs), responsável pela síntese de proteínas através do processo de tradução (Figura 1). O ribossomo bacteriano (*Escherichia coli*) tem coeficiente de sedimentação 70S e é formado pelas subunidades 30S (16S rRNA, 21 RPs) e 50S (23S, 5S rRNA, 34 RPs), enquanto o ribossomo eucariótico (*Saccharomyces cerevisiae*) sedimenta em gradiente de sacarose com coeficiente de 80S e é formado pelas subunidades 40S (18S rRNA, 33 RPs) e 60S (25S, 5.8S, 5S rRNA, 46 RPs).¹ As subunidades do ribossomo de Archaea têm tamanhos parecidos com o de bactérias, porém funcionalmente são mais parecidos com o ribossomo eucariótico. Apesar da função e estrutura molecular conservadas entre os ribossomos de bactéria, eucariotos e Archaea, o que indicam uma origem evolucionária comum, as diferenças em tamanho e complexidade podem refletir diferenças na regulação da tradução e também são a base para antibióticos que bloqueiam especificamente as subunidades procarióticas.¹⁻²

As duas subunidades do ribossomo executam diferentes papéis na síntese de proteínas. A subunidade menor é responsável pela localização do códon AUG inicial no RNA mensageiro (mRNA) e pela decodificação da informação carregada por ele por meio da mediação entre as interações dos anticódons do RNA transportador (tRNA) com os códons do mRNA. Sendo assim, a subunidade menor do ribossomo assegura que a ordem correta dos aminoácidos da proteína que está sendo sintetizada seja seguida. A subunidade maior contém o centro catalítico peptidil transferase (PTC) que catalisa a formação de ligações peptídicas no polipeptídeo em crescimento. Vários fatores de iniciação, alongação e de terminação são necessários para garantir a que a síntese proteica ocorra de forma progressiva e específica. Ambas as subunidades compartilham três sítios de ligação a tRNA (Figura 1): o sítio A que recebe o tRNA aminoacilado contendo o aminoácido a ser incorporado na cadeia polipeptídica; o sítio P que posiciona o tRNA ligado ao polipeptídeo nascente, e o sítio E, ou de saída, que abriga o tRNA desacetilado, após a ligação peptídica, antes de ser dissociado do ribossomo. Durante a fase de alongação da cadeia polipeptídica, um aminoácido é adicionado por vez ao polipeptídeo

nascente no sentido N- para o C- terminal e a terminação ocorre quando o códon de parada alcança o sítio A.³

Figura 1 – Representação esquemática das subunidades do ribossomo e da síntese de proteínas pelo processo de tradução, mostrando os três sítios de ligação a tRNA: sítio A (A), sítio P (P) e sítio E (E).

Fonte: Adaptada de RNA...⁴

1.2 Biogênese de ribossomos em eucariotos

A biogênese de ribossomos é um dos processos mais complexos, cruciais e energéticos de qualquer célula. Em eucariotos, esse processo é altamente regulado e ordenado, ocorrendo sequencialmente no nucléolo, no nucleoplasma e no citoplasma, onde os eventos finais de maturação acontecem (Figura 2). A biogênese do ribossomo enfrenta o desafio de coordenar o processamento e modificação do RNA ribossomal com a sua montagem com as proteínas ribossomais e, para isso, conta com mais de 200 fatores não ribossomais que conferem direcionalidade e precisão ao processo. Esses fatores não ribossomais transientes compreendem endo e exoribonucleases, RNA helicases dependentes de ATP, AAA-ATPases,

GTPases, quinases, “chaperonas” ou “fatores de montagem” e snoRNPs (do inglês, *small nucleolar ribonucleoprotein*).^{1,5}

Devido a facilidade experimental de ensaios genéticos, bioquímicos e celulares em levedura, a biogênese de ribossomo em eucariotos é melhor caracterizada em *S. cerevisiae* e as funções dos fatores transientes identificados em levedura foram atribuídos aos ortólogos de eucariotos em geral. Porém, recentemente foram descritas diferenças na via de processamento de rRNA e de fatores não ribossomais em mamíferos.^{6–11}

Figura 2 – Representação esquemática dos intermediários da síntese das subunidades 40S e 60S do ribossomo eucariótico de *S. cerevisiae*. Os complexos pré-ribossomais associados aos rRNAs correspondentes estão mostrados, assim como o número conhecido de fatores de montagem em cada etapa da maturação das subunidades.

Fonte: TALKISH et al.¹²

Em resumo, a biogênese das duas subunidades ribossômicas começa a partir da transcrição do pré-rRNA 35S (47S em humanos) e do rRNA 5S por RNA polimerases diferentes dentro do nucléolo. Os rRNAs maduros (5.8S, 25S/28S e 5S da subunidade 60S e 18S da subunidade 40S) são liberados dos rRNAs precursores através de clivagens endo e exonucleolíticas. Concomitantemente às clivagens, os pré-rRNAs são amplamente modificados e vinculados às proteínas ribossomais antes das subunidades pré-40S e pré-60S serem exportadas separadamente para o citoplasma. O complexo pré-40S continua a ser processado no citoplasma, enquanto a maturação do complexo pré-60S continua no núcleo, antes da exportação para o citoplasma. As modificações dos nucleotídeos dos rRNAs nascentes são

realizadas por cerca de 75 snoRNPs que medeiam a metilação da 2'-O-ribose e formação de pseudouridinas. Esses complexos snoRNPs são direcionados para seus substratos através do emparelhamento de bases entre o rRNA e o snoRNA (do inglês, *small nucleolar RNA*), enquanto as proteínas associadas catalisam a reação de modificação.^{1,5}

Os rRNAs e proteínas ribossomais são produzidos em quantidades equimolares e essa síntese é fortemente regulada por uma variedade de condições de crescimento. A formação de ribossomos eucarióticos não só requer a coordenação de eventos de processamento e montagem, mas também uma ordenação espaço-temporal dessas etapas, iniciando no nucléolo e terminando no citoplasma.⁵ Como os diferentes fatores determinam a exatidão do processo de montagem e de exportação dos complexos pré-ribossomais através dos poros nucleares ainda é uma questão que não está totalmente esclarecida.

1.2.1 A via de processamento do RNA

No modelo eucariótico de levedura, a síntese dos RNAs ribossomais se inicia quando a RNA polimerase I transcreve um grande RNA ribossomal policistrônico, o 35S pré-rRNA. Este transcrito primário contém os rRNAs 18S, 5.8S e 25S flanqueados por sequências espaçadoras externas (5' ETS e 3' ETS) e internas 1 (ITS1) e 2 (ITS2) não codificantes (Figura 3A). O rRNA 18S maduro irá compor a subunidade 40S, enquanto os rRNAs 5.8S e 25S junto com o rRNA 5S, sintetizado independentemente pela RNA polimerase III, irão compor a subunidade 60S. O pré-rRNA 35S é sucessivamente clivado na região 5'-ETS no sítio A₀ (gerando o pré-rRNA 33S), e no sítio A₁, que se localiza na extremidade 5' do 18S rRNA maduro (gerando o pré-rRNA 32S), e na região ITS1 no sítio A₂ (gerando os pré-rRNAs 20S e 27SA₂)(Figura 3B). Os próximos passos do processamento do pré-rRNA 20S ocorrem no citoplasma pela clivagem no sítio D para gerar o rRNA 18S maduro, enquanto o processamento do pré-rRNA 27SA₂ continua no núcleo. O pré-rRNA 27SA₂ é processado nos rRNAs 5.8S e 25S por duas rotas alternativas: cerca de 85% da população de 27SA₂ é clivada no sítio A₃ dentro da região ITS1 pela RNase MRP, e o seguimento 5' - 3' até o sítio B₂ é rapidamente aparado por Rat1; 15% de 27SA₂ é clivado diretamente no sítio B_{1L}. A clivagem no sítio B₂, na extremidade 3' do rRNA 25S, ocorre concomitantemente com a clivagem no sítio B₁. As duas formas de 27SB (27SB_S e 27SB_L) são processadas seguindo vias idênticas nos sítios C₂ → C₁ e pela digestão exonucleolítica da extremidade 3'a 5' até o sítio E pelo complexo exossomo.⁵

23S aberrante. Na ausência da clivagem em A₂, a subunidade 40S não é formada, enquanto a subunidade 60S pode ser formada por clivagens alternativas, por exemplo no sítio A₃.⁵

Em humanos, a biogênese de ribossomos se inicia com a transcrição do pré-rRNA 47S que é convertido no pré-rRNA 45S pela clivagem nos sítios A' dentro da região 5'-ETS e no sítio 6, na extremidade 3' do rRNA 28S maduro (Figura 4). O pré-rRNA 45S é então processado por três vias alternativas de processamento que são determinadas pela ocorrência da primeira clivagem, levando aos rRNAs maduros 18S, 5.8S e 28S. Na via de processamento A, a primeira clivagem do pré-rRNA 45S acontece no sítio 1, que corresponde à extremidade 5' do rRNA 18S maduro, gerando o pré-rRNA 41S. O pré-rRNA 41S é então processado no sítio 2c, separando os pré-rRNAs 21S e 32S. A extremidade 3' do pré-rRNA 21S é processada para gerar o rRNA 18S e o pré-rRNA 32S é clivado no sítio 4b, gerando o pré-rRNA 12S e o rRNA 28S. Por último, a extremidade 3' do pré-rRNA 12S é processada para gerar o rRNA 5.8S maduro. Na via de processamento B, a primeira clivagem do pré-rRNA 45S acontece no sítio 2c, gerando os pré-rRNAs 34S e 32S. O pré-rRNA 34S é processado no sítio 1, gerando o pré-rRNA 21S que é clivado em seguida na extremidade 3' para gerar o rRNA 18S maduro. E o pré-rRNA 32S segue o processamento descrito na via A. Na via de processamento C, a primeira clivagem do pré-rRNA 45S acontece no sítio 4b, gerando o pré-rRNA 37S e o rRNA 28S maduro. Em seguida, o pré-rRNA 37S é processado no sítio 1, gerando o pré-rRNA 26S, que é clivado no sítio 2, gerando o rRNA 18S maduro e o pré-rRNA 17S. Este pré-rRNA é então processado no sítio 3, gerando o pré-rRNA 12S e sua extremidade 3' é processada para gerar o rRNA 5.8S maduro.⁷

De acordo com a função celular fundamental dos ribossomos, várias mutações em proteínas ribossomais e em fatores não ribossomais, que atuam na biogênese do ribossomo, têm sido associadas à mais de 15 doenças genéticas humanas, entre elas a síndrome da anemia Diamond-Blackfan (DBA), disqueratose congênita (DC), síndrome de Shwachman-Bodian-Diamond (SDS), e hipoplasia cartilagem-cabelo (CHH). A maioria das síndromes já descritas estão relacionadas com o grupo de síndromes de insuficiência da medula óssea hereditárias, caracterizadas pela redução do número de células do sangue e predisposição para câncer. Aparentemente, as células precursoras das hemácias têm uma alta demanda por ribossomos, devido ao elevado número de divisões celulares e a necessidade de sintetizar grandes quantidades de hemoglobina, antes da perda do núcleo.¹

Essas ribossomopatias ressaltam a importância da síntese precisa dos ribossomos para uma função celular normal.

Figura 4 – Diagrama do processamento do pré-rRNA 47S em células humanas. O pré-rRNA 47S é convertido no pré-rRNA 45S que é processado por três vias alternativas de processamento dependendo de onde ocorre a primeira clivagem. Maiores detalhes sobre as vias estão descritos no texto.

Fonte: MORELLO et al.⁷

1.2.2 Fatores não ribossomais divergentes

Diversos fatores envolvidos na biogênese do ribossomo foram estudados inicialmente em *S. cerevisiae* e suas funções foram atribuídas para mamíferos, entretanto, as funções destes fatores nem sempre são conservadas. Estudos recentes mostram que as depleções condicionais de fatores envolvidos na biogênese de ribossomo humana produzem fenótipos diferentes dos observados em levedura, como por exemplo as proteínas humanas bystin e hTsr1, homólogas das proteínas de levedura Enp1 e Tsr1, respectivamente, são requeridas para a maturação do rRNA 18S e síntese da subunidade 40S.⁹⁻¹¹

Outro fator ribossomal divergente, é a proteína Nip7p de levedura que participa da biogênese da subunidade 60S do ribossomo, enquanto sua homóloga de humanos, HsNIP7, participa da biogênese da subunidade 40S.⁶⁻⁷

Devido à falta de conhecimento da biogênese de ribossomos em outros organismos eucarióticos, a homóloga de Nip7p/HsNIP7 da alga unicelular *Chlamydomonas reinhardtii*, chamada de CrNIP7, é o objeto principal de estudo deste trabalho, cujos resultados visam contribuir para caracterizar o papel de CrNIP7 e para esclarecer as diferenças de função apresentadas em leveduras e humanos. Assim, uma revisão sobre NIP7 é apresentada no tópico seguinte.

1.3 NIP7

A proteína NIP7 é altamente conservada em eucariotos, variando de 160 a 180 aminoácidos, e está envolvida na biogênese do ribossomo. A proteína Nip7p de *S. cerevisiae* possui cerca de 69% de similaridade com HsNIP7 humana e apresenta homologia a proteínas de eucariotos, *Caenorhabditis elegans* e *Arabidopsis thaliana*;⁶ e Archaeas, *Archaeoglobus fulgidus* e *Pyrococcus abyssi*.¹³ Entretanto, ainda não haviam sido caracterizadas proteínas homólogas a Nip7p em algas unicelulares, como a *C. reinhardtii*. Então, nesse trabalho, uma busca realizada no genoma de *C. reinhardtii* utilizando a sequência de Nip7p de levedura (<http://blast.ncbi.nlm.nih.gov/>) encontrou uma proteína homóloga que foi denominada CrNIP7 (NCBI: XP_001696971).

Em *S. cerevisiae*, Nip7p é codificada por um gene essencial de cópia única (*nip7*) e sua depleção interfere na formação da subunidade 60S do ribossomo, gerando o acúmulo do pré-rRNA não processado 27S e a diminuição da subunidade 60S.⁶ Ainda, foi verificado em levedura que Nip7p interage com as proteínas Nop8p, Nop53p e Sdo1, as quais também estão envolvidas na síntese da subunidade 60S do ribossomo.¹⁴⁻¹⁶ Nip7p também interage com a subunidade do exossomo Rrp43p, envolvida na maturação exonucleásica da extremidade 3' do rRNA 5.8.¹⁴

Entretanto, o *knockdown* de NIP7 em células humanas gerou uma deficiência no ribossomo 40S, com aumento na concentração dos pré-rRNAs 26S e 21S e diminuição do pré-rRNA 34S.⁷ Interessantemente, experimentos de complementação funcional mostram que HsNIP7 e PaNIP7 (de *Pyrococcus abyssi*) não complementam a função de NIP7p de levedura.⁷

Portanto, os fenótipos observados de estudos de depleção condicional e complementação de função sugerem que as proteínas Nip7p e HsNIP7 de mamíferos interagem com um grupo de parceiros diferentes. Entre os parceiros já conhecidos de humanos, estão a proteína SBDS, FTSJ3, Nop132, homólogas a Sdo1, Spb1 e Nop8p, de levedura, respectivamente.¹⁷⁻¹⁹

As proteínas ortólogas NIP7 compartilham uma arquitetura de dois domínios: o domínio C-terminal que contém o motivo conservado PUA (do inglês, *PseudoUridine Synthases and Archaeosine-specific transglycosylases*), que medeia a ligação ao RNA, e um domínio N-terminal que provavelmente medeia a interação com parceiros proteicos.²⁰ As estruturas cristalográficas de HsNIP7 (PDB: 1SQW e 1T5Y) e de PaNIP7 (PDB: 2P38) mostram uma arquitetura formada por estruturas de elementos α/β (Figura 5). A sequência específica do RNA em que as ortólogas NIP7 se ligam ainda não foi identificada, mas estudos mostraram que Nip7p, PaNIP7 e HsNIP7 se ligam aos oligonucleotídeos poli-U e poli-AU, porém HsNIP7 tem maior afinidade pela sequência poli-AU.^{7,20} Como o oligonucleotídeo poli-AU pode fazer interações intermoleculares através de pareamento de bases, é sugerido que NIP7 possa ter maior afinidade por RNAs estruturados.

A interação conservada entre o par Sdo1/Nip7p em levedura e HsSBDS/HsNIP7 em humanos merece destaque uma vez que mutações no gene *sbd*s em humanos causam a Síndrome Schwachman-Bodian-Diamond (SDS).^{16,21} A SDS é uma doença recessiva autossômica de predisposição à leucemia, caracterizada pela falência da medula óssea, disfunção hematológica, insuficiência pancreática exócrina, anormalidades esqueléticas e baixa estatura.²² Estudos genéticos em *S. cerevisiae* e em células de mamíferos mostram que Sdo1/HsSBDS participam da maturação e exportação da subunidade 60S para o citoplasma.^{23,24} Durante as etapas finais da montagem do ribossomo no citoplasma, fatores não ribossomais são liberados do complexo pré-60S e transportados de volta para o núcleo para se envolverem em novos ciclos de maturação da subunidade 60S. Então, Sdo1/HsSBDS em conjunto com o fator de alongação 1 (EFL1), uma GTPase, catalisam a liberação do fator de reciclagem Tif6/elf6, requerido na biogênese e transporte nuclear, das subunidades 60S nascentes.²³⁻²⁴ Portanto, a remoção dos fatores Tif6/elf6 das subunidades 60S nascentes é pré-requisito para a formação dos ribossomos 80S funcional.

Figura 5 – Representação da arquitetura de HsNIP7 de humanos. (A) Estrutura cristalográfica de HsNIP7. (B) Diagrama da topologia de HsNIP7. As α -hélices estão representadas por cilindros verdes, as fitas- β por setas azuis e as 3_{10} -hélices por círculos verdes. As estruturas secundárias das figuras A e B estão com a mesma numeração.

Fonte: Adaptada de LIU et al.²⁵

Outro fato interessante é que PaSBDS e PaNIP7 de *Pyrococcus abyssi* afetam a atividade do exossomo de Archaea. PaSBDS compete com o exossomo pela ligação a sequência de poli-A de RNA, diminuindo levemente sua degradação pelo complexo. PaNIP7 se liga preferencialmente a RNAs ricos em U (uracila) e AU (adenina e timina) e inibe fortemente o exossomo devido sua ligação ao complexo e ao RNA. Estes resultados indicam que PaSBDS e PaNIP7 são fortes candidatos a fatores reguladores do exossomo *in vivo*.²⁶

1.4 *Chlamydomonas reinhardtii*

Chlamydomonas reinhardtii é uma alga verde unicelular, pertencente à família Chlamydomonadaceae, da ordem Volvocales. É muitas vezes isolada a partir de amostras de solo, mas também pode ser encontrada em lagos. Cada célula de *Chlamydomonas* apresenta um formato oval de, aproximadamente, 5 a 10 µm de diâmetro, possui ainda múltiplas mitocôndrias, dois flagelos anteriores para motilidade e acasalamento e um cloroplasto único (Figura 6). A parede celular dessa alga não é formada por celulose, mas sim por glicoproteínas ricas em hidroxiprolina.²⁷⁻²⁸ *C. reinhardtii* é amplamente utilizada como modelo para estudos da fotossíntese eucariótica porque, diferente das angiospermas, esta alga pode crescer na ausência de luz através de uma fonte orgânica de carbono, permanecendo com seu aparato fotossintético funcional. Além disso, é também usada como modelo de estudo de estrutura e função de flagelos eucarióticos e corpos basais.^{27,29}

Figura 6 – A alga unicelular, *Chlamydomonas reinhardtii*. (A) Representação esquemática dos principais componentes de uma célula de *Chlamydomonas*: flagelos (F), corpos basais (B), núcleo (N), estigma (E) aparelho de Golgi (G), mitocôndria (M), cloroplasto (C), tilacóide (T), pirenóide (P). Adaptado de Gumen et al ²⁸ (B) Microscopia eletrônica de varredura de uma célula de *C. reinhardtii* (imagem cedida pelo Laboratório de Microscopia e Microanálise - CTMH). (C) Microscopia óptica de uma célula (12 µm x 17 µm) de *C. reinhardtii* (imagem cedida pelo *The Protist Information Server*, no endereço eletrônico http://protist.i.hosei.ac.jp/Protist_menuE.html).

Fonte: Elaborada pela autora.

As Clorófitas (algas verdes, incluindo *Chlamydomonas*) divergiram dos Streptófitos (plantas terrestres e parentes próximos) há cerca de 1 bilhão de anos. As algas verdes e plantas terrestres fazem parte do grupo de plantas verdes, Viridiplantae, que divergiu anteriormente do grupo Opisthokonte, o qual compreende animais, fungos e Choanozoa.²⁹ Com a disponibilização do genoma de *C. reinhardtii* em 2007, por Merchant e colaboradores, análises genômicas comparativas permitiram que muitos genes fossem rastreados e relacionados a seus ancestrais.²⁹ A partir dessas análises, observou-se que muitos genes de *Chlamydomonas* e de angiospermas são derivados de genes de um vegetal ancestral, incluindo aqueles associados às funções de fotossíntese. Por outro lado, há certos genes compartilhados entre *Chlamydomonas* e animais, atribuídos ao último ancestral vegetal-animal comum, que incluem os genes relacionados à flagelos e corpos basais.²⁹

1.4.1 Estrutura geral do genoma

O genoma nuclear de *C. reinhardtii*, de aproximadamente 121 Mb, está distribuído em 17 cromossomos em uma célula vegetativa normalmente haploide, mas diploides estáveis podem ser selecionados utilizando marcadores auxotróficos.²⁸⁻²⁹ *Chlamydomonas* ainda possui dois tipos de *mating*, que são governados por um *mating locus* nas células haploides, sendo MT+ ou MT-.²⁸

No trabalho de Merchant e colaboradores²⁹, predições genéticas *ab initio* e baseadas em homologia integradas com evidências de EST (do inglês, *Expressed Sequence Tag*) foram usadas para criar um conjunto referência de 15143 modelos de predições de genes codificantes de proteína. Desse conjunto, 8631 modelos (56%) são suportados por mRNA e sequências de EST e 35% estão sendo editados e anotados manualmente desde 2007.

O *cluster* genético que codifica os rRNAs citosólicos 18S/5.8S/28S) são encontrados em dois sítios, um na extremidade do cromossomo 14 e o outro na extremidade do cromossomo 8. O *cluster* único que codifica o rRNA 5S também está localizado no cromossomo 8, cerca de 200 kb do *cluster* 18S/5.8S/28S.²⁸ Ademais, o gDNA codifica 322 snoRNAs que estão envolvidos no processamento de rRNA e modificação de bases.^{28,29} É interessante observar que o conteúdo de bases CG do gDNA de *C. reinhardtii* é incomumente alto, em torno de 64%, o que pode dificultar a clonagem de genes.²⁷

O DNA mitocondrial (mtDNA) de *C. reinhardtii*, ao contrário de outras algas e maioria das plantas, é linear e relativamente pequeno, com 15,8 kb. O mtDNA codifica poucos genes: *cob*, *cox1*, cinco subunidades da NADH desidrogenase mitocondrial, e uma transcriptase reversa, além de rRNAs mitocondriais, e três tRNAs.²⁷⁻²⁸ Mutantes com o gene *cob* deletado não são capazes de crescer em acetato como fonte de carbono no escuro, mas são viáveis quando crescem por fotossíntese.²⁷ Outra característica interessante, é a fragmentação dos genes de rRNA mitocondriais que são transcritos separadamente e estão evoluindo rapidamente em relação aos outros genes mitocondriais e rRNAs citosólicos.²⁸

Já o genoma do cloroplasto (cpDNA) de *Chlamydomonas* é circular e possui 220 kb, sendo semelhante aos de outras algas verdes e plantas terrestres. Este cpDNA codifica 99 genes, incluindo tRNAs e rRNAs, e apresenta 80 cópias por célula, sendo empacotado em nucleóides que contém por volta de 10 cópias cada. A replicação do cpDNA não está sincronizada com a divisão nuclear, e ocorre continuamente durante o ciclo celular.²⁸

1.4.2 Ribossomos

C. reinhardtii possui três tipos de ribossomo que estão em compartimentos diferentes da célula: o citosólico, do cloroplasto e da mitocôndria. Este trabalho está focado na biogênese do ribossomo citosólico, que possui as características de ribossomo eucarioto. Existem poucas referências na literatura sobre a biogênese de ribossomos e processamento de rRNA ribossomal em *C. reinhardtii*.³⁰⁻³¹ Abaixo estão listadas, resumidamente, algumas das características importantes de cada tipo de ribossomo.

a) Ribossomo citosólico:

O ribossomo citosólico de *C. reinhardtii* apresenta coeficiente de sedimentação em sacarose de 80S e é bastante conservado em composição e estrutura em relação aos ribossomos 80S de humanos, leveduras e plantas superiores. É estimado por gel de eletroforese bidimensional que 96 proteínas compõem o ribossomo citosólico de *C. reinhardtii*, sendo 45 da subunidade menor, 48 da subunidade maior e três proteínas únicas do ribossomo 80S maduro.³²⁻³³ As proteínas ribossomais preditas de *Arabidopsis thaliana*, apresentam alta

similaridade (média de 67% de identidade de aminoácidos) com as de *C. reinhardtii* e as proteínas homólogas de alga e leveduras compartilham menor similaridade em comparação com proteínas de alga e mamíferos. As pequenas diferenças entre as proteínas ribossomais de alga e outros organismos aparecem nas duas subunidades e em proteínas que interagem tanto com rRNA, mRNA e fatores de tradução, mostrando que não são particulares de um grupo específico, sugerindo que a função dessas proteínas no contexto do ribossomo deve ser mantida.³²

A estrutura tridimensional do ribossomo 80S de *C. reinhardtii* foi resolvida por criomicroscopia com resolução de 25 Å. Nessa resolução, a estrutura geral do ribossomo da alga mostra alta conservação com o ribossomo de levedura. As similaridades estruturais se estendem para as duas subunidades e por toda parte do ribossomo observada em qualquer orientação, as diferenças de composição proteica e de rRNA não são visíveis na reconstrução do ribossomo nessa resolução.³²

b) Ribossomo do cloroplasto:

Os proteomas do ribossomo do cloroplasto da alga verde e de plantas terrestres já foram elucidados, sendo que a maioria dos componentes proteicos tem clara homologia com os ribossomos de bactéria, mas algumas proteínas e domínios únicos de cloroplastos também foram identificados.³⁴⁻³⁵ O ribossomo do cloroplasto de *C. reinhardtii* sedimenta em gradiente de sacarose com coeficiente de fracionamento 70S, e possui as subunidades de 30S e 50S, assim como o ribossomo de bactéria.³⁶ Na subunidade menor, foram identificadas 19 proteínas ortólogas de *Escherichia coli*, uma homóloga de espinafre da proteína ribossomal plastídeo-específica (PSRP), PSRP-3, e uma nova proteína PSRP-7, que possui dois domínios S1. Na subunidade maior, 27 proteínas ortólogas de *E. coli* foram identificadas, assim como a ortóloga de espinafre plastídeo-específica, PSRP-6. Dentre as proteínas de *C. reinhardtii* identificadas nas duas subunidades, algumas delas apresentam pequenas inserções, ou extensões, sendo maiores que suas homólogas de *E. coli* e espinafre. Ademais, duas proteínas, RAP38 e RAP41, foram identificadas a partir do ribossomo 70 S e não foram encontradas nas subunidades isoladas. Análises filogenéticas revelaram que essas duas proteínas são parálogas da CSP41 de espinafre, uma proteína do cloroplasto ligadora de RNA e com atividade endoribonucleásica.³⁴⁻³⁵

Apesar de compartilharem muitos componentes do ribossomo, a regulação da tradução no cloroplasto é mais complexa do que em bactéria e essa complexidade requer RNAs e proteínas que não são encontrados no sistema procariótico. Recentemente, um estudo de criomicroscopia eletrônica determinou a estrutura do ribossomo de *C. reinhardtii* com resolução de 15.5 Å. Este conhecimento é muito importante para estudos de tradução no plastídeo, já que o cloroplasto vem ganhando importância como plataforma biotecnológica para expressão heteróloga de proteínas.³⁷⁻³⁸

c) Ribossomo da mitocôndria

A composição do ribossomo mitocondrial de *C. reinhardtii* e como acontece a tradução dentro da mitocôndria são questões que ainda não foram elucidadas. Sabe-se que o coeficiente de sedimentação aparente em sacarose do ribossomo mitocondrial de *Chlamydomonas eugametos* é de aproximadamente 60-66S, distinto dos ribossomos citosólicos (80S) e do cloroplasto (70S) caracterizados em *C. reinhardtii*. Em termos de valores de sedimentação, o ribossomo mitocondrial de *C. eugametos* difere dos de plantas terrestres (77-78S) e aparece como intermediário entre os ribossomos mitocondriais de animais (55-60S) e de euglena e fungos (70-74S).³⁹⁻⁴⁰

1.4.3 Importância e aplicações

Durante os últimos anos, as microalgas surgiram como uma potencial plataforma biotecnológica, que atrai o interesse acadêmico e industrial, devido às suas particularidades biológicas, que incluem alta taxa de crescimento e baixo custo de cultivo (não precisa de uma fonte orgânica de carbono para crescer), aliada a características encontradas nos sistemas de bactérias e leveduras que são as plataformas biotecnológicas mais estudadas e utilizadas atualmente.⁴¹⁻⁴²

Os três genomas de *C. reinhardtii* (nuclear, cloroplastos e mitocôndria) são transformáveis e os protocolos de transformação estão bem estabelecidos. A vantagem da organização celular complexa da alga unicelular possibilita a compartimentalização das reações bioquímicas e facilita o fornecimento de precursores, o que não é possível em hospedeiros

bacterianos, por exemplo. A transformação nuclear de *C. reinhardtii* é randômica e pode ser feita por biobalística, por eletroporação, por *Agrobacterium* ou até mesmo por agitação com contas de vidro. Já a transformação do cloroplasto e mitocôndria ocorre por recombinação homóloga, possibilitando alteração de um alvo específico.^{28,41}

A transformação do cloroplasto tem sido mais utilizada, visto a quantidade de trabalhos realizados nos últimos anos. Interessantemente, o sistema do repressor lac, de bactéria, também funciona em *C. reinhardtii*.⁴¹ Mais de 50 proteínas recombinantes já foram produzidas com sucesso no cloroplasto de *C. reinhardtii*. Inicialmente eram produzidos marcadores e repórteres, contudo, o foco atual é usar esta organela como uma plataforma de baixo custo para a produção produtos de maior valor agregado.⁴² O interesse agora é utilizar a luz para guiar a síntese de novas moléculas de hidrocarbonetos para uso como biocombustíveis, proteínas terapêuticas (vacinas, hormônios, enzimas e anticorpos), vitaminas, carotenoides, e ácidos graxos poli insaturados.⁴¹⁻⁴² De fato, há trabalhos que reportaram a produção de proteínas humanas terapêuticas como eritropoietina e imunoglobulina A em cloroplasto de *C. reinhardtii*.⁴¹

Esses avanços destacam a importância da alga unicelular, *C. reinhardtii*, como uma plataforma biotecnológica, porém ainda há o desafio de desenvolver estratégias de engenharia metabólica para torná-la economicamente viável.

