



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS AGRÁRIAS
DEPARTAMENTO DE ENGENHARIA DE PESCA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PESCA

ANDRESSA ROCHA DE OLIVEIRA SOUSA

**ESTUDOS ESTRUTURAIS E DE EXPRESSÃO HETERÓLOGA DE LECTINAS DE
ORGANISMOS MARINHOS COM POTENCIAL BIOTECNOLÓGICO**

FORTALEZA

2026

ANDRESSA ROCHA DE OLIVEIRA SOUSA

ESTUDOS ESTRUTURAIS E DE EXPRESSÃO HETERÓLOGA DE LECTINAS DE
ORGANISMOS MARINHOS COM POTENCIAL BIOTECNOLÓGICO

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Pesca da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutor em Engenharia de Pesca. Área de concentração: Biotecnologia e Genética de Recursos Aquáticos.

Orientador: Prof. Dr. Celso Shiniti Nagano.
Coorientadora: Dra. Renata Pinheiro Chaves.

FORTALEZA

2026

ANDRESSA ROCHA DE OLIVEIRA SOUSA

ESTUDOS ESTRUTURAIS E DE EXPRESSÃO HETERÓLOGA DE LECTINAS DE
ORGANISMOS MARINHOS COM POTENCIAL BIOTECNOLÓGICO

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Pesca da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutor em Engenharia de Pesca. Área de concentração: Biotecnologia e Genética de Recursos Aquáticos.

Aprovada em: 26/02/2026.

BANCA EXAMINADORA

Prof. Dr. Celso Shiniti Nagano (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Alexandre Holanda Sampaio
Universidade Federal do Ceará (UFC)

Dr. Renato Cézar Farias Torres
Universidade Federal do Ceará (UFC)

Prof^ª. Dr^a. Denise Cavalcante Hissa
Universidade Federal do Ceará (UFC)

Prof. Dr. Ito Liberato Barroso Neto
Unichristus

A Deus,

Aos meus pais, Eugênia e Evandro

À minha irmã, Amanda

Às minhas avós, Antonina e Laura

AGRADECIMENTOS

À Deus, por ter sido minha fortaleza nos momentos difíceis e por ter me proporcionado tantas realizações pessoais e profissionais durante meu doutorado.

Aos meus pais Eugênia e Evandro, por sempre acreditarem em mim e me ajudarem a realizar meus sonhos, não medindo esforços para me apoiar ao longo desse caminho. Por ensinar desde sempre a importância do estudo.

À minha irmã Amanda por todo o apoio ao longo desses anos e sempre me lembrando do quanto sou capaz.

Às minhas avós Antonina e Laura, por todo o amor, apoio e preocupação com meus estudos.

Ao Bruno Lopes, por toda a ajuda durante o período mais difícil desses anos de doutorado. Obrigada por tudo.

Ao meu querido orientador Professor Dr. Celso Shiniti Nagano, pelos muitos anos de ensinamentos e pela oportunidade de ser sua orientanda. Obrigada por toda ajuda, por acreditar em mim e por todos os momentos compartilhados ao longo desses anos, além da vivência acadêmica. Serei sempre grata ao senhor e me recordarei com muito carinho do papel que o senhor desempenhou na minha trajetória acadêmica.

Ao Professor Dr. Alexandre Holanda por ter me encantado desde muito cedo pelo mundo da biotecnologia marinha. Obrigada por todos os ensinamentos desde a graduação, por toda a ajuda sempre que precisei, principalmente durante a realização do doutorado sanduíche. O senhor e a Prof. Dra. Silvana Saker serão sempre lembrados por mim com admiração.

À minha coorientadora Dra. Annabelle Varrot, por ter me recebido tão bem no seu laboratório, por ter me ensinado tanto e com tanta paciência e humildade, e por todo o cuidado e carinho para comigo durante o período do doutorado sanduíche. Obrigada pelo apoio e ajuda não apenas durante minha estadia em Grenoble, mas até hoje.

Ao Professor Dr. Rômulo Farias e à Dra. Renata Pinheiro pela contribuição no trabalho.

Aos meus queridos amigos que estiveram comigo até o fim, Philippe, Jéssica, Renato, Juliana, Pedro Arthur, Vinicius, Ellen Malveira e Aldeni. Vocês sabem o quanto foram essenciais para me dar forças quando mais precisei.

Aos amigos que fiz e permaneceram ao longo desses longos anos de BIOMAR, Alexandra, Livia, Victória Régia e Regivânio. Obrigada pelo apoio durante toda essa jornada.

Às minhas ICs, filhas queridas, Elany e Ellen Maria, por toda a ajuda durante a realização desse trabalho, que não foi pouca, e por todo o apoio não só na bancada.

Aos meus colegas do BIOMAR, especialmente ao Manoel que será sempre lembrado como meu primeiro IC, sempre se mostrando disponível para ajudar quando precisei. E a todos que dividiram esses 10 anos de convivência na bancada.

À toda equipe do GBMS team do CERMAV, Anne Imberty, Emilie Gillon, Serge Pérez e principalmente, à Valérie Chazalet, por todo o acolhimento no grupo e ensinamentos partilhados.

Aos meus companheiros de bancada do GBMS team, Noon, Luca, Palig, Batoul e Théodore por me acolherem tão bem, pelos momentos de lazer partilhados, pela enriquecedora troca cultural e pela ajuda no laboratório.

Aos meus amigos de Grenoble, Duda, Luara, Marina e Thales, por serem minha família na França e por toda ajuda durante minha temporada francesa.

Aos professores Prof. Dra. Denise Hissa e Prof. Dr. Ito Liberato, por aceitarem participar da banca avaliadora e contribuir neste trabalho.

A todos os professores e colaboradores que de forma direta e indireta, contribuíram para o desenvolvimento dessa tese e para o meu desenvolvimento profissional.

À CAPES pelo financiamento da bolsa e do programa de doutorado sanduíche no exterior.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

“A educação tem raízes amargas, mas
os seus frutos são doces.” (Aristóteles)

RESUMO

Os organismos marinhos possuem compostos bioativos com inúmeras atividades biológicas já descritas. Dentre esses compostos estão as lectinas, proteínas ou glicoproteínas que apresentam ligação específica e reversível a carboidratos. Os estudos estruturais das lectinas e de suas interações proteína-carboidrato são importantes, pois o padrão de glicosilação da superfície celular é a chave para compreender as funções dessas moléculas, fornecendo mecanismos para aplicações biotecnológicas. Porém, ainda são escassas as informações sobre a estrutura e a especificidade dessas moléculas oriundas de organismos marinhos. Nesse sentido, estratégias voltada à produção de proteínas recombinantes são ferramentas importantes para estudos estruturais e para a produção em larga escala de proteínas com enfoque biotecnológico. O presente trabalho teve como objetivo produzir, em sistema heterólogo, a lectina de alga marinha, AmL (*Amansia multifida* lectin) e as lectinas de ovos de invertebrados marinhos ELEL (*Echinometra lucunter* eggs lectin) e ADEL (*Aplysia dactylomela* eggs lectin), caracterizá-las bioquimicamente e estruturalmente, e avaliar o potencial antibacteriano da primeira associado a antibióticos. Por outra parte, o trabalho objetivou determinar as estruturas tridimensionais das galectinas oriundas de esponjas marinhas, CCL (*Chondrilla caribensis* lectin) e ALL (*Aplysina lactuca* lectin) e avaliar a interação dessas proteínas com seus respectivos ligantes. A lectina rEDEL não foi expressa com solubilidade e funcionabilidade, apesar das tentativas em diferentes vetores, inúmeras condições de expressão e estratégias de redobramento e cotransformação. A lectina recombinante rADEL foi expressa em sistema heterólogo de *Komagataella pastoris* biologicamente ativa, porém com um rendimento insatisfatório. A lectina rAmL foi expressa em sistema heterólogo de *Escherichia coli*, biologicamente ativa e bioquimicamente semelhante à proteína nativa. rAmL apresentou atividade antibacteriana associada ao antibiótico tetraciclina contra *Staphylococcus aureus* e *E. coli*. As duas galectinas, ALL e CCL, tiveram sua estrutura tridimensional determinada por cristalografia de raios X e a interação lectina-carboidrato analisada por cocristalização e calorimetria de titulação isotérmica. Portanto, esses resultados ampliam significativamente o conhecimento estrutural sobre lectinas de organismos marinhos, fornecendo novas ideias para a compreensão funcional das lectinas desses organismos, cujo estudos estruturais ainda são escassos.

Palavras-chave: lectina; produção recombinante; estrutura; especificidade; invertebrados marinhos; algas marinhas.

ABSTRACT

Marine organisms possess bioactive compounds with numerous described biological activities. Among these compounds are lectins, proteins or glycoproteins that exhibit specific and reversible binding to carbohydrates. Structural studies of lectins and their protein-carbohydrate interactions are important, as the glycosylation pattern of the cell surface is the key to understanding the functions of these molecules, providing mechanisms for biotechnological applications. However, information on the structure and specificity of these molecules from marine organisms is still scarce. In this sense, strategies focused on the production of recombinant proteins are important tools for structural studies and for the large-scale production of proteins with a biotechnological focus. This study aimed to produce, in a heterologous system, the marine algae lectin AmL (*Amansia multifida* lectin) and the marine invertebrate egg lectins ELEL (*Echinometra lucunter* eggs lectin) and ADEL (*Aplysia dactylomela* eggs lectin), to characterize them biochemically and structurally, and to evaluate the antibacterial potential of the former in combination with antibiotics. Furthermore, the study aimed to determine the three-dimensional structures of galectins derived from marine sponges, CCL (*Chondrilla caribensis* lectin) and ALL (*Aplysina lactuca* lectin), and to evaluate the interaction of these proteins with their respective ligands. The lectin rELEL was not expressed with solubility and functionality, despite attempts in different vectors, numerous expression conditions, and refolding and cotransformation strategies. The recombinant lectin rADEL was expressed in a heterologous system of *Komagataella pastoris*, which was biologically active, but with unsatisfactory yield. The lectin rAmL was expressed in a heterologous system of *Escherichia coli*, is biologically active, and biochemically similar to the native protein. rAmL showed antibacterial activity associated with the antibiotic tetracycline against *Staphylococcus aureus* and *E. coli*. The both galectins, ALL and CCL, had their three-dimensional structure determined by X-ray crystallography, and the lectin-carbohydrate interaction analyzed by cocrystallization and isothermal titration calorimetry. Therefore, these results significantly expand the structural knowledge about lectins from marine organisms, providing new ideas for the functional understanding of lectins from these organisms, whose structural studies are still scarce.

Keywords: lectin; recombinant production; structure; specificity; marine invertebrates; seaweed.

LISTA DE FIGURAS

Figura 01	– Tipos de <i>N</i> -glicanos.....	29
Figura 02	– Invertebrados marinhos portadores das lectinas do estudo.....	35
Figura 03	– Alga marinha <i>Amansia multifida</i>	35
Figura 04	– Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pET32a.....	44
Figura 05	– Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pEX-A128.....	49
Figura 06	– Representação dos vetores de expressão utilizados para subclonar a sequência gênica de rELEL.....	51
Figura 07	– Representação esquemática da proteína recombinante expressa Trx-rELEL.....	62
Figura 08	– Representação da proteína rELEL obtida após clivagem com a protease TEV.....	63
Figura 09	– Testes de expressão da proteína rELEL no vetor pET32a(+)......	64
Figura 10	– Alinhamento da sequência de aminoácidos da lectina Trx-rELEL com os peptídeos identificados por espectrometria de massas.....	67
Figura 11	– Novos testes de expressão da proteína rELEL no vetor pET32a(+)......	70
Figura 12	– Alinhamento da sequência de aminoácidos da lectina Trx-rELEL expressa na fração solúvel com os peptídeos identificados por espectrometria de massas.....	71
Figura 13	– Representação esquemática da proteína recombinante expressa no vetor pCOLD-TF-TEV-3G-Nco.....	78
Figura 14	– Representação da proteína rELEL obtida após clivagem com a protease TEV.....	79
Figura 15	– Análise da expressão de rELELpCOLD-TF-TEV-3G-Nco em SDS-PAGE.....	79
Figura 16	– Análise da purificação de rELELpCOLD-TF-TEV-3G-Nco em SDS-PAGE....	80
Figura 17	– Análise da purificação de rELEL clivada com TEV em SDS-PAGE.....	82

Figura 18 – Representação esquemática da proteína recombinante expressa no vetor pET39-TEV.....	84
Figura 19 – Representação da lectina rELEL obtida após clivagem com a protease TEV.....	85
Figura 20 – Análise da purificação periplasmática de rELELpET39-TEV em SDS-PAGE.....	87
Figura 21 – Análise da purificação citoplasmática de rELELpET39-TEV em SDS-PAGE.....	89
Figura 22 – Representação esquemática da proteína recombinante no vetor pET15b(+)......	90
Figura 23 – Análise da expressão de rELELpET15b(+) em SDS-PAGE.....	91
Figura 24 – Análise da expressão de rELELpET15b(+) com sistema CyDisCo em SDS-PAGE.....	93
Figura 25 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco.....	94
Figura 26 – Representação da proteína rELEL obtida após clivagem com a protease TEV.....	95
Figura 27 – Análise da purificação de rELELpCOLD-HAT-TEV-3G-Nco em SDS-PAGE.....	96
Figura 28 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pET21a.....	103
Figura 29 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pEX-A128.....	105
Figura 30 – Representação dos vetores de expressão utilizados para subclonar a sequência gênica.....	107
Figura 31 – Representação esquemática da proteína recombinante expressa rADEL.....	128
Figura 32 – Testes de expressão da proteína rADEL no vetor pET21a.....	130
Figura 33 – Géis de agarose referentes à linearização do vetor e confirmação da transformação em <i>K. pastoris</i>	135

Figura 34 – Representação esquemática da proteína recombinante expressa no vetor pPICZ α HISTEV.....	137
Figura 35 – Representação da proteína rADEL obtida após clivagem com a protease TEV.....	137
Figura 36 – Análise dos sobrenadantes e <i>pellets</i> da expressão de rADELpPICZ α HISTEV em SDS-PAGE.....	138
Figura 37 – Análise de <i>immunoblotting</i> dos sobrenadantes e <i>pellets</i> da expressão de rADELpPICZ α HISTEV.....	139
Figura 38 – Cromatograma da purificação de rADELpPICZ α HISTEV colônia 3.....	140
Figura 39 – Sobreposição dos cromatogramas do <i>screening</i> de melhor expressão.....	140
Figura 40 – Análise da purificação de rADELpPIC α HISTEV em SDS-PAGE.....	141
Figura 41 – Locais de <i>N</i> -glicosilação de ADEL.....	142
Figura 42 – Análise da purificação citoplasmática de rADELpPIC α HISTEV.....	145
Figura 43 – Representação esquemática da proteína recombinante no vetor pET25b(+)......	149
Figura 44 – Análise da expressão de rADELpET25b(+) em SDS-PAGE.....	150
Figura 45 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco.....	152
Figura 46 – Representação da proteína rADEL obtida após clivagem com a protease TEV.....	153
Figura 47 – Análise da purificação de rADELpCOLD-HAT-TEV-3G-Nco em SDS-PAGE.....	154
Figura 48 – Confirmação da transformação de rADELpPICZ α HisTEV em <i>E. coli</i>	157
Figura 49 – Géis de agarose referentes à linearização do vetor e a purificação do inserto.....	158
Figura 50 – Comparação do crescimento em zeocina das células transformadas e não transformadas.....	159
Figura 51 – Avaliação do melhor tempo de expressão para rADEL: Pureza da lectina purificada em SDS-PAGE.....	162

Figura 52 – Perfil de expressão e purificação da proteína rADEL no Brasil.....	165
Figura 53 – Purificação de rADEL com 175 mM e 270 mM de imidazol.....	167
Figura 54 – Estimativa da massa molecular rADEL.....	172
Figura 55 – Espectro de dicroísmo de rADEL.....	178
Figura 56 – Predição da estrutura secundária de rADEL.....	178
Figura 57 – Distribuição do tamanho hidrodinâmico por números de partículas da lectina rADEL nas duas concentrações.....	180
Figura 58 – Sequência de nucleotídeos codificantes (minúsculas) e a respectiva tradução em aminoácidos (maiúsculas) da lectina para inserção no vetor pEX-A258.....	186
Figura 59 – Representação dos vetores de expressão utilizados para subclonar a sequência gênica otimizada.....	188
Figura 60 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pET28a(+)......	197
Figura 61 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-TF-TEV-Nco.....	209
Figura 62 – Representação da proteína rAmL obtida após clivagem com a protease TEV....	210
Figura 63 – Análise da purificação de rAmLpCOLD-TF-TEV-Nco em SDS-PAGE.....	210
Figura 64 – Análise da purificação de rAmL clivada em SDS-PAGE.....	212
Figura 65 – Análise em SDS-PAGE da purificação de rAmL após segunda clivagem.....	214
Figura 66 – Representação esquemática da proteína rAmL no vetor pET-TEV.....	215
Figura 67 – Análise da purificação de rAmLpET-TEV em SDS-PAGE.....	216
Figura 68 – Representação esquemática da proteína recombinante rAmL expressa no vetor pCOLD-HAT-TEV-3G-Nco.....	218
Figura 69 – Representação da proteína rAmL obtida após clivagem com a protease TEV.....	219
Figura 70 – Análise da purificação de rAmLpCOLD-HAT-TEV-3G-Nco em SDS-PAGE.....	220

Figura 71	–	Análise da purificação de rAmL clivada em SDS-PAGE.....	
Figura 72	–	Cromatograma de exclusão molecular da proteína rAmL clivada.....	223
Figura 73	–	Análise em SDS-PAGE da purificação de rAmL após cromatografia de gel filtração.....	224
Figura 74	–	Análise em SDS-PAGE da purificação de rAmL após recromatografia com nova TEV otimizada.....	225
Figura 75	–	Frações de rAmL não clivadas e clivadas para uso posterior.....	228
Figura 76	–	Estabilidade térmica de rAmL em diferentes condições de tampão determinado por TSA.....	230
Figura 77	–	Representação esquemática da proteína recombinante expressa His-rAmL.....	233
Figura 78	–	Representação da proteína rAmL obtida após clivagem com a protease TEV....	234
Figura 79	–	Testes de expressão da proteína rAmL no vetor pCOLD-HAT-TEV-3G-Nco...	236
Figura 80	–	Testes de expressão da proteína rAmL no vetor pET28a(+)......	238
Figura 81	–	Cromatograma da purificação da proteína rAmLpCHT3GNco e rAmLpET28a.....	242
Figura 82	–	Análise da purificação de rAmLpCHT3GNco e rAmLpET28a em SDS- PAGE.....	243
Figura 83	–	Cromatograma da purificação da proteína rAmL clivada.....	244
Figura 84	–	Análise da purificação de rAmL clivada em SDS-PAGE.....	245
Figura 85	–	Perfil de expressão e purificação da proteína rAmLpCHT3GNco com <i>steps</i>	248
Figura 86	–	Estimativa da massa molecular de rAmLpCHT3GNco.....	250
Figura 87	–	Espectro de dicroísmo de rAmLpCHT3GNco.....	255
Figura 88	–	Predição da estrutura secundária de rAmLpCHT3GNco.....	256
Figura 89	–	Espectro de dicroísmo de rAmLpCHT3GNco em função da temperatura.....	257
Figura 90	–	Distribuição do tamanho hidrodinâmico por números de partículas da lectina rAmLpCHT3GNco.....	259

Figura 91 – Avaliação da pureza das lectinas ALL e CCL por SDS-PAGE.....	275
Figura 92 – Cristais obtidos de ALL.....	277
Figura 93 – Alinhamento das sequências de ALL determinada por espectrometria de massas e cristalografia.....	279
Figura 94 – Estrutura geral de ALL.....	281
Figura 95 – Pontes dissulfeto presentes na estrutura da ALL.....	282
Figura 96 – Análise das interações presentes no complexo ALL-lactose.....	286
Figura 97 – Cristais obtidos de CCL-3SL.....	287
Figura 98 – Alinhamento das sequências de CCL determinada por espectrometria de massas e cristalografia.....	289
Figura 99 – Estrutura geral da lectina CCL.....	291
Figura 100 – Análise das interações presentes no complexo CCL-3SL.....	295
Figura 101 – Alça componente do sítio de ligação a carboidratos.....	297
Figura 102 – Análise da interação ALL-Lactose e CCL-LacNac por calorimetria de titulação isotérmica.....	298