1.5 Motivação para os estudos funcionais de CrNIP7

Apesar de sua importância destacada como organismo modelo de estudo de flagelos e fotossíntese e ter grande potencial biotecnológico como mencionado, menos de 100 genes que codificam proteínas em *C. reinhardtii* foram validados funcionalmente até o momento. Este número é muito pequeno em comparação aos 6800 genes, aproximadamente, que já foram caracterizados experimentalmente em *Arabidopsis thaliana*.⁴¹ Desse modo, a proposta principal desse trabalho de mestrado foi ampliar os estudos sobre biogênese de ribossomos em outro eucarioto, neste caso, em *Chlamydomonas reinhardtii*, além de colaborar para a validação de genes preditos a partir do genoma de *C. reinhardtii*.

Como *C. reinhardtii* é um organismo evolutivamente ancestral a plantas, mas compartilha genes com animais, atribuídos ao último ancestral comum a vegetais e animais,

essa característica faz com que o entendimento do processamento de rRNA neste organismo seja relevante, sendo que o conhecimento gerado poderá, por exemplo, contribuir para esclarecer diferenças encontradas na biogênese de ribossomos em leveduras e humanos. Portanto, escolhemos como alvo central do trabalho a proteína CrNIP7, homóloga a Nip7p e HsNIP7, as quais foram caracterizadas por terem funções divergentes no processamento do rRNA em leveduras e humanos.

1.6 Objetivos

Os objetivos gerais do trabalho foram:

- ✓ Isolar a CDS de CrNIP7;
- ✓ Avaliar a complementação funcional de CrNIP7 em *Saccharomyces cerevisiae*;
- ✓ Encontrar potenciais parceiros de interação de CrNIP7 *in vivo* pelo sistema de duplo híbrido em levedura, utilizando uma biblioteca comercial de *C. reinhardtii*;
- ✓ Produzir as potenciais proteínas parceiras encontradas pelo ensaio de duplo híbrido;
- ✓ Validar as interações encontradas por *pull down*;
- ✓ Avaliar a potencial interação *in vivo*, por duplo híbrido dirigido, e *in vitro*, por *pull down*, de CrNIP7 e CrSBDS.

MATERIAIS E MÉTODOS

2 MATERIAIS E MÉTODOS

A seção de Materiais e Métodos está subdividida em métodos de biologia molecular para isolamento das sequências de DNA codificantes de interesse e outras técnicas complementares voltadas aos estudos funcionais. A Tabela 1 resume os experimentos relacionados com cada proteína estudada neste trabalho.

Tabela 1 – Experimentos relacionados com cada proteína.

	Amplificação	Clonagem	Complementação funcional	Duplo híbrido	Expressão heteróloga	Purificação	Co-expressão e co-purificação	Pull-down	Espectrometria de massas
CrNIP7	X	X	X	X	X	X	X	X	X
G-patch	X	X		X	X	X	X		
Predicted	X	X		X	X	X	X	X	X
CrSBDS	X	X		X	X	X	X	X	X

Fonte: Elaborada pela autora.

2.1 Isolamento dos Sequências de DNA codificantes (CDS)

2.1.1 Obtenção do DNA complementar (cDNA) total de *Chlamydomonas reinhardtii*

A linhagem utilizada foi a CC325 de *Chlamydomonas reinhardtii* que é um mutante dependente de arginina, com fenótipo [325 arg– (cw15 asl– mt+)], cedida pelo Dr. Andrew Bassett, da Universidade de Cambridge, UK.

2.1.1.1 Extração de RNA mensageiro (mRNA) total

Foram coletadas aproximadamente 90 mg (aproximadamente 10 mL de cultura) de células de *C. reinhardtii* de uma cultura não sincronizada crescida em meio Tris-acetato fosfato (TAP) a 25°C, sob luz contínua e agitação por 7 dias. A extração do mRNA foi realizada através do *μMACS™ mRNA Isolation kit* (MACS® Molecular) seguindo as especificações do fabricante.

2.1.1.2 Transcrição Reversa e Reação em Cadeia da Polimerase (RT-PCR)

A síntese de cDNA total foi realizada pela transcrição reversa do mRNA extraído anteriormente. A reação de RT-PCR foi realizada em um termociclador Mastercycler® gradiente 5331 (Eppendorf) utilizando-se o kit *Super Script®III First-Strand Synthesis Super Mix* (Life Technologies) e oligos-dT como iniciadores.

As sequências codificantes (CDS) das proteínas utilizadas nesse trabalho foram isoladas e amplificadas por reação de PCR em um termociclador Mastercycler® gradiente 5331 (Eppendorf) tendo o cDNA total de *C. reinhardtii* como molde para as CDS de CrNIP7, CrSBDS e G-patch. A CDS da proteína denominada Predicted foi amplificada a partir de um clone isolado no experimento de duplo híbrido.

Em um microtubo de 0,5 mL foram adicionados 0,5 μL de cDNA total de *C. reinhardtii*; 0,2 mM da mistura de desoxirribunucleotídeos trifosfatos (dNTPs); 0,2 μM de *primer Forward*; 0,2 μM de *primer Reverse*; 1X *High Fidelity PCR buffer with 15 mM MgCl₂*; 3% dimetilsulfóxido (DMSO); 2,5 U de *High Fidelity PCR Enzyme Mix* (ThermoScientific) para um volume final de reação de 50 μL.

As condições do ciclo de PCR foram determinadas de acordo com as especificações da enzima e das temperaturas *melting* (T_m) dos oligonucleotídeos iniciadores (*primers*). Os produtos das reações foram analisados por eletroforese em gel de agarose 0,8% preparado em tampão TAE 1X (Tris 40 mM, ácido acético glacial 0,11% e EDTA 1 mM).

As sequências dos oligonucleotídeos utilizados para a amplificação das sequências codificadoras, com os respectivos sítios de restrição e T_m , estão mostrados na Tabela 2 abaixo.

Tabela 2 – Oligonucleotídeos utilizados para amplificação das sequências codificadoras das proteínas estudadas neste trabalho.

CDS	Primer F	Sequência (5' - 3')	*T _m (°C)	Primer R	Sequência (5' - 3')	*T _m (°C)
<i>CrNip7</i>	<i>Bam</i> HI-F	GGATCCATGCGTCCGCTCACAGAAGAGGAGACG	69,6	<i>Pst</i> I-R	ACTGCAGTTATAGCTCGTCCTCGGC	56
	<i>Eco</i> RI-F	GAATTCATGCGTCCGCTCACAGAAGAGGAGACG	66,1	<i>Sal</i> I-R	GTCGACTTATAGCTCGTCCTCGGCACGCAGAAAC	68,1
	<i>Eco</i> RI-G-F	AGAATTC ^G ATGCGTCCGCTCACAGAAG	56	<i>Xba</i> I-R	TCTAGATTATAGCTCGTCCTCGGCACGCAGAAAC	65,7
	<i>Nde</i> I-F	CATATGATGCGTCCGCTCACAGAAGAGGAGAC	66,1			
<i>G-patch</i>	<i>Bam</i> HI- <i>Nde</i> I-F	AGGATCCCATATGGCGGCCCCAGAAGTAG	61	<i>Kpn</i> I- <i>Sal</i> I-R	AGTCGACGGTACCTTACTCCGCTAACGTCACGAAGTTG	64
	F-interno	ACCTGGACGCCATCTTGG	47	R-interno	ACAGCCTTTCTCCCCGAG	47
<i>predicted</i>	<i>Bam</i> HI- <i>Nde</i> I-F	AGGATCCCATATGGACTTTCTCCCTGTTGCGC	61	<i>Xho</i> I-R	ACTCGAGTTACACGCGCCCCTGCG	59
<i>CrSBDS</i>	<i>Sma</i> I- <i>Nde</i> I-F	ACCCGGGCATATGTCCAAGGGAGTCTTCCAGCCAGTC	66	<i>Xho</i> I-R	ACTCGAGTCATGCCGCCGCCTTGCTG	61
	<i>Sma</i> I-C-F	ACCCGGG ^C ATGTCCAAGGGAGTCTTCC	61			

*T_m = temperatura de *melting* fornecida pela empresa que sintetizou os oligonucleotídeos.

Letras em vermelho correspondem a base nitrogenada inserida na extremidade 3' do sítio de restrição para correção da fase de leitura.

Fonte: Elaborada pela autora.

Plasmídeos de propagação foram usados para clonagem das CDSs, a partir dos quais foram subclonados em outros vetores, conforme descrito no item 2.1.2.

2.1.2 Clonagens

Os produtos de PCR foram ligados em vetores de propagação linearizados pGEMT-Easy (Promega) ou pUC19 (previamente linearizado no sítio de *Sma*I). Para a ligação em pGEMT-Easy, foi necessário adicionar adeninas nas extremidades do fragmento amplificado com a enzima *High Fidelity PCR Enzyme Mix* devido à sua atividade exonucleásica 3'-5'. Resumidamente, em microtubo de 0,2 mL adicionou-se 200 ng de DNA, 1X *Taq Reaction Buffer*; 2,5 mM MgCl₂; 0,2 mM dATP; 5 U de *Taq DNA Polymerase* (ThermoScientific) para um volume final de reação de 10 µL e incubou-se a reação por 20 minutos a 70°C. Após a adenilação, o fragmento foi ligado ao vetor pGEMT-Easy na proporção 6:1 (inserto:vetor) seguindo as especificações do kit *pGEM®-T and pGEM®-T Easy Vector Systems –Quick Protocol* (Promega).

Para ligação dos fragmentos amplificados por *High Fidelity PCR Enzyme Mix* no pUC19, o vetor foi pré-linearizado pela digestão com *Sma*I, resultando em extremidades cegas. O produto de PCR foi ligado ao vetor linearizado na proporção 10:1 (inserto:vetor) utilizando a enzima T4 DNA ligase (ThermoScientific).

Após a confirmação por sequenciamento dos clones positivos, os fragmentos foram subclonados em vetores de expressão bacterianos ou em vetores de expressão em leveduras (vetor *shuttle*, propagados em bactérias e expressos em leveduras). Na subclonagem, a ligação foi direcional, utilizando um par de enzimas de restrição diferentes para cada inserto, e as ligações foram feitas na proporção 6:1 (inserto:vetor) utilizando a enzima T4 DNA ligase (ThermoScientific).

2.1.2.1 Transformação de bactéria por choque térmico

Células de *Escherichia coli*, da linhagem DH5α, utilizada para propagação plasmidial, e da linhagem *E. coli* Rosetta (DE3), utilizada para expressão heteróloga, foram quimicamente tratadas para tornarem-se competentes e serem capazes de receber DNA exógeno por meio de choque térmico através do protocolo descrito em Sambrook.⁴³

Em um microtubo de 1,5 mL adicionou-se uma alíquota de 50 µL da suspensão de células competentes à 30 ng de DNA plasmidial circularizado ou a todo o volume de uma reação de ligação. A mistura foi incubada por 15-30 minutos em banho de gelo, logo após foi dado um choque térmico de 1 minuto e 30 segundos a 42°C. Em seguida, adicionou-se 200 µL de meio LB (*Lysogenic broth*) às células e estas foram incubadas a 37°C por 1 hora sob agitação. As células foram plaqueadas em meio LB ágar, contendo os antibióticos de seleção apropriados.

2.1.2.2 Extração plasmidial de bactéria

As culturas bacterianas foram crescidas em 6 mL de meio LB, com os antibióticos apropriados, a 37°C por 16h sob agitação de 200 rpm, e então submetidas às extrações de DNA plasmidial pelo kit *Wizard®Plus SV Minipreps DNA Purification System* (Promega). Na eventual falta desse kit, a extração foi realizada pelo método de lise alcalina descrito em Sambrook.⁴³ A concentração do DNA plasmidial foi medida pelo equipamento NanoDrop 1000 (ThermoScientific).

2.1.2.3 Digestão e análise de restrição

Nas análises de restrição para confirmação da clonagem, foram utilizados aproximadamente 500 ng de DNA plasmidial. Para reações preparativas, foram utilizados 5 µg de DNA. Após a incubação adequada, as reações foram submetidas à eletroforese em gel de agarose 0,8% em tampão TAE (Tris 40 mM, ácido acético glacial 0,11% e EDTA 1 mM), sendo as bandas de interesse recortadas e o DNA extraído do gel com auxílio do *kit Wizard® SV Gel and PCR Clean-Up System* (Promega).

2.2 Ensaio de complementação funcional

O ensaio de complementação funcional visou avaliar se a proteína CrNIP7, de *Chlamydomonas reinhardtii*, era capaz de complementar a função da proteína Nip7p de *S. cerevisiae*. Para isso, dois experimentos de complementação, utilizando plasmídeos e linhagens de leveduras com *backgrounds* genéticos diferentes, foram realizados. A clonagem da CDS de CrNIP7 nos plasmídeos e a transformação das leveduras estão descritas em conjunto para as duas abordagens e as particularidades de cada experimento serão descritas em tópicos separados. As etapas gerais dos experimentos estão esquematizadas na Figura 7.

Figura 7 – Esquema do ensaio de complementação funcional. As linhagens de *S. cerevisiae* DG130, DG440 e DG442 estão especificadas na Tabela 3.

Fonte: Elaborada pela autora.

2.2.1 Linhagens de levedura e plasmídeos

As linhagens de *S. cerevisiae* e os plasmídeos utilizados nas construções necessárias para esse experimento foram cordialmente cedidos pelo Professor Dr. Nilson I. T. Zanchin e estão apresentados na Tabela 3.

Tabela 3 – Linhagens de *S. cerevisiae* utilizadas do ensaio de complementação funcional.

Linhagem	Genótipo	Origem
DG130	<i>MATa CYC1-NLS cyc7-67 lys5-10 ura3-52 nip7-1</i>	
DG440	<i>MATa ade2-1 leu2-3,112 trp1-1 ura3-1 nip7::HIS3 p[LEU2 ARS1 NIP7]</i>	Nilson I. T.
DG442	<i>MATa ade2-1 leu2-3,112 trp1-1 ura3-1 nip7::HIS3 p[LEU2 ARS1 GAL1::nip7-1]</i>	Zanchin

Fonte: ZANCHIN et al.⁶

Tabela 4 – Plasmídeos utilizados nos experimentos de complementação funcional.

Plasmídeo	Características relevantes	Tamanho	Origem
pJN40	<i>CUP1, URA3, 2 μm, Amp^R</i>	5,5 kb	J. Nelson (não publicado)
pJN40-CrNip7	<i>CUP1, URA3, 2 μm, Amp^R</i>	6 kb	Esse estudo
YCplac33-TEF	<i>TEF1, URA3, CEN/ARS, Amp^R</i>	6 kb	Nilson I. T. Zanchin (comunicação pessoal)
YCplac33-TEF-CrNip7	<i>TEF1, URA3, CEN/ARS, Amp^R</i>	6,5 kb	Esse estudo

Fonte: Elaborada pela autora.

2.2.2 Clonagem da CDS de CrNip7 nos plasmídeos de levedura

Duas amplificações da CDS de CrNIP7 foram realizadas, paralelamente, a partir de um clone positivo obtido anteriormente a fim de trocar os sítios de restrição para a subclonagem em vetores de levedura. Os pares de oligonucleotídeos utilizados para cada amplificação foram *CrNip7-EcoRI-F* e *CrNip7-SalI-R*; e *CrNip7-BamHI-F* e *CrNip7-XbaI-R* (Tabela 2). Ambos fragmentos obtidos foram ligados ao vetor de propagação pGEMT-Easy (Promega), as clonagens foram confirmadas por análises de restrição e sequenciamento. Em seguida, os fragmentos *CrNip7* (com extremidades *EcoRI/SalI*) e *CrNip7* (*BamHI/XbaI*) foram subclonados nos vetores pJN40 e YCplac33-TEF, respectivamente. As etapas de amplificação e clonagem foram realizadas conforme descrito em 2.1.2.

2.2.3 Transformação de levedura pelo método “*Quick and Easy Trafo*”

Quando poucos transformantes são necessários, o método “*Quick and Easy Trafo*”⁴⁴ é bastante rápido e flexível. Neste trabalho, foram feitas algumas adaptações no protocolo porque algumas das linhagens são sensíveis à temperatura e possuem crescimento lento.

Os estoques em glicerol das linhagens DG130, DG440 e DG442 foram estriados em placa contendo meio de cultura YPGal, para a DG442, e YPD (Tabela 5) para DG130 e DG440 (Tabela 2.3), as placas foram incubadas por 3-4 dias a 24 °C. Uma colônia (2-3 mm de diâmetro, com menos de 4 semanas estocadas) de cada linhagem foi inoculada em 5 mL de meio de cultura descrito e incubada a 24°C sob agitação de 200 rpm por 36 horas para DG130 ou 24 horas para DG440 e DG442. Um volume de 2 mL de cada cultura foi centrifugado por 30 segundos a 16000 x *g*. Após a remoção do sobrenadante, as células foram ressuspensas em 1 mL de solução de acetato de lítio 100 mM e então incubadas por 5 min a 30°C. As células foram coletadas por centrifugação de 1 minuto a 3000 x *g* e o sobrenadante foi removido. Ao precipitado celular, foram adicionados, na seguinte ordem: 240 µL de PEG 50%, 36 µL de acetato de lítio 1 M, 5 µL de DNA de espermatozoário de salmão (10 mg/mL) (previamente desnaturado por 15 minutos a 95°C e colocado imediatamente no gelo) e 5 µL de DNA plasmidial (de 100 ng a 5 µg) (Tabela 6), e 65 µL de água ultrapura estéril. O microtubo foi agitado vigorosamente (vórtex) por 1 minuto e incubado a 42°C por 20 min. As suspensões celulares foram centrifugadas por 30 segundos a 16000 x *g* e o sobrenadante foi removido com pipeta. As células foram ressuspensas cuidadosamente em 100 µL de água ultrapura estéril e então plaqueadas em meio seletivo apropriado SD-ura +Gal (Tabela 5) para a linhagem DG442 e SD -ura +Glu (Tabela 5) para as linhagens DG130 e DG440. As placas foram incubadas a 24°C por 4 a 5 dias ou até o aparecimento de colônias, sendo que as colônias resultantes continham o plasmídeo.

Tabela 5 – Meios de cultura utilizados no ensaio de complementação funcional.

Meio de cultura	Conteúdo
YPD	1% de extrato de levedura, 2% peptona, 2% glicose
YPGal	1% de extrato de levedura, 2% peptona, 1% galactose, 1% rafinose
SD –ura +Glu	6,7 g/L de YNB, 2g/L de DO –ura, 2% glicose
SD –ura +Gal	6,7 g/L de YNB, 2g/L de DO –ura, 1% galactose, 1% rafinose

Adicionar 2% de ágar para meio sólido. SD= meio de cultura mínimo. YNB = Difco yeast nitrogen base w/o amino acids. DO –ura = Drop-out Supplement –ura (Clontech).

Fonte: Elaborada pela autora.

Tabela 6 – Resumo das transformações realizadas e meios de cultura para seleção plasmidial.

Linhagem	DNA plasmidial	Meio seletivo
DG130	Água (controle)	SD -ura +Glu
DG130	pJN40	SD -ura +Glu
DG130	pJN40-CrNip7	SD -ura +Glu
DG440	Água (controle)	SD -ura +Glu
DG440	YCplac33-CrNip7	SD -ura +Glu
DG442	Água (controle)	SD-ura +Gal
DG442	YCplac33	SD-ura +Gal
DG442	YCplac33-CrNip7	SD-ura +Gal

Para o controle negativo de transformação foi realizado o mesmo procedimento substituindo os DNAs plasmidiais por água ultrapura estéril.

Fonte: Elaborada pela autora.

2.2.4 Ensaio de complementação em levedura utilizando a linhagem DG130

Três colônias transformantes de DG130 [pJN40-CrNip7] e 3 colônias transformantes de DG130 [pJN40] foram cultivadas em 5 mL de meio SD –ura +Glu a 24°C, sob agitação de 200 rpm, por 48 horas. A absorbância em 600 nm ($D.O._{600nm}$) das culturas celulares foi medida e 1 mL de cada cultura foi centrifugado por 30 segundos a 16.000 x *g*. O precipitado celular foi resuspenso em glicerol 20% estéril, normalizando as concentrações celulares (por exemplo, se a $D.O._{600nm}$ medida da cultura celular fosse igual 1.0, o precipitado celular seria resuspenso em 100 µL). Todo o volume da suspensão de células foi transferido para o primeiro poço de uma microplaca de 96 poços, e nos 5 poços seguintes foram adicionados 90 µL de glicerol 20% estéril. Cinco diluições seriadas da ordem de 10x foram feitas utilizando um pipetador

multicanal. Um volume de 4 μ L de cada diluição foi aplicado em 3 placas contendo meio de cultura SD -ura +Glu contendo 0,1 mM de CuSO_4 (Sulfato de Cobre II) e as placas foram incubadas nas temperaturas de 24°C, 30°C e 37°C por 3-4 dias. Esse experimento foi repetido no mínimo três vezes para confirmar o resultado.

2.2.5 Ensaio de complementação em levedura utilizando as linhagens DG440 e DG442

Três colônias transformantes de D440 [YCplac33] foram cultivadas em 5 mL de meio SD -ura +Glu, e 3 colônias transformantes de DG442 [YCplac33] e de DG442 [YCplac33-*CrNip7*] foram cultivadas em meio SD -ura +Gal a 24°C, sob agitação de 200 rpm por 36 horas. A absorbância em 600 nm ($\text{D.O.}_{600\text{nm}}$) das culturas celulares foi medida e 1 mL de cada cultura foi centrifugado por 30 segundos a 16000 x *g*. O precipitado celular foi resuspenso em glicerol 20% estéril normalizando as concentrações celulares. Todo o volume da suspensão de células foi transferido para o primeiro poço de uma microplaca de 96 poços, e nos 5 poços seguintes foram adicionados 90 μ L de glicerol 20% estéril. Cinco diluições seriadas da ordem de 10x foram feitas utilizando um pipetador multicanal. Um volume de 4 μ L de cada diluição foi aplicado em 3 placas contendo meio de cultura SD -ura +Glu e em 3 placas contendo meio de cultura SD -ura + Gal. As placas foram incubadas nas temperaturas de 24°C, 30°C e 37°C por 3-4 dias. Esse experimento foi repetido no mínimo três vezes para confirmar o resultado.

A parte experimental do estudo da *CrNIP7*, incluindo a amplificação da CDS, clonagens e ensaio de complementação funcional em levedura, foi realizada sob supervisão e auxílio da Dra. Juliana Sposto Avaca Crusca, colaboradora neste projeto.

2.3 Ensaio de Duplo Híbrido em levedura

O sistema de duplo híbrido em leveduras foi primeiramente desenvolvido por Fields e Song em 1989⁴⁵ como um ensaio genético *in vivo* para detectar a interação entre duas proteínas. Atualmente, existem variações otimizadas desse sistema que é amplamente utilizado, por exemplo, para fazer triagem de biblioteca de cDNA em buscas de novos parceiros de interação de uma proteína em estudo.

As proteínas ativadoras de transcrição geralmente possuem um domínio de ligação ao DNA (*Binding Domain* - BD) e um domínio de ativação da transcrição (*Activation Domain* - AD) que possuem funções independentes. Portanto, a ideia principal do ensaio de duplo híbrido consiste em explorar o fato de que esses dois domínios podem ser expressos separadamente, e posteriormente quando aproximados por interações não covalentes, eles restauram a função da proteína intacta, sendo capazes de ligar ao DNA e ativar a transcrição (Figura 8).⁴⁶⁻⁴⁷ Devemos ressaltar que a separação física dos domínios resulta em um fator de transcrição não funcional. Portanto, na abordagem clássica e resumida do sistema de duplo híbrido, uma proteína de interesse é fusionada ao domínio de ligação ao DNA (BD), e a proteína híbrida resultante é comumente chamada de “BD-isca”. Outra proteína de interesse é fusionada ao domínio de ativação de transcrição (AD), chamada de “AD-presa”. As duas construções são co-transformadas na mesma levedura, e se as proteínas isca e presa interagirem, os domínios BD e AD são fisicamente aproximados, mediados pela interação proteica, promovendo a transcrição de um gene repórter sob a atividade de um promotor reconhecido pelo domínio de ligação. Porém, se não houver interação entre as proteínas sob investigação, não há transcrição do gene repórter (Figura 8).