LISTA DE TABELAS

Tabela 01 – Características dos vetores utilizados para a expressão da lectina rEEL.....	53
Tabela 02 – Primers e condições de PCR para confirmação de vetores.....	55
Tabela 03 – Estratégias de purificação da lectina recombinante em cada vetor.....	57
Tabela 04 – Peptídeos sequenciados de Trx-rEEL na fração insolúvel.....	66
Tabela 05 – Peptídeos sequenciados de Trx-rEEL na fração solúvel.....	70
Tabela 06 – Características dos vetores utilizados para a expressão da lectina.....	108
Tabela 07 – Primers e condições de PCR para confirmação de vetores.....	110
Tabela 08 – Primers para subclonagem no vetor pCOLD-HAT-TEV-3G-Nco.....	111
Tabela 09 – Composição dos meios usados na expressão de proteínas em <i>K. pastoris</i>	114
Tabela 10 – Estratégias de purificação da lectina rADEL em cada vetor.....	117
Tabela 11 – Condições de cristalizações testadas para rADEL.....	120
Tabela 12 – Avaliação do melhor tempo de expressão para rADEL: Funcionabilidade da lectina recombinante expressa.....	163
Tabela 13 – Perfil de inibição da atividade hemaglutinante de rADEL.....	174
Tabela 14 – Parâmetros do espalhamento dinâmico de luz de rADEL a 0,250 mg/mL e 0,125 mg/mL.....	180
Tabela 15 – Características dos vetores utilizados para a expressão da lectina.....	189
Tabela 16 – Primers e condições de PCR para confirmação de vetores.....	190
Tabela 17 – Estratégias de purificação da lectina recombinante em cada vetor.....	191
Tabela 18 – Condições de cristalizações testadas para rAmL.....	196
Tabela 19 – Primers e condições de PCR para confirmação de vetores.....	199

Tabela 20 – Avaliação da melhor expressão para rAmLpCHT3GNco: Funcionabilidade da lectina recombinante expressa.....	240
Tabela 21 – Avaliação da melhor expressão para rAmLpET28a: Funcionabilidade da lectina recombinante expressa.....	241
Tabela 22 – Comparação do rendimento de expressão entre rAmLpCHT3GNco e rAmLpET28a.....	246
Tabela 23 – Título de hemaglutinação do ensaio para a lectina recombinante e nativa.....	251
Tabela 24 – Perfil de inibição da atividade hemaglutinante de rAmLpCHT3GNco e AmL.	252
Tabela 25 – Parâmetros do espalhamento dinâmico de luz de rAmLpCHT3GNco para as três concentrações sem cloreto de cálcio.....	259
Tabela 26 – Parâmetros do espalhamento dinâmico de luz de rAmLpCHT3GNco para as três concentrações com cloreto de cálcio.....	260
Tabela 27 – Efeito da lectina rAmLpCHT3GNco combinada aos antibióticos oxacilina, tetraciclina, ampicilina e estreptomicina contra <i>S. aureus</i> e <i>E. coli</i>	263
Tabela 28 – Dados de coleta e refinamento do complexo ALL-Lactose.....	278
Tabela 29 – Interações moleculares presentes na interface dimérica da lectina ALL.....	283
Tabela 30 – Interações moleculares presentes na interface tetramérica da lectina ALL.....	284
Tabela 31 – Interações moleculares entre ALL e lactose.....	285
Tabela 32 – Dados de coleta e refinamento dos complexos CCL-Lactose e CCL-3SL.....	288
Tabela 33 – Interações moleculares presentes na interface dimérica da lectina CCL.....	292
Tabela 34 – Interações moleculares presentes na interface tetramérica da lectina CCL.....	293
Tabela 35 – Interações moleculares entre CCL e 3SL (3'-sialilactose).....	294
Tabela 36 – Parâmetros da ligação entre a lectina e o ligante.....	299

SUMÁRIO

1	INTRODUÇÃO.....	24
1.1	Organismos marinhos como fonte de biomoléculas.....	24
1.2	Lectinas.....	25
1.3	Lectinas e o reconhecimento a carboidratos.....	27
1.3.1	<i>Carboidratos.....</i>	27
1.3.2	<i>Especificidade de ligação das lectinas.....</i>	30
1.4	Potencial biotecnológico das lectinas.....	31
1.5	Lectinas objetos do estudo.....	34
1.5.1	<i>Potencial biotecnológico das lectinas em estudo.....</i>	36
2	JUSTIFICATIVA.....	38
3	OBJETIVOS DA TESE.....	41
4	CAPÍTULO I: EXPRESSÃO HETRÓLOGA DA LECTINA rELEL PRESENTE NOS OVOS DO OURIÇO-DO-MAR <i>Echinometra lucunter</i>: DESAFIOS DE EXPRESSÃO E DOBRAMENTO.....	42
4.1	Objetivos.....	42
4.2	Metodologia.....	43
4.2.1	<i>Coleta das espécimes e purificação da lectina nativa.....</i>	43
4.2.2	<i>Produção da proteína recombinante rELEL em sistema heterólogo de Escherichia coli (Brasil).....</i>	43
4.2.2.1	<i>Desenho do gene sintético adquirido no Brasil.....</i>	43
4.2.2.2	<i>Expressão da lectina rELEL em sistema heterólogo de E. coli.....</i>	45
4.2.2.2.1	Testes de expressão.....	45
4.2.2.2.2	Estratégias para expressão da rELEL na fração solúvel.....	45
4.2.2.3	<i>Tentativa de purificação da proteína recombinante.....</i>	46
4.2.2.4	<i>Confirmação da expressão da lectina por espectrometria de massas sequencial..</i>	47
4.2.2.5	<i>Tentativa de refolding da lectina recombinante a partir da fração insolúvel.....</i>	47
4.2.3	<i>Avaliação da expressão da proteína recombinante por eletroforese (Brasil).....</i>	48
4.2.4	<i>Avaliação da funcionabilidade da proteína recombinante (Brasil).....</i>	48
4.2.5	<i>Produção da proteína recombinante rELEL em sistema heterólogo (França)...</i>	49
4.2.5.1	<i>Desenho do gene sintético adquirido na França.....</i>	49

4.2.5.1.1	Subclonagens do gene sintético para os vetores de expressão.....	50
4.2.5.2	<i>Expressão da lectina rELEL em sistema heterólogo.....</i>	56
4.2.5.3	<i>Purificação da lectina recombinante rELEL.....</i>	57
4.2.5.4	<i>Digestão da proteína fusionada e repurificação da lectina recombinante.....</i>	59
4.2.5.5	<i>Tentativa de cotransformação com cydisco.....</i>	60
4.2.6	Avaliação da expressão da proteína recombinante por eletroforese (França)....	60
4.2.7	Avaliação da funcionabilidade da proteína recombinante (França).....	61
4.3	Resultados e discussão.....	62
4.3.1	<i>Produção da proteína recombinante rELEL em sistema heterólogo de E. coli (Brasil).....</i>	62
4.3.1.1	<i>Desenho do gene sintético.....</i>	62
4.3.1.2	<i>Testes de expressão.....</i>	64
4.3.1.3	<i>Estratégias para expressão da proteína rELEL na fração solúvel.....</i>	68
4.3.1.4	<i>Tentativa de refolding da lectina recombinante a partir da fração insolúvel.....</i>	73
4.3.2	<i>Produção da proteína recombinante rELEL em sistema heterólogo de E. coli (França).....</i>	74
4.3.2.1	<i>Subclonagens do gene sintético para os vetores de expressão.....</i>	74
4.3.2.2	<i>Expressão da lectina rELEL em sistema heterólogo.....</i>	77
4.3.2.2.1	<i>Expressão da lectina rELEL no vetor pCOLD-TF-TEV-3G-Nco.....</i>	78
4.3.2.2.2	<i>Expressão da lectina rELEL no vetor pET39-TEV.....</i>	84
4.3.2.2.3	<i>Expressão da lectina rELEL no vetor pET15b(+)......</i>	90
4.3.2.2.4	<i>Expressão da lectina rELEL no vetor pCOLD-HAT-TEV-3G-Nco.....</i>	93
4.4	Conclusão.....	100
5	CAPÍTULO II: EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA LECTINA rADEL PRESENTE NOS OVOS DA LESMA-DO-MAR <i>Aplysia dactylomela</i>.....	101
5.1	Objetivos.....	101
5.2	Metodologia.....	102
5.2.1	<i>Coleta dos ovos e purificação da lectina nativa.....</i>	102
5.2.2	<i>Produção da proteína recombinante rADEL em sistema heterólogo de Escherichia coli (Brasil).....</i>	102
5.2.2.1	<i>Desenho do gene sintético adquirido no Brasil.....</i>	102
5.2.2.2	<i>Expressão da lectina rADEL em sistema heterólogo de E. coli.....</i>	103

5.2.2.3	<i>Confirmação da expressão da lectina por espectrometria de massas sequencial.</i>	104
5.2.3	<i>Produção da proteína recombinante rADEL em sistema heterólogo (França).</i>	105
5.2.3.1	<i>Desenho do gene sintético adquirido na França.....</i>	105
5.2.3.1.1	<i>Subclonagens do gene sintético para os vetores de expressão.....</i>	106
5.2.3.2	<i>Produção da proteína rADEL em sistema heterólogo de Komagataella pastoris.</i>	112
5.2.3.2.1	<i>Transformação de K. pastoris com o plasmídeo rADELpPICZαHisTEV.....</i>	112
5.2.3.2.2	<i>Expressão da lectina rADEL em K. pastoris.....</i>	113
5.2.3.2.3	<i>Purificação da lectina recombinante.....</i>	115
5.2.3.2.4	<i>Purificação da lectina recombinante a partir do concentrado celular.....</i>	116
5.2.3.3	<i>Produção da lectina rADEL em sistema heterólogo de E. coli.....</i>	116
5.2.3.3.1	<i>Expressão da lectina recombinante.....</i>	116
5.2.3.3.2	<i>Purificação da lectina recombinante.....</i>	117
5.2.4	<i>Avaliação da pureza e funcionabilidade da proteína recombinante (França)....</i>	118
5.2.4.1	<i>Estimativa da massa molecular e avaliação da pureza da lectina.....</i>	118
5.2.4.2	<i>Atividade hemaglutinante da proteína.....</i>	119
5.2.5	<i>Immunoblotting para detecção da proteína rADEL expressa (França).....</i>	119
5.2.6	<i>Ensaio de cristalização (França).....</i>	119
5.2.7	<i>Produção da proteína recombinante rADEL em sistema heterólogo de K. pastoris (Brasil).....</i>	121
5.2.7.1	<i>Recuperação do gene sintético francês no Brasil.....</i>	121
5.2.7.2	<i>Transformação de E. coli com o gene sintético rADELpPICZαHisTEV.....</i>	121
5.2.7.3	<i>Transformação de K. pastoris com o gene sintético rADELpPICZαHisTEV.....</i>	122
5.2.7.4	<i>Expressão da lectina rADEL em K. Pastoris.....</i>	123
5.2.7.5	<i>Purificação da lectina recombinante.....</i>	124
5.2.8	<i>Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil).....</i>	124
5.2.8.1	<i>Estimativa da massa molecular e avaliação da pureza da lectina.....</i>	125
5.2.8.2	<i>Ensaio de atividade hemaglutinante da proteína recombinante.....</i>	125
5.2.8.3	<i>Ensaio de inibição da atividade hemaglutinante com a proteína recombinante...</i>	125
5.2.8.4	<i>Dicroísmo circular.....</i>	126
5.2.8.5	<i>Espalhamento dinâmico de luz - DLS (Dynamic Light Scattering).....</i>	127
5.3	<i>Resultados e discussão.....</i>	128

5.3.1	<i>Produção da proteína recombinante rADEL em sistema heterólogo de E. coli (Brasil).....</i>	128
5.3.1.1	<i>Desenho do gene sintético.....</i>	128
5.3.1.2	<i>Testes de expressão.....</i>	129
5.3.2	<i>Produção da proteína recombinante rADEL em sistema heterólogo (França)..</i>	131
5.3.2.1	<i>Subclonagens do gene sintético para os vetores de expressão.....</i>	131
5.3.2.2	<i>Produção da proteína rADEL em sistema heterólogo de Komagataella pastoris..</i>	132
5.3.2.2.1	<i>Transformação do plasmídeo rADELpPICZαHisTEV em K. pastoris.....</i>	134
5.3.2.2.2	<i>Expressão da lectina rADEL em K. pastoris.....</i>	135
5.3.2.2.3	<i>Purificação da lectina recombinante rADEL a partir do concentrado celular.....</i>	144
5.3.2.3	<i>Produção da proteína rADEL em sistema heterólogo de E. coli.....</i>	148
5.3.2.3.1	<i>Expressão da lectina rADEL no vetor pET25b(+)......</i>	149
5.3.2.3.2	<i>Expressão da lectina rADEL no vetor pCOLD-HAT-TEV-3G-Nco.....</i>	151
5.3.3	<i>Ensaio de cristalização (França).....</i>	156
5.3.4	<i>Produção da proteína recombinante rADEL em sistema heterólogo de K. pastoris (Brasil).....</i>	156
5.3.4.1	<i>Transformação de E. coli com o gene sintético.....</i>	156
5.3.4.2	<i>Transformação de K. pastoris com o gene sintético.....</i>	157
5.3.4.3	<i>Expressão da lectina rADEL em K. pastoris.....</i>	160
5.3.5	<i>Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil).....</i>	171
5.3.5.1	<i>Estimativa da massa molecular.....</i>	172
5.3.5.2	<i>Ensaio de atividade hemaglutinante da proteína recombinante.....</i>	173
5.3.5.3	<i>Ensaio de inibição da atividade hemaglutinante.....</i>	173
5.3.5.4	<i>Dicroísmo circular.....</i>	177
5.3.5.5	<i>Espalhamento dinâmico de luz - DLS (Dynamic Light Scattering).....</i>	179
5.4	<i>Conclusão.....</i>	183
6	CAPÍTULO III: EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA LECTINA rAmL PRESENTE NA ALGA MARINHA <i>Amansia multifida</i>.....	184
6.1	<i>Objetivos.....</i>	184
6.2	<i>Metodologia.....</i>	185
6.2.1	<i>Coleta dos exemplares e purificação da lectina nativa.....</i>	185

6.2.2	<i>Produção da proteína recombinante rAmL em sistema heterólogo (França).....</i>	185
6.2.2.1	<i>Desenho do gene sintético francês.....</i>	185
6.2.2.1.1	<i>Subclonagens do gene sintético para os vetores de expressão.....</i>	187
6.2.2.2	<i>Expressão da lectina rAmL em sistema heterólogo a partir da construção gênica francesa.....</i>	190
6.2.2.3	<i>Purificação da lectina recombinante.....</i>	191
6.2.2.4	<i>Digestão da proteína fusionada e repurificação da lectina recombinante.....</i>	193
6.2.3	<i>Avaliação da pureza e funcionabilidade da proteína recombinante (França)....</i>	194
6.2.3.1	<i>Estimativa da massa molecular e avaliação da pureza da lectina.....</i>	194
6.2.3.2	<i>Atividade hemaglutinante da proteína.....</i>	194
6.2.4	<i>Ensaio de deslocamento térmico (TSA – Thermal Shift Assay) (França).....</i>	194
6.2.5	<i>Ensaio de cristalização (França).....</i>	195
6.2.6	<i>Produção da proteína recombinante rAmL em sistema heterólogo (Brasil).....</i>	196
6.2.6.1	<i>Desenho do gene sintético adquirido no Brasil.....</i>	196
6.2.6.2	<i>Recuperação do gene sintético francês no Brasil.....</i>	198
6.2.6.3	<i>Expressão da lectina rAmL em sistema heterólogo a partir das duas construções gênicas.....</i>	198
6.2.6.3.1	<i>Transformação bacteriana.....</i>	198
6.2.6.3.2	<i>Testes de expressão.....</i>	200
6.2.6.4	<i>Purificação da proteína rAmL.....</i>	201
6.2.6.5	<i>Digestão da proteína fusionada e repurificação da lectina recombinante.....</i>	201
6.2.7	<i>Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil).....</i>	202
6.2.7.1	<i>Estimativa da massa molecular e avaliação da pureza da lectina.....</i>	202
6.2.7.2	<i>Ensaio de atividade hemaglutinante da proteína recombinante e nativa.....</i>	202
6.2.7.3	<i>Ensaio de inibição da atividade hemaglutinante com a proteína recombinante e nativa.....</i>	203
6.2.7.4	<i>Dicroísmo circular.....</i>	203
6.2.7.5	<i>Espalhamento dinâmico de luz - DLS (Dynamic Light Scattering).....</i>	204
6.2.8	<i>Atividade antibacteriana combinada com antibióticos (Brasil).....</i>	204
6.2.8.1	<i>Microorganismos e Condições de Cultivo.....</i>	205
6.2.8.2	<i>Avaliação do efeito de rAmLpCHT3GNco combinado aos antibióticos.....</i>	205
6.3	<i>Resultados e discussão.....</i>	207

6.3.1	<i>Produção da proteína recombinante rAmL em sistema heterólogo de E. coli (França).....</i>	207
6.3.1.1	<i>Subclonagens do gene sintético para os vetores de expressão.....</i>	207
6.3.1.2	<i>Expressão da lectina rAmL em sistema heterólogo.....</i>	208
6.3.1.2.1	<i>Expressão da lectina rAmL no vetor pCOLD-TF-TEV-Nco.....</i>	209
6.3.1.2.2	<i>Expressão da lectina rAmL no vetor pET-TEV.....</i>	215
6.3.1.2.3	<i>Expressão da lectina rAmL no vetor pCOLD-HAT-TEV-3G-Nco.....</i>	217
6.3.2	<i>Ensaio de deslocamento térmico (TSA – Thermal Shift Assay) (França).....</i>	229
6.3.3	<i>Ensaaios de cristalização (França).....</i>	231
6.3.4	<i>Produção da proteína recombinante rAmL em sistema heterólogo (Brasil).....</i>	232
6.3.4.1	<i>Desenho do gene sintético adquirido no Brasil.....</i>	232
6.3.4.2	<i>Expressão da lectina rAmL em sistema heterólogo a partir das duas construções gênicas.....</i>	234
6.3.4.2.1	<i>Transformação de E. coli com os genes sintéticos.....</i>	235
6.3.4.2.2	<i>Testes de expressão.....</i>	235
6.3.5	<i>Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil).....</i>	249
6.3.5.1	<i>Estimativa da massa molecular.....</i>	249
6.3.5.2	<i>Ensaio de atividade hemaglutinante da proteína recombinante e nativa.....</i>	250
6.3.5.3	<i>Ensaio de inibição da atividade hemaglutinante com a proteína recombinante e nativa.....</i>	252
6.3.5.4	<i>Dicroísmo circular.....</i>	255
6.3.5.5	<i>Espalhamento dinâmico de luz - DLS (Dynamic Light Scattering).....</i>	258
6.3.6	<i>Avaliação da atividade antibacteriana combinada com antibióticos (Brasil).....</i>	262
6.4	<i>Conclusão.....</i>	268
7	CAPÍTULO IV: CARACTERIZAÇÃO ESTRUTURAL DAS LECTINAS ALL E CCL PRESENTES NAS ESPONJAS MARINHAS <i>Aplysina lactuca</i> e <i>Chondrilla caribensis</i>.....	269
7.1	<i>Objetivos.....</i>	269
7.2	<i>Metodologia.....</i>	270
7.2.1	<i>Purificação das lectinas nativas.....</i>	270
7.2.2	<i>Determinação da estrutura tridimensional das lectinas ALL e CCL por cristalografia de raios X.....</i>	271

7.2.2.1	<i>Cristalização.....</i>	
7.2.2.2	<i>Difração de raios X e resolução da estrutura.....</i>	272
7.2.3	<i>Análise de calorimetria por titulação isotérmica.....</i>	274
7.3	Resultados e discussão.....	275
7.3.1	<i>Purificação das lectinas nativas ALL e CCL.....</i>	275
7.3.2	<i>Determinação da estrutura tridimensional da galectina ALL.....</i>	276
7.3.2.1	<i>Cristalização da ALL.....</i>	276
7.3.2.2	<i>Estrutura geral da lectina ALL.....</i>	279
7.3.2.3	<i>Caracterização do sítio de ligação e análise da interação ALL-Lactose.....</i>	284
7.3.3	<i>Determinação da estrutura tridimensional da galectina CCL.....</i>	286
7.3.3.1	<i>Cristalização da CCL.....</i>	286
7.3.3.2	<i>Estrutura geral da lectina CCL.....</i>	288
7.3.3.3	<i>Caracterização do sítio de ligação e análise da interação CCL-3SL.....</i>	293
7.3.3.4	<i>Análise estrutural comparativa entre galectinas de esponjas.....</i>	295
7.3.4	<i>Análise de calorimetria por titulação isotérmica.....</i>	298
7.4	Conclusão.....	301
8	CONCLUSÃO DA TESE.....	302
	REFERÊNCIAS.....	304

1 INTRODUÇÃO

1.1 Organismos marinhos como fonte de biomoléculas

O oceano possui uma singularidade devido à sua alta diversidade biológica, química e ecológica. Por abranger mais da metade da biodiversidade global e possuir condições ambientais bastante diferentes, com variações de salinidade, radiação, temperatura e pressão, esse ambiente apresenta uma abundância de compostos que muitas vezes não são encontrados em outros lugares (Riccio *et al.*, 2020).

As moléculas químicas responsáveis pelo sucesso da sobrevivência dos organismos marinhos possuem inúmeras atividades biológicas, pois a produção e acúmulo desses metabólitos estão intimamente ligados com a defesa desses animais. Devido a isso, organismos sésseis e com baixa movimentação, como algas, invertebrados e microrganismos marinhos representam a principal fonte para esses produtos bioativos (Rigogliuso *et al.*, 2023; Youssef, 2021).

A diversidade de estruturas químicas encontradas nos organismos marinhos é alta, com características estruturais e funcionais únicas, o que sugere aplicações diversas em diferentes setores industriais. Porém, são as inúmeras atividades biológicas descritas com propriedades antitumorais, antivirais, anti-inflamatórias, antibacterianas, antioxidantes e antiincrustantes, dentre muitas outras, que estão atraindo os olhares do mundo para essas moléculas (Chbel *et al.*, 2021).

Por conta da abundância e das inúmeras propriedades terapêuticas, esses produtos naturais oriundo do ambiente marinho vêm ganhando destaque na área da biotecnologia. Vários projetos no mundo estão focando na exploração de organismos marinhos para identificar novos compostos visando a aplicação em diferentes setores industriais, principalmente farmacêutico. Somado a isso, esses projetos procuram abordagens mais ecológicas para a descoberta de medicamentos sem nenhum impacto negativo no ambiente marinho (Sousa *et al.*, 2025; Lauritano; Ianora, 2018). Portanto, se faz necessário a bioprospecção de novas moléculas com aplicações farmacológicas, mas sempre buscando métodos sustentáveis.

O *habitat* marinho, apesar do enorme potencial para este fim e da maior abundância em biodiversidade, ainda permanece inexplorado na sua grande maioria, principalmente em

comparação com o ambiente terrestre, atestando uma das principais vantagens de alavancar os estudos de organismos pertencentes a esse ambiente (Rigogliuso *et al.*, 2023).

Essa lacuna significativa na bioprospecção marinha é devida principalmente as dificuldades de acesso e amostragem em ambientes remotos e com condições extremas de pressão, temperatura e profundidade, limitando a exploração dessa biodiversidade. A investigação do ambiente marinho necessita de conhecimentos sobre a geografia e condições ambientais nessas regiões mais profundas, o que ainda é insuficientemente investigada. Além disso, as limitações tecnológicas para coleta de organismos marinhos, principalmente em ambientes mais profundos, é um fator que representa um entrave no estudo e isolamento de compostos bioativos de organismos marinhos (Bekiari, 2023; Moutinho *et al.*, 2024).

1.2 Lectinas

Lectinas são proteínas ou glicoproteínas que apresentam um domínio de ligação específica e reversível a carboidratos, porém não alteram as propriedades estruturais e nem estão necessariamente envolvidas no metabolismo dos seus ligantes, além disso, estas proteínas não pertencem a qualquer classe de imunoglobulinas (El-Maradny *et al.*, 2021; Kilpatrick, 2000). Assim, a característica mais importante para se classificar uma proteína como lectina é a presença de pelo menos um domínio que reconheça e se ligue reversivelmente a um carboidrato (El-Maradny *et al.*, 2021).

Primeiramente foram chamadas de hemaglutininas ou fitoaglutininas por possuírem a capacidade de aglutinar células sanguíneas observada em extrato de plantas (Sharon; Lis, 2004). Apesar das lectinas de animais serem observadas antes das lectinas de vegetais, a denominação de “lectina” foi usada inicialmente para designar essas moléculas de origem vegetal, pois as lectinas de animais, apesar de descobertas primeiro, não eram identificadas como proteínas que se ligavam a carboidratos (Kilpatrick, 2000; 2002).

Essas proteínas têm potencial para aglutinar eritrócitos, linfócitos e células microbianas, devido a sua especificidade a carboidratos presentes na superfície das células. A habilidade de aglutinar células sanguíneas foi por muito tempo determinante para o descobrimento de novas lectinas, e foi essa capacidade de distinguir entre os diferentes tipos de eritrócitos que nomeou essas proteínas com o termo atual lectina (do latim “legere” – escolher) (Sharon; Lis, 2004; Boyd; Shapleigh, 1954).

As lectinas estão presentes em todos os organismos, dos mais simples ao mais complexos, incluindo vírus, bactérias, fungos, algas, plantas e animais, e elas desempenham papel central em diversos processos biológicos, tais como defesa, simbiose, agregação celular e reprodução (Elumalai *et al.*, 2019).

As lectinas de animais estão presentes em todos os filos e possuem uma ampla variedade de funções, apresentando papéis cruciais em muitos processos biológicos como adesão celular, ligação entre células, diferenciação, sinalização celular e simbiose com microrganismos (Brooks, 2017). As lectinas podem estar localizadas em compartimentos intracelulares, mediando processos importantes como o *splicing* do RNA no núcleo até o enovelamento de proteínas e tráfego de glicoproteínas. Além disso, podem ser liberadas para o compartimento extracelular e permanecer na superfície celular ou estar presente nos fluidos biológicos como componentes solúveis (Wang *et al.*, 2020; Vasta *et al.*, 2011).

Entretanto, o principal papel dessas proteínas parece estar relacionado ao sistema imune inato. As lectinas atuam como moléculas de reconhecimento sendo capazes de distinguir entre as células próprias do organismo e células estranhas. Justamente essa propriedade é o princípio fundamental do sistema imunológico (Viswambari; Basilrose; Mercy, 2010).

Como o sistema de defesa dos invertebrados não envolve anticorpos, as lectinas encontradas no fluido desse grupo estão no centro da defesa desses animais e parecem agir como anticorpos naturais fornecendo uma defesa direta, além de atuarem no reconhecimento e tráfego dentro do sistema imunológico, regulação imune e prevenção de autoimunidade (Vasta; Wang, 2020).

Muitas lectinas encontradas tanto na hemolinfa de invertebrados como em vertebrados parecem agir como anticorpos fornecendo uma primeira linha de defesa, ligando e neutralizando patógenos e promovendo a fagocitose pelos hemócitos. Elas podem agir independentemente, aglutinando partículas estranhas ou podem agir como opsoninas, se ligando ao açúcar na superfície celular das células estranhas e facilitando a fagocitose ou encapsulamento pelas células de defesa (Vasta; Wang, 2020; Viswambari; Basilrose; Mercy, 2010). Além disso, essas proteínas são capazes de ativar reações em cascata, estimulando a secreção de moléculas diversas, tais como proteases serínicas, interleucinas, entre outras, que irão atuar na defesa do organismo (El-Maradny *et al.*, 2021).

Essas proteínas são consideradas um dos principais grupos de receptores de reconhecimento de padrões (*Pattern Recognition Receptors* - PRRs) que são responsáveis por identificar padrões moleculares associados a patógenos (*Pathogen-associated Molecular*

Patterns - PAMPs) expostos na superfície dos mesmos, como o peptidoglicano (PGN) na parede da célula bacteriana. Elas atuam reconhecendo células anormais e patógenos ligando-se às porções de açúcares presentes em suas superfícies através dos locais de ligação a carboidratos (Vasta; Wang, 2020; Wang; Wang, 2013).

A ativação do sistema complemento é resultante da ligação aos PAMPs, pois o DRC (Domínio de Reconhecimento a Carboidratos) das lectinas ajuda na identificação e na ligação de glicoproteínas, glicolipídeos e outros resíduos de açúcar presentes na superfície das células dos patógenos, provocando dessa forma a resposta imune (Vasta *et al.*, 2011). Essas células envolvidas na imunidade migram entre o sangue e a linfa para responder a sinais de invasão por patógenos. Atuam no papel de saída e recirculação de linfócitos, promovendo o extravasamento de neutrófilos para os locais de inflamação (Vasta; Wang, 2020).

A capacidade de reconhecer, aglutinar, imobilizar, levar a opsonização e fagocitose de possíveis patógenos, até o encapsulamento e lise de células estranhas, juntamente com a capacidade de ativar o sistema complemento, evidenciam algumas das ações efetivas que as lectinas exercem no sistema imune dos animais (Elumalai *et al.*, 2019).

As lectinas de algas são principalmente encontradas no filo Rhodophyta, devido ao maior número de espécies apresentando atividade hemaglutinante durante os trabalhos de levantamento/bioprospecção e à ausência de compostos que causam uma hemaglutinação falso-positivo, como é observada em algas pardas (Costa Filho *et al.*, 2025).

Diferentemente das lectinas de animais, em algas, essas moléculas ainda não possuem uma função biológica clara. Estudos já relataram o envolvimento dessas proteínas nos processos de agregação de organelas, regeneração de protoplastos e defesa contra bactérias em algas verdes (Klochkova *et al.*, 2005; Kim; Klochkova; West, 2002). Em algas vermelhas, as lectinas foram associadas ao processo de reprodução ao reconhecer os gametas feminino e masculino (Shim *et al.*, 2012).

Apesar das funções biológicas de lectinas de algas ainda não estarem claras, sabe-se que da mesma forma das lectinas de animais, elas desencadeiam processos fisiológicos por meio do reconhecimento dos resíduos de açúcar presentes na superfície das células (Elumalai *et al.*, 2019).

1.3 Lectinas e o reconhecimento a carboidratos

1.3.1 Carboidratos

Os glicanos são a classe mais abundante e diversa de biomoléculas na natureza, junto com proteínas e ácidos nucleicos, são os blocos de construção essenciais da vida e o repertório completo de glicanos de uma célula é chamado de glicoma (Reily *et al.*, 2019).

Os carboidratos apresentam uma grande diversidade estrutural devido à natureza e composição de monossacarídeos, à ligação entre esses monossacarídeos, à sequência das unidades constituintes, ao estado anomérico, ao número e tipo de ramificações das cadeias de hidratos de carbono e à posição das ligações glicosídicas. Além disso, os glicanos podem ganhar modificações, incluindo sulfatação e fosforilação, por exemplo, que aumentam ainda mais essa variação e é essa diversidade que faz com que essas moléculas desempenhem uma ampla gama de funções biológicas (Reily *et al.*, 2019).

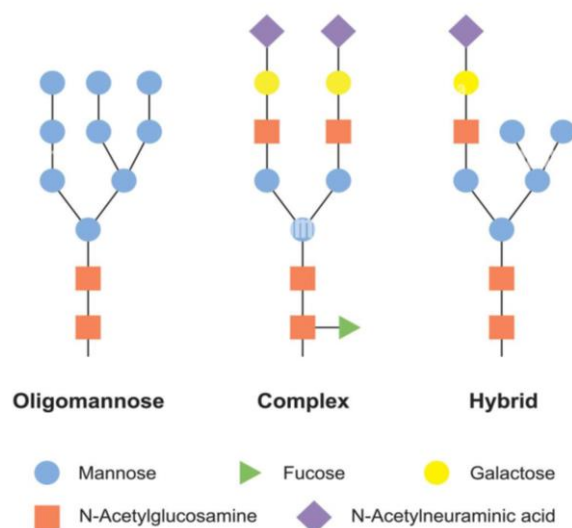
Os glicoconjugados são moléculas de carboidratos encontradas ligadas a proteínas, lipídios ou outros compostos revestindo o exterior das células da maioria dos organismos. Essas moléculas complexas são formadas por um processo pós-traducional chamado de glicosilação que ocorre no retículo endoplasmático e no aparelho de Golgi, e podem diferir em suas sequências de glicanos, nas conexões entre eles e em seu comprimento. Esses carboidratos desempenham papéis fundamentais em mecanismos biológicos através do reconhecimento de interações celulares e modulação de numerosos processos fisiológicos e patológicos (Tommasone *et al.*, 2019; Reily *et al.*, 2019).

Os oligossacarídeos (carboidratos que resultam da ligação glicosídica de mais de dois monossacarídeos) podem formar ligações glicosídicas com proteínas por dois tipos principais de ligações: oligossacarídeos ligados a *N* e oligossacarídeos ligados a *O*. A glicosilação do tipo *N* é caracterizada pela ligação da *N*-acetilglucosamina (GlcNAc) ao átomo de nitrogênio da cadeia lateral amida do aminoácido asparagina (Asn). Esse tipo de glicosilação só ocorre quando a asparagina está na sequência Asn-X-Ser/Thr (Serina/Treonina), onde X pode ser qualquer aminoácido menos prolina (Pro). Essa glicosilação é necessária para o enovelamento adequado de algumas proteínas, porém ela não ocorre obrigatoriamente se uma proteína possuir a presença desta sequência comum (Ghazarian; Idoni; Oppenheimer, 2011).

Os *N*-glicanos podem ser divididos em três classes: oligossacarídeos com alto teor de manose, oligossacarídeos complexos e oligossacarídeos híbridos (Figura 01). Apesar dessa diferenciação, todos os *N*-glicanos possuem uma estrutura básica central: um pentassacarídeo formado por dois resíduos de *N*-acetilglucosamina e três resíduos de manose (Man₃GlcNAc₂), chamado de núcleo. Os *N*-glicanos de alto teor de manose são aqueles que possuem vários

resíduos de manose em suas ramificações. Os *N*-glicanos complexos possuem resíduos de *N*-acetilglucosamina nas suas ramificações, podendo ter duas, três ou quatro antenas (ramificações) a partir do núcleo comum. E os *N*-glicanos híbridos possuem uma ramificação contendo resíduos de manose e outra ramificação contendo *N*-acetilglucosamina. Outros monossacarídeos como *N*-acetilglucosamina, *N*-acetilgalactosamina (GalNAc), galactose, fucose e ácido siálico podem ser adicionados alongando esses oligossacarídeos ligados a *N* (Reily *et al.*, 2019).

Figura 01 – Tipos de *N*-glicanos.
N-glycans Types



Fonte: STOCK, 2023.

A glicosilação do tipo *O* é a mais diversa e envolve a ligação do glicano ao átomo de oxigênio da hidroxila do aminoácido serina ou treonina e, assim como na glicosilação do tipo *N*, outros monossacarídeos podem ser adicionados ao açúcar inicial ligado à proteína (Ghazarian; Idoni; Oppenheimer, 2011). Os *O*-glicanos podem ser classificados com base no açúcar inicial ligado à proteína e nas estruturas de açúcar adicionadas ao glicano inicial. Os *O*-glicanos do tipo mucina possui o *N*-acetilgalactosamina (GalNAc) ligado ao aminoácido Ser/Thr da proteína e estes ainda podem ser classificados com base nos glicanos ligados ao GalNAc inicial, produzindo diferentes estruturas terminais que geralmente são fucosiladas (adição de fucose) e sialiladas (adição de ácido siálico). Outros tipos de *O*-glicanos são *O*-manose (manose ligada a Ser/Thr), *O*-fucose, *O*-galactose, *O*-GlcNAc (Pinho; Reis, 2015; Reily *et al.*, 2019).

Os glicanos estão envolvidos nos principais eventos biológicos, apresentando funções metabólicas, estruturais e físicas. Os carboidratos podem agir no reconhecimento extrínseco através de PAMPs, modulação do sistema imune do hospedeiro, captação e processamento de antígenos. Já no reconhecimento intrínseco participam da adesão e sinalização intercelular, interação célula-matriz, fertilização e reprodução, endocitose e fagocitose, dentre muitas outras funções (Varki, 2017).

Os glicanos também modulam eventos de reconhecimento hospedeiro-patógeno, além de contribuir para o enovelamento adequado de proteínas. Assim, o reconhecimento molecular de glicanos está no centro de processos patológicos, como adesão bacteriana, infecção viral, inflamação ou ativação do sistema imunológico. Devido a isso, diferenciações na glicosilação normal das células levam ao aparecimento e desenvolvimento de doenças (Smith; Bertozzi, 2021).

Essas informações biológicas que os glicanos guardam são chamadas de “código do açúcar” ou glicocódigo. Decifrar e compreender essas informações é um grande desafio da pesquisa biológica atual, uma vez que os carboidratos são estruturalmente muito complexos e como são moléculas que não são codificadas no genoma, suas funções são pouco compreendidas quando comparadas com proteínas, ácidos nucleicos e lipídeos. Ainda assim, sabe-se que os efeitos dos glicanos/glicoconjugados são dependentes do reconhecimento de moléculas que podem se ligar a essas estruturas com seletividade e afinidade (Brooks, 2017).

1.3.2 Especificidade de ligação das lectinas

Os carboidratos são capazes de transportar a informação biológica e as lectinas são moléculas capazes de decifrar o glicocódigo, pois se ligam a eles de forma específica e assim disparam diversas respostas. As interações entre proteínas e carboidratos estão intimamente relacionadas com as atividades biológicas, pois a glicosilação da superfície celular é a chave para a função das lectinas, e uma alteração mínima na estrutura do glicano pode ter grandes impactos nos processos biológicos associados. Por conta disso, o reconhecimento proteína-carboidrato é a solução para compreender as funções dessas moléculas e as mutações que levam a ocorrência de doenças, fornecendo mecanismos para aplicações biotecnológicas, diagnósticas e estratégias terapêuticas (Gabiús *et al.*, 2022).

As lectinas geralmente se ligam a açúcares simples (monossacarídeos) presentes na posição terminal nas cadeias de glicanos. Porém a especificidade não depende apenas dessas

moléculas de açúcar terminal, mas também do tipo de estrutura subterminal, do número e da distribuição dos sítios receptores e do impedimento estérico causado pelas moléculas circundantes (Gorakshakar; Ghosh, 2016).

O local de ligação a carboidratos nas lectinas (DRC) é tão especializado que permite a ligação apenas com o glicano específico, o que leva a uma especificidade muito alta. A afinidade entre o DRC da lectina e o carboidrato é muitas vezes baixa, com valores da constante de dissociação (K_D) na faixa de micromolar e milimolar, no entanto, essa afinidade é aumentada pela multivalência da lectina, onde uma lectina se oligomeriza para possuir múltiplos DRC. E muitas vezes essa multivalência é necessária para que a lectina possa desempenhar seu papel biológico com êxito. No caso de vários oligossacarídeos expostos na superfície celular, cada um pode ocupar um domínio da lectina, aumentando assim a interação. Além disso, quando as células expressam múltiplas lectinas, a força da ligação é intensificada (Nelson, Cox, 2022).

Outras interações menos específicas também contribuem para a ligação lectina-carboidrato. Os glicanos possuem um lado mais polar e outro mais apolar, o lado mais polar forma ligações de hidrogênio com a lectina, enquanto o lado mais apolar forma interações hidrofóbicas com resíduos de aminoácidos apolares. Assim, todas essas interações aumentam a afinidade de ligação e garante a elevada especificidade das lectinas a seus carboidratos, permitindo a transferência de informações entre essas duas moléculas que é de extrema importância em inúmeros processos biológicos (Nelson; Cox, 2022).

Como o reconhecimento de carboidratos pelas lectinas desempenham papéis diversos em inúmeros processos, compreender o que faz com que uma determinada lectina reconheça um glicano específico em meio a uma enorme variedade de outros açúcares e a natureza dessa interação pode nos ajudar a entender melhor esses processos e como monitorá-los e controlá-los (Brooks, 2017).

1.4 Potencial biotecnológico das lectinas

A especificidade dos carboidratos das lectinas são a base para a suas atividades, as quais incluem: ação mitogênica, modulação da imunidade, reconhecimento e eliminação de patógenos, detecção e destruição de células cancerígenas, modulação da atividade neural, potencial antiinflamatório, dentre muitas outras (Singh; Walia, 2018; Gardères *et al.*, 2015).

Devido à capacidade de se ligar a carboidratos presentes na superfície celular de patógenos, muitas lectinas são capazes de inibir a proliferação de vírus, fungos, protozoários e

bactérias. A interação da lectina com glicanos da superfície bacteriana pode resultar na aglutinação das bactérias, na morte celular por destruição da parede celular bacteriana pela alteração morfológica e formação de poros, ou impedindo a adesão das bactérias às superfícies, inibindo a formação de biofilme e a migração bacteriana (El-Maradny *et al.*, 2021).

Essa importante propriedade bactericida pode ser usada para desenvolver antibióticos contra infecções bacterianas resistentes. Desde que bactérias super-resistentes são causadoras de milhares de mortes anuais, o desenvolvimento de novas drogas terapêuticas se tornou uma questão que demanda urgência (Cheung *et al.*, 2015).

Os patógenos que desenvolveram resistência a um antimicrobiano específico podem não desenvolver resistência a seus derivados naturais com ações semelhantes, colocando os medicamentos antimicrobianos produzidos por espécies marinhas em relevância (Cheung *et al.*, 2015). Além disso, as bactérias também encontraram estratégias para vencer as defesas imunológicas do hospedeiro. Por isso, a descoberta de novos medicamentos que possam modular a resposta imune do hospedeiro após uma infecção bacteriana sem atingir necessariamente o patógeno, também é de grande importância (Mishra *et al.*, 2019).

As lectinas podem também interagir com quitina, carboidrato presente na parede celular dos fungos, induzindo alterações morfológicas que afetam a absorção de nutrientes, crescimento e formação de esporos. Além disso, elas podem penetrar na parede dos fungos e inibir a ação de enzimas vitais (Jandú *et al.*, 2017).