Geralmente, a interação é avaliada pela transcrição de mais de um gene repórter, que pode ser uma marca auxotrófica, *HIS3*, e colorimétrica, *LacZ*, para confirmar o resultado. O sistema de duplo híbrido *in vivo* possui algumas vantagens como: a busca de novos parceiros de interação por triagem de biblioteca de cDNA, a informação genética (cDNA) pode ser utilizada para outros estudos, a triagem é relativamente rápida e não precisa de ajustes de otimização como condições ótimas de ligação e lavagens como experimentos *in vitro*. A desvantagem é que existe chance de falsos positivos durante a triagem da biblioteca, os quais são analisados e descartados nas etapas de confirmação e validação da interação.

As etapas realizadas do ensaio de duplo híbrido estão resumidas e esquematizadas na Figura 9. A linhagem de levedura, os plasmídeos e as construções da “isca” e varredura da biblioteca de cDNA utilizadas nesse trabalho estão descritas no próximo tópico.

Figura 8 – Ilustração sobre o funcionamento do sistema de duplo híbrido. Em A, o fator de transcrição está intacto e a transcrição dos genes repórteres é ativada. Em B, os domínios independentes, BD e AD, estão fisicamente separados não havendo transcrição. Em C e D, as proteínas híbridas, BD-Isca e AD-Presa, expressas separadamente não ativam a transcrição. Em E, as proteínas são co-expressas e se interagem, elas mediam a aproximação dos domínios BD e AD ativando a transcrição dos genes repórteres.

Fonte: Elaborada pela autora.

2.3.1 Linhagem de levedura, plasmídeos e biblioteca de *C. reinhardtii*

A linhagem L40 da levedura *S. cerevisiae* (coleção do laboratório) foi utilizada neste ensaio, seu genótipo é *MATa his3Δ200 trp1-901 leu2-3,311 ade2 lys2-801am URA3::(lexAop)8-lacZ LYS2::(lexAop)4-HIS3*. Portanto, seu fenótipo é *Trp⁻, Leu⁻, Ade⁻* e a expressão dos genes repórteres *HIS3* e *LacZ* é controlada pelo promotor LexA.

As características relevantes dos plasmídeos utilizados no ensaio de duplo híbrido em leveduras estão listadas na Tabela 7. Esses plasmídeos foram adquiridos da empresa Dualsystems Biotech AG. O pLexA-N foi utilizado para a subclonagem da sequência codificadora da CrNIP7 nos sítios *EcoRI* (5') e *SalI* (3'). Portanto, a proteína CrNIP7 foi fusionada ao domínio de ligação ao DNA LexA (BD-isca). A combinação dos vetores pGAD-HA e pLexA-p53 foi utilizada como controle negativo de interação, e a combinação pLexA-p53 e pACT-largeT foi utilizada como controle positivo de interação forte.

A biblioteca de cDNA de *C. reinhardtii* utilizada nesse trabalho também foi adquirida da empresa Dualsystems Biotech AG e está clonada no vetor pGAD-HA. Assim, os insertos estão fusionados à sequência que codifica o domínio de ativação GAL4. Os dados de controle de qualidade da biblioteca indicam 10^7 de clones independentes contendo cDNAs com tamanhos de 0,6 a 10 kb e 95% de vetores com inserto clonado.

Figura 9 – Esquema do ensaio de duplo híbrido em levedura para busca de parceiros proteicos de interação.

Fonte: Elaborada pela autora.

Tabela 7 – Plasmídeos utilizados no ensaio de duplo híbrido.

Plasmídeo	Domínio	Características relevantes	Origem
pLexA-N	LexA BD	<i>ADH1, TRP1, 2μm, KanR, colE1 ori</i>	Adquiridos da Dualsystems Biotech AG
pGAD-HA	GAL4 AD	<i>ADH1, LEU2, 2μm, AmpR, pUC ori</i>	
pLexA-p53	LexA BD	<i>ADH1, TRP1, 2μm, KanR, colE1 ori</i>	
pACT-largeT	GAL4 AD	<i>ADH1, LEU2, 2μm, AmpR, pBR322 ori</i>	

Fonte: Elaborada pela autora.

2.3.2 Transformação de levedura em pequena escala

Nessa etapa do ensaio, a levedura L40 é transformada com plasmídeo BD-isca, ou seja, com a construção pLexA-N-CrNip7. Então, o protocolo rápido “*Quick and Easy Trafo*”⁴⁴ adaptado (descrito no item 2.2.3) foi utilizado para a transformação de L40 com algumas alterações. L40 foi cultivada em meio de cultura YPD a 30°C antes das transformações (Tabela 8). Após as transformações, as células foram plaqueadas em meio seletivo (Tabela 9) e as placas foram incubadas a 30°C por 4-5 dias até o aparecimento de colônias que continham o plasmídeo.

Tabela 8 – Resumo das transformações realizadas e meios de cultura para seleção plasmidial.

Linhagem	DNA plasmidial	Aplicação do transformante	Meio seletivo
L40	Água	Controle negativo de crescimento	SD – WL
	pLexA-N-CrNip7	BD-isca	SD – W
	pLexA-p53 + pGAD-HA	Controle negativo de interação (-)	SD – WL
	pLexA-p53 + pACT-largeT	Controle positivo de interação (+)	SD – WL

Para o controle negativo de transformação foi realizado o mesmo procedimento com água substituindo os DNAs plasmidiais.

Fonte: Elaborada pela autora.

Tabela 9 – Composição dos meios de cultura para leveduras.

Composição	SD-L	SD-W	SD-WL	SD-WH	SD-WLH
YNB	6,7 g/L	6,7 g/L	6,7 g/L	6,7 g/L	6,7 g/L
Glicose	20 g/L	20 g/L	20 g/L	20 g/L	20 g/L
Adenina (A)	0,02 g/L	0,02 g/L	0,02 g/L	0,02 g/L	0,02 g/L
Histidina (H)	0,02 g/L	0,02 g/L	0,02 g/L	-	-
Leucina (L)	-	0,1 g/L	0,1 g/L	0,1 g/L	0,1 g/L
Triptofano (W)	0,02 g/L	-	-	-	-

Adicionar 2% de ágar para meio sólido. SD = meio de cultura mínimo. YNB = *Difco yeast nitrogen base w/o amino acids* (Sigma).

Fonte: Elaborada pela autora.

2.3.3 Ensaio da β -galactosidade em papel de filtro

Primeiramente, o ensaio da atividade da enzima β -galactosidade em papel de filtro tem a finalidade de verificar se o gene repórter *LacZ* é ativado pela construção “BD-isca”. Em caso afirmativo, o sistema de duplo híbrido não é apropriado para busca de parceiros de interação. Em caso negativo, o ensaio de β -galactosidade é usado para avaliar a interação entre as proteínas “BD-isca” e “AD-presa”, pois se houver interação proteica, o gene repórter *LacZ* será transcrito.

Para realizar este ensaio, as colônias transformantes foram repicadas em meio sólido seletivo de acordo com a marca de seleção adequada e incubadas a 30°C por 3 dias até crescimento de colônias grandes. Um papel de filtro recortado em círculo foi friccionado sobre a placa com auxílio de uma alça de Drigalski para a transferência das células para o papel. Em seguida, o papel de filtro foi imerso em nitrogênio líquido para lisar as células (~10 minutos). Paralelamente, outro papel de filtro foi recortado e embebido com 5 mL de tampão Z (16,1 g/L de $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$; 5,5 g/L $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$; 0,75 g/L de KCl e 0,246 g/L de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, pH 7,0), 13,5 μL de β -mercaptoetanol e 83,5 μL de X-Gal a 20 mg/mL, em uma placa de Petri. Após a lise, o papel de filtro com as células voltadas para cima foi depositado em cima do papel umedecido com tampão Z, estes foram incubados a 37°C até o aparecimento de coloração azul. Em seguida, o papel foi colocado em uma capela para secar, devido ao β -mercaptoetanol. Os transformantes de pLexA-p53 + pGAD-HA foram utilizados como controle negativo de interação e os transformantes pLexA-p53 + pACT-largeT como controle positivo de interação.

2.3.4 Transformação de levedura em grande escala para varredura da biblioteca de cDNA

Uma vez que a proteína CrNIP7 não ativou o sistema de duplo híbrido, foi possível prosseguir com a varredura da biblioteca de cDNA de *C. reinhardtii* para a busca de parceiros de interação. A fim de diminuir a quantidade de falsos positivos, um teste preliminar de crescimento do transformante da levedura L40 [pLexA-N-CrNip7] em placas de meio sólido SD-WH (Tabela 9) contendo concentrações gradativas (0, 1, 2, 3, 4, 5, 10, 15 e 20 mM) de 3-AT (3-amino-1,2,4-triazol) foi realizado. O composto 3-AT modula a sensibilidade do gene repórter *HIS3* atuando com um inibidor competitivo da proteína his3p, portanto o aumento da concentração de 3-AT nos meios seletivos diminuem a sensibilidade de *HIS3*. A concentração de 3-AT determinada para a varredura da biblioteca foi a menor concentração de 3-AT que o transformante L40 [pLexA-N-CrNip7] conseguiu crescer (colônias menores que 1 mm) após uma semana de incubação a 30°C em meio SD-WH. Nesse caso a concentração determinada foi de 2 mM de 3-AT.

Três colônias de células da levedura L40 [pLexA-N-CrNip7] foram inoculadas em 75 mL de meio SD-W (Tabela 9) e a cultura foi crescida a 30°C sob agitação de 200 rpm por 24 horas até D.O._{600nm} maior ou igual a 1,5. A cultura foi centrifugada a 2500 x *g*, 21°C por 5 minutos, e as células foram ressuspensas em 500 mL de meio YPD. A cultura foi incubada a 30°C sob agitação de 200 rpm por 4 horas até atingir D.O._{600nm} entre 0,4 e 0,6. Após o crescimento, as células foram novamente centrifugadas a 2500 x *g*, 21°C por 5 minutos, o meio foi removido e as células foram lavadas com 200 mL de tampão TE 1X (Tabela 10). Novamente as células foram coletadas por centrifugação, e ressuspensas em 4 mL da solução TE/LiAc 1X (Tabela 10) (duas alíquotas de 200 µL foram separadas para os controles negativos) e à suspensão celular adicionaram-se 10 mg de DNA de esperma de salmão (previamente desnaturado por 15 minutos a 95°C e colocado imediatamente no gelo), 40 µg da biblioteca de cDNA de *C. reinhardtii* (Dualsystems Biotech) e 30 mL de PEG/LiAc (Tabela 10). A mistura foi vigorosamente homogeneizada (vórtex), em seguida incubada a 30 °C, sob agitação de 200 rpm por 30 minutos, depois submetida a choque térmico de 42°C por 15 minutos seguido de 3 minutos no gelo. A suspensão celular foi centrifugada por 2500 x *g* por 5 minutos e o precipitado de células foi suspenso em 5 mL de TE 1X. Alíquotas de 250 µL foram plaqueadas em 20 placas de Petri grandes, contendo 50 mL de meio SD-WLH sólido (Tabela 9) com 2 mM de 3-AT, e incubadas

por 7 dias a 30°C. A eficiência de transformação da biblioteca foi verificada pelo plaqueamento das diluições da transformação, 1:10, 1:100 e 1:1000, em meio SD-WL (Tabela 9).

As transformações de controle negativo de crescimento foram realizadas da seguinte forma: aos 200 µL de suspensão celular, previamente separados, foram adicionados 50 µg de DNA de espermatozoides de salmão, 600 µL de PEG/LiAc e 10 µL de água ultrapura ou um volume contendo 200 ng de pGAD-HA vazio. Os passos seguintes foram realizados conforme o protocolo descrito acima. Na última etapa, as células foram suspensas em 250 µL de TE 1X e plaqueadas em meio SD-WH sólido (água) e SD-WL sólido (pGAD-HA), as placas foram incubadas a 30°C por 7 dias.

O aparecimento de colônias foi monitorado durante sete dias e as colônias com tamanho aproximado de 2 mm ou mais foram repicadas em meio SD-WLH (Tabela 9) com 2 mM de 3-AT para confirmar a primeira triagem de interação, ou seja, a presença dos dois plasmídeos e a expressão do gene repórter *HIS3*. A placa foi incubada a 30°C por 3-4 dias, e os transformantes que não cresceram nesse tempo (falsos positivos) foram descartadas. As colônias positivas na primeira triagem foram submetidas ao teste da β-galactosidase (item 2.3.3), segunda triagem de falsos positivos, para verificar a expressão do gene repórter *LacZ*.

Os transformantes positivos nas duas triagens seguiram nas próximas etapas do ensaio, passando pela extração plasmidial de leveduras, transformação plasmidial em bactérias, extração plasmidial de bactérias, sequenciamento e, por fim, confirmação da interação por co-transformação de levedura, teste da β-galactosidase e teste de força de interação.

Tabela 10 – Soluções utilizadas na transformação da biblioteca de cDNA.

Solução*	Composição
TE 10X	100 mM Tris-HCl, 10 mM EDTA, pH 7,5
TE 1X	1 mL de 10X TE, 9 mL água ultrapura estéril
PEG 3350 50% (p/v)	50 g de PEG 3350 em água ultrapura para volume final de 100 mL
LiAc 10X (Acetato de lítio)	1M LiAc, pH 7,5 (ajuste do pH com ácido acético)
PEG/LiAc	PEG 3350 40%, TE 1X, LiAc 1X
TE/LiAc	1X TE, 1X LiAc

*Todas as soluções devem ser estéreis.

Fonte: Elaborada pela autora.

2.3.5 Extração plasmidial de levedura

As colônias transformantes positivas no teste da β -galactosidase contém os plasmídeos pLexA-N-*CrNip7* e pGAD-HA-presa que conferem auxotrofia às células de triptofano (W) e leucina (L), respectivamente. Portanto, para a extração do plasmídeo pGAD-HA-presa de uma cultura de leveduras L40, primeiro precisamos promover a perda do plasmídeo pLexA-N-*CrNip7* por crescimento em meio com ausência de pressão seletiva para triptofano. Para isso, uma colônia de cada transformante foi crescida em 5 mL de meio SD-L líquido (Tabela 9) a 30°C sob agitação de 200 rpm por 24 horas. A extração DNA plasmidial de leveduras foi realizada com adaptações do protocolo descrito em Ling.⁴⁸

Todo o volume da cultura foi coletado em um microtubo por centrifugação a 16000 x *g* por 1 minuto, o meio de cultura foi removido e as células foram resuspensas em 60 μ L de 50 mM TRIS-HCl, pH 8,0. Um volume de 20 μ L da enzima liticase (5 mg/mL) foi adicionado à suspensão celular seguida de agitação vigorosa (vórtex) e incubação a 37°C (em banho-maria) por uma hora (com agitação a cada 10 minutos). Após a incubação com liticase, 20 μ L de SDS 20% foram adicionados ao microtubo seguida de agitação vigorosa (vórtex) por um minuto. Os microtubos foram expostos a 20 segundos em nitrogênio líquido intercalados com 20 segundos em banho-maria a 42°C por 10 vezes. Logo após, 100 μ L de água ultrapura e 200 μ L de fenol:clorofórmio:álcool isoamílico (25:24:1 v/v) foram adicionados aos microtubos e estes foram agitados vigorosamente por um minuto e centrifugados a 16000 x *g* por 1 minutos. A fase aquosa foi transferida para um novo microtubo e 10 μ L de LiCl 8 M e 300 μ L de etanol absoluto gelado foram adicionados. As misturas foram homogeneizadas, incubadas a -20°C por 20 minutos e centrifugadas a 16000 x *g* para a precipitação do DNA. O precipitado foi lavado com 500 μ L de etanol 70% gelado, seguido de centrifugação por 10 minutos a 16.000 x *g* para a retirada do excesso de sal. O sobrenadante foi removido e microtubos foram deixados com a tampa aberta à temperatura ambiente para o precipitado de DNA secar mais rápido. O precipitado de DNA foi suspenso em 30 μ L de água ultrapura.

2.3.6 Transformação de bactérias, extração plasmidial de bactérias, sequenciamento e análise das sequências “AD-Presa”

Células de *E. coli* DH5 α competentes foram transformadas (item 2.1.2.1) com um volume de 15 μ L dos DNAs plasmidiais (construção AD-Presa) extraídos de levedura (como descrito na seção anterior) e plaqueadas em meio LB ágar contendo ampicilina (100 μ g/mL). Uma colônia de cada transformante foi crescida em 6 mL de meio LB líquido, contendo o mesmo antibiótico descrito acima, a 37°C por 16 horas sob agitação de 200 rpm. O DNA plasmidial foi extraído dessas células com auxílio do kit *Wizard®Plus SV Minipreps DNA Purification System* (Promega) (seção 2.1.2.2) e suas sequências foram determinadas pelo sequenciador automático *3130 Genetic Analyzer* (Applied Biosystems®). Os oligonucleotídeos utilizados no sequenciamento foram específicos para o vetor pGAD-HA, no qual a biblioteca de cDNA foi clonada e suas sequências são: primer F (*foward*) (5'-CTATTCGATGATGAAGATACCCACC-3') e primer R (*reverse*) (5'-TCGTAGATACTGAAAAACCCGCAAGTTCAC-3').

A fim de identificar as interações encontradas, as sequências de nucleotídeos resultantes do sequenciamento foram então comparadas às sequências do banco de dados geral nr (*Non-redundant protein sequences*) dos programas BLASTX e BLASP⁴⁹ do *National Center for Biotechnology Information* (NCBI) no endereço eletrônico <http://www.ncbi.nlm>. Apenas as sequências que se apresentaram em fase aberta de leitura +1, (FRAME +1) com a sequência codificante do domínio de ativação GAL4 foram consideradas. A fase de leitura foi analisada identificando a sequência do adaptador DSM4 (5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3') utilizado na produção da biblioteca de cDNA de *C. reinhardtii*. A sequência proteica foi traduzida a partir da sequência de nucleotídeos pela ferramenta *Translate Tool* disponível no portal de bioinformática ExPASy⁵⁰ (<http://www.expasy.org>). Os parceiros proteicos positivos no ensaio de duplo foram então submetidos a análises de bioinformática para busca de domínios conservados, predição de estrutura secundária e parâmetros teóricos (item 2.4).

2.3.7 Confirmação das interações das proteínas encontradas e teste de força de interação

Após a identificação dos candidatos positivos de interação, a levedura L40 foi co-transformada (seção 2.3.2) com os DNAs plasmidiais pLexA-N-*CrNip7* e pGAD-HA-presca (vetor contendo as sequências dos parceiros encontrados), plaqueada em meio de cultura SD-WL e incubada a 30°C por 3-4 dias. Três colônias de cada transformação foram repicadas em meio SD-WLH e incubadas a 30°C por 3 dias e depois foram submetidas ao teste da β -galactosidase (ver seção 2.3.3) para verificar a expressão do gene repórter *LacZ*. Para o teste de força de interação, 3 colônias de cada transformante foram crescidas em 5 mL de meio SD-WL líquido a 30°C sob agitação de 200 rpm por 24 horas. A absorbância em 600 nm ($D.O_{600nm}$) das culturas celulares foi medida, e 1 mL de cada cultura celular foi centrifugado por 30 segundos a 16000 x *g*. O precipitado celular foi ressuspenso em glicerol 20% estéril, normalizando as concentrações celulares. Todo o volume da suspensão de células foi transferido para o primeiro poço de uma microplaca de 96 poços, e nos 5 poços seguintes foram adicionados 90 μ L de glicerol 20% estéril. Cinco diluições seriadas da ordem de 10x foram feitas utilizando um pipetador multicanal. Um volume de 4 μ L de cada diluição foi aplicado em placas de meios-WLH contendo concentrações gradativas de 3-AT (0, 1, 2, 3, 5, 7 e 10 mM). O mesmo procedimento foi realizado com os controles de interação (Tabela 8).

O ensaio de duplo híbrido teve a colaboração da Dra. Joci Neuby Alves Macedo.

2.4 Análises de bioinformática

Todas as proteínas estudadas nesse trabalho passaram por análises de bioinformática. O programa Protparam⁵¹ disponível no servidor da internet ExPASy⁵⁰ (<http://www.expasy.org>) foi utilizado para a obtenção de parâmetros teóricos, como massa molar, *pI* e coeficiente de extinção molar. A análise de domínios conservados foi realizada pelo programa Pfam⁵² encontrado no endereço eletrônico <http://pfam.xfam.org/>. As predições de estrutura secundária foram realizadas pelos programas de GOR4⁵³ e PSIPRED⁵⁴⁻⁵⁵ cujos links podem ser acessados pelo servidor ExPASy. Os alinhamentos múltiplos foram realizados pelo programa ClustalOmega⁵⁶ pelo endereço eletrônico <http://www.clustal.org/omega/>. A predição de regiões intrinsecamente desordenadas (IDRs) foi realizada pelos programas Globplot⁵⁷ e

DisEMBL⁵⁸ cujos links podem ser encontrados no servidor da internet DisProt (do inglês, *Disordered Protein Database*) em www.disprot.org. A predição de regiões de ligação em proteínas intrinsicamente desordenadas foi realizada com auxílio do programa ANCHOR,⁵⁹⁻⁶⁰ disponível no endereço eletrônico <http://anchor.enzim.hu/>.

2.4.1 Anotação das sequências codificadoras das proteínas CrSBDS e G-patch

As regiões codificantes das proteínas CrSBDS (XP_001689481.1) e G-patch (XP_001692447.1) não estão anotadas corretamente por tradução conceitual no banco de dados aberto do *National Center for Biotechnology Information* (NCBI).

A análise das extremidades N- e C-terminal incompletas de CrSBDS foi realizada pelo alinhamento (BLAST⁴⁹) da CDS parcial anotada de CrSBDS contra o banco de dados SRA (do inglês, *Sequence Read Archive*) para *C. reinhardtii*. O banco de dados SRA é uma plataforma de sequenciamento de alto rendimento de transcritos de RNA disponível no servidor do NCBI. Os oligonucleotídeos iniciadores para amplificação da CDS correta de CrSBDS foram desenhados a partir da sequência completa.

Já a sequência de nucleotídeos de G-patch anotada por tradução conceitual não corresponde à sequência de nucleotídeos amplificada por PCR a partir do cDNA de *C. reinhardtii*. Portanto, a sequência correta foi obtida através da comparação entre o sequenciamento obtido do fragmento amplificado e o *locus* genômico que codifica para o gene *G-patch*. A sequência completa de G-patch também foi verificada por alinhamento no banco de dados SRA.

As duas análises foram realizadas com auxílio e supervisão do Prof. Dr. Ricardo De Marco, docente do Grupo de Biofísica Molecular “Sérgio Mascarenhas”, do Instituto de Física de São Carlos - USP, especialista em bioinformática.

2.5 Verificação da interação entre CrNIP7 e CrSBDS por duplo híbrido dirigido

A interação entre CrNIP7 e CrSBDS não foi encontrada pela varredura da biblioteca de cDNA de *C. reinhardtii* pelo ensaio de duplo híbrido, porém a interação entre Nip7p/Sdo1 e HsNIP7/HsSBDS, homólogas à CrNIP7 e CrSBDS, respectivamente, é descrita na literatura.^{16,21}

Então, a sequência codificante de CrSBDS foi amplificada a partir do cDNA de *C. reinhardtii*, segundo as técnicas de biologia molecular descritas nos métodos de clonagem, para ser subclonada no vetor de duplo híbrido pGAD-HA nos sítios *Sma*I (5') e *Xho*I (3'). Essa construção resultou na expressão da proteína CrSBDS fusionada ao domínio de ativação GAL4 ("AD-presa"). A linhagem de *S. cerevisiae* L40 foi co-transformada (item 2.3.2) com as construções pLexA-N-CrNip7 e pGAD-HA-CrSBDS e a interação foi avaliada pela expressão dos genes repórteres através do teste da β -galactosidade (item 2.3.3) e pela força de interação pelo crescimento em concentrações gradativas de 3-AT (item 2.3.7). Após confirmação *in vivo*, a interação foi submetida à validação *in vitro*, por *pull down*, como descrito no item 2.6.

Toda a parte experimental do estudo de CrSBDS desde a amplificação da CDS, duplo híbrido dirigido, expressão heteróloga e *pull down*, foi realizada em conjunto com a aluna de iniciação científica Angélica Luana Carrillo Barra, contemplada com uma bolsa PIBIC/CNPq.

2.6 Validação das interações proteicas encontradas nos ensaios de duplo híbrido

Depois de encontrar as interações proteicas pelo sistema de duplo híbrido, torna-se necessário validar a interação por outra técnica. Nesse trabalho optou-se por tentar a co-expressão das proteínas em vetor pETDuet-1, além da mistura das proteínas expressas separadamente seguidos de experimentos de *pull down* (co-precipitação das proteínas). Para isso, foi preciso amplificar e clonar as CDS *predicted* e *G-patch*. A CDS de Predicted foi amplificada a partir do plasmídeo pGAD-*predicted* isolado no duplo híbrido, e a CDS de G-patch foi amplificada a partir do cDNA de *C. reinhardtii*. As CDS de CrNIP7 e CrSBDS tiveram que ser amplificadas a partir de construções anteriores para trocar os sítios de restrição. Todas as amplificações e clonagens foram realizadas segundo as técnicas descritas no item 2.1 e estão listadas na Tabela 11. Testes de expressão heteróloga foram realizados com todas as construções, e as expressões positivas foram submetidas à testes de purificação e *pull down*.

Tabela 11 – Construções em plasmídeos de expressão realizadas neste trabalho.

Construção	Proteínas expressas	Clonagem direcional		Inserto (pb)	Vetor-inserto	Resistência
		5'	3'			
pETDuet-1- <i>CrNip7</i>	6His-CrNIP7	<i>EcoRI</i>	<i>PstI</i>	540	~5960	<i>bla</i>
pETDuet-1- <i>CrNip7-G-patch</i>	6His-CrNIP7	<i>EcoRI</i>	<i>PstI</i>	540	~7772	
	G-patch	<i>NdeI</i>	<i>KpnI</i>	1812		
pETDuet-1- <i>CrNip7-predicted</i>	6His-CrNIP7	<i>EcoRI</i>	<i>PstI</i>	540	~6440	
	predicted	<i>NdeI</i>	<i>XhoI</i>	480		
pETDuet-1- <i>CrNip7-CrSBDS</i>	6His-CrNIP7	<i>EcoRI</i>	<i>PstI</i>	540	~7100	
	CrSBDS	<i>NdeI</i>	<i>XhoI</i>	1140		
pETSUMO- <i>CrNip7</i>	SUMO-CrNIP7	<i>BamHI</i>	<i>Sall</i>	540	~6140	<i>KanR</i>
pETSUMO- <i>G-patch</i>	SUMO-G-patch	<i>BamHI</i>	<i>Sall</i>	1812	~7412	
pETSUMO- <i>predicted</i>	SUMO-predicted	<i>BamHI</i>	<i>XhoI</i>	480	~6080	
pET28a- <i>SBDS</i>	6His-CrSBDS	<i>NdeI</i>	<i>XhoI</i>	1140	~6509	
pGEX-4T-1- <i>CrNip7</i>	GST-CrNIP7	<i>BamHI</i>	<i>Sall</i>	540	~5509	<i>AmpR</i>
pGEX-4T-1- <i>predicted</i>	GST-predicted	<i>BamHI</i>	<i>XhoI</i>	480	~5449	

Tamanho dos plasmídeos vazios: pETDuet-1 (5420 pb), pETSUMO (~5600); pET28a (5369pb) e pGEX-4T-1 (4969 pb).

Fonte: Elaborada pela autora.