De forma geral, a ligação da lectina com os glicanos presentes na parede celular dos mais diversos parasitas pode impedir o ciclo de vida dos mesmos e ativar o sistema imunológico, induzindo a ação de anticorpos contra o parasita (Mishra *et al.*, 2019; Jandú *et al.*, 2017).

Como as lectinas mostram alta afinidade por estruturas específicas de carboidratos, elas podem ser usadas como sondas específicas para glicoconjugados da superfície celular em aplicações histológicas. Algumas moléculas devido a sua estrutura menos complexa não desencadeiam a produção de anticorpos no organismo, o que vem colocando as lectinas em uso nos estudos histoquímicos ao invés de anticorpos tradicionais (Cheung *et al.*, 2015). Por conta disso, essas proteínas são usadas em várias aplicações farmacêuticas como diagnóstico histoquímico, administração de vacinas e controle da liberação de drogas (Meiers *et al.*, 2019; Juan *et al.*, 2017).

No direcionamento de drogas para o local de ação, as lectinas se mostraram efetivas pois são capazes de diferenciar o tecido normal do tecido doente, devido a alteração de glicanos

que ocorre nas células anormais. Assim, essas moléculas podem transportar e direcionar a droga para locais específicos (Juan *et al.*, 2017).

As lectinas podem ser ferramentas importantes para obter informações sobre o desenvolvimento, estágio de diferenciação e capacidade de metástase das células tumorais, pois as células transformadas apresentam variações quantitativas e qualitativas no padrão de glicosilação da membrana e possuem aglomerados de carboidratos que ocupam áreas restritas da superfície celular, enquanto que, nas células normais, essa distribuição de carboidratos é homogênea (Catanzaro *et al.*, 2020; Pinho; Reis, 2015).

Portanto, as lectinas podem ser usadas no diagnóstico como um marcador, identificando as células transformadas, no mapeamento dos tecidos suspeitos de neoplasias e no estudo de estruturas glicídicas aberrantes sintetizadas pelas células tumorais, além de modular a apoptose por diferentes vias. Apresentando um grande potencial no tratamento, prevenção e diagnósticos do câncer (Cheung *et al.*, 2015; Brooks, 2017).

No campo da ciência farmacêutica, as lectinas também vem ganhando destaque na efetiva ação contra um amplo espectro de vírus devido a sua especificidade a carboidratos presentes nos envelopes virais, se tornando um dos agentes antivirais mais promissores, pois os antivirais tradicionais possuem ação de suprimir o ciclo de vida viral, enquanto que as lectinas previnem a entrada do vírus na célula hospedeira e evitam a propagação viral (Singh; Walia, 2018; Cheung *et al.*, 2015).

Devido ao elevado potencial biotecnológico já comprovado, as lectinas vêm se consolidando na área de estudos da saúde humana, (El-maradny *et al.*, 2021). E os organismos marinhos, tornaram-se o foco de pesquisas sobre aplicações biotecnológicas de lectinas, pois essas moléculas, possuem menor imunogenicidade quando comparadas às lectinas terrestres, por possuírem menor peso molecular, facilitando seu uso para fins farmacêuticos (Cheung *et al.*, 2015).

As lectinas oriundas dos filos Porifera, Mollusca e Echinodermata estão ganhando destaque como moléculas anticancerígenas (Nascimento-Neto *et al.*, 2018; Datta; Talapatra; Swarnakar, 2016) e antimicrobiana, sendo eficazes contra bactérias e protozoários (Marques *et al.*, 2018; Zhang *et al.*, 2020; Wang *et al.*, 2011; Sousa *et al.*, 2021).

Além disso, essas moléculas possuem ação mitogênica contra células do sistema imunológico e quimiotaxia em neutrófilos (Dresh *et al.*, 2012; Xue *et al.*, 2015; Datta; Talapatra; Swarnakar, 2016). Essa quimiotaxia poderia representar um potencial pró-

inflamatório e ser usado como ferramenta molecular no estudo de eventos celulares relacionados à inflamação (Dresch *et al.*, 2008).

Dessa forma, as lectinas desses invertebrados podem atuar como opsoninas, ativando a multiplicação de células imunes, atraindo neutrófilos, estimulando a produção de citocinas e levando a sua endocitose. Apesar das esponjas não possuírem células imunes especializadas, suas lectinas são capazes de atuar como imunomoduladoras em vertebrados, uma propriedade que pode ser usada contra doenças que afetam o sistema imunológico (Gardères *et al.*, 2015).

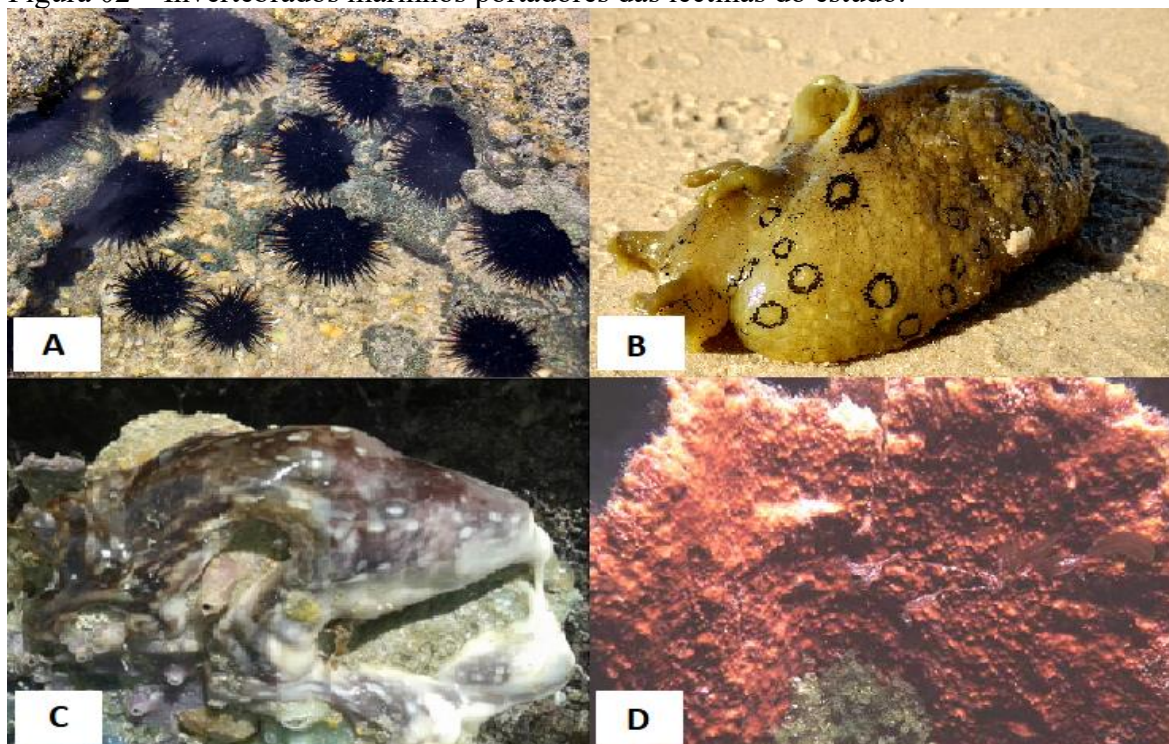
As lectinas de algas marinhas ganham muito destaque com suas propriedades antivirais. Elas são capazes de interromper as infecções causadas pelo vírus da imunodeficiência humana (HIV), evidenciada por pesquisas bastante avançadas e promissoras como a lectina GRFT da alga marinha vermelha *Griffithsia* sp. (Barton *et al.*, 2016). Além disso, elas possuem ação contra os vírus da hepatite, do herpes, coronavírus, dentre outros (Meuleman *et al.*, 2011; Levendosky *et al.*, 2015; Millet *et al.*, 2016).

Somada a ação antiviral, essas lectinas também apresentam ação anticâncer (Anam *et al.*, 2017), anti-inflamatória e antinociceptiva (Abreu *et al.*, 2016), antimicrobiano (Boonsri *et al.*, 2017) e são capazes de promover a cicatrização (Nascimento-Neto *et al.*, 2012).

1.5 Lectinas objetos do estudo

Quatro lectinas em estudo são oriundas de invertebrados marinhos (Figura 02). A ELEL (*Echinometra lucunter* eggs lectin) foi isolada dos ovos do ouriço-do-mar *Echinometra lucunter*, pertencente ao filo Echinodermata. ADEL (*Aplysia dactylomela* eggs lectin) foi purificada dos ovos da lesma-do-mar *Aplysia dactylomela*, invertebrado que pertence ao filo Mollusca. E as duas lectinas CCL (*Chondrilla caribensis* lectin) e ALL (*Aplysina lactuca* lectin) são oriundas das esponjas marinhas *Chondrilla caribensis* e *Aplysina lactuca*, pertencentes ao filo Porifera.

Figura 02 – Invertebrados marinhos portadores das lectinas do estudo.



Fonte: Autora. Adaptado de ALÉSSIO, 2023; SOUSA, 2021; PINHEIRO; HAJDU; CUSTODIO, 2007.

A: *Echinometra lucunter*, B: *Aplysia dactylomela*, C: *Chondrilla caribensis* e D: *Aplysina lactuca*.

A quinta lectina em estudo, AmL (*Amansia multifida* lectin) é oriunda da alga marinha vermelha *Amansia multifida* (Figura 03).

Figura 03 – Alga marinha *Amansia multifida*.



Fonte: AlgaeBase. (GUIRY; GUIRY, 2025).

A lectina da *Chondrilla caribensis* é uma proteína homotetramérica formada por subunidades de 15.445 ± 2 Da, não possuindo cisteína em sua estrutura (Marques *et al.*, 2018). ALL é formada por um monômero de aproximadamente 15 kDa que se oligomeriza formando

uma proteína homotetramérica, e que em condições não redutoras ocorre a formação de ponte dissulfeto entre as subunidades, apresentando uma massa molecular de 31.810 ± 2 Da (Carneiro *et al.*, 2017a; Duarte *et al.*, 2023).

Ambas as lectinas de esponjas marinhas, CCL e ALL, fazem parte do restrito grupo de lectinas de esponjas com a estrutura primária elucidada. As duas lectinas tiveram sua sequência de aminoácidos determinada por espectrometria de massas em tandem (MS/MS) e pertencem a família das galectinas. As estruturas tridimensionais das duas lectinas foram preditas por programas computacionais, e a interação lectina carboidrato também foi analisada por ancoragem molecular (Sousa *et al.*, 2021; Duarte *et al.*, 2023).

A lectina dos ovos de *Aplysia dactylomela*, ADEL, possui uma massa molecular determinada de 57.228 ± 2 Da e teve sua estrutura primária resolvida por MS/MS totalizando uma massa de 24.450 Da. A lectina possui 11 cisteínas envolvidas em cinco ligações dissulfeto intracadeias e uma intercadeia e possui três glicanos de massas diferentes ligados à sua estrutura, sendo, portanto, uma proteína glicosilada. ADEL teve sua estrutura tridimensional predita por programas computacionais, e ainda não pertence a nenhuma família definida (Carneiro *et al.*, 2017b).

ELEL é um homodímero ligado por uma ponte dissulfeto com subunidades de aproximadamente 11 kDa cada, que teve sua estrutura primária determinada por MS/MS e a estrutura 3D, assim como o sítio de reconhecimento a carboidratos da lectina, foram preditos por programas computacionais. Cada subunidade de ELEL possui nove cisteínas envolvidas em quatro pontes dissulfeto intracadeias e uma ponte dissulfeto intercadeia. Aminoácidos conservados em lectinas de ligação a ramnose (RBLs) foram encontrados na estrutura (Carneiro *et al.*, 2015).

AmL é uma proteína monomérica pertencente à família OAAH, com a estrutura primária determinada por combinação de MS/MS e clonagem de cDNA. A lectina possui quatro domínios repetidos, totalizando 270 aminoácidos, com aproximadamente 29 kDa. AmL não possui pontes dissulfeto em sua estrutura (Silva, 2016).

1.5.1 Potencial biotecnológico das lectinas em estudo

As proteínas no foco deste trabalho são bastante promissoras na área da saúde. A lectina da esponja marinha *Chondrilla caribensis* (CCL) possui ação antibacteriana pois é capaz de reduzir significativamente a biomassa total em biofilmes de *Staphylococcus aureus*,

Staphylococcus epidermidis e *Escherichia coli*, além de ser capaz também de aglutinar células de *S. aureus*, *S. epidermidis* e *E. coli* (Marques *et al.*, 2018). CCL também é uma molécula leishmanicida por ser capaz de inibir o crescimento dos promastigotas do protozoário *Leishmania infantum* induzindo a produção de espécies reativas de oxigênio (Sousa *et al.*, 2021). Além disso, mostrou resultados promissores contra células cancerígenas por ser citotóxica em células de câncer de mama e de próstata, sem afetar células saudáveis (Sousa, 2021).

ALL da esponja marinha *Aplysina lactuca* é capaz de aglutinar bactérias gram-negativas e gram-positivas e reduzir a biomassa do biofilme bacteriano de maneira dose-dependente (Carneiro *et al.*, 2017a). Além disso, a lectina mostrou efeito sinérgico quando associada aos antibióticos tetraciclina e ampicilina, sendo capaz de aumentar seus efeitos bactericida contra cepas de *Staphylococcus aureus*, *Staphylococcus epidermidis* e *Escherichia coli* (Duarte *et al.*, 2023). ELET do invertebrado *Echinometra lucunter* também é uma proteína antibacteriana por ser capaz de aglutinar *S. aureus* (Carneiro *et al.*, 2015), assim como a ADEL de *Aplysia dactylomela* que além de aglutinar células gram-positivas também foi capaz de reduzir a formação de biomassa no biofilme bacteriano (Carneiro *et al.*, 2017b).

A lectina de *A. multifida* foi capaz de reduzir a viabilidade celular de células de câncer colorretal e de reduzir o número de células viáveis de *S. aureus* e *E. coli* na formação de biofilme bacteriano. Além disso, apresentou atividade contra o vírus HIV e demonstrou ter propriedade antioxidante (Silva, 2016). Neves e colaboradores (2007) também destacaram a propriedade antinociceptiva dessa lectina, por conta da migração de neutrófilos *in vivo* e *in vitro*.

Devido ao enorme potencial dessas lectinas, os estudos estruturais, moleculares e de interação proteína-ligante são fundamentais para complementar e permitir um futuro promissor para o uso dessas biomoléculas no âmbito da saúde humana.

2 JUSTIFICATIVA

Nos últimos anos, o interesse em lectinas de organismos marinhos tem aumentado, devido ao potencial destas proteínas em aplicações diversas no campo da saúde humana (Sivakamavalli *et al.*, 2021), evidenciando que a biodiversidade marinha representa uma rica fonte para a busca de novos compostos com potencial terapêutico (Catanzaro *et al.*, 2020).

Para os estudos de lectinas e o uso delas nas mais diferentes áreas atingir o seu potencial, a interação lectina-carboidrato precisa ser bem caracterizada e as funções biológicas envolvidas nessa interação ainda são pouco compreendidas. Poucas lectinas de organismos marinhos foram caracterizadas experimentalmente em relação à sua caracterização estrutural e à especificidade de ligação (Mattox; Bailey-Kellogg, 2021).

Portanto é necessário o enfoque em conjunto dos estudos estruturais e de glicobiologia das lectinas de organismos marinhos para elucidar a estrutura dessas moléculas. Os estudos estruturais de lectinas são importantes para o entendimento do mecanismo de reconhecimento a carboidratos e a especificidade a ligantes dessas proteínas, para então compreender o papel fisiológico dessas moléculas e fornecer informações para o uso biotecnológico. Além de ajudar nos estudos taxonômicos e de evolução das espécies, através da identificação de características estruturais comuns, e na classificação dessas proteínas em famílias (Fujimoto; Tatenos; Hirabayash, 2014; Sharon; Lis, 2004; El-Maradny *et al.*, 2021).

As lectinas de plantas foram as primeiras a possuírem suas estruturas primárias e tridimensionais determinadas, e apesar dos estudos estruturais de lectinas de animais terem sido impulsionados há décadas, poucas lectinas de algas e invertebrados marinhos possuem sua estrutura primária totalmente determinada e são ainda mais escassas as informações disponíveis sobre as características de suas estruturas, principalmente tridimensionais, em comparação com as lectinas oriundas de organismos terrestres (Singh; Walia, 2018; Gardères *et al.*, 2015). Como suas funções são baseadas no reconhecimento de ligantes, a determinação de suas estruturas tridimensionais é de importante relevância por fornecerem uma base sólida para caracterização e estudos lectina-carboidrato (Nelson; Cox, 2022).

Apesar do filo Porífera ser o grupo de animais que mais contribui com novos produtos bioativos isolados nos últimos anos (Youssef, 2021), as esponjas pertencem a um dos filos de invertebrados que menos possuem dados sobre a estrutura tridimensional de lectinas, demonstrando ser um grupo com conhecimentos escassos na área de biologia estrutural e que necessita impulsionar os estudos neste segmento.

Nesse sentido, estratégias biotecnológicas voltada à produção de proteínas recombinantes são importantes para estudos estruturais e para a produção em larga escala de proteínas com enfoque biotecnológico, proporcionando maiores rendimentos com funcionabilidade. As evidências do potencial biotecnológico das lectinas vem aumentando o interesse na purificação e caracterização dessas moléculas, porém estudos estruturais e funcionais mais específicos requerem a produção de quantidades suficientes de proteínas, e a produção pelas fontes naturais possui muitas desvantagens (Martínez-Alarcón; Blanco-Labra; García-Gasca, 2018).

A extração de lectinas por técnicas convencionais é um processo laborioso exigindo uma grande quantidade de biomassa; além de apresentar um rendimento baixo e possibilidade de apresentar impurezas, pois as lectinas purificadas podem ser contaminadas com classes de (bio)moléculas; além da variação da mesma lectina que pode acontecer devido ao isolamento em organismos diferentes, causando uma heterogeneidade nas propriedades de ligação da proteína (Oliveira; Teixeira; Domingues, 2013).

Além disso, existe a presença de isoformas da mesma proteína que devido à heterogeneidade genética da fonte, as propriedades físico-químicas, funcionais e atividades biológicas das lectinas acabam sendo determinadas para a mistura de proteínas, porque é difícil separar essas isolectinas usando técnicas convencionais. O problema é que as isoformas podem possuir características diferentes, incluindo atividade biológica, e essa variabilidade indesejada gera dificuldades no uso como ferramentas biomédicas (Oliveira; Teixeira; Domingues, 2013).

Assim, uma das alternativas é a produção dessas moléculas de forma recombinante e produzir proteínas com um maior rendimento e de forma pura, apresentando uma caracterização mais precisa e um espectro mais amplo de aplicação (Martínez-Alarcón; Blanco-Labra; García-Gasca, 2018). Além disso, a produção de proteínas em sistemas heterólogos otimiza o tempo e as operações, melhora o custo, e permite um maior controle nas propriedades finais de uma fabricação e aplicação em larga escala (Ahmad *et al.*, 2018).

Muitas lectinas de algas e animais são produzidas em sua forma recombinante, com promissoras atividades biológicas (Ahmmed *et al.*, 2022). Potanto, a expressão de lectinas em sistema heterólogo tem se consolidado como uma estratégia fundamental para o avanço de estudos estruturais e produção de proteínas para fins biotecnológicos.

As lectinas objetivo do estudo já foram estudadas anteriormente, porém algumas lacunas permaneceram presentes como a caracterização estrutural e funcional dessas lectinas. A especificidade por carboidratos dessas lectinas foi avaliada exclusivamente por ensaios

indiretos, que não permitem determinar com precisão a preferência de ligação, a intensidade e os aminoácidos envolvidos na interação lectina-carboidrato. De forma semelhante, a compreensão estrutural dessas proteínas é baseada em predições computacionais. Mesmo que as ferramentas *in silico* sejam fundamentais e amplamente utilizadas, predição estrutural não anula a necessidade da determinação experimental. A ausência de estruturas tridimensionais determinadas experimentalmente para diversas lectinas marinhas, incluindo aquelas investigadas neste trabalho, representa uma lacuna significativa no campo, limitando a compreensão de suas funções e dificultando a exploração biotecnológica de suas aplicações.

Outro aspecto crítico está relacionado à obtenção dessas proteínas em sua forma nativa, uma vez que os métodos convencionais de purificação frequentemente resultam em baixos rendimentos, limitando sua disponibilidade para estudos mais aprofundados. Nesse contexto, a expressão em sistemas heterólogos surge como uma alternativa estratégica, permitindo a produção em maiores quantidades, sem a necessidade de coleta contínua dos organismos no ambiente natural, e favorecendo uma caracterização bioquímica e estrutural mais precisa.

Nesse contexto, o estudo integrado de proteínas isoladas de organismos marinhos como rADEL, rAmL, rEEL e das galectinas ALL e CCL permite preencher diferentes níveis dessas lacunas existente não só para essas proteínas, mas para as lectinas desses grupos de animais. Os estudos de expressão em sistema heteróloga permitem explorar as possibilidades de obtenção dessas proteínas, enquanto as análises estruturais determinam suas conformações e mecanismos de interação. Dessa forma, este estudo propõe um avanço no conhecimento para a exploração e aplicação biotecnológica dessas proteínas promissoras na área da saúde.

O presente trabalho foi realizado em colaboração entre os laboratórios INCT-BIOMAR, Fortaleza, Brasil e CERMAV, Grenoble, França. A estrutura do trabalho está dividida em capítulos de acordo com as lectinas em estudo. As metodologias empregadas ao longo da pesquisa são apresentadas em ordem cronológica, visando facilitar a compreensão da sequência experimental. Para fins de clareza, os ensaios realizados em cada instituição são identificados ao longo da metodologia como “Brasil” ou “França”.

3 OBJETIVOS DA TESE

Objetivo geral

Caracterizar estruturalmente, bioquimicamente e funcionalmente lectinas oriundas de organismos marinhos por meio de estratégias de expressão heteróloga, análises moleculares e estudos de interação proteína-carboidrato visando a ampliação do conhecimento sobre essas proteínas para sua utilização como ferramenta biotecnológica.

Objetivos específicos

- Produzir em sistema heterólogo e otimizar a produção recombinante de lectinas oriundas de organismos marinhos, avaliando sua solubilidade, funcionabilidade e rendimento;
- Caracterizar bioquimicamente e estruturalmente as lectinas recombinantes, com ênfase na especificidade de ligação a carboidratos e na organização estrutural;
- Determinar estruturas tridimensionais e perfis de interação molecular relevantes para o reconhecimento lectina-carboidrato;
- Investigar o potencial biotecnológico das lectinas estudadas por meio da investigação de suas atividades biológicas;

4 CAPÍTULO I: EXPRESSÃO HETERÓLOGA DA LECTINA rELEL PRESENTE NOS OVOS DO OURIÇO-DO-MAR *Echinometra lucunter*: DESAFIOS DE EXPRESSÃO E DOBRAMENTO

4.1 Objetivos

Objetivo geral

Expressar em sistema heterólogo, a lectina presente nos ovos do ouriço-do-mar *Echinometra lucunter*, rELEL (recombinante *Echinometra lucunter* eggs lectin).

Objetivos específicos

- Selecionar clones positivos de *Escherichia coli* transformados com o gene sintético codificador de rELEL;
- Produzir em sistema heterólogo de *E. coli* a lectina recombinante rELEL, testando diferentes vetores e condições de expressão para obter a lectina na fração solúvel e potencialmente ativa;
- Purificar a lectina recombinante rELEL;
- Obter a proteína ativa por meio de *refolding* da fração insolúvel;

4.2 Metodologia

4.2.1 Coleta das espécimes e purificação da lectina nativa

A lectina nativa ELEL, oriunda dos ovos do ouriço-do-mar *Echinometra lucunter*, foi obtida através do protocolo previamente estabelecido por Carneiro e colaboradores (2015). A coleta e o acesso ao patrimônio genético da espécie foram autorizados e regulados pelos órgãos competentes (SISBIO ID: 33913-12 e SISGEN ID:A1792FE).

Os exemplares do ouriço-do-mar foram coletados durante a maré de sizígia na zona entre marés da praia do Pacheco, no município de Caucaia, Ceará. Os animais foram transportados até o laboratório, e os ovos foram obtidos a partir de uma injeção celômica de KCl 0,5 M. A camada gelatinosa externa foi removida por sucessivas lavagens com água do mar ácida, pH 5, e em seguida foram desgordurados com três lavagens de acetona fria.

Os ovos tratados foram suspensos em cinco volumes de tampão Tris 20 mM pH 7,6; NaCl 150 mM (TBS), contendo fluoreto de fenilmetilsulfonil (PMSF) 0,1 mM, e sonicados. Após centrifugação, o sobrenadante foi submetido a uma cromatografia de afinidade em uma matriz de lactosil-agarose, previamente equilibrada com TBS. A proteína de interesse foi eluída com L-ramnose 0,3 M em TBS.

A pureza das frações obtidas foi confirmada por eletroforese (SDS-PAGE) (Laemmli, 1970), e a funcionabilidade da proteína foi avaliada por atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina conforme preferência da lectina. As frações puras e funcionalmente ativas foram reunidas, dialisadas contra água destilada e liofilizadas para uso posterior.

4.2.2 Produção da proteína recombinante rELEL em sistema heterólogo de *Escherichia coli* (Brasil)

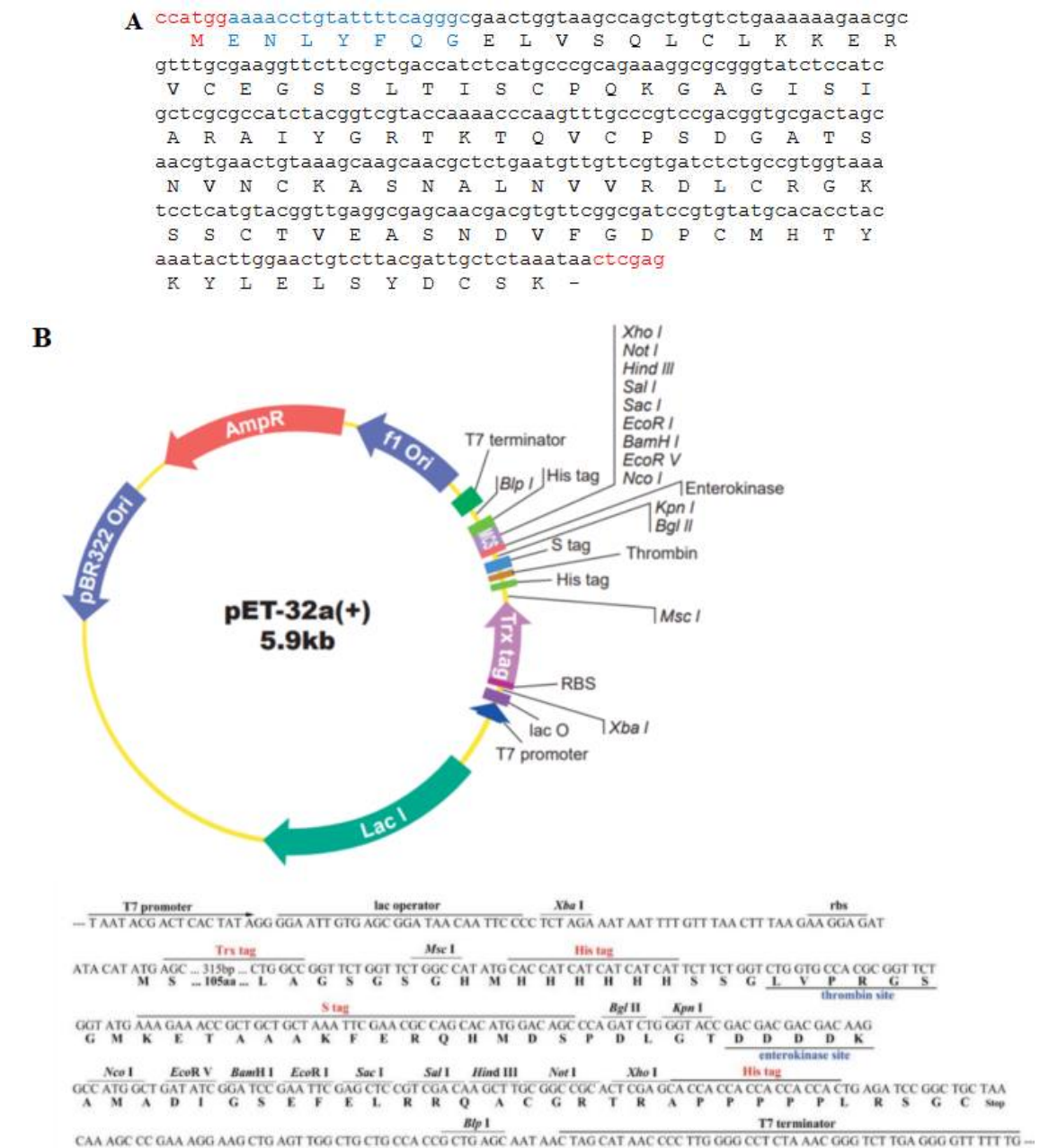
4.2.2.1 Desenho do gene sintético adquirido no Brasil

A sequência de aminoácidos da lectina nativa (C0HK24.1 - UniProt: <https://www.uniprot.org/>) foi utilizada para desenhar o gene sintético (Carneiro *et al.*, 2015).

O gene sintético foi adquirido pela empresa FastBio (<http://www.fastbio.com.br/>, SP, Brasil) no vetor de expressão pET32a(+) modificado para ter um sítio de clivagem para a

protease TEV, com as regiões de clonagem delimitadas pelas enzimas de restrição *NcoI* e *XhoI*, sem a otimização de códons raros em *Escherichia coli* (Figura 04).

Figura 04 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pET32a.



Fonte: Autora. A – Representação da sequência de nucleotídeos e a respectiva tradução da lectina com a inserção da sequência específica de clivagem para a protease TEV. A sequência em letras minúsculas representa os nucleotídeos e a sequência em letras maiúsculas representa os aminoácidos traduzidos. As letras em vermelho e azul representam os sítios de clivagem *NcoI* (início) e *XhoI* (final); e o sítio da protease TEV, respectivamente. O símbolo – representa o códon de parada da tradução. B – Representação do vetor de expressão pET32a e seu respectivo mapa. Disponível em: <https://fastbio.com.br/wp-content/uploads/2023/09/GenScript-Lista-e-mapa-de-vetores-bacteria.pdf>.

4.2.2.2 Expressão da lectina rELEL em sistema heterólogo de *E. coli*

4.2.2.2.1 Testes de expressão

O gene sintético foi diluído para uma concentração de $50 \text{ ng} \cdot \mu\text{L}^{-1}$ e inserido, por choque térmico, na cepa de clonagem *Escherichia coli* DH5 α , e nas cepas de expressão: *E. coli* BL21(DE3), *E. coli* Rosetta-gami, *E. coli* Origami B (DE3) plysS e *E. coli* ArcticExpress, seguindo o protocolo de transformação descrito por Sambrook (2002). A manipulação e uso de organismos geneticamente modificados tem autorização pelo certificado CQB-R007-2016 CTNBio.

As colônias crescidas em LB ágar e selecionadas pelo antibiótico de seleção (ampicilina a $100 \mu\text{g/mL}$) foram inoculadas em meio LB *Broth* contendo o mesmo antibiótico e incubadas a 37°C , *overnight*, sob agitação constante. As células crescidas foram diluídas na proporção de 1:100 em meio LB contendo ampicilina na mesma proporção anterior e a cultura foi crescida a 37°C e induzida de acordo com cada condição testada.

Os testes para avaliar as melhores condições de expressão foram feitos em relação à concentração de IPTG ($0,05 \text{ mM}$; $0,1 \text{ mM}$; $0,5 \text{ mM}$ e $1,0 \text{ mM}$) a 25°C por 20h, a variação de temperaturas (37°C por 4 h, 30°C por 16h, 25°C por 20h, 18°C por 24 horas e 10°C por 24 horas apenas para *E. coli* ArcticExpress) e a melhor cepa de expressão. Em todos os experimentos, a indução ocorreu quando a cultura atingiu a densidade óptica a 600 nm ($\text{DO}_{600\text{nm}}$) entre 0,4 e 0,5.

Após o tempo de indução em cada teste, as células foram recuperadas por centrifugação a $6000 \times g$, por 5 minutos a 4°C . O sobrenadante foi descartado e o concentrado celular (*pellet*) foi pesado e ressuspenso em tampão Tris 20 mM pH 8,0 com NaCl 150 mM , glicerol 5% e imidazol 20 mM , na proporção de 1g de *pellet* para 10 mL de tampão. As células foram rompidas por sonicação com amplitude de 30% durante 15s com 30s de intervalo. Em seguida, o lisado celular foi centrifugado a $9000 \times g$, por 30 min, a 4°C , e o sobrenadante (fração solúvel) e o precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE.

4.2.2.2.2 Estratégias para expressão da rELEL na fração solúvel

Para obter a lectina recombinante na fração solúvel, novas condições de expressão foram testadas: redução da concentração do indutor, aumento da densidade óptica de indução e adição de aditivos no meio de cultura.

As células foram crescidas em uma pré cultura, e em seguida foram diluídas na proporção de 1:100 em meio LB contendo ampicilina, conforme descrito anteriormente. A cultura foi mantida a 37 °C e induzida de acordo com cada condição avaliada.

As novas concentrações de IPTG testadas foram: 0,05 mM; 0,025 mM e 0,010 mM em cada cepa de expressão (*E. coli* BL21(DE3), *E. coli* Rosetta-gami, *E. coli* Origami B (DE3) plysS e *E. coli* ArcticExpress). Em todos os experimentos, a indução ocorreu quando a cultura atingiu a DO_{600nm} de 0,7, e a expressão foi conduzida a 25°C, por 20 horas.

A indução também foi testada na presença dos aditivos PEG (Polietilenoglicol) a 3g/L e NaCl a 10 g/L no meio de cultura. Nessa estratégia, a cepa *E. coli* Rosetta-gami foi induzida com 0,025 mM de IPTG na DO_{600nm} de 0,7 e a expressão ocorreu a 25°C, por 20 horas.

Após o tempo de indução em cada teste, as células foram recuperadas por centrifugação e lisadas, conforme protocolo detalhado acima. A fração solúvel e a fração insolúvel foram coletadas e analisadas por SDS-PAGE.

Para avaliar se a proteína estaria sendo secretada para o meio extracelular, após a recuperação da cultura por centrifugação, o sobrenadante foi precipitado com sulfato de amônio a 70%. O precipitado foi recuperado por centrifugação a 9000 x g, 30 minutos a 4 °C, e analisado em SDS-PAGE.

4.2.2.3 Tentativa de purificação da proteína recombinante

Com o objetivo de purificar a lectina recombinante, uma cromatografia de afinidade foi realizada, conforme o método empregado para a purificação da lectina nativa (Carneiro *et al.*, 2015). A fração solúvel dos testes que apresentaram confirmação da expressão da lectina por espectrometria de massas foi aplicada em uma coluna de lactosil-agarose, previamente equilibrada com Tris 20 Mm pH 7,6 e NaCl 150 mM. As proteínas possivelmente retidas a coluna foram eluídas com L-ramnose 0,3 M em TBS. A cromatografia foi conduzida por gravidade a um fluxo de 2 mL/min.

Como a construção no vetor pET32a possui cauda de histidina, a fração solúvel também foi submetida à cromatografia de afinidade a metais imobilizados (IMAC), paralelamente. A fração foi aplicada em uma coluna HisTrapTM FF de 5mL, e as proteínas

possivelmente retidas a coluna foram eluídas com tampão Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 270 mM.

Para avaliar se a proteína presente nas frações solúveis estava ativa, a amostra foi submetida a um teste de atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina a 3%.

4.2.2.4 Confirmação da expressão da lectina por espectrometria de massas sequencial

A expressão da lectina recombinante rELEL foi analisada por espectrometria de massas sequencial (MS/MS), com o objetivo de verificar se a proteína produzida correspondia com a estrutura primária da lectina nativa.

Inicialmente, as frações insolúveis e solúveis dos testes de expressão, após a lise celular, foram submetidas a SDS-PAGE 12% (Laemmli, 1970). As bandas referentes às lectinas foram excisadas e transferidas para tubos limpos. As proteínas foram reduzidas e alquiladas e, posteriormente, digeridas com a enzima tripsina, de acordo com a metodologia de Shevchenko e colaboradores (2007).

As soluções contendo os peptídeos foram injetadas em um sistema nanoAcquity (Waters Corp) conectado a nESI no espectrômetro de massas híbrido. O espectrômetro foi calibrado na faixa de 50 a 1600 m/z com os fragmentos do íon de fosfato. O analito foi aplicado em uma coluna de fase reversa C18 (75 µm x 100 mm) e os peptídeos foram eluídos através de gradiente linear de 10 % a 85 % de ACN contendo AF 0,1 % a um fluxo de 0,5 µL.min⁻¹.

O espectrômetro de massas operou como descrito por Carneiro e colaboradores (2017b). Os espectros coletados foram processados pelos programas MassLynx v4.1 e ProteinLynx v2.4 (Waters Corp.) e os dados interpretados.

4.2.2.5 Tentativa de refolding da lectina recombinante a partir da fração insolúvel

As frações contendo a lectina na fração insolúvel, identificadas por espectrometria de massas, foram reunidas e submetidas a um processo de redobramento (*refolding*) para recuperar a proteína solúvel e ativa.

O concentrado celular foi lavado com tampão de lavagem contendo Tris 50 mM pH 8, NaCl 500 mM, Ureia 4M e EDTA 1 mM, para remover contaminantes. Em seguida, a fração

insolúvel foi solubilizada em tampão desnaturante contendo Ureia 8M, DTT 10 mM e Tris 50 mM pH 8, sob agitação até a solubilização completa.

A fração desnaturada foi então submetida a redobramento usando o tampão 50 mM Tris, pH 7,5 com 150 mM NaCl e 1 mM DTT. A troca do tampão foi realizada por diálise a cada 24 horas, reduzindo progressivamente a concentração de ureia no tampão de redobramento (6 M, 4 M, 2 M, 0 M). As frações renaturadas foram analisadas por SDS-PAGE e sua funcionabilidade avaliada por atividade hemaglutinante.

4.2.3 Avaliação da expressão da proteína recombinante por eletroforese (Brasil)

A presença da lectina recombinante nas frações obtidas nos diferentes testes de expressão, e na fração renaturada foi avaliada por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) (Laemmli, 1970). O gel de 12% foi submetido a uma corrente elétrica de 25 mA e tensão de 150 V durante 1 h.

As alíquotas de amostras líquidas foram diluídas em tampão de amostra 5x contendo Tris pH 6,8 (0,06M), glicerol (25%), SDS (2%) e azul de bromofenol (1mg), para concentração final de 1%. As alíquotas de concentrado celular foram ressuspensas no mesmo tampão de amostra 1x.

A presença da lectina foi avaliada em condições redutoras, contendo β mercaptoetanol, comparando a migração da proteína com a migração de marcadores moleculares. Os marcadores moleculares usados foram um marcador molecular preparado com as proteínas conhecidas: HSA – 66 kDa (NP_0000468); rSFL - 30 kDa (7UMJ_AA; Chaves *et al.*, 2023) e CCL – 15 kDa (C0HLX7.1; Marques *et al.*, 2018).

4.2.4 Avaliação da funcionabilidade da proteína recombinante (Brasil)

O ensaio de hemaglutinação foi realizado com as frações solúveis dos testes de expressão e com a proteína renaturada, a fim de avaliar se a lectina rEEL identificada estava ativa e funcional.

A atividade hemaglutinante foi executada contra eritrócitos de coelho tratados com enzima proteolítica tripsina. Os eritrócitos de coelho foram lavados, tratados e preparados em uma suspensão a 3%, e o experimento foi executado pelo método da dupla diluição seriada em

placas de microtitulação de fundo V (Sampaio *et al.*, 2002). A atividade hemaglutinante foi expressa em título de hemaglutinação (U.H.).

A lavagem dos eritrócitos, a preparação da suspensão a 3% e o tratamento com a enzima proteolítica foi conduzido conforme descrito por Sampaio e colaboradores (2002). A coleta e utilização dos eritrócitos foi autorizada pelo órgão competente: CEUAP nº2211202101.

4.2.5 Produção da proteína recombinante rEEL em sistema heterólogo (França)

4.2.5.1 Desenho do gene sintético adquirido na França

A sequência de aminoácidos da lectina nativa (C0HK24.1 - UniProt: <https://www.uniprot.org/>) foi utilizada para desenhar o gene sintético (Carneiro *et al.*, 2015).

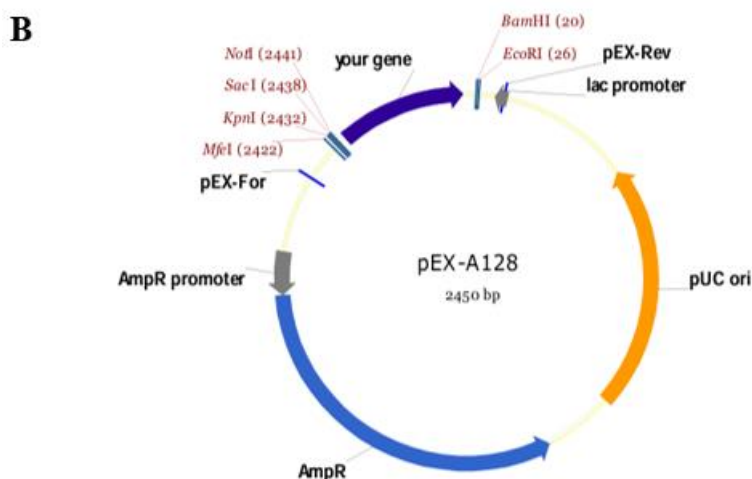
O gene sintético foi adquirido pela empresa Eurofins Genomics (França) no vetor *backbone* pEX-A128 com a sequência gênica otimizada para *Escherichia coli* e delimitada pelas enzimas de restrição *NcoI* e *XhoI* (Figura 05).