2.6.1 Expressão heteróloga de proteínas e purificação por cromatografia de afinidade a níquel

Células competentes de *E. coli* Rosetta (DE3) foram transformadas (item 2.1.2.1) com DNA plasmidial de uma construção em vetor de expressão e plaqueadas em LB ágar contendo cloranfenicol (34 µg/mL) e o antibiótico apropriado para cada vetor e incubadas a 37°C por 16 horas. Um pré-inóculo foi cultivado a partir de uma colônia transformante, em meio de cultura LB líquido contendo os antibióticos descritos acima, a 37°C sob agitação de 200 rpm por 16 horas. A cultura foi diluída na proporção 1:50 (inóculo:meio de cultura) em 500 mL de meio LB com antibióticos apropriados em um frasco de 2 L para aeração 1:5 (meio:ar) e incubada a 37°C, sob agitação de 200 rpm até atingir absorbância a 600 nm ($D.O_{600nm}$) entre 0,6 e 0,8. Então, a expressão da proteína foi induzida com 0,5 mM de IPTG (isopropil-β-D-tiogalactopiranosídeo), a 18°C por 16 horas. A cultura foi dividida em 5 frascos de 100 mL e as células foram coletadas por centrifugação a 5600 x *g*, 10°C, por 15 minutos. As alíquotas de precipitado celular foram armazenadas no freezer a -80°C até a realização da purificação (em até um mês). Para a purificação, o precipitado de células (de 100 mL de cultura) foi suspenso em 5 mL de tampão 50 mM TRIS-HCl; 0,3 M NaCl, e pH 8 (tampão A). Em seguida, as células em suspensão foram lisadas por 8 ciclos de sonicação de 30 segundos de pulso ultrassom e 30 segundos em banho de gelo. As células lisadas foram centrifugadas a 18500 x *g*, 10°C por 25 minutos para a separação da fração solúvel e restos celulares. A fração solúvel foi incubada com 1 mL de resina Ni-NTA *super flow* (Qiagen), previamente equilibrada com 10 volumes de tampão A, na geladeira (4°C) por 30 minutos. Todos os passos de purificação foram realizados a 4°C. Após a incubação, a fração não ligada foi coletada e a resina passou por 4 lavagens, sendo a primeira lavagem com 10 volumes de tampão A, a segunda lavagem com 10 volumes de tampão A acrescido de 0,1% de TRITON X-100 (ou com 0,2% de Tween-20), a terceira lavagem com 10 volumes de tampão A acrescido de 10 mM de imidazol, e por fim a quarta lavagem com 10 volumes de tampão A. A eluição da proteína com 6xHis-tag foi realizada em duas etapas com 2 volumes de tampão A acrescido de 500 mM de imidazol.

Alíquotas de todas as etapas, antes da indução, após a indução, da fração solúvel, da fração insolúvel, da fração não ligada à coluna, das etapas de lavagens e da eluição foram coletadas para análise em gel SDS-PAGE. Quando foi necessário retirar o imidazol da amostra

eluída, a troca de tampão foi realizada na coluna Hi-Trap *Desalting* (GE Healthcare) acoplada ao sistema *Äkta purifier* (GE Healthcare).

Para alguns experimentos de interação, o protocolo geral acima foi modificado quanto à concentração de sal (NaCl) e pH do tampão, sendo que estas modificações serão apresentadas nos resultados.

2.6.2 Teste de clivagem com SUMO protease na coluna

Nas construções em vetor pETSUMO, a proteína de interesse é expressa fusionada com a proteína SUMO e por isso, a SUMO-fusão resultante foi submetida a um teste de clivagem na resina de níquel para a obtenção da proteína isolada. A expressão e purificação da SUMO-fusão em resina de níquel foi realizada conforme o protocolo padrão descrito acima. Porém, após a quarta lavagem com 10 volumes de tampão A, uma alíquota de 250 µL de SUMO protease (0,5 mg/mL) em 2 volumes de tampão A foram adicionados à resina e a mistura foi incubada sob agitação na geladeira por uma noite. Em seguida, o eluato foi coletado e a resina foi lavada com 2 volumes por 2 vezes e, por fim, o que ficou retido na coluna foi eluído com 500 mM de imidazol. Alíquotas de todas as etapas foram coletadas para análise em gel SDS-PAGE.

2.6.3 Purificação da GST-CrNIP7

A proteína de fusão GST-CrNIP7 foi expressa da mesma maneira descrita no item 2.6.1, pela indução da construção pGEX-4T-1-*CrNip7* em *E. coli* Rosetta (DE3) e purificada por afinidade à resina Glutathione Sepharose 4 Fast Flow (GE). Este ensaio foi apenas um teste preliminar de expressão e purificação. Então, 100 mL de uma cultura induzida foi submetida a lise celular e o extrato solúvel foi incubado em 0,5 mL da resina equilibrada com 5 volumes de coluna de tampão PBS (140 mM NaCl, 2,7 mM KCl, 10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, pH 7,3). O extrato solúvel foi misturado à coluna e incubado por aproximadamente 30 minutos sob agitação a 4°C. A coluna foi submetida a 3 lavagens antes da eluição, a primeira com 10 volumes de PBS, a segunda com 5 volumes de PBS com 0,1% de TRITON X-100 e a terceira com 10 volumes de PBS. Em seguida, a proteína recombinante foi eluída por 2 vezes com 2 volumes de

tampão 50 mM TRIS-HCl, 10 mM de glutatona reduzida, pH8. As alíquotas de cada etapa foram recolhidas para análise SDS-PAGE.

2.6.4 Co-expressão e Co-purificação

A co-expressão e co-purificação das proteínas 6xHis-CrNIP7 e CrSBDS expressas pela construção pETDUET-1-CrNip7-CrSBDS em *E. coli* Rosetta (DE3) foi realizada da mesma maneira descrita no item 2.6.1. Porém, nesse experimento foram utilizados 200 mL de cultura celular e a composição do tampão A, nesse caso, foi 20 mM TRIS-HCl, 150 mM NaCl e pH 7.

2.6.5 Pull down

Neste ensaio de *pull down*, duas proteínas cuja interação estava sob investigação foram expressas separadamente, purificadas, misturadas e depois submetidas a uma nova purificação. A ideia é que se uma das proteínas tiver um *tag* (6xHis-tag, neste caso), e a outra proteína não tiver a marcação (presa), se houver interação, a proteína 6xHis-isca carrega a presa durante a eluição da resina de afinidade. Esse ensaio foi realizado para verificar a interação entre as proteínas 6xHis-CrSBDS e CrNIP7 e entre as proteínas 6xHis-CrNIP7 e Predicted.

Preparo da amostra com 6xHis-tag

A expressão das proteínas foi realizada em células de *E. coli* Rosetta (DE3), pela transformação com o vetor pET28a-CrSBDS ou pelo vetor pETDuet-1-CrNip7 (Tabela 11). As proteínas com 6xHis-tag foram expressas e purificadas como descrito em 2.6.1 com Tampão A (20 mM TRIS, 500 mM NaCl pH7) e a troca de tampão foi realizada pela coluna Hi-Trap *Desalting* (GE Healthcare) acoplada ao sistema *Äkta purifier* (GE Healthcare) utilizando o Tampão B (20 mM TRIS, 150 mM NaCl pH 7).

Preparo da amostra sem 6xHis-tag

Já as proteínas Predicted e CrNIP7 foram expressas em pETSUMO (Tabela 11) em células *E. coli* Rosetta (DE3) (item 2.6.1), o que resultou em proteínas de fusão. As proteínas de fusão foram purificadas e clivadas com SUMO protease como descrito no item 2.6.2, utilizando o Tampão A de composição 20 mM TRIS, 500 mM NaCl, pH7. O eluato (contendo a proteína de interesse) foi submetido à troca de tampão na coluna Hi-Trap *Desalting* (GE Healthcare), acoplada ao sistema *Äkta purifier* (GE Healthcare) utilizando o Tampão B (20 mM TRIS, 150 mM NaCl pH7).

Ensaio de interação

Após a obtenção das proteínas purificadas no tampão B, as potenciais parceiras de interação foram misturadas em um microtubo e incubadas sob agitação a 4°C por uma noite. Em seguida, a mistura foi centrifugada para a retirada de precipitados, e o sobrenadante foi transferido para outro microtubo contendo 50 µL de resinal Ni-NTA *super flow* (Qiagen). O microtubo foi novamente incubado a 4°C por 20 minutos e depois centrifugado, e o sobrenadante foi removido (este sobrenadante contém as proteínas não ligadas à resina). Em seguida, a resina passou por 3 lavagens com 1 mL de Tampão B por 5 minutos à velocidade máxima da centrífuga de bancada (Eppendorf). As proteínas que ficaram ligadas na resina foram eluídas com 60 µL de tampão B acrescido de 500 mM de imidazol. Alíquotas das etapas foram coletadas para análise por gel SDS-PAGE.

2.7 Espectrometria de massas

O protocolo utilizado para extração proteica a partir do gel de SDS-PAGE 16% de tricina é uma adaptação do protocolo descrito em Shevchenko.⁶¹

A banda referente à proteína de interesse foi recortada do gel de poliacrilamida e colocada em um microtubo de 1,5 mL, os passos a seguir foram realizados à temperatura ambiente. Ao microtubo, adicionou-se 0,5 mL da solução metanol 50%/ácido acético 2,5% (solução 1) e este foi incubado por 2 horas para a descoloração e remoção do SDS. Essa etapa foi repetida com tempo de incubação de 1 hora. Após a remoção da solução 1, a amostra foi

desidratada com 200 µL de acetonitrila (100%) por 5 minutos, a acetonitrila foi removida e o mesmo passo foi repetido. Um evaporador foi utilizado para evaporar o que restou da acetonitrila. Para a redução da amostra, 30 µL da solução de DTT (ditiotreitól) 10 mM/bicarbonato de amônio 100mM foi adicionado ao microtubo e este foi incubado por 30 minutos. A solução de DTT foi removida e a alquilação foi feita com 30 µL da solução Iodoacetamida (IAA) 50 mM/bicarbonato de amônio 100 mM por 30 minutos. A solução de IAA foi removida e o gel foi lavado com 100 µL de bicarbonato de amônio 100 mM por 10 minutos. A solução de bicarbonato de amônio foi removida e o gel foi desidratado com 200 µL de acetonitrila 100% por 5 minutos e após a remoção da acetonitrila foi reidratado com 200 µL de bicarbonato de sódio 100 mM por 10 minutos. A amostra foi novamente desidratada com 200 µL de acetonitrila em dois passos de incubação de 5 minutos. O restante da acetonitrila foi evaporado por aproximadamente 2 minutos em um evaporador. A tripsinização da proteína foi realizada adicionando-se 30 µL da solução gelada de tripsina (20 ng/µL) em bicarbonato de amônio 50 mM ao microtubo e este foi incubado por 30 minutos em banho de gelo. Depois de remover o excesso da solução de tripsina, adicionou-se 20 µL de bicarbonato de amônio para cobrir o gel e este foi incubado durante 16 horas a 37°C. No dia seguinte, a extração foi realizada adicionando-se 10 µL da solução de ácido fórmico 5% e incubando por 5 minutos à temperatura ambiente. O sobrenadante foi coletado em um novo microtubo, e mais 12 µL da solução ácido fórmico 5%/acetonitrila 50% foi adicionada ao gel, este foi incubado novamente por 10 minutos. O sobrenadante foi recuperado e adicionado no mesmo microtubo da primeira extração. A amostra foi entregue para a técnica de laboratório Andressa Alves Pinto, que fez a análise no espectrômetro de massas micrOTOF-QII (Bruker Daltonics). Os peptídeos identificados foram analisados pela ferramenta BLASTP⁴⁹ do NCBI (<http://www.ncbi.nlm.nih.gov/>).

2.8 Análise do perfil de expressão de CrNIP7, G-patch, Predicted e CrSBDS

Recentemente, Zones e colaboradores⁶² disponibilizaram dados do transcriptoma de uma cultura de *C. reinhardtii* com ciclo de vida sincronizado. A fim de relacionar e interpretar o perfil de expressão das proteínas CrNIP7, G-patch, Predicted e CrSBDS durante o ciclo celular da alga, os dados de expressão gênica correspondentes a estas proteínas foram localizados pelo nome do gene (Tabela 12) no conjunto de dados suplementares (número 1) do trabalho acima referido e foram plotados em um gráfico para comparação.

Tabela 12 – Loci gênico correspondente às proteínas de interesse.

Proteína	Nome do <i>locus</i> gênico
CrNIP7	Cre12.g528900
G-pacth	Cre07.g347400
Predicted	Cre01.g037250
CrSBDS	Cre01.g017500

Fonte: Elaborada pela autora.

RESULTADOS E DISCUSSÃO

3 RESULTADOS E DISCUSSÃO

Para maior clareza, a seção de resultados foi dividida em quatro partes referentes às proteínas estudadas nesse trabalho. Uma discussão geral, relacionando todos os resultados, é apresentada ao final da seção.

3.1 CrNIP7

Ao realizarmos uma busca por proteínas homólogas a Nip7p de levedura (GenBank: EIW07327.1) no banco de proteínas nr (do inglês, *Non-redundant protein sequences*) de *C. reinhardtii* pelo algoritmo BLASTP⁴⁹, encontramos apenas um *hit* correspondente à uma proteína hipotética anotada por tradução conceitual do genoma de código de acesso XP_001696971 (Apêndice A – Figura 40). Esta sequência proteica, de 179 aminoácidos, foi nomeada de CrNIP7 (Apêndice A – Figura 41).

3.1.1 Amplificação e clonagens da CDS *CrNip7*

Oligonucleotídeos específicos foram desenhados a partir da CDS de 540 pb (NCBI: XM_001696919.1) que codifica a proteína CrNIP7. Desse modo, a CDS *CrNip7* foi amplificada e isolada com sucesso a partir do cDNA total de *C. reinhardtii* utilizando o par de oligonucleotídeos iniciadores, *CrNip7-BamHI-F* e *CrNip7-Sall-R* (Tabela 2), para clonagem no vetor pGEMT-*Easy* e posterior subclonagem no vetor pETSUMO nos sítios *BamHI* e *Sall*. Após a confirmação da clonagem por sequenciamento automático, a construção pGEMT-*CrNip7*-BS serviu de molde para novas amplificações da CDS com outros pares de oligonucleotídeos (Tabela 2 e Tabela 13) que foram utilizados para clonagem em outros vetores (pETDuet-1, pGEX-4T-1, pJN40, YCplac33 e pLexA-N). As características importantes de cada plasmídeo estão descritas na seção dos Materiais e Métodos. As Figuras 10, 11, 12 e 13 abaixo mostram as bandas de amplificações da CDS *CrNip7* e as análises de restrição confirmando as clonagens.

Tabela 13 – Descrição dos oligonucleotídeos utilizados para amplificação de *CrNip7* e construções em vetores de clonagem e de expressão

Construção em vetor de clonagem	Par de oligonucleotídeos utilizados na PCR	Subclonagem	Clonagem direcional nos sítios:
pGEMT-Easy- <i>CrNip7</i> -BS	<i>CrNip7</i> - <i>Bam</i> HI-F <i>CrNip7</i> - <i>Sall</i> -R	pETSUMO- <i>CrNip7</i> -BS pGEX-4T-1- <i>CrNip7</i> -BS	<i>Bam</i> HI e <i>Sall</i>
pGEMT-Easy- <i>CrNip7</i> -ES	<i>CrNip7</i> - <i>Eco</i> RI-F <i>CrNip7</i> - <i>Sall</i> -R	pLexA-N- <i>CrNip7</i> -ES pJN40- <i>CrNip7</i> -ES	<i>Eco</i> RI e <i>Sall</i>
pGEMT-Easy- <i>CrNip7</i> -BX	<i>CrNip7</i> - <i>Bam</i> HI-F <i>CrNip7</i> - <i>Xba</i> I-R	YCplac33- <i>CrNip7</i> -BX	<i>Bam</i> HI e <i>Xba</i> I
pUC19- <i>CrNip7</i> -EGP	<i>CrNip7</i> - <i>Eco</i> RI-G-F <i>CrNip7</i> - <i>Pst</i> I-R	pETDuet-1- <i>CrNip7</i> -EGP	<i>Eco</i> RI e <i>Pst</i> I

Ver sequência dos oligonucleotídeos na Tabela 2.

Fonte: Elaborada pela autora.

Figura 10 – Análise da amplificação por PCR de *CrNip7* e das clonagens visando a expressão heteróloga: (A) Produto de PCR (540pb) amplificado a partir do cDNA de *C. reinhardtii* (1) Tm= 62°C; (2) Tm= 64°C. (B) Clonagem de *CrNip7* no vetor pGEMT-Easy: (3) plasmídeo não digerido, (4) análise de restrição com as enzimas *Bam*HI e *Sall*. (C) Subclonagem de *CrNip7* no vetor pETSUMO: (5) plasmídeo não digerido, (6) análise de restrição com as enzimas *Bam*HI e *Sall*. (D) Subclonagem de *CrNip7* no vetor pGEX-4T-1: (7) plasmídeo não digerido, (8) análise de restrição com as enzimas *Bam*HI e *Sall*. (E) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborada pela autora.

Figura 11 – Análise da amplificação por PCR de *CrNip7* e subclonagens para os ensaios de complementação funcional e duplo híbrido: (A) Produto de PCR (540pb) amplificado a partir do molde pGEMT-*CrNip7*-BS: (1) Tm= 62°C; (2) Tm= 64°C. (B) Verificação da clonagem de *CrNip7* no vetor pGEMT-Easy: (3) clone negativo, (4 e 5) clones positivos. (C) Análise da clonagem de *CrNip7* no vetor pJN40: (6) clone negativo, (7,8) clones positivos. (D) Análise da clonagem de *CrNip7* no vetor pLEXA-N: (9 e 10) clones positivos. (E) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Todas as análises de restrição foram realizadas com as enzimas *EcoRI* e *Sall*. Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborado pela autora.

Figura 12 – Análise da amplificação por PCR de *CrNip7* e subclonagens para os ensaios de complementação funcional: (A) Produto de PCR (540 pb) a partir da construção pGEMT-*CrNip7*-BS: (1) Tm= 62°C. (B) Análise da clonagem de *CrNip7* no vetor pGEMT-Easy: (2) plasmídeo não digerido, (3 e 4) análise de restrição com as enzimas *Bam*HI e *Xba*I. (C) Análise da clonagem de *CrNip7* no vetor YCplac33: (5,6) análise de restrição com as enzimas *Bam*HI e *Xba*I. (D) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborado pela autora.

Figura 13 – Análise da amplificação por PCR de *CrNip7* e subclonagens para fins de co-expressão em *E. coli*: (A) Produto de PCR (540pb) amplificado a partir da construção pGEMT-*CrNip7*-BS: (1) Tm= 62°C. (B) Análise de restrição da construção de *CrNip7* no vetor pUC19, usando as enzimas *EcoRI* e *PstI*: (2, 3 e 4) clones positivos. (C) Análise da clonagem de *CrNip7* no vetor pETDuet-1: (5 e 7): plasmídeos não digeridos, (6 e 7) análise de restrição com as enzimas *EcoRI* e *PstI*. (D) M = marcador molecular GeneRuler DNA 1kb (Fermentas). Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborado pela autora.

3.1.2 Ensaio de complementação funcional

O primeiro objetivo deste trabalho foi avaliar se a proteína CrNIP7, de *C. reinhardtii*, seria capaz de complementar a função da proteína Nip7p em levedura. Elucidar esta questão, era crucial para localizar CrNIP7 na via de processamento de RNA durante a biogênese de ribossomos.

Sabe-se que Nip7p, de levedura, e HsNIP7, de humanos, compõem a lista crescente de fatores comuns aos dois organismos que não tem atuação conservada na biogênese de ribossomos. Estudos da literatura mostram que as Nip7p e HsNIP7 são essenciais para a formação dos ribossomos, porém os dados mostram que apesar da alta identidade sequencial entre Nip7p e HsNIP7 (56,67%), elas possuem funções diferentes no processamento de RNA ribossomal. Nip7p participa na formação da subunidade 60S do ribossomo, enquanto que HsNIP7 atua na formação da subunidade 40S⁶⁻⁷ (detalhes sobre a divergência de função entre as duas proteínas estão detalhadas na Introdução). Experimentos de complementação funcional em levedura, realizados no trabalho de Morello *et al* (anexo suplementar),⁷ mostram que HsNIP7 não complementa a função de Nip7p, assim como a ortóloga mais distante PaNIP7, de *Pyrococcus abyssi*, também não é capaz de complementar a função de Nip7p.

Para responder a questão inicial, um ensaio de complementação funcional foi realizado utilizando três linhagens de *S. cerevisiae* diferentes: a linhagem DG130 (derivada da linhagem

B-8106)⁶³, que possui o gene *nip7* mutado (*nip7-1*) no cromossomo, o que a tornou termossensível e de baixo crescimento; e as linhagens DG440 e DG442 (com *background* genético da levedura W303), que possuem *nip7* genômico nocauteado, sendo que para sobreviver, a DG440 possui um plasmídeo com *NIP7* selvagem sob regulação do seu próprio promotor e a DG442 possui um plasmídeo com a CDS *nip7-1* sob regulação do promotor GAL (Tabela 3). As células que possuem o alelo *nip7-1* exibem fenótipos associados a defeitos no aparato de tradução, incluindo a sensibilidade à paromomicina e acúmulo de *halfmers* no perfil polissomal.⁶ Lembrando que *nip7* é um gene essencial, de cópia única no genoma de levedura e as linhagens utilizadas nesse estudo são haploides,⁶ duas abordagens experimentais diferentes foram desenhadas para verificar a função de CrNIP7.

Na primeira, a CDS *CrNip7* foi clonada no plasmídeo pJN40 sob o promotor *CUP1* e a linhagem DG130 foi transformada com a construção pJN40-*CrNip7*. Nesse caso, a levedura DG130 [pJN40-*CrNip7*] foi crescida em meio SD –Ura +Glu contendo sulfato de cobre II e, assim, expressou Nip7p-1 (proteína mutada de *S. cerevisiae*) e CrNIP7 (proteína selvagem de *C. reinhardtii*). Se houver complementação de função, DG130 expressando Nip7p mutada e CrNIP7 deve crescer mais que DG130 expressando apenas Nip7p mutada. O resultado desse experimento está apresentado na Figura 14, e mostra que a levedura DG130 [pJN40-*CrNip7*] cresceu até a sexta casa de diluição após 3 dias de incubação a 30°C, enquanto que a DG130 [pJN40] cresceu até a quarta diluição nas mesmas condições. Além de demonstrar a complementação de função, esse ensaio também revelou o aumento de resistência à temperatura de DG130 [pJN40-*CrNip7*] em relação à DG130 [pJN40].

Na segunda abordagem, a linhagem DG440 transformada com o plasmídeo YCplac33 vazio foi utilizada como controle positivo de fenótipo normal de crescimento em meio SD –Ura +Glu, pois expressa Nip7p sob seu próprio promotor. E a linhagem DG442 transformada com plasmídeo YCplac33 vazio foi utilizada como controle negativo de crescimento em meio SD –Ura +Glu, pois nesse caso Nip7p nativa não é expressa pela ausência de galactose no meio de cultura. Então a verificação de complementação de função se deu pela linhagem DG442 transformada com a construção YCplac33-*CrNip7*, na qual a CDS *CrNip7* está clonada sob o promotor constitutivo TEF, crescendo em meio SD –Ura +Glu. Nesse meio de cultura, a levedura DG442 [YCplac33-*CrNip7*] expressa apenas CrNIP7 e, se houver complementação funcional, deve crescer com o fenótipo normal de crescimento semelhante à linhagem DG440. A Figura

14 mostra que CrNIP7 foi capaz de restaurar o crescimento igual ao fenótipo normal em 24°C e 30°C, além de também crescer a 37°C.

Figura 14 – Ensaio de complementação funcional mostrando que a proteína CrNIP7, de *C. reinhardtii*, complementa a função de Nip7p, de levedura, nas duas abordagens: (A) A levedura DG130 [pJN40-*CrNip7*] expressando CrNIP7 e Nip7p-1 cresce melhor que a levedura DG130 [pJN40], controle de crescimento deficiente e termossensível, que expressa apenas a mutante Nip7p-1. (B) A levedura DG442 [YCplac33-*CrNip7*] expressando CrNIP7 tem crescimento igual ao controle positivo de fenótipo normal, DG440 [YCplac33]. Como esperado a levedura DG442 [YCplac33], utilizada como controle negativo, não cresceu no meio contendo apenas glicose.

Fonte: Elaborada pela autora.

Portanto, os dois experimentos utilizando linhagens de *backgrounds* genéticos diferentes e plasmídeos com diferentes promotores revelaram que a proteína CrNIP7 complementa a função da proteína Nip7p de levedura, uma vez que as linhagens termossensíveis expressando *CrNip7* recuperaram o crescimento do fenótipo normal. Esse resultado é a primeira evidência de que CrNIP7 tem papel importante na biogênese de ribossomos, a princípio na formação da subunidade 60S, assim como Nip7p.

3.1.3 Ensaio de Duplo Híbrido

Sabendo que a proteína CrNIP7 complementa a função de Nip7p em levedura, a próxima questão do trabalho foi investigar se CrNIP7 interage com os mesmos parceiros proteicos que foram reportados na literatura para suas homólogas. Desse modo, o ensaio de duplo híbrido é uma técnica que permite uma triagem direta de cDNAs de *C. reinhardtii*, e foi escolhido para fazer uma varredura inicial de possíveis parceiros de interação de CrNIP7. Para

isso, a CDS *CrNip7* foi clonada no vetor pLEXA-N pelos métodos de biologia molecular (item 2.1) e tem sua confirmação apresentada no item 3.1.1, dos resultados. A construção pLEXA-N-*CrNip7* produz a proteína de fusão BD-CrNIP7, que foi utilizada como a isca no rastreamento. Três clones transformantes da levedura, L40 [pLEX-N-*CrNip7*], foram testados quanto a ativação do sistema pela isca (BD-CrNIP7) através do teste da β -galactosidade. Este teste de auto ativação é necessário pois, sabe-se que várias proteínas que contém domínio de ativação (AD) quando fusionadas ao domínio de ligação ao DNA (BD) são capazes de ativar a transcrição dos genes repórteres, e assim não são apropriadas para a varredura de interações por duplo híbrido. De fato, é descrito na literatura que aproximadamente 10% dos cDNAs clonados randomicamente ao domínio de ligação ao DNA auto ativam o sistema.⁴⁶ Desse modo, as bibliotecas de cDNA comerciais estão clonadas convenientemente fusionadas ao domínio de ativação e não ao domínio de ligação ao DNA.

Como nesse caso, a fusão BD-CrNIP7 não ativou a expressão do gene repórter *LacZ*, ou seja, os clones testados não ficaram azuis no teste da β -galactosidade em papel de filtro (Figura 15 -1), o ensaio pode prosseguir.

Sabe-se também que a expressão basal do gene repórter *HIS3* existe e, então, com o objetivo de eliminar falsos positivos e selecionar as interações positivas, testou-se os clones L40 [pLEXA-N-*CrNip7*] para o crescimento em meio restritivo SD -WH com concentrações crescentes de 3-AT (inibidor competidor na biossíntese de histidina). A concentração de 2 mM de 3-AT foi escolhida para ser adicionada ao meio de cultura para selecionar os transformantes da biblioteca de cDNA.

Após a transformação da levedura L40 [pLEXA-N-*CrNip7*] com a biblioteca de cDNA de *C. reinhardtii* foram selecionadas 125 colônias que atendiam aos requisitos do ensaio e estas foram repicadas em novas placas. Essas colônias foram então submetidas ao primeiro teste de interação para ativação do gene *LacZ* (teste da β -galactosidade em papel de filtro) e *HIS3* (por crescimento em meio restritivo SD-WLH contendo 2 mM de 3-AT). Dessa primeira triagem, 22 duas colônias foram positivas para os dois genes repórteres e foram escolhidas para a extração de DNA plasmidial (contendo a presa) para sequenciamento e identificação.