Figura 05 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pEX-A128.

```

A ccatggaactcgtcagccagttatgcctgaagaaagagcgcgtttgtgaaggcagttcactg
  M E L V S Q L C L K K E R V C E G S S L
accatctcttgtccgcagaaaggagcaggcatttagcattgctcgtgccatctatggtcgc
  T I S C P Q K G A G I S I A R A I Y G R
accaaaacccaagtgtgtccgagtgatgggtgcgacttccaacgtgaactgcaaagcgtct
  T K T Q V C P S D G A T S N V N C K A S
aatgcgctgaacgtagttcgcgatctttgccgtggcaaaagctcctgtacagtcgaagcc
  N A L N V V R D L C R G K S S C T V E A
agcaatgacgtgtttgggatccatgcatgcatacgtacaagtacctggagttgtcgat
  S N D V F G D P C M H T Y K Y L E L S Y
gactgctcgaaataactcgag
  D C S K -

```



MCS of pEX-A128

```

GGAGCAGACAAGCCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTCTG
GGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGAAAGG
caattgGGTACCgagctcGCGGCCGAAGC>your_gene>ACCTGCTT
TTGCTCGCTTGGATCCgaattcAAAGGTGAAATTGTTATCCGCTCACAA
TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTG

```

Start point: gene insertion site

(A of ACCT 3' of "your gene" is bp no. 1)

Multiple Cloning Site: bases 19-30 / 2421-2446

lac Promoter: bases 65-95

pUC origin: bases 419-1007

Ampicillin resistance ORF: bases 1178-2038

pEX-For priming site: bases 2323-2343

pEX-Rev priming site: bases 75-100

Fonte: Autora.

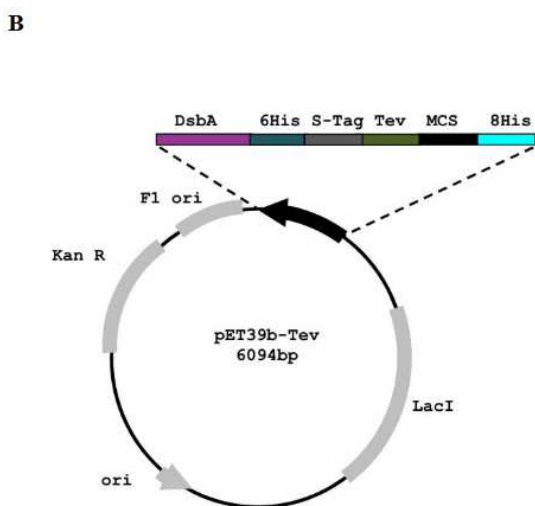
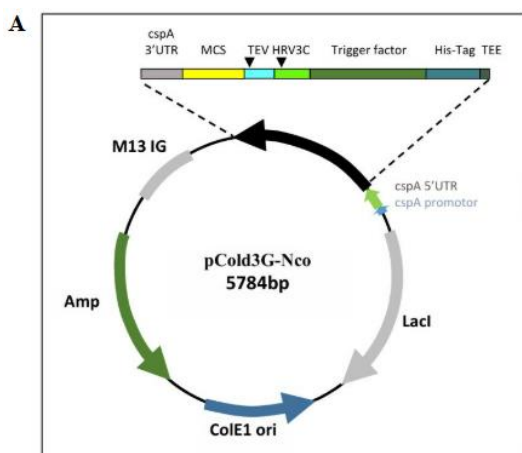
A – Representação da sequência de nucleotídeos e a respectiva tradução da lectina. A sequência em letras minúsculas representa os nucleotídeos e a sequência em letras maiúsculas representa os aminoácidos traduzidos. As letras em vermelho representam os sítios de clivagem *NcoI* (início) e *XhoI* (final). O símbolo – representa o códon de parada da tradução. B – Representação do vetor de expressão pEX-A128 e seu respectivo mapa. Disponível em: https://eurofinngenomics.eu/media/1587668/pex-a128_map_seq_v10.pdf.

4.2.5.1.1 Subclonagens do gene sintético para os vetores de expressão

O gene sintético inicialmente inserido no vetor *backbone* pEX-A128 foi subclonado nos vetores de expressão pCOLD-TF-3G-Nco, pET39-TEV, pET15b(+) e pCOLD-HAT-TEV-3G-Nco. O plasmídeo pET15b(+) foi adquirido pela empresa Novagen (França) e os demais foram desenvolvidos pelo grupo GBMS, disponível na coleção do CERMAV- CNRS, Grenoble, França.

Os mapas e as particularidades dos plasmídeos usados para as subclonagens encontram-se na Figura 06 e Tabela 01, respectivamente.

Figura 06 – Representação dos vetores de expressão utilizados para subclonar a sequência gênica de rELEL.



pCold3G-NcoI

Derived from pCold TF-TEV-Nco by inactivation of Thrombin site and addition of 2 more Glycine.

[illegible]

pET39b-Tev

Enterokinase site replaced by TEV cleavage site and deletion of *SrfI* et *BseRI* sites before *NcoI* site

```

PET upstream primer #69214-3
ATGCGTCCGGCGTAGAGATCGAGATCGATCTCGATCCCGCGAAATTAATACGACTCATTATAGGCAATATGCG
TAAATACGACTCATTATAGGCAATATGCG
XbaI Rbs NdeI
CTCTAGAAATAATTTTGTTTAACTTTAAGACGCAATATACAAATATGAAAAAGATTTCG
MetLysLysIleTrp
DsbA
CTGSGCGTGGCTGGTTAGTTTTCAGTCATCGGCGGGC...522bp...CAGCAGTAIGCTGATACA
LeuAlaLeuAlaGlyLeuValLeuAlaPheSerAlaSerAlaAla 174aa GlnGlnTyrAlaAspThr
Signal peptidase ↑
GTGAAATACTTAAAGCAGAAAAAAGGTCACACTAGTGGTTCTGGTTCGCGGGGTCTG
ValLysTyrLeuSerGluLysLysGlySerThrSerGlySerGlyHisHisHisHisHisSerAlaGlyLeu
ScaI STagFor S-Tag NspV
GTGCGACGCGGTACTGACTGCAATTGGTATCAAGAACCCCGCTGCTGTAATTCGACGCCACACATGGACAGC
ValProArgGlySerThrAlaIleGlyMetLysGluThrAlaAlaAlaLysPheGluArgGlnHisMetAspSer
Thrombin ↑
KpnI NcoI EcoRV BamHI EcoRI SacI
CCAGATCTGGGTACCAAACTCTGTATTTTCAGGGCGCATGGCGATATCGATCCGAATTCGAGCTCC
ProAspLeuGlyThrGluAsnLeuTyrPheGlnGlyAlaMetAlaIleSerAspProAsnSerSerSer
Tev ↑
SalI HindIII NotI XhoI Histag
GTCGACAAAGCTTCGGGCGCATCTCAGCACCACCAACCCACCACTAATGTGATTAATACCTAGGCTGCT
ValAspLysLeuAlaAlaLeuGlyLysHisHisHisHisHisHisHisHisTop

```


Tabela 01 – Características dos vetores utilizados para a expressão da lectina rELEL.

VETOR	ESPECIFICIDADE	PROMOTOR (Proteína)	ANTIBIÓTICO DE SELEÇÃO
pCOLD-TF-3G-Nco	Derivado do pCOLD-TF-TEV-Nco, desativando sítio para trombina e aumentando duas glicinas após o sítio de clivagem para a TEV. Com o sítio de clivagem protease 3C do rinovírus humano. Presença da <i>His-tag</i> , chaperona <i>trigger factor</i> e do TEE (<i>Translation Enhancing Element</i>).	cspA	Ampicilina
pET39-TEV	Derivado do pET39b(+), com a substituição do sítio de clivagem da enteroquinase para um sítio TEV e remoção dos dos sítios de restrição <i>SrfI</i> e <i>BseRI</i> antes do <i>NcoI</i> . Presença de <i>His-tag</i> , <i>S-tag</i> e da proteína DsbA	T7 Lac	Canamicina
pET15b(+)	Presença de <i>His-tag</i> e sítio de clivagem para a trombina.	T7 Lac	Ampicilina
pCOLD-HAT-TEV-3G-Nco	Derivado do pCOLDIII contendo TEE, com a inserção da <i>HAT tag</i> , sítio de clivagem TEV, adição de + 3 Gly após o sítio da TEV e a substituição do sítio de restrição <i>NdeI</i> por <i>NcoI</i> .	cspA	Ampicilina

Fonte: Autora. As enzimas de restrição *NcoI* e *XhoI* foram empregadas em todos os vetores utilizados. DsbA: Proteína de ligação dissulfeto A.

O gene sintético adquirido foi diluído para uma concentração de 50 ng.µL⁻¹ e inserido, por choque térmico, na cepa de clonagem *Escherichia coli* NEB5α, seguindo o protocolo de transformação bacteriana descrito por Sambrook (2002).

Após a transformação, as colônias crescidas em LB ágar e selecionadas pelo antibiótico de seleção do vetor foram inoculadas a 37 °C, *overnight*, em meio LB *Broth*

contendo o antibiótico para extração do DNA plasmidial. A miniprep foi realizada utilizando o kit de purificação em pequena escala QIAprep® *Spin Miniprep* (Qiagen), seguindo as recomendações do fabricante.

O DNA plasmidial purificado foi quantificado em espectrofotômetro *NanoDrop* (Thermo Fisher Scientific) e avaliado qualitativamente por eletroforese em gel de agarose a 1%, corado com SYBR Safe 1x (Invitrogen).

O plasmídeo extraído foi digerido a 37 °C, por 3 horas, usando as enzimas *NcoI* (BioLabs) e *XhoI* (BioLabs), de acordo com a configuração do vetor. A reação foi preparada seguindo as recomendações do fabricante (New England BioLabs), utilizando uma proporção de 10 U de enzima/μg de DNA. Após a digestão, os produtos resultantes foram separados por eletroforese em gel de agarose a 1% e o fragmento correspondente ao inserto da proteína foi identificado pelo tamanho em pares de bases (pb) em comparação ao marcador comercial O'GeneRuler 1kb DNA Ladder (*Thermo Scientific*) com intervalo de 250 a 10.000 pb. O fragmento foi purificado usando o kit Monarch® DNA Gel Extraction (New England BioLabs), conforme protocolo disponibilizado pelo fabricante, quantificado em espectrofotômetro *NanoDrop* (Thermo Fisher Scientific).

Os vetores pCOLD-TF-3G-Nco, pET39-TEV, pET15b(+) e pCOLD-HAT-TEV-3G-Nco sem inserto e previamente transformados em cepas de *E. coli* NEB5α foram extraídos por miniprep e submetidos ao mesmo protocolo de digestão enzimática com *NcoI* e *XhoI*, obtendo extremidades coesas compatíveis com o inserto. Os vetores linearizados foram purificados a partir do gel de agarose a 0,7% usando o mesmo kit Monarch e quantificado em espectrofotômetro *NanoDrop*.

O inserto purificado e o vetor linearizado foram submetidos a uma reação de ligação usando o DNA *Ligation Kit* (TAKARA). A quantidade de inserto e de vetor usado na reação foi calculada pelo NEBioCalculator (<https://nebiocalculator.neb.com/#!/ligation>) usando o parâmetro de inserto/vetor = 3/1. A reação de ligação foi preparada seguindo as recomendações do fabricante e incubada a 16 °C, por 1 hora. Em seguida, 2 μL da reação de ligação foram usados para transformar as células de clonagem *E. coli* NEB5α, por choque térmico (Sambrook, 2002).

Algumas colônias isoladas no meio LB ágar contendo o antibiótico de seleção do vetor foram selecionadas para confirmação da transformação por PCR em colônia (Reação em Cadeia da Polimerase) usando o kit *Taq'Ozyme Purple Mix*. A reação de PCR foi montada de acordo com as recomendações do fabricante, onde uma colônia foi ressuspensa em 10 μL de

água MiliQ estéril e 5 µL da suspensão bacteriana foram adicionados a um igual volume de solução de PCR mix (*Purple Mix 2x*) juntamente com os primers *forward* e *reverse* específicos para cada vetor na concentração final de 0,2 µM.

A reação de PCR ocorreu em um termociclador *Eppendorf* programado para uma desnaturação inicial de 2 min a 95 °C, seguido por 34 ciclos de 30 segundos a 95 °C (desnaturação), 30 segundos a uma temperatura de anelamento e 60 segundos a 72 °C (extensão). Depois do último ciclo, uma fase de extensão prolongada foi realizada (72 °C por 5 min). Após a reação de PCR, o termociclador foi programado para incubar as amostras a 16 °C.

A T °C de anelamento foi ajustada de acordo com a temperatura de *melting* (Tm) para o par de primer específico do vetor utilizado. Os primers e a temperatura de anelamento usados para cada vetor estão descritos na Tabela 02.

Tabela 02 – Primers e condições de PCR para confirmação de vetores.

<u>Vetor</u>	<u>Primers</u>	<u>T °C de anelamento</u>
pCOLD-TF-3G-Nco	pCOLDTF-F1 (F)	52 °C
	pCOLD-R(R)	
pET39-TEV	S TagFor (F)	50 °C
	T7 TER (R)	
pET15b(+)	T7 PRO (F)	50 °C
	T7 TER (R)	
pCOLD-HAT-TEV-3G-Nco	pCOLDTF-F (F)	52 °C
	pCOLD-R(R)	

Fonte: Autora. (F): primer *forward* e (R): primer *reverse*.

Os produtos da PCR foram visualizados por eletroforese em gel de agarose a 1%. Um clone positivo, identificado de acordo com o tamanho esperado do fragmento amplificado, foi selecionado e cultivado em LB ágar e LB *Broth*. A alíquota restante (5 µL) da suspensão bacteriana em água MiliQ foi suplementada com 50 µL de LB *Broth* e plaqueada por estriamento em LB ágar contendo o antibiótico de seleção. O restante da suspensão foi crescido em 5mL de LB *Broth*, a 37 °C, *overnight*, contendo o mesmo antibiótico, e o DNA plasmidial foi extraído por miniprep conforme descrito anteriormente.

Em seguida, o plasmídeo extraído foi quantificado e submetido novamente a reação de digestão a 37 °C, por 1 hora, usando as mesmas endonucleases *NcoI* e *XhoI*, para

confirmação. Os produtos resultantes da digestão foram analisados em gel de agarose a 1% e a presença do fragmento esperado foi confirmada com base no tamanho em pares de bases. O DNA plasmidial extraído do clone transformado e confirmado por digestão foi enviado para sequenciamento de confirmação pela EUROFINS (França).

4.2.5.2 Expressão da lectina rELEL em sistema heterólogo

Após a confirmação por sequenciamento, o DNA plasmidial do clone positivo foi transformado, na cepa de expressão: *E. coli* Shuffle B, por choque térmico (Sambrook, 2002).

No grupo parceiro (CERMAV), as condições de indução (concentração do indutor, temperatura de indução e tempo de indução) para a expressão heteróloga nos vetores utilizados já haviam sido extensivamente trabalhadas e as melhores condições previamente estabelecidas para proteínas com características estruturais semelhantes ao da lectina de interesse. Por isso, optou-se pela não realização de novos testes de expressão, adotando as condições padrões recomendadas pelo grupo, visando otimização do tempo e dos recursos disponíveis.

Para expressão da proteína, uma colônia de *E. coli* Shuffle B transformada com o plasmídeo de interesse e isolada em LB ágar contendo o antibiótico de seleção foi pré-cultivada em 20 mL de LB Broth contendo o antibiótico correspondente ao vetor (ampicilina a 100 µg/mL ou canamicina a 50 µg/mL) a 30 °C, *overnight*, sob agitação constante de 180 *rpm*.

Todo o volume de pré-cultura foi inoculado em 1L de LB Broth em *erlenmayer* de 3L contendo o mesmo antibiótico. As células foram cultivadas a 30 °C, em agitação constante de 180 *rpm* até atingir a densidade ótica a 600nm (DO_{600nm}) de 0,4. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu até atingir DO = 0,8, quando o indutor IPTG (isopropil-beta-D-tiogalactopiranosídeo) foi adicionado ao meio na concentração de 0,2 mM e a expressão continuou a 16 °C, 180 *rpm*, por 18 horas.

Para o vetor pCOLD-HAT-TEV-3G-Nco, a expressão foi realizada em meio de autoindução LB NYZ. A transformação em *E. coli* Shuffle B e a pré-cultura seguiram o mesmo protocolo utilizado para os demais vetores. Para a cultura, todo o volume de pré-cultura foi inoculado em 1L de LB NYZ suplementado com 10 mL/L de glicerol (PA) em *erlenmayer* de 3L contendo o antibiótico correspondente. As células foram cultivadas a 30 °C, 180 *rpm* até atingir a DO_{600nm} superior a 5,0. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu a 180 *rpm*, por 18 horas.

Após o tempo de expressão para todos os vetores utilizados, as células foram recuperadas por centrifugação a 5000 x g, por 10 minutos a 25 °C. O sobrenadante foi descartado e o concentrado celular (*pellet*) foi pesado e congelado a -20 °C.

4.2.5.3 Purificação da lectina recombinante rEEL

A estratégia de purificação da lectina foi conduzida de acordo com as características da construção inserida em cada vetor de expressão. A Tabela 03 apresenta as diferentes abordagens de purificação empregadas em cada caso.

Tabela 03 – Estratégias de purificação da lectina recombinante em cada vetor.

Vetor (Sítio de restrição)	Cauda/tag	Estratégia	Cromatografia (Matriz)
pCOLD-TF-3G-Nco (<i>Nco-XhoI</i>)	<i>His-tag</i>	Purificação por afinidade via <i>tag</i> pela sequência de histidinas	Afinidade – IMAC (Hisrap – Ni ²⁺)
pET39-TEV (<i>Nco-XhoI</i>)	DsbA <i>His-tag</i>	Extração periplasmática; Purificação por afinidade via <i>tag</i> pela sequência de histidinas	Afinidade – IMAC (Hisrap – Ni ²⁺)
pET15b(+) (<i>Nco-XhoI</i>)	Ausente	Purificação seguindo o método da lectina nativa	Afinidade (α -Lactose- Sephrose; ramnose)
pCOLD-HAT-TEV- 3G-Nco (<i>Nco-XhoI</i>)	<i>HAT-tag</i>	Purificação por afinidade via <i>tag</i> pela presença de histidinas	Afinidade – IMAC (Hisrap – Ni ²⁺)

Fonte: Autora. IMAC: Immobilized Metal Affinity Chromatography – Cromatografia de afinidade a metais imobilizados.

Para o vetor pET15b(+), cada 1g de *pellet* celular foi ressuspensionado em 5 mL de tampão Tris 20 mM pH 7,5 com NaCl 150 mM. A enzima *denarase* foi adicionada na proporção de 1 µL para cada 25 mL de suspensão celular, seguida de incubação em misturador rotativo por 20 minutos. As células foram lisadas usando um disruptor celular mecânico (prensa francesa) da *One Shot Cell Disrupter* (Constant System) a 1,9 Kbars. Em seguida, o lisado celular foi centrifugado a 24.000 x g, por 30 minutos a 10 °C e o sobrenadante obtido filtrado em filtro de 0,45 µM. Aliquotas do sobrenadante (fração solúvel) e do precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE (Laemmili, 1970).

Como a construção para o vetor pET15b(+) não possui cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade, conforme o método usado para purificação da lectina nativa (Carneiro *et al.*, 2015). A fração solúvel foi aplicada em matriz de ramnose e matriz de α -Lactose-Sepharose, separadamente, para tentar diferentes afinidades. As colunas foram previamente equilibradas e lavadas com o mesmo tampão de lise. As proteínas possivelmente retidas a coluna foram eluídas com 0,3 M de ramnose no mesmo tampão. A cromatografia foi conduzida por gravidade a um fluxo de 2 mL/min e alíquotas de cada passo de purificação foram coletadas e analisadas por SDS-PAGE.

Para os vetores pCOLD-TF-3G-Nco e pCOLD-HAT-TEV-3G-Nco, o *pellet* celular foi ressuspensionado na mesma proporção em tampão Tris 20 mM pH 8, contendo NaCl 500 mM e imidazol 10 mM (Tampão A). A suspensão celular foi submetida ao mesmo procedimento de incubação e lise descrito anteriormente. Após a lise, alíquotas da fração solúvel e da fração insolúvel obtidas para ambos os vetores foram coletadas por centrifugação e analisadas por SDS-PAGE.

Como a construção dos vetores pCOLD-TF-3G-Nco e pCOLD-HAT-TEV-3G-Nco possui cauda com histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados (IMAC). A fração foi aplicada em uma coluna HisTrapTM FF de 1mL, previamente equilibrada e lavada com o mesmo tampão A. As proteínas retidas a coluna foram eluídas em duas etapas. Inicialmente, as proteínas mais fracamente ligadas a coluna foram removidas com 5% do mesmo tampão de equilíbrio contendo imidazol 500 mM (tampão B), enquanto as proteínas fortemente ligadas a coluna foram eluídas com um gradiente linear de 5% a 100% do tampão B. A cromatografia foi realizada em um sistema NGC (Bio-Rad) e alíquotas de cada passo de purificação foram coletadas e analisadas por SDS-PAGE. A funcionabilidade foi avaliada por atividade hemaglutinante contra eritrócitos de coelho a 4% (Sampaio *et al.*, 2002).

Para o vetor pET39-TEV, primeiramente foi feito uma extração periplasmática. Cada 1g de *pellet* celular foi ressuspensionado gentilmente em 5 mL de tampão de extração contendo Tris 50 mM pH 7,5 com NaCl 150 mM, EDTA 5 mM e 1 mg/mL de polimixina B sulfato, e a solução foi agitada lentamente por 30 minutos a 37 °C. Em seguida, as células foram centrifugadas a 11.000 x g, por 20 minutos a 4 °C, e o sobrenadante foi recuperado e filtrado a 45 µM.

A fração periplasmática foi diluída 1:1 (v/v) com tampão Tris 50 mM pH 8, contendo NaCl 300 mM, e foi submetida a uma cromatografia de afinidade em coluna cOmplete™ His-tag. A coluna foi equilibrada e lavada com o mesmo tampão de diluição, e as proteínas retidas a coluna foram eluídas com um gradiente linear de imidazol em tampão Tris 50 mM pH 8, NaCl 300 mM contendo imidazol 500 mM. A cromatografia foi conduzida em sistema de cromatografias NGC (Bio-Rad), a um fluxo de 1 mL/min.

O pellet resultante da extração periplasmática foi ressuspensionado em tampão A e submetido a uma extração citoplasmática usual, conforme descrito anteriormente. A enzima denarase foi adicionada, as células foram lisadas, e o lisado celular foi centrifugado. A fração solúvel foi aplicada em uma HisTrap™ FF de 1mL e as proteínas retidas a coluna foram eluídas com 5% e com um gradiente linear de 5% a 100% do tampão B.

Alíquotas da fração solúvel, fração insolúvel, e de cada passo de purificação para cada construção plasmidial foram coletados e analisados por SDS-PAGE.

4.2.5.4 Digestão da proteína fusionada e repurificação da lectina recombinante

Para remoção da cauda fusionada à lectina, as frações eluídas com 5% de tampão B, contendo baixa concentração de imidazol, foram reunidas e diluídas na proporção de 1:1 com tampão Tris 20 mM pH 8, visando reduzir a concentração do NaCl. Já as frações eluídas com o gradiente de tampão B, contendo alta concentração de imidazol, tiveram o tampão trocado por centrifugação usando concentradores de 10 kDa, realizando 3 repetições de diluição de 1:10 com tampão Tris 20 mM pH 8, contendo NaCl 100 mM.

Após a diluição ou a troca de tampão, a proteína foi concentrada até 5 mg/mL e a protease TEV foi adicionada na razão enzima:substrato de 1:100 juntamente com 0,5 mM de EDTA. A reação de digestão foi incubada a 20 °C por 18 horas.

Para a repurificação da proteína clivada, a concentração de imidazol na amostra digerida foi ajustada para 10 mM e a solução reaplicada em coluna HisTrap™ FF de 1mL,

previamente equilibrada com tampão Tris 20 mM pH 8, contendo NaCl 100 mM e imidazol 10 mM (tampão C). A fração não retida a coluna, contendo a proteína clivada (sem *tag*), foi lavada com o mesmo tampão e a fração das proteínas não clivadas juntamente com a *tag*, ligadas a coluna, foi eluída com um step de 50% do tampão B. Todas as frações foram analisadas por SDS-PAGE e a funcionabilidade da proteína após a clivagem foi confirmada por atividade hemaglutinante contra eritrócitos de coelho a 4%.

Para melhorar a purificação, a fração não retida a coluna foi submetida a cromatografia em matriz de ramnose e a fração retida a coluna foi aplicada em matriz de α -Lactose-Sepharose. As proteínas potencialmente ligadas a coluna foram eluídas com 0,3 M de ramnose conforme descrito acima.

Todas as frações foram analisadas por SDS-PAGE e a funcionabilidade da proteína após a clivagem foi avaliada por atividade hemaglutinante contra eritrócitos de coelho a 4%.

4.2.5.5 Tentativa de cotransformação com *cydisco*

A construção rELELpET15b foi cotransformada, na cepa de expressão *E. coli* Shuffle B, com o plasmídeo CyDisCo pMJS205 a 33 ng/ μ L (*Cytoplasmic Disulfide bond formation in E. coli* - Formação de pontes dissulfeto no citoplasma de *E. coli*), seguindo a mesma metodologia de choque térmico (Sambrook, 2002). A expressão seguiu o protocolo descrito na seção 3.2.5.2, com a adição do cloranfenicol (100 μ g/mL) como antibiótico de seleção.

A possível purificação da lectina, obtida por extração citoplasmática convencional, foi conduzida em matriz de α -Lactose-Sepharose, de acordo com o exposto na seção 3.2.5.3 para o pET15b(+).

4.2.6 Avaliação da expressão da proteína recombinante por eletroforese (França)

A presença da lectina recombinante nas frações obtidas das expressões e das etapas de purificação foi avaliada por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) (Laemmli, 1970). O gel de 12% foi submetido a uma corrente elétrica de 25 mA e tensão de 150 V durante 1 h, e corado com InstantBlue™ Protein Stain, Bio-Rad.

As alíquotas de concentrado celular foram ressuspensas em tampão de amostra 5x contendo Tris pH 6,8 (0,25 M), glicerol (50%), SDS (10%), DTT (0,25 M), e azul de

bromofenol (5mg). As alíquotas de amostras líquidas foram diluídas no mesmo tampão para a concentração final de 1% de tampão de amostra. Em seguida, as amostras foram aquecidas a 100 °C, 5 minutos.

A massa molecular foi estimada em condições redutoras, comparando a migração da proteína com a migração do marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad, com intervalo de 10 kDa a 250 kDa.

4.2.7 Avaliação da funcionabilidade da proteína recombinante (França)

O ensaio de hemaglutinação foi realizado com as frações das etapas de purificação nas quais a proteína de interesse foi detectada, a fim de avaliar se a lectina estava ativa e funcional.

Os eritrócitos de coelho fornecidos pela empresa bioMérieux foram preparados em uma suspensão a 8%. O experimento foi executado pelo método da dupla diluição seriada em placas de microtitulação de fundo V, resultando em uma suspensão final de 4% de hemácias não tratadas com enzimas proteolíticas (Ribeiro *et al.*, 2016). A atividade hemaglutinante foi expressa em título de hemaglutinação (U.H.).

4.3 Resultados e discussão

4.3.1 Produção da proteína recombinante rELEL em sistema heterólogo de *E. coli* (Brasil)

4.3.1.1 Desenho do gene sintético

O gene sintético da lectina rELEL, com 344 pb, incluindo os sítios de restrição, foi inicialmente clonado no vetor pET32a(+), delimitado pelos sítios de restrição *NcoI* e *XhoI* e sem a otimização de códons raros para *E. coli*. O quadro aberto de leitura (ORF – *Open Reading Frame*) possui 333 pb, codificando uma proteína de 111 aminoácidos. Devido a construção do vetor pET32a (+), o gene foi expresso em fusão N-terminal à proteína tiorredoxina, resultando na expressão da proteína de fusão Trx-rELEL com peso molecular previsto de 29.186 Da.

A proteína expressa Trx-rELEL foi construída para ser clivada com a protease TEV. Como o pET32a(+) não possui o sítio de reconhecimento específico para essa enzima, o gene sintético de rELEL foi modificado para conter um sítio de clivagem para a protease, inserido antes da sequência codificadora da lectina.

A construção da proteína recombinante expressa no pET32a(+) contendo a proteína tiorredoxina como proteína de fusão, as *tags* de purificação e os sítios de clivagem está representada na Figura 07.

Figura 07 – Representação esquemática da proteína recombinante expressa Trx-rELEL.

> Trx-rELEL

```

      10      20      30      40      50      60
MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY QGKLTVAKLN
      70      80      90     100     110     120
IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL KEFLDANLAG SGSGHMHMHH
      130     140     150     160     170     180
HHSSGLVPRG SGMKETAAAK FERQHMDSPD LGTDDDDKAM ENLYFQGELV SQLCLKKERV
      190     200     210     220     230     240
CEGSSLTISC PQKGAGISIA RAIYGRKTQ VCPSDGATSN VNCKASNALN VVRDLCRGKS
      250     260     270
SCTVEASNDV FGDPCHMTYK YLELSYDCSK

```

Cauda de Tiorredoxina

Cauda de Histidina

Sítio para a protease trombina

S Tag

Sítio de enteroquinase

Sítio de clivagem TEV

Lectina

Número de aminoácidos: 270
Peso molecular: 29186,0570
pI: 5,92

Coefficiente de extinção:
 O coeficiente de extinção é: 21430
 Em 1mg/ml (34,263 uM) a solução da proteína tem Abs280nm de: 0,73

Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 22055
 Em 1mg/ml (34,263 uM) a solução de proteína tem Abs280nm de: 0,76

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

A Figura 08 mostra a representação da proteína de interesse rEEL, após remoção da tiorredoxina por clivagem da protease TEV.

Figura 08 – Representação da proteína rEEL obtida após clivagem com a protease TEV.

>rEEL clivada

```

      10      20      30      40      50      60
GELVSQLCLK KERVCEGSSL TISCPQKGAG ISIRAIYGR TKTQVCPSDG ATSNVNCKAS
      70      80      90     100    104
NALNVVRDLC RKGSSCTVEA SNDVFGDPCM HTYKYLELSY DCSK

```

Sítio de clivagem TEV residual

Lectina

Número de aminoácidos: 104
Peso molecular: 11131,6781
pI: 8,40

Coefficiente de extinção:
 O coeficiente de extinção é: 5960
 Em 1mg/ml (89,834 uM) a solução da proteína tem Abs280nm de: 0,54

Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 6460
 Em 1mg/ml (89,834 uM) a solução de proteína tem Abs280nm de: 0,58

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

O índice de adaptação de códons (CAI – *Codon Adaptation Index*) da sequência não otimizada foi calculado para *E. coli* pela *GenScript*. O valor foi de 0,85, indicando boa adaptação ao sistema hospedeiro, pois valores acima de 0,8 são considerados bons, e quanto mais perto de 1,0, melhor a expressão naquele sistema. Por esse motivo, a sequência permaneceu não otimizada.

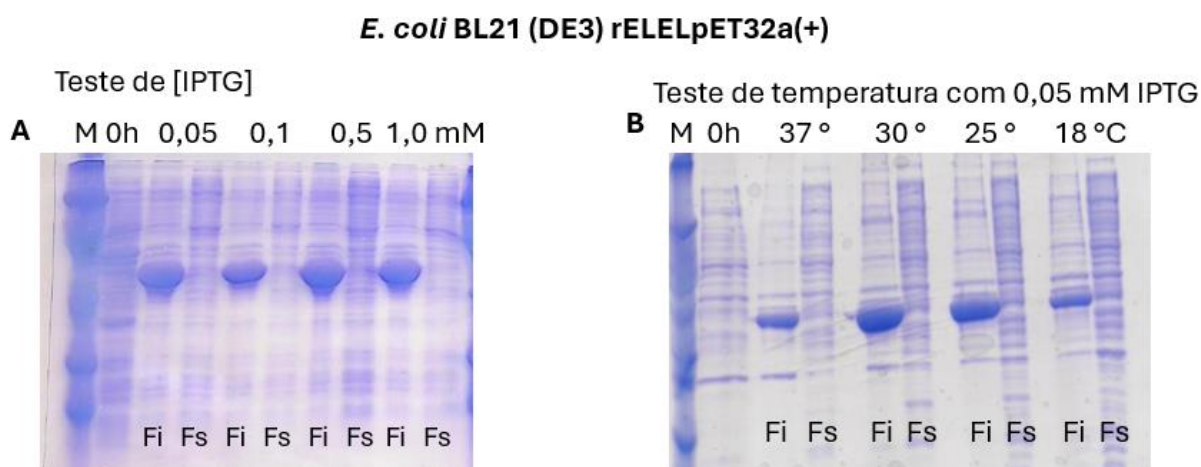
4.3.1.2 Testes de expressão

A cepa *E. coli* DH5 α foi utilizada como cepa de clonagem para replicação do vetor. As cepas *E. coli* BL21(DE3), *E. coli* Rosetta-gami, *E. coli* Origami B (DE3) plysS e *E. coli* ArcticExpress foram usadas para expressão da proteína recombinante, pois cada uma apresenta características específicas que podem favorecer a expressão e o dobramento correto da proteína de interesse. Esses ensaios de expressão consistem na variação controlada de parâmetros que influenciam diretamente a taxa de síntese e o dobramento da proteína. Os principais fatores a se avaliar são a concentração do indutor IPTG, que regula o nível de transcrição, e a temperatura de expressão, que afeta a velocidade de tradução (Mital; Christie; Dikicioglu, 2021; Simas; Pessoa Junior; Long, 2023).

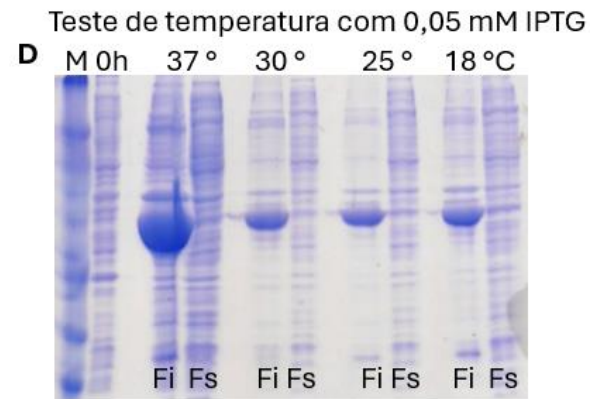
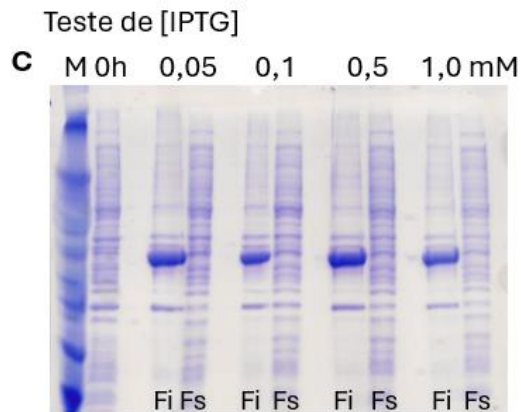
Para rELEL, os efeitos de diferentes concentrações do indutor IPTG, temperaturas de indução e cepas de expressão também foram avaliados.

Após o tempo de indução em cada teste, as células foram recuperadas, lisadas e as frações solúveis e insolúveis foram analisadas por SDS-PAGE (Figura 09).

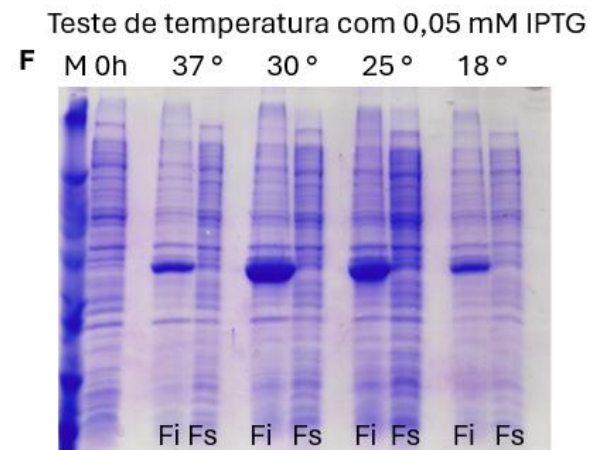
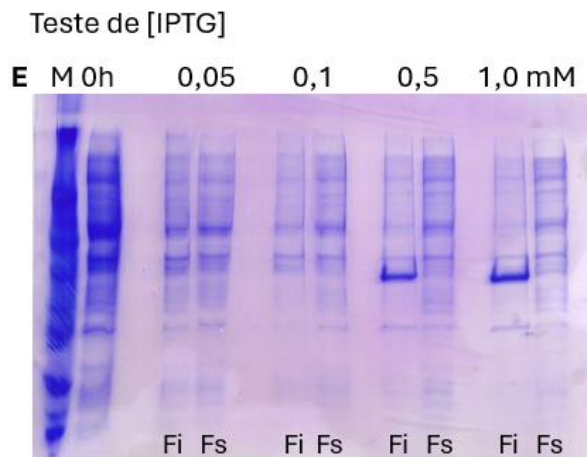
Figura 09 – Testes de expressão da proteína rELEL no vetor pET32a(+).



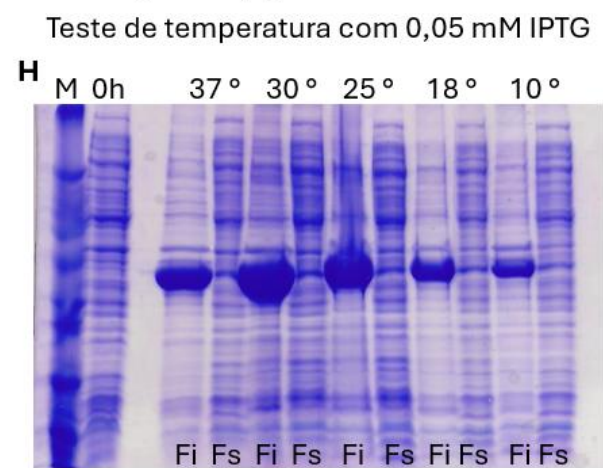
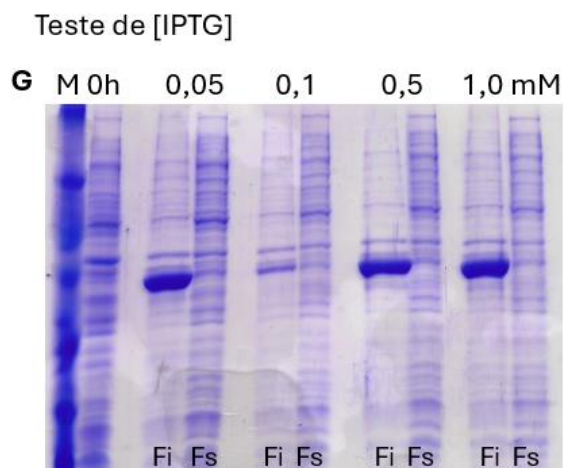
***E. coli* Rosetta-gami rELELpET32a(+)**



***E. coli* Origami B(DE3)pLysS rELELpET32a(+)**



***E. coli* ArcticExpress rELELpET32a(+)**



Fonte: Autora. Géis de eletroforese SDS-PAGE. M: Marcador molecular. 0h: alíquota do *pellet* celular retirada antes da indução. Fi: Fração insolúvel. Fs: Fração solúvel. **A:** Teste de concentração de IPTG na cepa de *E. coli* BL21(DE3) a 25 °C por 20h. **B:** Teste de temperatura de com 0,05 mM de IPTG na cepa de *E. coli* BL21(DE3). **C:** Teste de concentração de IPTG na cepa de *E. coli* Rosetta-gami a 25 °C por 20h. **D:** Teste de temperatura de com 0,05 mM de IPTG na cepa de *E. coli* Rosetta-gami. **E:** Teste de concentração de IPTG na cepa de *E. coli* Origami(DE3)pLysS a 25 °C por 20h. **F:** Teste de temperatura de com 0,05 mM de IPTG na cepa de *E. coli* Origami(DE3)pLysS. **G:** Teste de concentração de IPTG na cepa de *E. coli* ArcticExpress a 25 °C por 20h. **H:** Teste de temperatura de com 0,05 mM de IPTG na cepa de *E. coli* ArcticExpress.

A proteína de interesse rEEL foi expressa apenas na fração insolúvel em todas as condições testadas, incluindo diferentes temperaturas (37°C por 4 h, 30°C por 16h, 25 °C por 20h, 18°C por 24 horas e 10°C por 24 horas apenas para ArcticExpress) e diferentes concentrações de IPTG (0,05 mM; 0,1 mM; 0,5 mM e 1,0 mM) em todas as quatro cepas de expressão testadas. A proteína foi detectada pela banda visível no gel que não está presente na amostra coletada antes da indução.

A expressão da lectina rEEL na fração insolúvel foi confirmada por espectrometria de massas, após a digestão da lectina com a enzima tripsina. Os peptídeos identificados e analisados estão listados na Tabela 04.

Tabela 04 - Peptídeos sequenciados de Trx-rEEL na fração insolúvel.