As sequências que supostamente interagiram foram identificadas por alinhamento no BLASTX⁴⁹ e 11 clones que continham a CDS fora do quadro de leitura ou que apresentavam segmentos clonados sem similaridade com proteínas de *C. reinhardtii* foram eliminados. Dentre os clones com CDS no *frame* +1 e que apareceram em duplicidade, estão as sequências

correspondentes a um segmento da proteína álcool desidrogenase de *S. cerevisiae*, e ao citocromo b (GenBank: CAA39011.1) de *C. reinhardtii*, porém estes clones são falsos positivos. O citocromo b, por exemplo, faz parte de um grupo de proteínas mitocondriais que são frequentemente “pescadas” em ensaios de duplo híbrido e, portanto, foi excluído. As subunidades ribossomais, subunidades de proteossomos e componentes de citoesqueleto também são frequentemente “pescadas” nestes ensaios ⁶⁴; entretanto, nesse estudo em particular, proteínas ribossomais estão no contexto de função esperada para CrNIP7, uma vez que esta participa da biogênese de ribossomos.

Para confirmar as possíveis interações positivas das proteínas que estavam no quadro de leitura correto, a levedura L40 foi co-transformada com o plasmídeo expressando BD-CrNIP7 e cada um dos plasmídeos codificando os 7 potenciais parceiros. Os transformantes foram testados novamente para a ativação dos genes repórteres (Figura 15 e 16). Nesse teste, mais três proteínas falharam em ativar o gene repórter *LacZ*, foram elas: uma proteína hipotética (NCBI: XP_001701328.1) de 87 aminoácidos de *C. reinhardtii*; uma proteína anotada como “major light-harvesting complex II protein m1” (GenBank: AAM18057.1) que participa do complexo antena no processo de fotossíntese; e a proteína frutose-1,6-bisfosfatase, que catalisa a hidrólise de frutose-1,6-bisfosfato em frutose-6-fosfato e é crítica na via de gliconeogênese. O clone expressando a proteína fosforibosilformilglicinamida ciclo ligase (NCBI: XP_001700829.1), que participa na biossíntese de purinas, apresentou uma fraca ativação do gene repórter *LacZ*, mas conseguiu crescer em meio restritivo SD-WLH até 10 mM de 3-AT (Figura 15, linha 5 e Figura 16, linhas 5 e 6). Porém, por não encontrar um contexto de interação em biogênese de ribossomos, esse parceiro foi desconsiderado como positivo.

Outra interação considerada falsa positiva, mesmo ativando os dois genes repórteres, correspondeu à proteína anotada como “flagellar outer dynein arm-docking complex protein 2” (NCBI: XP_001701436.1), que possui uma região Smc, descrita como sendo uma ATPase de segregação de cromossomos. Apesar de aparentar ser uma interação forte em comparação ao controle (Figura 15, linha 3 e Figura 16, linhas 7 e 8), componentes do citoesqueleto são frequentemente pescados em experimentos de duplo híbrido. Como não foi possível identificar um contexto fisiológico para tal interação entre as duas proteínas, os ensaios não seguiram adiante com esse parceiro.

Portanto, as interações consideradas positivas e relevantes no contexto de biossíntese de ribossomos corresponderam a duas proteínas preditas no genoma de *C. reinhardtii*:

Predicted (XP_001690112.1), por não existir nenhuma informação sobre ela na literatura; e a proteína chamada de G-patch (XP_001692447.1) por conter o domínio G-patch conservado. Então, essas duas proteínas foram selecionadas para iniciar os experimentos de validação da interação *in vitro*. Interessantemente, Predicted e G-patch tem várias regiões preditas como intrinsecamente desordenadas (detalhes estão descritos nos próximos capítulos), assim como a proteína “flagellar outer dynein arm-docking complex protein 2” (dado não mostrado), sugerindo que CrNIP7 tenha uma tendência em interagir com proteínas que possuem regiões desordenadas. É importante salientar que todas as sequências analisadas nesse ensaio estavam apenas conceitualmente anotadas no banco de dados do NCBI.

Figura 15 – Teste da β - galactosidase em papel de filtro para verificação da ativação do gene repórter *LacZ* pela isca BD-CrNIP7 (1) e pelas potenciais interações encontradas no rastreamento da biblioteca comercial de *C. reinhardtii* (2, 3, 4 e 5). Controle negativo de interação (6) e controle positivo de interação forte (7).

Fonte: Elaborada pela autora.

Em um panorama geral dos resultados, apenas dois clones, de 125 inicialmente selecionados, foram considerados como interações positivas e relevantes no contexto fisiológico. Além do fato do rastreamento retornar poucas interações, não foi encontrada nenhuma interação de CrNIP7 com possíveis ortólogos, considerando os parceiros de interação já descritos para Nip7p e HsNIP7. Esses fatos, combinados à “pescagem” da proteína álcool desidrogenase de *S. cerevisiae* clonada no vetor da biblioteca, além de sequências sem

similaridade, colocaram em dúvida a qualidade da biblioteca comercial de cDNA utilizada. Então, um ensaio de duplo híbrido direcionado visando a verificação da interação entre CrNIP7 e CrSBDS, uma homóloga de Sdo1, de levedura e SBDS humana, que são parceiras de Nip7p/HsNIP7 respectivamente, foi realizado e seu resultado está apresentado no capítulo dos resultados de CrSBDS (item 3.4).

Assim, com o experimento de duplo híbrido em levedura foi possível identificar a interação de CrNIP7 com duas proteínas relevantes, mas que ainda precisam ser estudadas e validadas *in vitro*. Portanto, a partir desses resultados, esforços se concentraram em tentar expressar e purificar tanto CrNIP7 quanto os seus potenciais parceiros proteicos de para realização de ensaios de *pull down* e caracterização.

Figura 16 – Teste de força de interação das potenciais interações por crescimento em meio restritivo, SD-WLH, com adição gradativa de 3-AT. O experimento foi realizado com dois transformantes diferentes para cada interação.

Fonte: Elaborada pela autora.

3.1.4 Testes de expressão de CrNIP7 em sistema heterólogo

A CDS de CrNIP7 foi clonada em três vetores diferentes de expressão bacterianos, visando obter fusões proteicas diferentes de CrNIP7 para experimentos futuros. CrNIP7, de 179 aminoácidos, tem massa molecular calculada de 20287,6 Da, pI teórico igual a 9,32 e possui apenas um triptofano, portanto tem baixa absorção a 280 nm (dados obtidos pelo programa Protparam⁵¹).

3.1.4.1 Expressão, purificação e teste de clivagem de 6xHis-SUMO-CrNIP7

A fusão 6xHis-SUMO-CrNIP7, com migração esperada em 37 kDa, foi altamente expressa pela indução da construção pETSUMO-*CrNip7* (Tabela 11) em *E. coli* Rosetta (DE3). A Figura 17 mostra o gel da expressão e purificação a partir de uma cultura induzida em 250 mL de meio de cultura e pode-se observar que a banda superexpressa corresponde ao tamanho previsto para a fusão e concentra-se nas frações solúveis, embora apareça também na fração insolúvel. O rendimento obtido nessa expressão foi de aproximadamente 25 mg da proteína 6xHis-SUMO-CrNIP7/L de meio LB, no entanto esse número não é preciso, uma vez que este experimento não foi repetido sistematicamente, e CrNIP7 eluiu com alguns contaminantes significativos.

Testes de clivagem com 6xHis-SUMO-protease (~30 kDa) foram realizados para a fusão 6xHis-SUMO-CrNIP7 imobilizada na resina de afinidade a fim de verificar se CrNIP7 sem marcação era recuperada. A Figura 18 apresenta o melhor resultado do teste de clivagem alcançado. A digestão foi praticamente total, porém para conseguir desprender a proteína CrNIP7 (~20 kDa) de interações inespecíficas com a resina e/ou com a própria *tag* 6xHis-SUMO (~17 kDa), foi necessário adicionar de 500 mM de NaCl. Mesmo com a alta força iônica do tampão, após algumas lavagens, CrNIP7 continuou retida na resina como mostrado na amostra EI (eluição com o competidor) da Figura 18. O rendimento de CrNIP7 recuperada após clivagem não foi estimado, mas foi suficiente para os ensaios propostos.

Figura 17 – Teste de expressão de 6xHis-SUMO-CrNIP7 e purificação. Expressão a partir de *E. coli* Rosetta (DE3) pETSUMO-CrNip7 e purificação por cromatografia de afinidade utilizando 2 mL de resina Ni-NTA (Qiagen) para um extrato proteico proveniente de 250 mL de meio de cultura LB. Análise por SDS-PAGE 15%.

Fonte: Elaborada pela autora.

Figura 18 – Teste de clivagem da 6xHis-SUMO-CrNIP7 imobilizada na resina de afinidade (Ni-NTA) utilizando 6xHis-SUMO protease. Análise por SDS-PAGE 15%.

Fonte: Elaborada pela autora.

3.1.4.2 Expressão e purificação de 6xHis-CrNIP7

A CDS *CrNip7* foi clonada no vetor de co-expressão, pETDuet-1, no sítio que permite a expressão da proteína de interesse fusionada a *tag* 6xHis no seu N-terminal. O segundo sítio de clonagem do vetor foi utilizado, posteriormente, para subclonar a CDS de possíveis parceiras de interação, objetivando avaliar a interação por co-expressão e co-purificação. Como a obtenção de 6xHis-CrNIP7 isolada é estratégica para outros ensaios de *pull down*, testamos a expressão da construção pETDuet-1-*CrNip7* em *E. coli* Rosetta (DE3). A Figura 19 mostra o resultado da expressão de 6xHis-CrNIP7 (~22kDa) em uma cultura de 500 mL de LB e purificação por cromatografia de afinidade. O rendimento preliminar dessa expressão foi de aproximadamente 8 mg de proteína por litro de cultura em LB. Neste sistema heterólogo, 6xHis-CrNIP7 é bem expressa e também se concentra predominantemente nas frações solúveis, sendo um resultado vantajoso para futuros ensaios de caracterização proteica.

Figura 19 – Expressão de 6xHis-CrNIP7 pela indução da construção pETDuet-1-*CrNip7* em *E. coli* Rosetta (DE3) e purificação por cromatografia de afinidade. Foram utilizadas células provenientes de 500 mL de cultura em meio LB e 2 mL de resina Ni-NTA (Qiagen). Análise por SDS-PAGE 15%.

Fonte: Elaborada pela autora.

3.1.4.3 Expressão e purificação de GST-CrNIP7

Paralelamente, visando outras estratégias de *pull down*, a CDS de CrNIP7 também foi clonada no vetor pGEX-4T-1 para produzir a proteína de fusão GST-CrNIP7. Então, a expressão da construção pGEX-4T-1-*CrNip7* em *E. coli* Rosetta (DE3) foi analisada. A Figura 20 mostra um teste preliminar de expressão de 100 mL de cultura induzida e a purificação por cromatografia

de afinidade utilizando a resina Glutathione Sepharose 4 Fast Flow (GE). A fusão GST-CrNIP7 (~46 kDa) migra em ~45 kDa, sendo compatível com o tamanho esperado, mas eluiu com alguns contaminantes. Este foi um teste preliminar e que ainda deve ser otimizado e padronizado.

Figura 20 – Expressão de GST-CrNIP7 pela indução da construção pGEX-CrNip7 em *E. coli* Rosetta (DE3) e purificação por cromatografia de afinidade. Foram utilizados 100 mL de cultura em LB e 0,5 mL de resina Glutathione Sepharose 4 Fast Flow (GE).

Fonte: Elaborada pela autora.

3.1.5 Análise de bioinformática

Sequências de NIP7 são encontradas apenas nos domínios Archaea e Eukarya, porém o domínio conservado PUA (do inglês, *PseudoUridine synthase and Archaeosine transglycosylase*) presente em NIP7 pode ser encontrado em algumas proteínas de bactérias.⁶⁴⁻⁶⁵ O domínio PUA, caracterizado por ser um motivo de ligação ao RNA altamente conservado, está presente em enzimas que catalisam modificações em bases do tRNA, modificações pós-transcricionais de RNA ribossomal, proteínas envolvidas na biogênese de ribossomo e tradução, e também em enzimas envolvidas na biossíntese de prolina.^{20,65-67}

Um alinhamento realizado entre as sequências homólogas a CrNIP7 dos organismos *S. cerevisiae* (Nip7p), *Homo sapiens* (HsNIP7), *Pyrococcus abyssi* (PaNIP7), *Arabidopsis thaliana* (AtNIP7) e *Caenorhabditis elegans* (CeNIP7) (Figura 21) mostra a alta conservação sequencial de NIP7 de eucariotos, sendo a ortóloga PaNIP7, de Archaea, a que mais difere sequencialmente. A matriz de identidade (Tabela 14) indica que CrNIP7 possui 55,31% de identidade com Nip7p e 63,69%, com HsNIP7. Um fato interessante é que apesar da maior identidade sequencial de CrNIP7 com HsNIP7, os ensaios funcionais de complementação

mostraram que CrNIP7 possui a função biológica conservada de Nip7p, o que não é observado para HsNIP7.⁶⁻⁷

```

PaNIP7 1 MSGELRVRRASSWELDL-----ILKEAEKYG-ELLHEFFCVCVEGKYRDVY 44
Nip7p 1 -----MRQLTEETKVVFEKLAGYIGRNISFLVDNKE-----LPHVFRLQK-DRVY 45
AtNIP7 1 -----MRPLDETETTVVFEKLFKFVGNNLKKIVENPSDEGPSTPGSYCFRLQK-NRVY 53
CeNIP7 1 -----MRPLTEETSLVFAKLASFIGDNVSMILDRND-----GDYCFRNHK-ERVY 45
CrNIP7 1 -----MRPLTEETKAVFEKLYKFIGKNIKNLIDRPN-----DPHCFRLHK-NKVF 45
HsNIP7 1 -----MRPLTEETTRVMFEKIAKYIGENLQLLVDRPD-----GTTCFRLHN-DRVY 45
          :* . *                :... . . . : *:

PaNIP7 45 AVNEEVWKIIEDINM-RPYSLGTFVGTIRVDENLVEKFYPNLEFFS-LIKLEKNYVILGP 102
Nip7p 46 YVPDHVAKLATSVARPNLMSLGICLGKFTKTGK----FRLHITS�TVLAKHAKYKIWIKP 101
AtNIP7 54 YVSEALVKRATNISRKNLVSFGTCIGKYTHAGS----FHLTİMSLİNILAANAKHKVWLKP 109
CeNIP7 46 YCSENLMRQAACIAREPLLSFGTCIGKFTKSKK----FHLQİTİLDYLİPYİKİVWLKP 101
CrNIP7 46 YVKERLMKKATNVARDKLVALGTCIGKLTHTGK----FRLTİGİLDLİSİYİKİVWLKP 101
HsNIP7 46 YVSEKİMKLAANİSGDKLVSLGTCİFGKFTKTHK----FRLHVTİLDYLİPYİKİVWLKP 101
          : : : : : : * . * . . * : : : * : : *

PaNIP7 103 KASFLFTTGKDAPKEAVREIKWQG--SKRVVLNDLGDİİGİGLİNPK--S-----DRR 152
Nip7p 102 NGEMPFLYGNHVLKAHVGMSSDİPEHİGİVİFİAMNDVPLGFGVİSAKİSTESİRNMQPTGİ 161
AtNIP7 110 TSEMSFLYGNHVLKGGİGRİTDSİVPGDGİVİVİSMSDVPLGFGİAİAKİSTQDCRKLDPNGİ 169
CeNIP7 102 NİEQQFLYGNİILKSGİİARMTDGTPTİHİGİVİVİSMTDVPLGFGVİSAKİTSDSKRADPTİL 161
CrNIP7 102 SİEMQFLYGNİVLKSGİLİRTİTİPAYTGVİVİYNMİDİPLGFGVTİAKİSTNECRNLDPNİV 161
HsNIP7 102 GİEQSFLYGNHVLKSGİGRİTİTİSİYQGVVİYSMİDİPLGFGVİAİAKİSTQDCRKVDPİAİ 161
          .. * * : . * : . . : : : : * : * : .

PaNIP7 153 FİKNLKDVGİEFLRR----- 166
Nip7p 162 VİFRQADİGİYLİRDEDTLFT 181
AtNIP7 170 VVLHQADİGİYLİRGEDDL-- 187
CeNIP7 162 VVLHQCDLGEYLİRNİSHLA- 180
CrNIP7 162 VTFHQADCGEFLİRAEDEL-- 179
HsNIP7 162 VVFHQADİGİYVRHEETLT- 180
          . . * * : : *

```

Figura 21 – Alinhamento gerado pelo programa ClustalO comparando as sequências homólogas de CrNIP7 (XP_001696971.1): PaNIP7 (CAB49179.1), Nip7p (EIW07327.1), AtNIP7 (AEE83650.1), CeNIP7CCD65161.1, e HsNIP7 (NP_057185.1). Os aminoácidos conservados estão representados por asteriscos, os aminoácidos com alta similaridade com dois pontos e aminoácidos com baixa similaridade com um ponto.

Fonte: Elaborada pela autora.

Tabela 14 – Matriz de identidade entre as sequências homólogas a CrNIP7 (programa Clustal O).

Organismo	<i>P. abyssi</i>	<i>S. cerevisiae</i>	<i>A. thaliana</i>	<i>C. elegans</i>	<i>C. reinhardtii</i>	<i>H. sapiens</i>
<i>P. abyssi</i>	100	20.95	23.87	21.62	20.95	22.30
<i>S. cerevisiae</i>	20.95	100	55.31	53.33	55.31	56.67
<i>A. thaliana</i>	23.87	55.31	100	54.19	57.54	62.01
<i>C. elegans</i>	21.62	53.33	54.19	100	55.87	61.11
<i>C. reinhardtii</i>	20.95	55.31	57.54	55.87	100	63.69
<i>H. sapiens</i>	22.30	56.67	62.01	61.11	63.69	100

Fonte: Elaborada pela autora.

No banco de dados do Pfam, CrNIP7 está classificada na superfamília PUA/ASCH (acesso: CL0178) dentro da família UPF0113, que compreende sequências de eucariotos e Archaeas envolvidas na biogênese de subunidades ribossomais, sendo HsNIP7 e PaNIP7, os membros caracterizados. Essa família também compreende proteínas ainda não caracterizadas que apresentam a arquitetura UPF0113 em conjunto com outros domínios como o domínio Sec39 (PF08314), envolvido no transporte retrógrado Golgi-Retículo Endoplasmático; o domínio AAA_22 (PF13401), de ATPases associadas com diversas atividades celulares; e o domínio DEAD (PF00270), presente nas helicases do tipo *DEAD/DEAH box* que estão envolvidas no metabolismo de RNA incluindo *splicing*, biogênese de ribossomos, transcrição nuclear, tradução e expressão gênica. Em vista de todo conhecimento na literatura sobre Nip7p, HsNIP7 e PaNIP7 até o momento, a descrição e referências sobre NIP7 no Pfam estão desatualizadas.

Atualmente, existem duas estruturas cristalográficas de HsNIP7 (PDB:1SQW e 1T5Y) e uma de PaNIP7 (PDB:2P38) depositadas no *Protein Data Bank* (<http://www.rcsb.org>). Então, foi realizada uma comparação entre a composição dos elementos estrutura secundária do modelo cristalográfico resolvido para HsNIP7 com as predições feitas para CrNIP7 pelos programas GOR4 e PSIPRED (Figura 23). HsNIP7 é composta por dois domínios mistos α/β , sendo que o domínio N-terminal (resíduos 1-92) tem uma arquitetura nova, sem similaridade com domínios já conhecidos, e o domínio C-terminal (resíduos 93-176) corresponde ao domínio PUA. É possível especular que as estruturas tridimensionais de CrNIP7 e HsNIP7 sejam muito parecidas, uma vez que elas têm alta identidade sequencial. As predições para CrNIP7, no geral, coincidem com o modelo resolvido para HsNIP7, porém as estruturas β_1 , α_2 , α_4 , β_5 , α_6 , β_8 , α_7 , β_9 , e α_8 não foram preditas pelos dois programas ou por algum deles. Porém, observa-se que a combinação dos resultados entre as duas predições chega muito próxima ao modelo cristalográfico de HsNIP7.

3.1.6 Conclusão parcial

A proteína CrNIP7 apresenta alta identidade com as proteínas Nip7p (*S. cerevisiae*) e HsNIP7 (humanos) e mostrou-se capaz de complementar a função biológica de Nip7p no ensaio em levedura, apesar de sua maior identidade ser com a homóloga de humanos. O resultado experimental aponta na direção de que CrNIP7 tenha papel na formação da subunidade 60S do

ribossomo eucariótico. As proteínas G-patch e Predicted, “pescadas” como parceiras de interação de CrNIP7 *in vivo* no ensaio de duplo híbrido, não tem homólogas descritas como parceiros de NIPs de outros organismos. Esse fato amplia a lista de parceiros para NIPs e a continuação dos estudos visando a validação *in vitro* destes parceiros está descrita nos próximos capítulos. Além dos experimentos funcionais, a clonagem da CDS *CrNip7* em três construções para expressão heteróloga permitiu a obtenção das proteínas solúveis: 6xHis-SUMO-CrNIP7, 6xHis-CrNIP7 e GST-CrNIP7. Finalmente, foi possível obter CrNIP7 sem *tag*, pela clivagem da fusão 6xHis-SUMO-CrNIP7 com SUMO-protease.

Figura 22 – Comparação entre a estrutura secundária resolvida por cristalografia de raio X da proteína HsNIP7 (PDB:1SQW) (linha 1, em vermelho) e às predições de estrutura secundária pelos programas GOR4 (linha 4, em azul) e PSIPRED (linha 5, em azul) de CrNIP7 (linha 3). As α-hélices estão representadas por um cilindro e as fitas-β por setas. Os aminoácidos realçados em cinza fazem parte do domínio PUA.

Fonte Elaborada pela autora.

3.2 Predicted

A interação *in vivo* entre a proteína CrNIP7 e Predicted foi encontrada através da varredura da biblioteca de cDNA de *C. reinhardtii* pelo sistema de duplo híbrido em leveduras. O vetor pGAD-HA-predicted (AD-isca) foi isolado de um transformante positivo para a expressão dos dois genes repórteres, *LacZ* e *HIS3*. A sequência codificadora de Predicted foi determinada por sequenciamento automático e analisada contra o banco de dados geral do programa BLASTX⁴⁹ no servidor do NCBI. A busca retornou dois *hits* significativos pertencentes a proteínas hipotéticas anotadas por tradução conceitual do genoma de duas algas unicelulares da ordem Volvocales. O primeiro *hit* (NCBI: XP_001690112.1), com 100% de cobertura e identidade com a sequência contrastada pertence a uma proteína de 159 aminoácidos de *C. reinhardtii*. Como não há nenhum tipo de informação funcional/estrutural sobre esta proteína até o momento, será referida ao longo do texto como Predicted. O segundo *hit* (GenBank: KXZ55342.1) também significativo, mas com 67% de cobertura e 47% de identidade com a sequência contrastada, pertence a uma proteína de 180 aminoácidos de *Gonium pectorale*.

Uma busca por similaridade no genoma de *Volvox carteri*, também da ordem Volvocales, com auxílio do Professor Ricardo De Marco, resultou na sequência que codifica a proteína predita de identificação Vocar.0001s0712.1 no banco de dados Phytozome11.

Como essas proteínas são relativamente pequenas, predições gênicas podem excluí-las, mascarando a existência de homólogas em outros organismos. As proteínas preditas de *G. pectorale* e *V. carteri* possuem 180 aminoácidos enquanto a de *C. reinhardtii* possui 159 aminoácidos. O alinhamento múltiplo entre as três sequências (Figura 23) mostra os aminoácidos conservados e revela alta similaridade entre elas (Tabela 15).

Para iniciar o estudo sobre a proteína Predicted nesse trabalho, sua CDS foi amplificada com intuito de cloná-la em vetores de expressão para estudos de *pull down* e caracterização. Uma análise de bioinformática também foi realizada e está apresentada nos tópicos seguintes.

```

G.pectorale      1 MDLGLYQPPARVVGFTYSHPERDSTLNCSETATSKLVTQRSLLSELLAAQLLGDEDEER 60
V.carteri       1 ----MDSGLRVVGFTFSNMGD--STLNSEVSNAKLQTQRSLLSELLAIHAVGDDSLNSD 53
C.reinhardtii   1 --MDFLPVAPRVVGFSATEAP---LQHEAAHEVCKLQTQRSLLSEMLSQCAQEGEYSESE 55
                *****: :. : .** ******: : :.

G.pectorale      61 SGS-YDDEETGSDEDEDVEDPANAPPARQSTYQQLMPSPSEPISIPCSGSMANQRVKDAYKR 119
V.carteri       54 PASDVSGEDDD---VDEYDAGAEPVPVAR-ADTFGAQSNPIPIPCPESLANARARDAYRR 109
C.reinhardtii   56 ----VDDEESY---SDDDDAGAPIVRQDTGSSMEPARSEPIAIPCGESFANQRAKDAYRR 108
                . *: : * * . . : **: * * * *: * * *: * *: *

G.pectorale      120 LLHQRQMEQLYQQQLRLRLQQQ-----RLAQQQQAMYYQQ-----SGHVPNVGHAAQGP 166
V.carteri       110 LLYQRKVEHLYYQQQQQWQYNAQQQLLIQQQRRRLHYLQQQQNQVAPPTHQPGQERDV--A 167
C.reinhardtii   109 LLYQKQVEHLYQKQQQQVQQ-----IALAQQRQAYLVQQQRLQQAQAQARAQGRV---- 159
                **: *: *: * * : * : *: * : : :

G.pectorale      167 AMPQGPQCHPAGAY 180
V.carteri       168 VQQHGPPTA-TVPVY 180
C.reinhardtii   -----

```

Figura 23 – Alinhamento múltiplo realizado pelo programa ClustalO das proteínas preditas das algas verdes *C. reinhardtii* (XP_001690112.1), *G. pectorale* (KXZ55342.1) e *V. carteri* (Vocar.0001s0712.1). Os aminoácidos conservados estão representados por asteriscos, os aminoácidos com alta similaridade com dois pontos e aminoácidos com baixa similaridade com um ponto.

Fonte: Elaborada pela autora.

Tabela 15 – Matriz de identidade entre as proteínas preditas de *C. reinhardtii*, *G. pectorale* e *V. carteri* pelo programa ClustalO.

Organismo	<i>G. pectorale</i>	<i>V. carteri</i>	<i>C. einhardtii</i>
<i>G. pectorale</i>	100	37,95	43,33
<i>V. carteri</i>	37,95	100	44,52
<i>C. einhardtii</i>	43,33	44,52	100

Fonte: Elaborada pela autora.

3.2.1 Amplificação da CDS *Predicted* e clonagem em vetores de expressão

A CDS de 480 pb que codifica a proteína Predicted (Apêndice A- Figura 42 e 43) estava com a sequência completa no plasmídeo “pescado” na biblioteca de cDNA de *C. reinhardtii* pelo ensaio de duplo híbrido. Assim, a construção pGAD-HA-*predicted* foi utilizada como molde para a amplificação com o par de oligonucleotídeos *Predicted-Bam*HI-*Nde*I-F e *Predicted-Xho*I-R (Tabela 2). A CDS de Predicted foi amplificada com sucesso e foi clonada no vetor de propagação pUC19 (previamente digerido com *Sma*I), sequenciada e depois inserida no vetor de co-expressão pETDuet-1-*CrNip7*, nos sítios de *Nde*I e *Xho*I, e nos vetores de expressão pETSUMO e pGEX-4T, nos sítios *Bam*HI e *Xho*I (Figura 24).

Figura 24 – Análise da amplificação e subclonagens de *predicted* em vetores de expressão: (A) “M” = Marcador molecular GeneRuler DNA 1kb (Fermentas). (B) Amplificação da CDS *predicted* (~480 pb) a partir de pGAD-HA-*predicted* (1), utilizando Tm= 54°C. (C) Clones positivos em vetor de propagação pUC19: (2,4) análise de restrição com *NdeI* e *XhoI*, liberando o inserto correspondente à *predicted* e uma banda de ~230 pb; (3,5) plasmídeos não digeridos. (D) Clone positivo de *predicted* em pETDuet-1-*CrNip7*; (6) plasmídeo não digerido; (7) análise de restrição com *NdeI* e *XhoI*. (E) Clones positivos de *predicted* em pETSUMO; (8,10) plasmídeos não digeridos; (9,11) análise de restrição com *BamHI* e *XhoI*. (F) Clones positivos de *predicted* inseridos em pGEX-4T; (12,14) plasmídeos não digeridos; (13, 15) análise de restrição com *BamHI* e *XhoI*. Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborada pela autora.

3.2.2 Teste de expressão

Como a estrutura de Predicted é predita por ter regiões intrinsecamente desordenadas (IDRs) e é composta por muitos aminoácidos carregados e hidrofílicos (item 3.2.4), essa pode ser uma característica que interfira na sua expressão em sistema heterólogo de bactéria. Assim, a abordagem foi expressá-la em três sistemas diferentes. A Figura 25B mostra a expressão da construção pETDuet-*CrNip7-predicted* em *E. coli* Rosetta (DE3), que deveria resultar na co-expressão das proteínas 6xHis-CrNIP7 (~22 kDa) e *predicted* (~18 kDa). Porém, apenas 6xHis-CrNIP7 foi expressa, como indicado pela seta, pois a banda de Predicted (tem migração aparente em aproximadamente 26 kDa, ver item 3.2.3) não aparece. Então, a validação da interação entre CrNIP7 e Predicted por co-expressão e co-purificação através desse sistema não foi possível de ser realizada.

A Figura 25C mostra a expressão da fusão “GST-predicted” (~ 44 kDa) no vetor pGEX-4T em *E. coli* Rosetta (DE3). Como a proteína GST (~26 kDa) tem migração aparente de 30 kDa e Predicted (~18 kDa, mas com migração anômala de 26 kDa), esperava-se uma proteína de fusão “GST-predicted” com migração próxima de 56kDa. Pode-se observar que uma banda de aproximadamente 45kDa, presente na amostra induzida, já está presente na amostra da cultura não induzida. Embora o resultado seja promissor, os experimentos com essa construção ainda não foram realizados, uma vez que optamos por testar a próxima construção descrita.

Figura 25 – Teste de expressão da proteína Predicted em três sistemas diferentes: (A) “M” = Marcador de massa molecular de proteínas; “MB”= Marcador de baixa massa molecular de proteínas. (B) Teste de co-expressão de 6xHis-CrNIP7 e Predicted pelo vetor pETDuet-1; (1) cultura não induzida; (2) cultura induzida. (C) Teste de expressão da proteína de fusão “GST-predicted”; (3) Cultura não induzida; (4) Cultura induzida; (5) Fração insolúvel; (6) Fração solúvel. (D) Teste de expressão da fusão “6xHis-SUMO-predicted”; (7) Cultura não induzida; (8) Fração insolúvel; (9) Fração solúvel. SDS-PAGE Tricina 10%.

Fonte: Elaborada pela autora.

Dentre os sistemas testados, a maior expressão obtida foi o da fusão 6xHis-SUMO-predicted (~ 31 kDa) no vetor pETSUMO em *E. coli* Rosetta (DE3), mostrada na Figura 25D. O sistema de SUMO-fusão é conhecido por aumentar a solubilidade das proteínas fusionadas e por estabilizá-las em solução.⁶⁸ A tag 6xHis-SUMO (~13kDa) tem migração aparente de 17 kDa, assim a banda da fusão superexpressa foi obtida em 43 kDa e se mostrou solúvel após a lise celular.

3.2.3 Teste de clivagem de 6xHis-SUMO-predicted e *pull down* com CrNIP7

O teste preliminar de clivagem da fusão 6xHis-SUMO-predicted imobilizada na resina de níquel com SUMO-protease (Figura 26) mostrou que a fusão foi completamente clivada. Porém, nas condições do tampão utilizado, a proteína Predicted não se desprende da resina de níquel, provavelmente por interações inespecíficas com a resina ou com a própria *tag* 6xHis-SUMO que ficou imobilizada. Pode-se observar que a maior parte de Predicted foi eluída apenas com a adição de imidazol no tampão, junto com a SUMO e os contaminantes que permaneciam ligados à resina.

As quatro bandas mais fortes da eluição com imidazol foram recortadas e analisadas por espectrometria de massa para identificação das proteínas. A banda de aproximadamente 40 kDa se revelou como um fator de alongação de *E. coli* (NCBI: WP_032251098.1), a banda de 26 kDa foi identificada como sendo a proteína Predicted (confirmando sua migração anômala) e a banda de 17 kDa foi identificada como a *tag* 6xHis-SUMO. A banda forte de aproximadamente 20 kDa entre Predicted e a *tag* 6xHis-SUMO trouxe 3 *hits* na busca: Predicted, CrNIP7 e uma proteína de *E. coli* chamada YRDA (GenBank: ABE09161.1), da família de anidrase carbônica. Esse último resultado pode ter sido fruto de contaminação durante o manuseio das amostras, uma vez que CrNIP7 não foi expressa nesse experimento.

Figura 26 – Teste preliminar de clivagem da fusão 6xHis-SUMO-predicted com SUMO protease na resina de níquel. SDS-PAGE Tricina 10%.

Fonte: Elaborada pela autora.

Um ensaio preliminar de *pull down* foi realizado com a amostra da eluição de Predicted após a clivagem com SUMO protease e a proteína 6xHis-CrNIP7 imobilizada na resina de níquel como mostrado na Figura 27. Observa-se que na eluição com imidazol, aparece uma banda acima de 6xHis-CrNIP7 que provavelmente corresponde à Predicted (~ 26 kDa). Mas como as amostras purificadas de 6xHis-CrNIP7 e de Predicted contém muitas impurezas, se faz necessária a otimização da purificação para novos ensaios de *pull down* e comprovação da interação. Outra estratégia que poderá ser usada para a validação da interação é o *pull down* entre a fusão GST-CrNIP7 (~46 kDa) com Predicted. Essa abordagem tem a vantagem de eliminar o problema das duas proteínas parceiras migrarem em tamanho próximo em SDS-PAGE e ainda permite o uso de outro tipo de resina de afinidade.

Figura 27 – Teste preliminar de *pull down* das proteínas CrNIP7 e Predicted. SDS-PAGE Tricina 10%.

Fonte: Elaborada pela autora.

3.2.4 Análise de bioinformática

Predicted é uma proteína codificada por uma sequência de nucleotídeos de 480 pares de base, possui 159 aminoácidos, massa molecular predita de 17922,7 Da e *pI* teórico igual 4,86 calculados pelo programa Protparam. Por não possuir triptofano (Trp), a proteína não pode ser fielmente quantificada por absorção em 280 nm. Além disso, não foram identificados domínios conservados conhecidos, até o momento, através de análise pelo programa Pfam.

A composição de Predicted é de 24,5% de aminoácidos descritos como “provedores de ordem” como Ile, Leu, Val, Trp, Tyr, Phe, Cys, e Asn e de 64,5% de aminoácidos “provedores de

desordem” como Ala, Arg, Gly, Gln, Ser, Glu, Lys e Pro ⁶⁸⁻⁶⁹. Essa informação, combinada ao alto conteúdo de aminoácidos polares e carregados em relação à quantidade de aminoácidos hidrofóbicos, e a existência de regiões de baixa complexidade sugeriu que Predicted faz parte das chamadas proteínas intrinsecamente desordenadas (IDPs) ou com regiões intrinsecamente desordenadas (IDRs).

Neste trabalho, comparamos os resultados de dois programas de predição de estrutura secundária com os resultados de dois programas de predição de proteínas intrinsecamente desordenadas, que utilizam métodos diferentes, capazes de identificar padrões nas sequências primárias das proteínas. A predição de estrutura secundária resultou em 61,6% (GOR4) e 56,6% (PSIPRED) de elementos de estrutura secundária, sendo que os dois programas concordam na predição de duas α -hélices com tamanho médio de 27 e 58 resíduos, e uma fita- β estendida (3-4 resíduos) (Figura 28, linha 2 e 3). O programa GOR4 prediz três fitas- β a mais que o programa PSIPRED.

As predições de regiões intrinsecamente desordenadas pelos programas Globplot e DisEMBL complementam a predição de estruturas secundária (ou seja, as regiões ordenadas coincidem com elementos de estrutura secundária e as regiões desordenadas coincidem com a predição de *coils*) (Figura 28, linha 4 e 5). O programa ANCHOR (Figura 28, linha 6) foi utilizado para predizer sítios de ligação na sequência de Predicted, e resultou em 5 possíveis regiões de ligação, sendo três delas situadas dentro de regiões preditas como ordenadas (duas α -hélices e uma fita- β) e as outras duas em regiões preditas como intrinsecamente desordenadas.

Considerando as análises de bioinformática, a proteína Predicted parece se comportar como uma IDR, e provavelmente não forma uma estrutura compacta e globular em solução aquosa já que possui poucos aminoácidos hidrofóbicos para formar o núcleo hidrofóbico em relação ao alto teor de aminoácidos polares e carregados que levam a uma forte repulsão eletrostática.⁶⁹

A ausência de estrutura tridimensional estável e definida em solução das IDPs não significa que não possuem função, pelo contrário, as IDPs podem passar por uma reestruturação transitória quando ligadas a outras moléculas e, portanto, possuem grande importância em vias de sinalização celular e em processos celulares como regulação de transcrição e tradução. Muitas IDPs também estão envolvidas em ligação a ácidos ribonucleicos e em interações com outras proteínas. As IDPs também tem papel fundamental na montagem de ribossomos, na organização da cromatina, no transporte de pequenas moléculas, no

funcionamento de chaperonas, e na separação de domínios funcionais de proteínas, atuando como *linkers* flexíveis.^{69,71-73}

Figura 28 – Sequência de aminoácidos de Predicted (linha 1). As previsões de estrutura secundária feitas pelos programas GOR4 (linha 2) e PSIPRED (linha 3) utilizam o código linha = *coil*, seta = fita β e cilindro = α -hélice. As previsões de regiões intrinsecamente desordenadas pelos programas Globplot (linha 4), DisEMBL (linha 5), tem o código: “D”= desordenado, e “-” = ordenado. Previsões de possíveis sítios de ligação pelo programa ANCHOR estão representadas em laranja (linha 6).

Fonte: Elaborada pela autora.

O fato de Predicted ser encontrada como parceira de interação de uma proteína envolvida na biogênese de ribossomos, especificamente no processamento de rRNA, sugere que ela também participe, de alguma forma, do mesmo processo. Por ser uma proteína predita com regiões intrinsecamente desordenadas, pode-se especular que esteja envolvida na mediação de interações proteína-proteína e/ou proteína-ácido nucléico, já que sua parceira de interação, CrNIP7, tem um domínio PUA de ligação ao RNA.

3.2.5 Conclusão parcial

A interação entre CrNIP7 e Predicted encontrada pelo sistema de duplo híbrido em leveduras sugere que Predicted também participe da biogênese de ribossomos em *C. reinhardtii*. Neste trabalho, foi possível amplificar a sequência codificadora de Predicted e subcloná-la em alguns sistemas de expressão. A fusão 6xHis-SUMO-predicted foi produzida de forma heteróloga, porém a clivagem da fusão deve ainda ser otimizada para ensaios de *pull down* com GST-CrNIP7. Embora ainda preliminar, os resultados sugerem que as proteínas CrNIP7 e Predicted interagem também *in vitro*, mas esse fato necessita de confirmação.

A ausência de similaridade com outras proteínas, combinada com a predição de IDR e a aparente exclusividade em Volvocales, torna Predicted uma proteína muito interessante para estudos futuros.

3.3 G-patch

A interação entre CrNIP7 e uma proteína com um domínio conservado chamado G-patch em sua extremidade N-terminal foi encontrada pela varredura da biblioteca de cDNA de *C. reinhardtii*, utilizando a técnica de duplo híbrido em leveduras. O plasmídeo pGAD-HA-G-patch (AD-isca) foi isolado de um transformante positivo para a expressão dos dois genes repórteres, *LacZ* e *HIS3*. A sequência codificadora da proteína “pescada” foi, então, determinada por sequenciamento automático e analisada contra o banco de dados geral do programa BLASTX⁴⁹ no servidor do NCBI. O melhor *hit* da busca (NCBI: XP_001692447.1), com 97% de identidade sequencial dentro de 65% de cobertura com a sequência contrastada e parâmetro de confiabilidade (*e-value*) igual a 2e-85, pertence à uma proteína predita e hipotética de 478 aminoácidos de *C. reinhardtii*. Um alinhamento foi realizado entre a sequência de nucleotídeos do plasmídeo pGAD-HA-isca e a sequência conceitual da CDS predita (NCBI: XM_001692395.1) de 1437 pb. Este revelou que a CDS da proteína predita estava parcialmente clonada no vetor de duplo híbrido (apenas os 730 pb a partir do códon inicial ATG) (Apêndice A – Figura 44 e 45). Logo, é interessante notar que apenas os 243 primeiros aminoácidos da extremidade N-terminal, que carrega o domínio conservado G-patch, foram suficientes para a interação com CrNIP7.

3.3.1 Amplificação da sequência codificadora de G-patch

O par de oligonucleotídeos, *G-patch-BamHI-NdeI-F* e *G-patch-KpnI-SalI-R* (Tabela 2), utilizado para a amplificação da CDS da proteína G-patch foi desenhado a partir da sequência XM_001692395.1, anotada no banco de dados do NCBI. A PCR utilizando o cDNA total de *C. reinhardtii* como molde foi realizada algumas vezes, em diferentes condições, e em nenhuma apresentou amplificação da banda de 1437pb esperada, porém houve a amplificação de duas bandas mais fortes em 1800 pb e 900 pb, aproximadamente (Figura 29). A reprodutibilidade desse padrão de amplificação levou ao sequenciamento das bandas para identificação. A banda de aproximadamente 900 pb foi resultado de um anelamento inespecífico dos oligonucleotídeos na sequência codificadora (NCBI: XM_001693948.1) de uma proteína predita de *C. reinhardtii*, anotada como subunidade catalítica alfa de uma Serina/Treonina Quinase (XP_001694000.1) e foi, por isso, descartada. O produto de aproximadamente 1800 pb foi identificado como a CDS (NCBI: XM_001692395.1) da proteína (NCBI: XP_001692447.1) que encontramos pelo duplo híbrido, porém com 375 pb a mais do que a sequência anotada por tradução conceitual do genoma de *C. reinhardtii*.

Figura 29 – Amplificação da CDS da proteína G-patch, apresentando bandas amplificadas por PCR, usando $T_m = 54^\circ \text{C}$.

Fonte: Elaborada pela autora.

O sequenciamento do fragmento amplificado foi comparado com o *locus* do gene no genoma de *C. reinhardtii* e, assim, foi possível completar a CDS real da proteína predita com tamanho total de 1812 pb. A sequência encontrada neste trabalho é a mesma que a depositada no banco de dados Phytozome 11 (<https://phytozome.jgi.doe.gov/pz/portal.html>) com identificação de acesso Cre07.g347400 (Apêndice A). Esta proteína predita foi denominada G-

patch, devido ao nome do único domínio presente na extremidade N-terminal encontrado pela análise do programa Pfam (Figura 31).

3.3.2 Clonagens de G-patch e teste preliminar de expressão

A CDS *G-patch* completa (1812 pb) foi inicialmente clonada no vetor de propagação pUC19 (previamente digerido com *Sma*I) (Figura 30B) e sequenciada. Após a confirmação da sequência, esta foi subclonada nos sítios de *Nde*I e *Kpn*I no vetor de co-expressão pETDuet-1-*CrNip7* (Figura 30C) e, paralelamente, foi subclonada nos sítios de *Bam*HI e *Sal*I no vetor de expressão pETSUMO (Figura 30D).

Figura 30 – Análise da amplificação e subclonagens de *G-patch*: (A) “M” = Marcador molecular GeneRuler DNA 1kb (Fermentas). (B) Clones positivos de *G-patch* em vetor de propagação pUC19: (1 e 4) plasmídeo não digerido; (2 e 5) análise de restrição com *Nde*I e *Sal*I; (3 e 6) análise de restrição com *Bam*HI, as manchas ao fim do gel correspondem a RNA remanescente da extração plasmidial pelo método lise alcalina. (C) Clone positivo de *G-patch* em pETDuet-1-*CrNip7*; (7,9) plasmídeo não digerido; (8, 9) restrição com *Bam*HI e *Xho*I. (D) Clones positivos de *G-patch* em pETSUMO; (12,14) plasmídeo não digerido; (13,15) análise de restrição com *Bam*HI e *Sal*I. Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborado pela autora.

A expressão da CDS *G-patch* produz, por tradução conceitual, uma proteína de 603 aminoácidos, massa molar calculada de 62924 Da e pI teórico de 9,13, conforme calculado pelo programa Protparam.

Os testes preliminares de co-expressão de 6xHis-*CrNIP7* e *G-patch* (em pETDuet-1-*CrNip7-G-patch*) e de expressão de 6xHis-SUMO-*G-patch* (em pETSUMO-*G-patch*) foram realizados em *E. coli* Roseta (DE3), em meio de cultura LB, com indução da expressão com 0,5 mM de IPTG a 18°C por 16 horas. Nessa condição, as proteínas recombinantes não foram

expressas em nenhum dos dois sistemas de expressão (dados não mostrados). Já haviam expectativas de que fosse difícil expressar a proteína G-patch em função de seu tamanho aliado a predição de muitas IDR's, como descrito a seguir.

3.3.3 Análise de bioinformática

A sequência proteica de G-patch apresenta alto teor de aminoácidos carregados e hidrofílicos em relação aos aminoácidos hidrofóbicos, regiões de baixa complexidade com aminoácidos repetidos e é composta por 74,4% de aminoácidos considerados “desestruturantes” (Ala, Arg, Gly, Gln, Ser, Glu, Lys e Pro) e apenas 14,8 % de aminoácidos “provedores de ordem” (Ile, Val, trp, Tyr, Phe, Cys e Asn). Desse modo, a propensão de G-patch formar elementos de estrutura secundária, IDRs e possíveis regiões de ligação a outras biomoléculas foi investigada. As predições de estrutura secundária, pelos programas GOR4 e PSIPRED, concordam na maioria das α -hélices e fitas- β preditas, mas há divergências no tamanho e na predição dos dois tipos de elementos em alguns casos (Figura 31, linha 2 e 3). Os programas de predição de IDRs, Globplot e DisEMBL, resultaram em grandes extensões desestruturadas da proteína G-patch ao longo de toda sequência, concordando entre si na maioria dos casos (Figura 31, linhas 4 e 5). Porém, ao comparar as predições de IDRs e estrutura secundária, observa-se que algumas regiões desordenadas coincidem com elementos de estrutura. Assim, pode-se especular que G-patch seja uma IDP com tendência em formar elementos de estrutura ao interagir com outras proteínas e/ou ácidos nucleicos. O programa ANCHOR predisse 14 regiões de tamanhos entre 9 e 55 aminoácidos de possíveis sítios de ligação. Essas regiões estão ao longo de toda extensão da proteína, porém o programa não destacou a sequência do domínio conservado G-patch como possível sítio de ligação.

Figura 31 – Sequência de aminoácidos de G-patch (linha1), o domínio conservado está realçado cinza. As linhas 2 e 3 representam as predições de estrutura secundária pelos programas GOR4 e PSIPRED, respectivamente. As linhas 4 e 5 representam as predições de IDR pelos programas Globplot e DisEMBL, respectivamente. A linha 6 mostra os possíveis sítios de ligação (em laranja) preditos pelo programa ANCHOR.

Fonte: Elaborada pela autora.

3.3.4 O domínio G-patch

O domínio G-patch, de 45-50 aminoácidos e rico em glicinas, foi identificado por alinhamento de sequências e descrito pela primeira vez por Aravind & Koonin.⁷⁴ Este domínio pode ser encontrado tanto nas extremidades N- e C- terminais de proteínas ou flanqueados por regiões de baixa complexidade.⁷³⁻⁷⁴ G-patch apresenta um motivo de 6 glicinas altamente conservadas; e além disso, o resíduo que segue a primeira glicina conservada é quase sempre um resíduo aromático.⁷⁴ O padrão consenso de G-patch pode ser encontrado em banco de dados de proteínas através de busca por $hhx(3)Gax(2)GxGhGx(4)G$, onde “h” se refere a um resíduo hidrofóbico volumoso (I, L, V, M), “a” refere-se a um resíduo aromático (F, Y, W) e “x” refere-se a posições ocupadas por resíduos não conservados. A predição de estrutura secundária indica que o domínio G-patch é formado por duas α -hélices com quatro das seis glicinas conservadas localizadas no *loop* de ligação entre as duas hélices.⁷⁴ O domínio G-patch foi encontrado em proteínas de eucariotos que contém outros domínios já descritos de ligação a RNA, como RRM, dsRBD, SWAP, R3H; no entanto, também pode ser encontrado como domínio único. G-patch não foi encontrado em proteínas de Archaea e bactéria.⁷⁴

O fato de ser encontrado junto com outros domínios de ligação a RNA e em proteínas envolvidas no processo *splicing* (como o fator de *splicing* 45, SPF-45) ou transporte de mRNA, levou à hipótese de que o domínio G-patch também estivesse envolvido em interações proteína-RNA. De fato, estudos recentes mostram que o domínio G-patch pode ser um domínio de interação proteína-RNA e proteína-proteína. G-patch foi encontrado em proteinases e transcriptase reversa de β -retrovírus, na proteína repressora de transcrição ZIP (do inglês, *zinc finger and G-patch domain-containing protein*), e está presente em proteínas parceiras de interação de helicases da família DEAH/RHA.⁷³⁻⁷⁴

Ensaio de mobilidade eletroforética demonstraram que o domínio G-patch presente na extremidade C- terminal das proteinases de β -retrovírus, como a do vírus Mason-Pfizers (MPMV), é responsável pela ligação em sequências inespecíficas com ácidos nucleicos de fita única (DNA e RNA). Além disso, G-patch tem papel importante na infecção por esse vírus, porém não demonstrou estar ligado à atividade de protease dessas proteínas.⁷⁵⁻⁷⁷

Estudos com a proteína repressora de transcrição ZIP apontam que 18 resíduos do domínio G-patch fazem parte dos resíduos cruciais para a formação de dímeros, mostrando que o domínio G-patch medeia interações proteína-proteína.^{75,78}

Atualmente, existem seis proteínas G-patch conhecidas que interagem com helicases da família DEAH/RHA: RBM5, Ntr1/TFIP11, Gno1/PinX1, Spp2, Pfa1/Sqs1, e GPATCH2.⁷⁵ Para esse trabalho, o fato mais interessante é a helicase bifuncional Prp43, envolvida no processo de *splicing* e na biogênese de ribossomos, ser recrutada por proteínas G-patch diferentes. No *splicing*, Prp43 é recrutada pelo complexo Ntr1/Ntr2 e ativada por Ntr1 para a desmontagem da estrutura laço-spliceossomo.^{74,78-79} Na biogênese de ribossomos, Prp43 é ativada pelas proteínas Pfa1 e Gno1 (PinH1 em humanos), mas não é recrutada por elas.⁸⁰⁻⁸¹ Estudos mostram que Pfa1 e Gno1 são fundamentais no processamento e maturação de pré-rRNAs.⁸²⁻⁸⁵

Estudos de dicroísmo circular mostram que o domínio G-patch é desestruturado em solução, mas adquire elementos de estrutura secundária na presença de $\geq 20\%$ de trifluoretanol (TFE), suportando a ideia de que ele pode se estruturar durante a interação com outras proteínas e/ou ácido nucléico.⁸⁶ Um modelo de estrutura tridimensional de Prp43 já foi resolvido por cristalografia de raio-X mostrando o domínio OB-*fold* que é responsável pelas interações com RNA.⁸⁶⁻⁸⁷ Ensaio de interação com proteínas truncadas de Prp43 e Pfa1 revelam que a ligação e ativação de Prp43 por Pfa1 se dá através do contato entre o domínio G-patch e o domínio OB-*fold*.⁸⁸ Até o momento, não há modelos de estrutura atômica do domínio G-patch ou do complexo entre proteínas G-patch e alguma helicase da família DEAH/RHA para a elucidação do mecanismo de ativação.

Interessantemente, foi possível detectar no presente trabalho uma proteína ainda não caracterizada de *C. reinhardtii* que possui o domínio G-patch e que potencialmente também pode ter função na biogênese de ribossomo por ter sido encontrada como parceira de interação de CrNIP7.

3.3.5 Conclusão parcial

A proteína G-patch predita no genoma de *C. reinhardtii* foi “pescada” pelo sistema de duplo híbrido como parceira de interação de CrNIP7. Os fatos de o domínio G-patch estar presente em diversas proteínas envolvidas na ligação a ácido nucléico e da helicase Prp43, que participa no processo de *splicing* e biogênese de ribossomos, ser ativada por proteínas que possuem o domínio G-patch, reforçam a potencial interação entre a proteína G-patch,

encontrada nesse estudo, com CrNIP7 no processo de biogênese de ribossomo em *C. reinhardtii*. A CDS *G-patch* amplificada apresentou um tamanho maior do que a anotação conceitual do banco de dados do NCBI, o que permitirá a correção do dado depositado no NCBI. Interessantemente, apenas a sequência codificadora dos primeiros 243 da proteína, contendo o domínio G-patch conservado, estava clonada no vetor de duplo híbrido, mostrando os outros 360 resíduos não são necessários para a interação com CrNIP7. A proteína G-patch completa não foi expressa nos sistemas de co-expressão e de SUMO-fusão. Assim, construções truncadas que incluam apenas o domínio G-patch são uma alternativa promissora para estudos futuros de interação.

3.4 CrSBDS

A interação entre as proteínas CrNIP7 e CrSBDS não foi verificada no experimento inicial de duplo híbrido realizado nesse trabalho, mas era esperada uma vez que as homólogas, NIP7 e SBDS, interagem tanto em leveduras, como em humanos.^{16,21} Frente a esse fato, surgiu a motivação para realizar um experimento de duplo híbrido direcionado para investigar se a parceria entre CrNIP7 e CrSBDS também existia. Para isso, uma busca por homólogas a SBDS foi realizada no banco de dados de proteínas nr (do inglês, *Non-redundant protein sequences*) de *C. reinhardtii* no NCBI. Essa busca resultou numa sequência proteica parcial (NCBI: XP_001689481.1) traduzida de uma CDS incompleta (XM_001689429.1) anotada conceitualmente a partir do genoma de *C. reinhardtii*. Então, à essa sequência foi dada o nome de CrSBDS e os trechos faltantes foram completados por uma análise do banco de dados SRA de *C. reinhardtii*. A sequência de 1140 pares de nucleotídeos de *CrSBDS*, que codificam uma proteína de 379 aminoácidos encontrada nesse trabalho é a mesma que está anotada pelo banco de dados Phytozome 11 (<https://phytozome.jgi.doe.gov/pz/portal.html>) com o nome de Cre01.g017500 (Apêndice A – Figura 46 e 47).

3.4.1 Amplificações e clonagens da CDS *CrSBDS*

O fragmento de 1140 pb referente a CDS de *CrSBDS* foi amplificado por PCR a partir do cDNA de *C. reinhardtii* (Figura 32A) utilizando o par de oligonucleotídeos *CrSBDS-SmaI-NdeI-F* e *CrSBDS-XhoI-R* (Tabela 2) para posterior subclonagem em pETDuet-1 e pET28a (Tabela 11). Para a clonagem em pGAD-HA, uma nova amplificação (Figura 32F) foi realizada com os oligonucleotídeos *CrSBDS-SmaI-C-F* e *CrSBDS-XhoI-R* (Tabela 2). Os produtos de PCR foram clonados no vetor de propagação pUC19 (previamente digerido com *SmaI*) e sequenciados. A CDS de *CrBDS* foi, então, subclonada nos vetores pETDuet-1 e pET28a nos sítios *NdeI* e *XhoI*, e no vetor pGAD-HA nos sítios de *SmaI* e *XhoI* (Figura 32B, C, D, G, H).

Figura 32 – Análise da amplificação por PCR e subclonagens de *CrSBDS*: (A) Produto de PCR (1140 pb) amplificado a partir do cDNA de *C. reinhardtii* utilizando $T_m = 57^\circ\text{C}$; (1) com DMSO, (2) sem DMSO. (B) Análise da clonagem de *CrSBDS* no vetor pUC19: (3), plasmídeo não digerido (4) análise de restrição com *SmaI* e *XhoI*; (C) Análise da clonagem de *CrSBDS* em pET28a; (5) plasmídeo não digerido; (6) análise de restrição com *NdeI* e *XhoI*. (D) Análise da clonagem de *CrSBDS* em pETDuet-1-CrNip7; (7) plasmídeo não digerido; (8) análise de restrição com *NdeI* e *XhoI*. (E) “M” = Marcador molecular GeneRuler DNA 1kb (Fermentas); (F) Análise do produto de PCR para troca de sítios de restrição a partir de um clone positivo, $T_m = 55^\circ\text{C}$. (G) Clones positivos de *CrSBDS* em pUC19; (10) plasmídeo não digerido; (11) análise de restrição com *SmaI* e *XhoI*, mostrando plasmídeo não digerido (ND), plasmídeo sem o inserto (pUC19) e inserto (*CrSBDS*). (H) Análise da clonagem de *CrSBDS* em pGAD-HA: (12) plasmídeo não digerido; (13) Análise de restrição com *SmaI* e *XhoI*. Eletroforese em gel de agarose 0,8% em TAE 1X.

Fonte: Elaborada pela autora.

3.4.2 Duplo Híbrido direcionado para análise da interação CrNIP7 e CrSBDS

A interação entre as proteínas CrNIP7 e CrSBDS foi demonstrada pelo ensaio de duplo híbrido direcionado (item 2.5), sendo os testes de ativação da expressão dos genes repórteres *LacZ* e *HIS3* positivos (Figura 33). Porém, a levedura L40 transformada com os plasmídeos pLexA-N- *CrNip7* e pGAD-HA-*CrSBDS* demonstrou ser sensível à adição do composto 3-AT no meio SD-WLH. Em vista disso, a interação entre CrNIP7 e CrSBDS não é considerada forte. Assim, pode-se concluir que a interação entre as homólogas NIP7/SBDS, Nip7p/Sdo1p, e CrNIP7/CrSBDS é conservada entre humanos, leveduras e, também, na alga unicelular, *C. reinhardtii*.

Figura 33 – Teste de interação *in vivo* utilizando duplo híbrido dirigido para verificar a interação das proteínas BD-CrNIP7 e AD-CrSBDS. (A) Ensaio da β -galactosidase em papel de filtro mostrando a expressão do gene repórter *LacZ*. (B) Teste de força de interação por gradiente de 3AT, mostrando que o clone expressão BD-CrNIP7 e AD-CrSBDS é capaz de crescer em meio sem histina, porém é sensível à dição de 3AT no meio de cultura.

Fonte: Elaborada pela autora.

3.4.3 Expressão heteróloga e purificação de 6xHis-CrSBDS

CrSBDS é uma proteína de 379 aminoácidos, com massa molecular predita de 41079,8 Da e ponto isoelétrico teórico igual a 7,72 calculados pelo programa Protparam. A proteína 6xHis-CrSBDS (~42,5 kDa) foi altamente expressa no sistema de expressão heterólogo em bactéria, utilizando o vetor pET28a. A Figura 34 mostra os resultados da expressão e purificação em coluna de afinidade do extrato proteico proveniente de uma cultura de 250 mL. Nota-se que a proteína 6xHis-CrSBDS migra com aproximadamente 45 kDa e o rendimento estimado foi de aproximadamente 50 mg de proteína por litro de cultura em meio LB. Esse rendimento pode estar subestimado, uma vez CrSBDS não possui triptofano e a medida da concentração foi realizada por método de absorbância em 280 nm. Além disso, o experimento não foi realizado repetidas vezes para precisão do rendimento.

Pode-se observar que CrSBDS concentra-se majoritariamente nas frações solúveis, embora seja eluída da cromatografia de afinidade ainda com alguns contaminantes. O alto rendimento obtido para a CrSBDS é um fator favorável para a realização de experimentos futuros de caracterização estrutural.

Figura 34 – Análise por SDS-PAGE 15% da expressão e purificação da proteína 6xHis-CrSBDS.

Fonte: Elaborada pela autora.

3.4.4 Validação da interação por *Pull Down*

Após evidência positiva da interação *in vivo* das proteínas CrNIP7 e CrSBDS, buscou-se confirmar a mesma interação *in vitro*. Para isso, as proteínas 6xHis-CrNIP7 e CrSBDS foram co-expressas pela indução do vetor pETDuet-1 em *E. coli* Rosetta (DE3). A Figura 35 mostra que 6xHis-CrNIP7 e CrSBDS foram expressas na fração solúvel e em proporções parecidas (intensidade da banda). Desse modo, o extrato solúvel da lise celular foi incubado com resina de afinidade, na tentativa de co-purificar as proteínas. Após as lavagens, observou-se que as duas proteínas co-eluíram quando o competidor imidazol foi adicionado. As bandas dos tamanhos referentes à 6xHis-CrNIP7 e CrSBDS foram identificadas e confirmadas por espectrometria de massas. O fato das proteínas co-eluírem junto com muitos outros contaminantes trouxe a dúvida se a interação entre CrNIP7 e CrSBDS era direta ou não. Assim, um novo experimento de *pull down* (item 2.6.5), desta vez colocando o *His-tag* na proteína CrSBDS, foi realizado. A proteína 6His-CrSBDS, expressa pelo vetor pET28a, foi imobilizada na resina de níquel e incubada com CrNIP7 sem *tag* (proveniente da clivagem da fusão SUMO-CrNIP7). Apesar deste experimento não ter apresentado equimolaridade das proteínas co-eluídas em função da diferença na quantidade de cada proteína presente na mistura (a banda imobilizada de CrSBDS está mais forte do que a de CrNIP7), a interação *in vitro* foi confirmada (Figura 36). Conjuntamente, os experimentos mostram que independentemente de qual das duas proteínas possua o *His-tag*, elas são co-purificadas na resina de afinidade. No entanto, os resultados desses experimentos não descartam a possibilidade de que a interação entre CrNIP7 e CrSBDS seja estabilizada de alguma maneira por RNA, como relatado para o par Nip7p/Sdo1p em leveduras.¹⁶

Figura 35 – Análise mostrando a co-expressão e co-purificação de CrSBDS e 6xHis-CrNIP7 por indução do vetor pETDuet-1 em *E. coli* Rosetta (DE3). SDS-PAGE Tricina 16%

Fonte: Elaborada pela autora.

Figura 36 – Análise do experimento de *pull down* das proteínas 6xHis-CrSBDS com CrNIP7. SDS-PAGE Tricina 16%.

Fonte: Elaborada pela autora.

3.4.5 Análise de bioinformática

O alinhamento múltiplo (Apêndice A - 48), realizado com o programa Clustal O, entre as sequências de SBDS dos organismos *Archaeoglobus fulgidus* (WP_010877998.1), *Pyrococcus abyssi* (WP_010867732.1), *Trypanosoma cruzi* (ACD88761.1), *Saccharomyces cerevisiae* (NP_013122.1), *C. reinhardtii* (Cre01.g017500), *Arabidopsis thaliana* (NP_564488.1), *Homo sapiens* (NP_057122.2) e *Caenorhabditis elegans* (NP_497226.2.), mostra a conservação

sequencial entre elas. A matriz de identidade (Apêndice A – Tabela 16) revela a alta identidade de CrSBDS com suas homólogas de *S. cerevisiae* (38,21%), *H. sapiens* (48,40%), *C. elegans* (46,64%) e *A. thaliana* (40,29%).

Apesar das homólogas de SBDS apresentarem em média 250 aminoácidos, várias espécies possuem a extremidade C-terminal alongada. CrSBDS também apresenta 129 aminoácidos a mais na região C-terminal que a média. Nas homólogas de plantas com flores (Magnoliophyta) é possível encontrar um domínio putativo *zinc finger* C2H2, como o encontrado na espécie *A. thaliana*.⁸⁹ O alinhamento da Figura 37 mostra que CrSBDS é muito conservada em relação à sua homóloga de *A. thaliana*, inclusive na extensão C-terminal, porém CrSBDS não apresenta o domínio *zinc finger* C2H2. Alguns tripanossomatídeos, incluindo duas espécies de *Leishmania* e três espécies de tripanossoma, também possuem extensões C-terminais. Essas extensões variam de 199 a 269 aminoácidos ricos em glicina e aminoácidos carregados.⁸⁹ Estudos da SBDS de *T. cruzi* (TcSBDS) mostram que a extensão adicional é uma região intrinsecamente desordenada e que interage com RNA.⁹⁰

Assim, a presença dessa extensão C-terminal da proteína SBDS em várias espécies pode indicar uma função adicional da SBDS nesses organismos, inclusive em *C. reinhardtii*.

```

C. reinhardtii 1 MSKGVFPVPGQKRLTNIAVRLKKNRGERFEIACYKNKVGDWRSKIEKDLDEVLTQTTTIFH 60
A.thaliana    1 MSKTLVQPVGQKRLTNVAVRLKKQGNRFEIACYKNKVLWSRSGVEKDIDEVLTQSHSTVYS 60
               *** :.*****:*****:*****:*****:*****:*****:*****:
C. reinhardtii 61 NVGKGVVAKDKELQAAFGTTDQHAICLEILAKGELQVGERERKQEYEHFKDVAGVLEK 120
A.thaliana    61 NVSKGVLAKSKDLMKSFSGDDHTKICIDILEKGELQVAGKERESQFSSQFRDIATIVMQK 120
               **.***.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*
C. reinhardtii 121 CVNPNTNRPYTLTMLERALRDIHFNLDPKKSAKQQAALPLQLKEFPFERARMRLRLTV 180
A.thaliana    121 TINPETQRPYTISMVERLMHEIHFAVDPHSNSKKQALDVIRELQKHFPKRSPMRLRLTV 180
               :*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
C. reinhardtii 181 PVEAQQLEQLLQREAADVAGLDVAAGMCVVVAQVEPGVFRDLNFMQTARGAGRLEV 240
A.thaliana    181 PVQNFPSLLEKLKEWDGSVVSKDE-SGTQMSTVCEMEPGLFRECDSHVRS---IQGRLEI 236
               **:*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*
C. reinhardtii 241 LSLAVMAEGATADEF----ANMGAAATRPVAAAQPEASSSGHVEHRPVGVR-----VG 290
A.thaliana    237 LAVSVHAEGDTSMDHYDEHDDMALQTHKPLLPAETETKD---LTPVVELSKKLQKQEIS 293
               *:::* **.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*
C. reinhardtii 291 ERPTVAAPSGGGSGGAVVYARGPVSGLPPEEHASRRERFAELDTLQPGWTVELR-TKGDV 349
A.thaliana    294 TTDNIKQEGGEEKGKTKCSTCNTFVGEAK-----QY--REHCKSDWHKHNLRKTRKI 344
               .:.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*
C. reinhardtii 350 -----EAVFFHPSGERVGAFANARRMALQASKAAA 379
A.thaliana    345 PPISADECMSEIDMDDSRADLKDYSF----- 370
               . : :. * : : :

```

Figura 37 – Alinhamento gerado pelo programa Clustal O das sequências proteicas de SBDS de *C. reinhardtii* (Cre01.g017500), *Arabidopsis thaliana* (NP_564488.1). Os aminoácidos conservados estão representados por asteriscos, os aminoácidos com alta similaridade com dois pontos e aminoácidos com baixa similaridade com um ponto. Os aminoácidos sombreados em cinza fazem parte do domínio *zinc finger* C2H2.

Fonte: Elaborada pela autora.

A análise de domínios conservados pelo Pfam encontrou dois domínios com as referências PF01172 e PF09377, que correspondem aos 3 domínios típicos das SBDSs: N-terminal, central e C-terminal, mas não encontrou nenhum outro domínio conservado na porção C-terminal estendida (Figura 38, linha 3). Uma comparação entre a estrutura resolvida de SBDS humana (HsSBDS) (PDB: 2KDO)¹⁸ por NMR e a predição de estrutura secundária de CrSBDS pelo programa PSIPRED (Figura 38, linha 1,2,3 e 4) mostra a coincidência dos elementos preditos para CrSBDS em relação às estruturas de SBDS humana, mas com diferenças no tamanho de alguns elementos e a ausência de $\beta 4$ na predição de CrSBDS. A extremidade C-terminal estendida de CrSBDS, a partir do aminoácido 250, é predita por ter alguns elementos de estrutura secundária, mas também apresenta uma grande extensão que compreende 64 aminoácidos (entre os aminoácidos 257 e 321) predita como *coil* com regiões IDR (Figura, linha 4,5 e 6). Além disso, uma análise de possíveis sítios de ligação em proteínas desordenadas, pelo programa ANCHOR, resultou em quatro regiões possíveis de ligação, sendo três delas incluídas na extremidade C-terminal estendida de CrSBDS e um possível sítio de ligação dentro de uma fita beta predita por participar do domínio C-terminal (Figura 38, linha 7). Os resultados das predições reforçam a hipótese de que a cauda C-terminal alongada de CrSBDS tenha, provavelmente, uma função adicional, que pode ou não estar relacionada à biogênese de ribossomos em *C. reinhardtii*. Dessa maneira, os estudos funcionais e de caracterização estrutural de CrSBDS se tornam muito interessantes.

3.4.6 Conclusão parcial

A CDS correta da proteína CrSBDS foi amplificada com sucesso a partir do cDNA de *C. reinhardtii*.

A interação entre as proteínas CrNIP7 e CrSBDS foi avaliada e confirmada *in vivo* e *in vitro*, por experimentos de duplo híbrido e *pull down*, revelando a conservação evolucionária desse par de interação em *C. reinhardtii*. A presença de RNA mediando essa interação ainda precisa ser investigada. A expressão heteróloga de CrSBDS mostrou que ela é muito bem expressa e solúvel em *E. coli*, o que permitirá estudos de caracterização estrutural e outros experimentos futuros.

Figura 38 – Comparação entre a estrutura secundária resolvida de HsSBDS (PDB:2KDO) (linha 1 e 2) e a predição feita pelo programa PSIPRED para CrSBDS (linha 3 e 4, os aminoácidos realçados fazem parte do domínio SBDS conservado encontrado pelo Pfam). As fitas- β estão representadas por setas, as α -hélices por cilindros e coils por linha. Estruturas em vermelho fazem parte do domínio N-terminal, em amarelo, do domínio central; em azul, do domínio C-terminal e em verde, são elementos preditos para a extensão C-terminal de CrSBDS. As linhas 5 e 6 correspondem às predições de IDRs na sequência de CrSBDS pelos programas Globplot e DisEMBL, e a linha 7 corresponde aos possíveis sítios de ligação preditos (em laranja) pelo programa ANCHOR.

Fonte: Elaborada pela autora

3.5 Discussão final

Este trabalho é pioneiro no estudo de fatores não ribossomais envolvidos na biogênese de ribossomos na alga unicelular *C. reinhardtii*. A complementação de função da proteína CrNIP7 em levedura foi o primeiro indício de que CrNIP7 está envolvida na maturação da subunidade ribossomal 60S. A interação entre CrNIP7 e CrSBDS ainda reforça a participação dessas duas proteínas na biogênese de ribossomo, uma vez que esse par de interação é conservado em leveduras (Sdo1/Nip7p) e humanos (HsNIP7/HsSBDS). Dessa maneira, podemos sugerir que as novas parceiras de interação de CrNIP7, Predicted e G-patch, encontradas por duplo-híbrido, ampliam a lista de proteínas que podem ter um papel atuante na biogênese de ribossomos.

Em um trabalho recente, Zones e colaboradores⁶² disponibilizaram os dados do transcriptoma de células de *Chlamydomonas reinhardtii* crescidas sob sincronização do ciclo de vida. A partir desses dados, foi possível avaliar o perfil de expressão das proteínas envolvidas nessa dissertação na tentativa de estabelecer uma relação entre elas. O gráfico da Figura 39 traz esses dados mostrando que, exceto pelos picos a 0 hora de CrNIP7 e CrSBDS, de forma geral todas as proteínas têm um perfil de expressão aumentado no período de escuro (das 12h às 24h), que é coincidente com o período em que ocorre de divisão celular (aproximadamente das 13h às 17h). O trabalho de Zones também mostra que apesar do nível de expressão das proteínas ribossomais citosólicas, em geral, não cair abaixo da metade do seu valor de pico durante o ciclo celular, elas mostram um aumento significativo de expressão durante o período de escuro com um pico de expressão em ZT15 (zona temporal 15h), quando as células estão saindo da fase S/M do ciclo celular dando origem às células filhas.⁶² Sendo assim, pode-se observar que a expressão de CrNIP7, CrSBDS, Predicted e G-patch tem correspondência temporal com a produção de proteínas ribossomais citosólicas e montagem do ribossomo citosólico.

Outro ponto interessante a considerar, é que apesar da proteína G-patch ter sido encontrada como parceira de interação de CrNIP7, podemos especular que G-patch possa interagir com CrSBDS também. Essa especulação vem do fato que Sdo1 interage com a proteína helicase bifuncional Prp43, envolvida no processo de *splicing* e síntese de ribossomos, e com a proteína Spp382p/Ntr1p, que possui um domínio G-patch.¹⁶ Sabe-se também que a helicase

Prp43 interage com proteínas G-patch diferentes nos dois processos em que participa. No *splicing*, Prp43 é recrutada pelo complexo Ntr1/Ntr2 e ativada por Ntr1 para a desmontagem da estrutura laço-spliceossomo.^{75,79,80} Portanto, em estudos futuros, seria interessante verificar a possível interação entre CrSBDS e a proteína G-patch de *C. reinhardtii*, estendendo a rede de conexões entre os fatores não ribossomais requeridos para a biogênese de ribossomos.

Para a proteína Predicted são necessários estudos mais aprofundados de caracterização funcional devido ao fato curioso de aparentemente estar presente apenas no grupo restrito das algas Volvocales e não possuir similaridade com nenhuma outra proteína ou domínio já caracterizados.

Figura 39 – (A) Representação esquemática do experimento de sincronização de *C. reinhardtii*. O diagrama mostra a zona temporal (ZT) dos períodos de luz (ZT0 a ZT12) e escuro (ZT12 a ZT24/0) de uma cultura de células sincronizadas com a fase de crescimento (G1), divisão (S/M) e descanso (G0) demarcadas pelas setas em cinza. As amostras da cultura foram coletadas a cada uma hora com amostras adicionais a cada meia hora durante a divisão celular (linhas em marrom). Esquema reproduzido de Zones et. al.⁶² (B) Nível de expressão de CrNIP7, CrSBDS, G-patch e Predicted ao longo do ciclo celular de *C. reinhardtii*. RRKM, unidade de expressão gênica utilizada em Zones et. al, do inglês reads per kilobase per million mapped.

Fonte: Elaborada pela autora.

CONCLUSÕES E PERSPECTIVAS

4 CONCLUSÕES E PERSPECTIVAS

Conclusões:

- ✓ A proteína CrNIP7, de *Chlamydomonas reinhardtii*, complementou a função da proteína Nip7p de *Saccharomyces cerevisiae*, sugerindo uma participação na biogênese da subunidade 60S do ribossomo.
- ✓ As proteínas G-patch e Predicted foram encontradas como potenciais parceiras de interação de CrNIP7 pelo rastreamento da biblioteca de cDNA de *C. reinhardtii*.
- ✓ O teste preliminar de interação entre CrNIP7 e Predicted *in vitro* sugeriu que a interação ocorra de fato, mas esse resultado ainda precisa ser confirmado.
- ✓ A proteína G-patch não foi expressa nos sistemas de expressão bacterianos utilizados nesse trabalho, impossibilitando a realização do teste de interação *in vitro* com CrNIP7.
- ✓ A interação entre CrNIP7 e CrSBDS foi verificada, *in vivo*, por duplo híbrido dirigido e validada, *in vitro*, por *pull down*, seguida da confirmação da identidade proteica por espectrometria de massas.
- ✓ O isolamento dos cDNAs das proteínas permitiu a obtenção das sequências codificantes completas, validando as predições genômicas e possibilitando a correção dos depósitos no banco de dados.

Perspectivas:

- ✓ Otimizar a purificação de Predicted para novos ensaios de *pull down* com CrNIP7.
- ✓ Realizar testes de expressão de construções truncadas de G-patch para ensaios de validação da interação *in vitro* com CrNIP7.
- ✓ Verificar a interação por duplo híbrido dirigido entre CrSBDS e construções truncadas de G-patch;
- ✓ Realizar estudos estruturais das proteínas individuais e de suas interações por técnicas biofísicas.

REFERÊNCIAS

REFERÊNCIAS

- 1 KRESSLER, D.; HURT, E.; BAßLER, J. Driving ribosome assembly. **Biochimica et Biophysica Acta - molecular cell research**, v. 1803, n. 6, p. 673–683, 2010.
- 2 RODNINA, M. V.; WINTERMEYER, W. Recent mechanistic insights into eukaryotic ribosomes. **Current Opinion in Cell Biology**, v. 21, n. 3, p. 435–443, 2009.
- 3 STEITZ, T. A. A structural understanding of the dynamic ribosome machine. **Nature Reviews - molecular cell biology**, v. 9, n. 3, p. 242–253, 2008.
- 4 RNA Structure and Functions. 2013. Disponível em: <<http://upendrats.blogspot.com.br/2013/08/rna-structure-and-functions.html>>. Acesso em: 10/08/2016.
- 5 FROMONT-RACINE, M. et al. Ribosome assembly in eukaryotes. **Gene**, v. 313, n. 1-2, p. 17–42, 2003.
- 6 ZANCHIN, N. I. et al. *Saccharomyces cerevisiae* Nip7p is required for efficient 60S ribosome subunit biogenesis. **Molecular and cellular biology**, v. 17, n. 9, p. 5001–5015, 1997.
- 7 MORELLO, L. G. et al. The NIP7 protein is required for accurate pre-rRNA processing in human cells. **Nucleic Acids Research**, v. 39, n. 2, p. 648–665, 2011.
- 8 HENRAS, A. K. et al. The post-transcriptional steps of eukaryotic ribosome biogenesis. **Cellular and Molecular Life Sciences**, v. 65, n. 15, p. 2334–2359, 2008.
- 9 CHEN, W. et al. Enp1, a yeast protein associated with U3 and U14 snoRNAs, is required for pre-rRNA processing and 40S subunit synthesis. **Nucleic Acids Research**, v. 31, n. 2, p. 690–699, 2003.
- 10 GELPERIN, D. et al. Bms1p, a novel GTP-binding protein, and the related Tsr1p are required for distinct steps of 40S ribosome biogenesis in yeast. **RNA**, v. 7, n. 9, p. 1268–1283, 2001.
- 11 CARRON, C. et al. Analysis of two human pre-ribosomal factors, bystin and hTsr1, highlights differences in evolution of ribosome biogenesis between yeast and mammals. **Nucleic Acids Research**, v. 39, n. 1, p. 280–291, 2011.
- 12 TALKISH, J. et al. Hierarchical recruitment into nascent ribosomes of assembly factors required for 27SB pre-rRNA processing in *Saccharomyces cerevisiae*. **Nucleic Acids Research**, v. 40, n. 17, p. 8646–8661, 2012.
- 13 COLTRI, P. P. et al. Expression, crystallization and preliminary X-ray analysis of the *Pyrococcus abyssi* protein homologue of *Saccharomyces cerevisiae* Nip7p. **Acta Crystallographica D - biological crystallography**, v. 60, n. 10, p. 1925–1928, 2004.
- 14 ZANCHIN, N. I.; GOLDFARB, D. S. Nip7p interacts with Nop8p, an essential nucleolar protein required for 60S ribosome biogenesis, and the exosome subunit Rrp43p. **Molecular and Cellular Biology**, v., n. 2, p. 1518–1525, 1999.
- 15 GRANATO, D. C. et al. Nop53p, an essential nucleolar protein that interacts with Nop17p and Nip7p, is required for pre-rRNA processing in *Saccharomyces cerevisiae*. **FEBS Journal**, v. 272, n. 17, p. 4450–4463, 2005.

- 16 LUZ, J. S. et al. Sdo1p , the yeast orthologue of Shwachman–Bodian–Diamond syndrome protein , binds RNA and interacts with nuclear rRNA-processing factors. **Yeast**, v. 26, p. 287–298, 2009. doi: 10.1002/yea.1668.
- 17 SEKIGUCHI, T. et al. A novel human nucleolar protein, Nop132, binds to the G proteins, RRAG A/C/D. **Journal of Biological Chemistry**, v. 279, n. 9, p. 8343–8350, 2004.
- 18 OLIVEIRA, J. F. et al. Structure, dynamics, and RNA interaction analysis of the human SBDS protein. **Journal of Molecular Biology**, v. 396, n. 4, p. 1053–1069, 2010.
- 19 MORELLO, L. G. et al. The human nucleolar protein FTSJ3 associates with NIP7 and functions in pre-rRNA processing. **PloS One**, v. 6, n. 12, p. e29174, 2011.
- 20 COLTRI, P. P. et al. Structural insights into the interaction of the Nip7 PUA domain with polyuridine RNA. **Biochemistry**, v. 46, n. 49, p. 14177–14187, 2007.
- 21 HESLING, C. et al. The Shwachman-Bodian-Diamond syndrome associated protein interacts with HsNip7 and its down-regulation affects gene expression at the transcriptional and translational levels. **Experimental Cell Research**, v. 313, n. 20, p. 4180–4195, 2007.
- 22 BOOCCOCK, G. R. B. et al. Mutations in SBDS are associated with Shwachman-Diamond syndrome. **Nature Genetics**, v. 33, n. 1, p. 97–101, 2003.
- 23 MENNE, T. F. et al. The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. **Nature Genetics**, v. 39, n. 4, p. 486–495, 2007.
- 24 FINCH, A. J. et al. Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes shwachman-diamond syndrome. **Genes and Development**, v. 25, n. 9, p. 917–929, 2011.
- 25 LIU, J. F. et al. Crystal structure of KD93, a novel protein expressed in human hematopoietic stem/progenitor cells. **Journal of Structural Biology**, v. 148 , n. 3, p. 370–374, 2004.
- 26 LUZ, J. S. et al. Identification of archaeal proteins that affect the exosome function in vitro. **BMC Biochemistry**, v. 11, p. 22, 2010. doi: 10.1186/1471-2091-11-22.
- 27 HARRIS, E. H. As a model organism. **Molecular Biology**, v. 52, n. 1, p. 363–406, 2001.
- 28 OLSON, B. J. S. C.; UMEN, J. G. Genomics of volvocine algae. **Advances in Botanical Research**, v. 64, p. 185-243, 2012.
- 29 MERCHANT, S. S. et al. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. **Science**, v. 318, n. 5848, p. 245–250, 2007.
- 30 HOLLOWAY, S. P.; HERRIN, D. L. Processing of a composite large subunit Rrna: studies with *Chlamydomonas* mutants deficient in maturation of the 23s-like rRNA. **Plant Cell**, v. 10, n. 7, p. 1193–1206, 1998.
- 31 HARRIS, E. H.; BOYNTON, J. E.; GILLHAM, N. W. Chloroplast Ribosome Biogenesis in *Chlamydomonas*. **Journal of Cell Biology**, v. 63, n. 1, p. 160–179, 1974.

- 32 MANUELL, A. L. et al. Composition and structure of the 80 S ribosome from the green alga *Chlamydomonas reinhardtii*: 80 S ribosomes are conserved in plants and animals. **Journal of Molecular Biology**, v. 351, n. 2, p. 266–279, 2005.
- 33 FLEMING, G. H.; BOYNTON, J. E.; GILLHAM, N. W. Cytoplasmic ribosomal proteins from *Chlamydomonas reinhardtii*: characterization and immunological comparisons. **Molecular Genetics and Genomics**, v. 206, n. 2 p. 226–37, 1987.
- 34 YAMAGUCHI, K. et al. Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome: identification of proteins unique to the 70 S ribosome. **Journal of Biological Chemistry**, v. 278, n. 36, p. 33774–33785, 2003.
- 35 YAMAGUCHI, K. et al. Proteomic characterization of the small subunit of *Chlamydomonas reinhardtii* chloroplast ribosome: identification of a novel S1 domain-containing protein and unusually large orthologs of bacterial S2, S3, and S5. **Plant Cell Online**, v. 14, n. 11, p. 2957–2974, 2002.
- 36 SAGER, R.; HAMILTON, M. G. Cytoplasmic and chloroplast ribosomes of *Chlamydomonas*: ultracentrifugal characterization. **Science**, v. 157, n. 3789, p. 709–711, 1967.
- 37 MANUELL, A. L.; QUISPE, J.; MAYFIELD, S. P.. Structure of the chloroplast ribosome: novel domains for translation regulation. **PLoS Biology**, v. 5, n. 8, p. 1785–1797, 2007.
- 38 WAGNER, V.; BOESGER, J.; MITTAG, M. Sub-proteome analysis in the green flagellate alga *Chlamydomonas reinhardtii*. **Journal of Basic Microbiology**, v. 49, n. 1, p. 32–41, 2009.
- 39 BOER, P. H.; GRAY, M. W. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. **Cell**, v. 55, n. 3, p. 399–411, 1988.
- 40 DENO VAN-WRIGHT, E. M.; LEE, R. W. Evidence that the fragmented ribosomal RNAs of *Chlamydomonas* mitochondria are associated with ribosomes. **FEBS Letters**, v. 370, n. 3, p. 222–226, 1995.
- 41 SCAIFE, M. A. et al. Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host. **Plant Journal**, v. 82, n. 3, p. 532–546, 2015.
- 42 PURTON, S. et al. Genetic engineering of algal chloroplasts: progress and prospects. **Russian Journal of Plant Physiology**, v. 60, n. 4, p. 491–499, 2013.
- 43 SAMBROOK, J.; RUSSEL, D. W. **Molecular Cloning: a laboratory manual**. New York: Cold Spring Harbor Laboratory Press, 2001.
- 44 AGATEP, R. et al. Transformation of *Saccharomyces cerevisiae* by the lithium acetate/single-stranded carrier DNA/polyethylene glycol protocol. **Technical Tips Online**, v. 3, n. 1, p. 133–137, 1998.
- 45 FIELDS, S.; SONG, O. A novel genetic system to detect protein-protein interactions. **Nature**, v. 340, n. 6230, p. 245–246, 1989.
- 46 FASHENA, S. J.; SEREBRIISKII, I.; GOLEMIS, E. A.. The continued evolution of two-hybrid screening approaches in yeast: How to outwit different preys with different baits. **Gene**, v. 250, n. 1-2, p. 1–14, 2000.

- 47 VIDAL, M.; FIELDS, S. The yeast two-hybrid assay: still finding connections after 25 years. **Nature Methods**, v. 11, n. 12, p. 1203–1206, 2015.
- 48 LING, M.; MERANTE, F.; ROBINSON, B. H.. A rapid and reliable DNA preparation method for screening a large number of yeast clones by polymerase chain reaction. **Nucleic Acids Research**, v. 23, n. 23, p. 4924–4925, 1995.
- 49 ALTSCHUL, S. F. et al. Basic Local Alignment Search Tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990.
- 50 ARTIMO, P. et al. ExPASy : SIB bioinformatics resource portal. **Nucleic Acids Research**, v. 40, p. 597–603, 2012. doi: 10.1093/nar/gks400.
- 51 GASTEIGER, E. et al. Protein identification and analysis tools on the ExPASy Server. **Proteomics Protocols Handbook**, p. 571–607, 2005. doi: 10.1385/1-59259-890-0:571.
- 52 FINN, R. D. et al. The Pfam protein families database: towards a more sustainable future. **Nucleic Acids Research**, v. 44, p. 279–285, 2015. doi: 10.1093/nar/gkv1344.
- 53 COMBET, C. et al. NPS@: network protein sequence analysis. **Trends in Biochemical Sciences**, v. 25, n. 3, p. 147–150, 2000.
- 54 JONES, D. T. Protein secondary structure prediction based on position-specific scoring matrices. **Journal of Molecular Biology**, v. 292, n. 2, p. 195–202, 1999.
- 55 BUCHAN, D. W. A. et al. Scalable web services for the PSIPRED protein analysis workbench. **Nucleic Acids Research**, v. 41, p. 349–357, 2013. doi: 10.1093/nar/gkt381.
- 56 SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, v. 7, n. 539, p. 1–6, 2011.
- 57 LINDING, R. et al. GlobPlot: exploring protein sequences for globularity and disorder. **Nucleic Acids Research**, v. 31, n. 13, p. 3701–3708, 2003.
- 58 LINDING, R. et al. Protein disorder prediction: implications for structural proteomics. **Structure**, v. 11, n. 11, p. 1453–1459, 2003.
- 59 DOSZTÁNYI, Z.; MÉSZÁROS, B.; SIMON, I. ANCHOR: Web server for predicting protein binding regions in disordered proteins. **Bioinformatics**, v. 25, n. 20, p. 2745–2746, 2009.
- 60 MÉSZÁROS, B.; SIMON, I.; DOSZTÁNYI, Z. Prediction of Protein Binding Regions in Disordered Proteins. **PLoS Computational Biology**, v. 5, n. 5, p. e1000376, 2009.
- 61 SHEVCHENKO, A. et al. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. **Analytical Chemistry**, v. 68, n. 5, p. 850–858, 1996.
- 62 ZONES, J. M. et al. High-resolution profiling of a synchronized diurnal transcriptome from *chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. **Plant Cell**, v.27, n. 10, p. 2743–2769, 2015.
- 63 GU, Z. et al. *NIP1*, a gene required for nuclear transport in yeast. **Proceedings of the National Academy of Sciences of the United States of America**, v. 89, n. 21, p. 10355–10359, 1992.

- 64 SEREBRIISKII, I. et al. Approaches to detecting false positives in yeast two-hybrid systems. **BioTechniques**, v. 28, n. 2, p. 328–336, 2000.
- 65 ARAVIND, L.; KOONIN, E. V. Novel predicted RNA-binding domains associated with the translation machinery. **Journal of Molecular Evolution**, v. 48, n. 3, p. 291–302, 1999.
- 66 PÉREZ-ARELLANO, I.; GALLEGÓ, J.; CERVERA, J. The PUA domain - a structural and functional overview. **FEBS Journal**, v. 274, n. 19, p. 4972–4984, 2007.
- 67 LI, L.; YE, K. Crystal structure of an H/ACA box ribonucleoprotein particle. **Nature**, v. 443, n. 7109, p. 302–307, 2006.
- 68 MALAKHOV, M. P. et al. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. **Journal of Structural and Functional Genomics**, v. 5, n. 1-2, p. 75–86, 2004.
- 69 UVERSKY, V. N. Intrinsically disordered proteins from A to Z. **International Journal of Biochemistry and Cell Biology**, v. 43, n. 8, p. 1090–1103, 2011.
- 70 CAMPEN, A. et al. TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. **Protein and Peptide Letters**, v. 15, n. 9, p. 956–963, 2008.
- 71 WRIGHT, P. E.; DYSON, H. J. Intrinsically disordered proteins in cellular signalling and regulation. **Nature Reviews Molecular Cell Biology**, v. 16, n. 1, p. 18–29, 2014.
- 72 WRIGHT, P. E.; DYSON, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. **Journal of Molecular Biology**, v. 293, n. 2, p. 321–331, 1999.
- 73 HSU, W. L. et al. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. **Protein Science**, v. 22, n. 3, p. 258–273, 2013.
- 74 ARAVIND, L.; KOONIN, E. V. G-patch: A new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. **Trends in Biochemical Sciences**, v. 24, n. 9, p. 342–344, 1999.
- 75 ROBERT-PAGANIN, J.; RÉTY, S.; LEULLIOT, N. Regulation of DEAH/RHA helicases by G-patch proteins. **BioMed Research International**, v. 2015, p. 931857, 2015. doi: 10.1155/2015/931857.
- 76 ŠVEC, M. et al. Proteinases of betaretroviruses bind single-stranded nucleic acids through a novel interaction module, the G-patch. **FEBS Letters**, v. 576, n. 1-2, p. 271–276, 2004.
- 77 BAUEROVÁ-ZABRANSKA, H. et al. The RNA binding G-patch domain in retroviral protease is important for infectivity and D-type morphogenesis of Mason-Pfizer monkey virus. **Journal of Biological Chemistry**, v. 280, n. 51, p. 42106–42112, 2005.
- 78 GUI, B. et al. Dimerization of ZIP promotes its transcriptional repressive function and biological activity. **International Journal of Biochemistry and Cell Biology**, v. 44, n. 6, p. 886–895, 2012.
- 79 TANAKA, N.; ARONOVA, A.; SCHWER, B. Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. **Genes and Development**, v. 21, n. 18, p. 2312–2325, 2007.

- 80 TSAI, R.-T. et al. Dynamic interactions of Ntr1-Ntr2 with Prp43 and with U5 govern the recruitment of Prp43 to mediate spliceosome disassembly. **Molecular and Cellular Biology**, v. 27, n. 23, p. 8027–8037, 2007.
- 81 LEBARON, S. et al. The splicing ATPase prp43p is a component of multiple preribosomal particles. **Molecular and Cellular Biology**, v. 25, n. 21, p. 9269–82, 2005.
- 82 CHEN, Y. L. et al. The telomerase inhibitor Gno1p/PINX1 activates the helicase Prp43p during ribosome biogenesis. **Nucleic Acids Research**, v. 42, n. 11, p. 7330–7345, 2014.
- 83 GUGLIELMI, B.; WERNER, M. The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition. **Journal of Biological Chemistry**, v. 277, n. 38, p. 35712–35719, 2002.
- 84 LEBARON, S. et al. The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. **EMBO Journal**, v. 28, n. 24, p. 3808–3819, 2009.
- 85 PERTSCHY, B. et al. RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. **Journal of Biological Chemistry**, v. 284, n. 50, p. 35079–35091, 2009.
- 86 CHRISTIAN, H. et al. Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. **Nucleic Acids Research**, v. 42, n. 2, p. 1162–1179, 2014.
- 87 HE, Y.; ANDERSEN, G. R. NIELSEN, K. H.. Structural basis for the function of DEAH helicases. **EMBO Reports**, v. 11, n. 3, p. 180–186, 2010.
- 88 WALBOTT, H. et al. Prp43p contains a processive helicase structural architecture with a specific regulatory domain. **EMBO Journal**, v. 29, n. 13, p. 2194–2204, 2010.
- 89 BOOCOCK, G. R. B.; MARIT, M. R.; ROMMENS, J. M. Phylogeny, sequence conservation, and functional complementation of the SBDS protein family. **Genomics**, v. 87, n. 6, p. 758–771, 2006.
- 90 OLIVEIRA, J. F. et al. Characterization of the *Trypanosoma cruzi* ortholog of the SBDS protein reveals an intrinsically disordered extended C-terminal region showing RNA-interacting activity. **Biochimie**, v. 91, n. 4, p. 475–483, 2009.

APÊNDICE A

APÊNDICE A – Informações adicionais sobre as proteínas tratadas no trabalho

Informações sobre a proteína CrNIP7

ATGCGTCCGCTCACAGAAGAGGAGACGAAAGCTGTGTTTCGAGAAGCTTTACAAATTCATAGGCAAGAACATAAAGAACCTC
ATCGACCGGCCAAACGACCGCACTGTTTCCGCCTGCACAAAAACAAGGTTTCTATGTGAAGGAGAGGCTGATGAAGAAG
GCCACAAACGTTGCACGAGACAAGCTGGTGGCGCTGGGAACGTGCATCGGCAAGCTGACGCACACGGGGAAGTCCGCCTG
ACCATTGGGGCGCTGGACCTCATTTTCGCAATACGCGAAGTACAAGATCTGGGTGAAGCCTTCCGCGGAGATGCAGTTCCCTG
TACGGCAACAACGTGCTCAAAAGCGGCCTGGCGCGCATCACCGAGAACACGCCCCGCTACACGGGTGTGGTCATCTACAAC
ATGGCTGACATCCCGCTGGGCTTCGGCGTGACCGCCAAGTCCACCAACGAGTGCCGGAATCTGGACCCCAACGCCGTCGTC
ACCTTCCACCAGGCCGACTGCGGCGAGTTTCTGCGTGCCGAGGACGAGCTATAA

Figura 40 – Sequência de nucleotídeos da CDS CrNip7(NCBI: XM_001696919.1).

Fonte: Elaborada pela autora.

MRPLTEETKAVFEKLYKFIGKNIKNLIDRPNDPHCFRLHKNKVIFYVKERLMKKATNVARDKLVALGTCIGKLTHTGKFR
TIGALDLISQYAKYKIWVKPSAEMQFLYGNNVLKSLARITENTPAYTGVIYNMADIPLGFGVTAKSTNECRNLDPNAV
TFHQADCGEFLRAEDEL

Figura 41 – Sequência de aminoácidos da proteína CrNIP7 (NCBI: XP_001696971.1).

Fonte: Elaborada pela autora.

Informações sobre a proteína Predicted

ATGGACTTTCTCCCTGTTGCGCCGCGCGTGGTTCGGCTTCTCTGCCACGGAGGCGCCCTGCAGCACGAGGCAGCTCACGAG
GTTTGCAAGCTGCAGACGACGCGAGCTTGCTCAGCGAGATGCTGAGCCAGTGCGCGCAAGAAGGCGAGTACTCGGAGAGT
GAGGTCGACGACGAGGAGAGCTACAGCGACGATGATGATGCTGGGGCGCCCATCGTGCGCCAGGACACCGGCTCAAGTATG
GAGCCAGCGCGCTCCGAGCCCATCGCAATCCCTTGCGGAGAGTCTTTTCGCAAACCAGCGTGCCAAGGACGCCTACCGTCGC
CTGCTGTACCAGAAGCAGGTTGAGCACCTTTACCAGAAGCAGCAGCAGCAAGTGCGAGCAGATCGCACTGGCGCAACAACGG
CAGCAGGCGTACCTGGTGCAACAGCAGCGCTGCAGCAGGCGCAGGCGCAGGCCCGCGCGCAGGGGCGCGTGTAA

Figura 42 – Sequência de nucleotídeos codificadores (CDS) da proteína Predicted (NCBI: XM_001690060.1).

Fonte: Elaborada pela autora.

MDFLPVAPRVVGFSAEAPLQHEAAHEVCKLQTQRSLLSEMLSQCAQEGEYSESEVDDEESYSDDDDAGA
PIVRQDTGSSMEPARSEPIAIPCGESFANQRAKDAYRRLLYQKQVEHLYQKQQQVQQIALAQQRQAYL
VQQQRLQQAQAQARAQGRV

Figura 43 – Sequência de aminoácidos da proteína Predicted (NCBI: XP_001690112.1) de *C. reinhardtii*.

Fonte: Elaborada pela autora.

Informações sobre a proteína CrSBDS

ATGTCCAAGGGAGTCTTCCAGCCAGTCCGGCCAGAAGCGGCTGACAAACATCGCGGTGGTGC GGCTAAAGAAGAATGGAGAG
 CGGTTTCGAGATTGCGTGTATATAAAATAAAGTGGGCGACTGGCGCAGCGGCATTGAGAAGGACCTGGATGAGGTCCTTCAG
 ACCACCACCATTTTCCACAACGTGGGCAAGGGCGTGGTGGCCAAGGACAAGGAGCTGCAGGCAGCATTCCGGCACACGGAC
 CAGCACGCCATCTGCCTGGAAATCCTGGCAAAGGGAGAGCTGCAGGTCGGGGAGCGCGAGCGAAAAACAGGAGTATGAACAC
 CTTTTTAAAGATGTTCGCAGGCGTGTTAGTAGAAAAATGCGTGAATCCAAACACCAACCGGCCGTACACGCTGTCAATGCTG
 GAGCGCGCTCTGCGGGACATCCACTTCAACCTGGATCCAAAGAAGTCTGCCAAGCAGCAGGCGCTGGAGGCGCTGCCGCTG
 CTGCAGAAGGAGTTCCCCATCGAGCGTGC GCGCATGCGGCTGCGGCTGACGGTGCCGGTGGAGGCGCAGCAGGAGCTGGAG
 CAGCTGCTGCAGCGCAGGCGGCGGACGTGGCTGGGCTTGACGTGGCGGCGGCGGCATGTGTGTGGTGGTGGCGCAGGTG
 GAGCCGGGCGTGTTCCGGGACCTCAACAACCTTCATGCAGACCACGGCGCGCGGTGCGGGGCGGCTGGAGGTGCTGAGCCTG
 GCGGTGATGGCGGAGGGCGCCACGGCCGACGAGTTTCGCAAACATGGGCGCGGCGGCCACGCGGCCGGTGGCGGCGGCGCAG
 CCGGAGGCCAGCTCGTCAGGCCATGTGGAGCACCGGCCAGTGGGTGTTGGGCGGGTGGGCGAGCGGCCACCGTGGCGGCG
 CCCAGCGCGGTGGCAGTGGCGGCGCGGTGGTGTACGCGCGAGGGCCGGTCAGCGGGTTACCCGAGGAGCACGCCAGCCGG
 CGGGAGCGCTTTGCGGAGCTGGACACGCTGCAGCCCGGCTGGACTGTGGAGCTGCGCACCAGGGGCGACACGGTGGAGGCG
 GTGTTCTTCCACCCAGCGGCGAGCGCGTGGGCGCATTCGCCAACGCACGCCGCATGGCGCTGCAGGCCAGCAAGGCGGCG
 GCATGA

Figura 46 – Sequência 1140 pares de nucleotídeos da CDS de CrSBDS. Os nucleotídeos grifados em cinza foram encontrados por análise do banco de dados SRA em *C. reinhardtii* e correspondem à sequência anotada no banco de dados Phytozome 11 com identificação Cre01.g017500.

Fonte: Elaborada pela autora.

MSKGVFPVQKRLTNIAVRLKKNGERFEIACYKNKVGDWRSIEKDLDEVLQTTTIFHNVGKGVVAKDKELQAAFGTTD
 QHAICLEILAKGELQVGERERKQEYEHLEFKDVAGVLVEKCVNPNTNRPYTLMLERALRDIHFNLDPKKSAAQQALEALPL
 LQKEFPIERARMRLRLTVPVEAQQELEQLLQREAADVAGLDVAAGGMCVVVAQVEPGVFRDLNNFMQTTARGAGRLEVL
 AVMAEGATADEFANMGAAATRPVAAAQPEASSSGHVEHRPVGVRVGERPTVAAPSGGSGGAVVYARGPVSGLPPEEHASR
 RERFAELDTLQPGWTVELRTKGDTVEAVFFHPSGERVGAFAFANARRMALQASKAAA

Figura 47 – Sequência correta de aminoácidos da proteína CrSBDS (Cre01.g017500).

Fonte: Elaborada pela autora.

```

A.fulgidus      1  -----MVSLDKAVIARLRKGGEFEVLVDPYLARDLKEGKEVNFEDLLAAEEVFKDAKKGERASVDELRKIFGTDDV-FEIAKRIILEGEVQITAEQRREMLEAKRKQIINFISR 109
P.abysssi      1  -----MPISVDKAVIARLKVHGETFEILVDPYLARDFKEGKEVPIEELATPYVFKDAHKGDKASEKEMEKIFGTSDP-YEVAKIILRKGEVQLTAQQRREMLEEKKRQIATI IHR 110
T.cruzi        1  MSARIQVPLSQRRHTNVAVVRYTKNGVRLEIACYKNKVISYRGGIETRMDEVLQVERVFTNVSRGLYSSEKEIEAVFGKGTSEKEALQYILDHGELQVAQQERAAEI DQMFIDIAVIISQ 120
S.cerevisiae   1  --MPINQPSGQIKLTNVSLVRLKKARKRFEVACYQNKVQDYRKGIEKDLDEVLQIHQVFMNVSKGLVANKEDLQKCFGTTNV-DDVIEEIMHKGEIQLSEKERQLMLNKVNNEMLTIVSA 117
C.reinhardtii  1  MSKGVFQPVGQKRLTNIAVRLKKNGERFEIACYKNKVGDRSGIEKDLDEVLQTTTI FHNVGKGVVAKDKELQAAF GTTDQ-HAICLEILAKGELQVGERERKQYEHLFKDVAGVLVE 119
A.thaliana     1  MSKTLVQPVGQKRLTNVAVRLKKQGNRFEIACYKNKVLWSRSGVEKDIDEVLQSHTVSVNSKGVLAQSKDLMSFGSDDH-TKICIDILEKGELQVAGKERESQFSSQFRDIATIVMQ 119
H.sapiens      1  --MSIFTPTNQIRLTNVAVVRMKRAGKRFEIACYKNKVVGWRSGVEKDLDEVLQTHSVFVNVSKGVAKKEDLISAFGTDDQ-TEICKQILTKGEVQVSDKERHTQLEQMFRDIATIVAD 117
C.elegans      1  MSKNIKTPTNQKVLTNVAVVRMKKTGKRFEIACYKNKVVNRNKSEKDIDEVLQTHTVFSNVSKGQLSKKEELIAAFGIEDQ-LEICKIILDKGDLQVSEKERQAASDQSLKEVSQLIAS 119
               :  :.*      :*      .      :      *      :::*      ::  :. :*      :.  :.:      **      * : .*:*:      :.*      .      :.  :.

A.fulgidus     110 NTIDPRTNAPHPPSRIERALEEAKVHIDIFKSVEAQVKDIVKALKP--ILPLKFEEMEIAIKIPPEHTGRAISALY--NFGG----VTREEWQ-RDGSWICVMRIPSGMYGDLMDLLGK 219
P.abysssi     111 HAVDPRTGYPHVDRILRAMEEVGVRVDIFKDAEAQVDVIKAI--ILPLRIEMKVIKIPSEYVGRAYGEVR--KFGR----IKKEEWA-SDGSWLFLEIIPGGVEEEFYEKLNA 220
T.cruzi       121 KCVNEVTQRFPFSPQVIEQALHSIGAAVKLDQPVKKQALAFIHQLIDAKTIPIARARMKLRVVPDEPSLEKLVEWCETNGTSILQKVVEAGGSDAALQIHSLLILLPHLFRDIERFVK 240
S.cerevisiae  118 KCINPVSKKRYPTMIHKALQELKFSPIVINKPAKLQALEAIKLLVSKQIPIVRAKMKVVAISEPSRQPELIEKISKLIA-----SPGESTKPELDPWCTCTGLIDPVNYDLMTLCDK 232
C.reinhardtii 120 KCVNPNTNRPYTLMLERALRDIHFNLDPKKSQKQALEALPLLQK--EFPIERARMRLRLTVPE-AQQELEQLLQREAAD----VAGLDVA-AGGMCVVVAQVEPGVFRDLNFMQT 230
A.thaliana    120 KVINPETQRPYTISMVERLMHEIHFAVDPHNSKKQALDVIRELQK--HFPIKRSPMRRLRLTVPVQ-NFPSLLEKLKEWDGS----VVSKE--SGTQMSVCEMEPGLFRECDSHVRS 229
H.sapiens     118 KCVNPETKRPTYVILIERAMKDIHYSVKTNKSTKQQALEVIKQLKE--KMKIERAHMRLRFILPVN-EGKKLKEKLKPLIKV----IESEDY--G-QQLEIVCLIDPGCFREIDELIKK 226
C.elegans     220 MVVNPETKRPPVPSVIDKALQEMHFSKPNRSSKQQAALDAIPKLE--TLKIERAKMKIRVAIPTK-EAKSVHTKLKTLFSD----VEVDDW--QDGSLEMVGLIEPGSFRALDDLVRN 229
               ::  :      :  :..      : *.      :  :      :  :      :  :      :

A.fulgidus     220 VAKGEALTKVLRRIG----- 234
P.abysssi     221 LTKGNAQTKLIERKGL----- 236
T.cruzi       241 EMPPGGSVHVIEAAMDVGEGDVMDAELIARANAHMTGASGVGNAD-KSHCTGSSSNPHAESNRGRKKGKGSRKPPQGVQVRLGKDDGTDQQTVTTAELSSTPAVVGENDDE---LKVAL 356
S.cerevisiae  233 ----KGTQVQLDMAVIDNTTHN----- 250
C.reinhardtii 231 TARGAGRLEVLSLAVMAEGATADEFA---NMGAAA--TRPVAAAQPEASSSGHVEH-----RPVGVGR-----VGERPTVAAPSGGSGGAVVYA 310
A.thaliana    230 ---IQGRLEILAVSVHAEGDTSMDHYDEHDDMALQT--HKPLLPAETETKD--LTD-----PVVELSKKLQKQ-----EISTTDNIKQEGGEEKGKTKCST 313
H.sapiens     227 ETKGKGSLEVLNLKDVEEGDEKFE----- 250
C.elegans     230 ETKGHGRLEILSLKDVVEGELQIS----- 253
               .  :.:

A.fulgidus     -----
P.abysssi     -----
T.cruzi       357 SKLV--LDETNGSRDEDYDGKGRGKKKAKRRQGEFPHQKQQQQPATPQKTVEVDEQQGSDEEVLVNRKQKQAAARVRVKDDKRNEYWNADDDDDYDGYEDEDAGEVRD--- 464
S.cerevisiae  -----
C.reinhardtii 311 RGPVSGLPPEEHASRRERFAELDT-----LQPGWTVELR-----TKGDTV-----EAVFFHPSGERVGAFANARMALQASKAAA 379
A.thaliana    314 CNTFVGEAK-----QY--REH-----CKSDWHKHNL-----KRKTRKL--PPISADECMSEIDMDDSRADLKDYSF----- 370
H.sapiens     -----
C.elegans     -----

```

Figura 48 – Alinhamento múltiplo (Clustal O) de SBDS dos organismos *A. fulgidus* (WP_010877998.1), *P. abyssi* (WP_010867732.1), *T. cruzi* (ACD88761.1), *S. cerevisiae* (NP_013122.1), *C. reinhardtii* (Cre01.g017500), *A. thaliana* (NP_564488.1), *H. sapiens* (NP_057122.2) e *C. elegans* (NP_497226.2.). Os aminoácidos conservados estão representados por asterisco, os aminoácidos com alta similaridade; com dois pontos e aminoácidos com baixa similaridade; com um ponto.

Fonte: Elaborada pela autora

Tabela 16 - Matriz de identidade entre as sequências de SBDS (programa Clustal O).

Organismo	<i>A. fulgidus</i>	<i>P. abyssi</i>	<i>T. cruzi</i>	<i>S. cerevisiae</i>	<i>C. reinhardtii</i>	<i>A. thaliana</i>	<i>H. sapiens</i>	<i>C. elegans</i>
<i>A. fulgidus</i>	100	54.27	17.95	26.09	25.32	21.40	24.24	25.43
<i>P. abyssi</i>	54.27	100	20.34	21.98	21.70	22.94	28.33	24.36
<i>T. cruzi</i>	17.95	20.34	100	31.60	32.43	31.04	38.80	36.76
<i>S. cerevisiae</i>	26.09	21.98	31.60	100	38.21	34.29	42.62	38.78
<i>C. reinhardtii</i>	25.32	21.70	32.43	38.21	100	40.29	48.40	46.64
<i>A. thaliana</i>	21.40	22.94	31.04	34.29	40.29	100	51.42	45.20
<i>H. sapiens</i>	24.24	28.33	38.80	42.62	48.40	51.42	100	56
<i>C. elegans</i>	25.43	24.36	36.76	38.78	46.64	45.20	56	100

Fonte: Elaborada pela autora.