Peptídeo	(m/z)	Sequência de aminoácidos	Massa		Δ (Da)
			Observada	Calculada	
rELEL					
T1	866.38	LLHLTDDSFDTDVLK	1730.74	1730.87	0.13
T2	903.42	MLAPLLDELADEYQGK	1804.80	1804.89	0.09
T3	886.97	GIPTLLLFQDGEVAATK	1771.92	1771.97	0.05
T4	563.54	QHMDSPDLGTDDDDK	1686.69	1687.66	0.97
T5	783.32	VCEGSSLTLSCPQK	1564.64	1564.72	0.08
T6	372.71	GAGISIAR	743.40	743.42	0.02
T7	870.84	TQVCPSDGATSNVNCK	1736.63	1736.74	0.11
T8	472.22	ASNALNVVR	942.42	942.52	0.10
T9	802.64	SSCTVEASNDVFGDPCMHTYK	2403.87	2403.97	0.10
T10	639.27	YLELSYDCSK	1276.52	1276.56	0.04

Fonte: Autora. Peptídeos sequenciados de Trx-rEEL transformada em *E. coli* BL21 resultantes da digestão com tripsina. (m/z): massa carga; Δ : massa calculada – massa observada; T: tripsina.

A Figura 10 mostra o alinhamento da sequência da proteína expressa Trx-rEEL com os peptídeos trípticos identificados por espectrometria de massas, confirmando a identidade da proteína da lectina recombinante.

Figura 10 – Alinhamento da sequência de aminoácidos da lectina Trx-rELEL com os peptídeos identificados por espectrometria de massas.

> Trx-rELEL

```

MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY
|-----T866.38-----| |-----T903.42-----|
QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL
-| |-----T886.97-----|
KEFLDANLAG SGSGHMHMHH HHSSGLVPRG SGMKETAAAK FERQHMDSPD
|-----T563.54-----|
LGTDDDDKAM ENLYFQG ELVSQLCLKK ERVCEGSSLT ISCPQKGAGI
-----| |-----T783.32-----| |T372.71-
SIARAIYGR KTQVCPDGA TSNVNCKASN ALNVVRDLCK GKSSCTVEAS
---| |-----T870.84-----| -T472.22---| |-----T802.64-----|
NDVFGDPCMH TYKYLELSYD CSK
-----| |-----T693.27-----|

```

Fonte: Autora. Letras sublinhadas: cauda de fusão (tiorredoxina, *tags* de purificação e sítios para as proteases trombina, enteroquinase e TEV). Letras em itálico: Tiorredoxina. Aminoácidos marcados com o tracejado abaixo da sequência, em vermelho, indicam peptídeos tripticos com o valor massa/carga obtidos por espectrometria de massas (MS/MS).

A bactéria *Escherichia coli* é o hospedeiro mais utilizado para produção de proteínas recombinantes. Esse sistema de expressão é predominante devido à capacidade de atingir rapidamente altas densidades celulares, a simplicidade, o baixo custo, é geneticamente bem caracterizada e aceita bem DNA exógeno (Fathi-Roudsari; Akhavian-Tehrani; Maghsoud, 2016; Kaur; Kumar, Kaur, 2018). No entanto, apesar de possuir inúmeras vantagens já consolidadas, ainda apresenta muitas limitações na expressão de proteínas solúveis oriundas de outros organismos (Pouresmaeil; Azizi-Dargahlou, 2023).

A expressão das proteínas em qualquer sistema heterólogo é controlada por diversos fatores intrínsecos do próprio hospedeiro. A *E. coli* tem um ambiente citoplasmático redutor, por isso não consegue expressar algumas proteínas com muitas ligações dissulfeto, como é o caso de rELEL que possui quatro pontes dissulfeto intracadeia e uma intercadeia (Lobstein *et al.*, 2012; Kuar; Kumar; Kuar, 2018; Carneiro *et al.*, 2015).

Outras lectinas de equinodermos que possuem pontes dissulfeto em suas estruturas também foram expressas na fração insolúvel, como observado com as lectinas do tipo C do pepino-do-mar *Apostichopus japonicus*, AjCTL-1 com duas pontes dissulfeto intracadeias, AjCTL-2 e AjSJL-1 com duas pontes dissulfeto intracadeias e duas cisteínas adicionais características do tipo de domínio *long-form* (Wei *et al.*, 2015; Wang *et al.*, 2018; Zhao *et al.*, 2021). E com a lectina MBL-AJ também isolada de *Apostichopus japonicus* com duas pontes dissulfeto intracadeia e uma intercadeia (Vasilenko *et al.*, 2012). As lectinas do pepino-do-mar *Cucumaria echinata*, CEL-III com 12 pontes dissulfeto intracadeias, e CEL-IV com uma ponte

dissulfeto intercadeia e duas intracadeia, também foram expressas na fração insolúvel (Shimizu *et al.*, 2012; Hatakeyama *et al.*, 2006).

4.3.1.3 Estratégias para expressão da proteína rEEL na fração solúvel

Para a expressão de proteínas recombinantes, o ideal é obter a proteína expressa de forma solúvel, porém não existe uma metodologia universal que seja eficaz para uma produção recombinante eficiente, com solubilidade e consequentemente atividade biológica. Cada proteína e sistema hospedeiro terão suas características e por isso diversas estratégias podem ser testadas para evitar a agregação proteica, como a coexpressão de chaperonas moleculares na bactéria hospedeira, uso de proteínas de fusão, a otimização da concentração do indutor, o ajuste da densidade óptica no momento da indução, a redução da temperatura e a adição de aditivos no meio de cultura (Kaur; Kumar; Kaur, 2018).

O uso de proteínas de fusão já foi bastante documentado como uma abordagem capaz de contornar os problemas de agregação. A adição de uma proteína de fusão ajuda na solubilidade pois são proteínas pequenas que se dobras rapidamente e corretamente no citoplasma da *E. coli*, no momento do dobramento a proteína de fusão carrega a proteína de interesse fusionada, favorecendo o dobramento correto. Além disso, a proteína de fusão pode cobrir parcialmente as regiões hidrofóbicas durante a síntese da proteína recombinante, evitando a interação dessas regiões e consequentemente a agregação (Costa *et al.*, 2013; (Rosano; Ceccarelli, 2014).

A proteína de fusão tiorredoxina está entre as mais usadas, apresentando ótimos resultados na solubilidade da proteína de interesse. Essa molécula pode melhorar a solubilidade atuando diretamente no dobramento correto de proteínas que possuem pontes dissulfeto pois é uma proteína redox ativa. Isso significa que ela participa de reações de oxidação e redução no ambiente celular, transferindo elétrons através de seus resíduos de cisteínas e auxiliando na correção e rearranjo das pontes dissulfeto formadas incorretamente, favorecendo a obtenção da conformação nativa (Rosano; Ceccarelli, 2014; Matsuzawa, 2017).

Além do uso de proteínas de fusão, outra estratégia bem estabelecida para melhorar a solubilidade de proteínas recombinantes é a redução da temperatura de expressão. A diminuição da temperatura reduz a taxa de tradução, levando a uma redução da velocidade da síntese proteica. Com menos proteína sendo expressa em um curto período de tempo, as moléculas recém-sintetizadas têm mais tempo para enovelar corretamente, diminuindo a chance

da formação de corpos de inclusão (Kaur; Kumar; Kaur, 2018). Para essa estratégia o uso da cepa *E. coli* ArcticExpress pode diminuir ainda mais a temperatura de expressão pois possui chaperonas e co-chaperonas que são adaptadas ao frio e exibem uma alta capacidade de dobramento correto em baixas temperaturas (Rosano, Ceccarelli, 2014).

Em outros estudos, a adição da tiorreoxina como proteína de fusão se mostrou crucial para a expressão de proteínas na fração solúvel. Quando combinadas com outras estratégias como o uso de baixas temperaturas de expressão e o uso de cepas geneticamente modificadas para ter um citoplasma mais oxidativo, como *E. coli* Rosetta-gami e *E. coli* Origami, essa abordagem demonstrou ser efetiva para melhorar a solubilidade da proteína expressa (Ganjave *et al.*, 2025). Porém, para rEEL essas estratégias não foram suficientes para expressar uma proteína solúvel e consequentemente ativa.

Outras abordagens foram aplicadas para tentar melhorar a solubilidade de rEEL, a redução da concentração do indutor, o aumento da densidade óptica de indução e a adição de aditivos no meio de cultura.

A concentração do indutor está relacionada a sistemas de promotores fortes, pois altas concentrações do indutor leva a maior síntese da proteína recombinante sobrecarregando o dobramento e levando a proteína para a fração insolúvel. Por isso, diminuir a concentração do indutor, principalmente em sistema de promotor forte como o T7, diminui a produção da proteína e consequentemente a molécula terá mais tempo para se enovelar corretamente (Kaur; Kumar; Kaur, 2018).

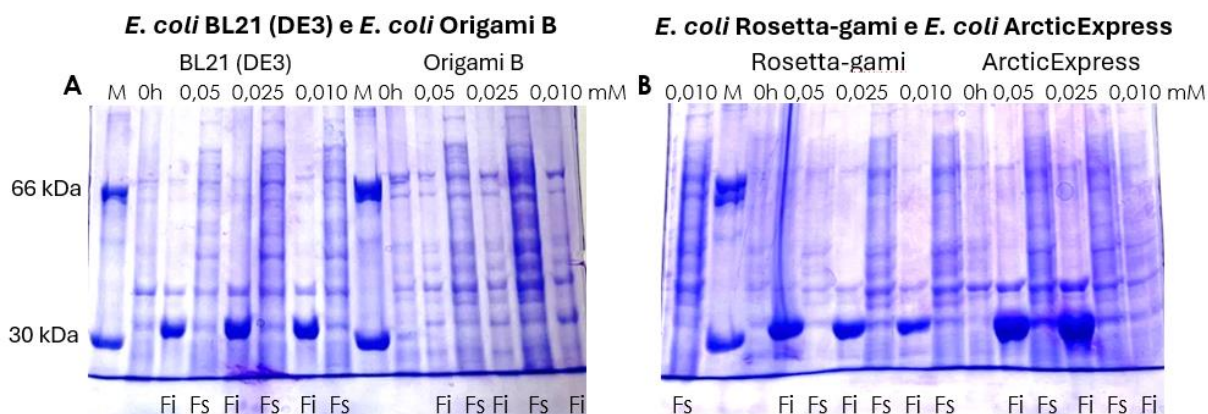
O aumento da densidade óptica de indução segue o mesmo princípio da redução da temperatura. Com uma densidade maior, as células entram na fase estacionária que é metabolicamente mais lenta, por isso sintetizam a proteína mais lentamente, diminuindo a agregação (Galloway; Sowden; Smith, 2003).

As novas concentrações de IPTG testadas foram: 0,05 mM; 0,025 mM e 0,010 mM em cada cepa de expressão (*E. coli* BL21(DE3), *E. coli* Rosetta-gami, *E. coli* Origami B (DE3) plysS e *E. coli* ArcticExpress). A densidade óptica de indução aumentou para a DO_{600nm} de 0,7, a expressão foi conduzida a 25°C, por 20 horas.

Ao se utilizar as estratégias de diminuição da concentração do indutor e aumento da densidade óptica de indução, a proteína recombinante rEEL foi detectada na fração solúvel das expressões em *E. coli* Rosetta-gami com as concentrações de 0,025 mM e 0,010 mM de IPTG e na cepa *E. coli* ArcticExpress com as concentrações de 0,05 mM e 0,025 mM de IPTG. A proteína foi detectada pela banda visível no gel que não está presente na amostra coletada

antes da indução. Entre as condições avaliadas, a expressão em *E. coli* ArcticExpress com a concentração de 0,05 mM de IPTG teve, aparentemente, uma maior abundância de proteína na fração solúvel, evidenciada por uma banda de maior intensidade no gel (Figura 11).

Figura 11 – Novos testes de expressão da proteína rEEL no vetor pET32a(+).



Fonte: Autora. Géis de eletroforese SDS-PAGE. M: Marcador molecular. 0h: alíquota do *pellet* celular retirada antes da indução. Fi: Fração insolúvel. Fs: Fração solúvel. **A:** Teste das novas concentrações de IPTG testadas 0,05 mM; 0,025 mM e 0,010 mM nas cepas de *E. coli* BL21(DE3) e *E. coli* Origami(DE3)pLysS. **B:** Primeiro poço: Fração solúvel do teste de 0,010 mM de IPTG na cepa *E. coli* Origami(DE3)pLysS. Poços restantes: Teste das novas concentrações de IPTG testadas 0,05 mM; 0,025 mM e 0,010 mM nas cepas de *E. coli* Rosetta-gami e *E. coli* ArcticExpress.

Tabela 05 - Peptídeos sequenciados de Trx-rEEL na fração solúvel.

Peptídeo	(m/z)	Sequência de aminoácidos	Massa		Δ (Da)
			Observada	Calculada	
rEEL_AE					
0.05 mM					
T1	634.30	LNLDQNPGTAPK	1266.6044	1266.6567	0.05
T2	472.25	ASNALNVVR	942.4644	942.5247	0.06
T3	866.93	LLHLTDDSFDTDVLK	1730.8844	1730.8727	-0.01
rEEL_RG					
0.025 mM					
T1	472.24	ASNALNVVR	942.4644	942.5247	0.06
T2	866.93	LLHLTDDSFDTDVLK	1730.8243	1730.8727	0.05

Fonte: Autora. Peptídeos sequenciados de Trx-rEEL na fração solúvel resultantes da digestão com tripsina. rEEL_AE: Expressão na cepa *E. coli* ArcticExpress na concentração de 0,05 mM de IPTG. rEEL_RG: Expressão na cepa *E. coli* Rosetta-gami na concentração de 0,025 mM de IPTG. (m/z): massa carga; Δ : massa calculada – massa observada; T: tripsina.

A expressão da lectina rELEL na fração solúvel foi confirmada por espectrometria de massas, após a digestão da lectina com a enzima tripsina. Os peptídeos identificados e analisados estão listados na Tabela 05.

A Figura 12 mostra o alinhamento da sequência da proteína expressa Trx-rELEL expressa na fração solúvel com os peptídeos tripticos identificados por espectrometria de massas, confirmando a identidade da lectina recombinante.

Figura 12 – Alinhamento da sequência de aminoácidos da lectina Trx-rELEL expressa na fração solúvel com os peptídeos identificados por espectrometria de massas.

>rELEL

10	20	30	40	50	60
MSDKIIHLTD	DSFDTDVLKA	DGAILVDFWA	EWCGPCKMIA	PILDEIADEY	QGKLTVAKLN
--T3-866.93----					--
--T2-866.93----					
70	80	90	100	110	120
IDQNPGTAPK	YGIRGIPTLL	LFKNGEVAAT	KVGALSKGQL	KEFLDANLAG	SGSGHMHMHH
T1-634.30					
130	140	150	160	170	180
HHSSGLVPRG	SGMKETAAAK	FERQHMDSPD	LGTDDDDKAM	ENLYFQGELV	SQLCLKKERV
190	200	210	220	230	240
CEGSSLTISC	PQKGAGISIA	RAIYGRKTQ	VCPSDGATSN	VNCKASNALN	VVRDLCRGKS
				T2-472.25	
				T1-472.24	
250	260	270			
SCTVEASNDV	FGDPCMHTYK	YLELSYDCSK			

Fonte: Autora. Letras sublinhadas: cauda de fusão (tiorredoxina, *tags* de purificação e sítios para as proteases trombina, enteroquinase e TEV). Letras em itálico: Tiorredoxina. Aminoácidos marcados com o tracejado abaixo da sequência indicam peptídeos tripticos com o valor massa/carga obtidos por espectrometria de massas (MS/MS). Vermelho: peptídeos oriundos da rELEL expressa em *E. coli* ArcticExpress com 0,05 mM de IPTG. Azul: peptídeos oriundos da rELEL expressa em *E. coli* Rosetta-Gami com 0,025 mM de IPTG.

Com o objetivo de purificar a lectina recombinante, a fração solúvel dos testes em *E. coli* ArcticExpress com 0,05 mM de IPTG e em *E. coli* Rosetta-Gami com 0.025 mM de IPTG, que apresentaram confirmação da expressão da lectina por espectrometria de massas, foi aplicada em uma coluna de lactosil-agarose, conforme o método empregado para a purificação da lectina nativa (Carneiro *et al.*, 2015). No entanto, nenhuma ligação a coluna foi observada.

As mesmas frações solúveis foram submetidas à cromatografia de afinidade em uma coluna HisTrap™ FF, onde deveria ocorrer a ligação da cauda de histidina a matriz de níquel imobilizado. Novamente, nenhuma interação com a coluna foi observada. Para avaliar

se a proteína presente nas frações solúveis estava ativa, a amostra foi submetida a um teste de atividade hemaglutinante contra eritrócitos de coelho e nenhuma atividade hemaglutinante foi detectada (dados não mostrados).

A redução da concentração do indutor e o aumento da densidade óptica de indução foram eficientes em expressar rEEL na fração solúvel. Porém, apesar de solúvel, a proteína recombinante não apresentou atividade hemaglutinante e nem interagiu com as colunas específicas para a especificidade de ligação a carboidratos da lectina ou para a cauda de histidina, mostrando que provavelmente a lectina não obteve a sua conformação nativa funcional.

A cromatografia de afinidade para lectinas é baseada entre a interação específica da lectina e os carboidratos imobilizados na matriz cromatográfica. Portanto, proteínas capazes de se ligar ao seu carboidrato específico na matriz, demonstra funcionabilidade. Se a proteína recombinante não consegue se ligar na coluna de afinidade ao carboidrato específico significa que ela não está funcionalmente ativa e não é capaz de realizar o reconhecimento a carboidratos, possivelmente por dobramento incorreto. A falta da capacidade funcional foi confirmada pela ausência da atividade hemaglutinante (O'Connor; Monaghan; Cawley, 2017; Deb; Manfield; Carpenter, 1997).

A ausência da ligação à coluna de afinidade pela histidina também valida a hipótese de dobramento incorreto pois no caso a cauda de histidina está estruturalmente inacessível, uma vez que em condições de dobramento correto, a purificação por afinidade da histidina à coluna de níquel é eficiente.

Estudos anteriores já descreveram a expressão de proteínas na fração solúvel, mas sem apresentar atividade biológica. Isso pode acontecer pois as caudas de fusão podem ajudar na solubilidade, porém nem sempre irão atuar no dobramento correto da proteína (DE Coninck; Vanhaeren; Van Damme, 2025). Esses resultados demonstram que a expressão da proteína na fração solúvel observada por SDS-PAGE não é garantia de um dobramento correto para manter a funcionabilidade da proteína nativa.

Uma estratégia alternativa é a suplementação do meio de cultura com cofatores essenciais para o enovelamento correto e atividade funcional da proteína. A presença desses aditivos durante a expressão recombinante favorece o dobramento adequado da cadeia polipeptídica recém-sintetizada (Kaur; Kumar; Kaur, 2018).

A indução também foi testada na presença dos aditivos PEG (Polietilenoglicol) a 3g/L e NaCl a 10 g/L no meio de cultura. O PEG pode ajudar a manter a estabilidade da proteína,

enquanto o NaCl ajuda a manter a força iônica do meio, reduzindo interações eletroestáticas indesejadas e consequentemente, reduzindo a formação de agregados (Kaur; Kumar; Kaur, 2018; Chopra *et al.*, 1994). Novamente, nenhuma expressão na fração solúvel foi detectada (dados não mostrados). Para avaliar se a proteína estaria sendo secretada para o meio extracelular, o sobrenadante foi precipitado com sulfato de amônio a 70% e analisado em SDS-PAGE, e nenhuma proteína do tamanho esperado para rEEL foi detectada (dados não mostrados).

4.3.1.4 Tentativa de refolding da lectina recombinante a partir da fração insolúvel

As proteínas mal enoveladas que estão nos corpos de inclusão não possuem atividade biológica, pois o funcionamento da molécula depende do dobramento correto. Assim, essas proteínas precisam passar por um processamento para serem solubilizadas e redobradas a partir dos corpos de inclusão (Kaur; Kumar; Kaur, 2018).

O processo de recuperação de proteínas a partir de corpos de inclusão envolve as etapas de purificação dos corpos de inclusão, a solubilização, o *refolding* das proteínas solubilizadas com um tampão otimizado, e a purificação da proteína renaturada. Todas essas etapas podem ser otimizadas para cada molécula, promovendo uma maior recuperação da proteína bioativa a partir dos corpos de inclusão (Singh *et al.*, 2015).

Apesar de existirem inúmeros relatos de proteínas redobradas com sucesso por *refolding*, recuperar as proteínas dos corpos de inclusão é um processo que possui muitas desvantagens. A principal delas é a baixa recuperação da proteína ativa devido a perda da estrutura secundária e a agregação durante a solubilização e o redobramento. Além disso, os procedimentos de redobramento podem não restaurar completamente a conformação nativa da proteína e podem reduzir a função da proteína recombinante (Fathi-Roudsari; Akhavian-Tehrani; Maghsoud, 2016).

Por isso, o processo de refolding é considerado apenas após inúmeras tentativas sem sucesso de obter a proteína da fração solúvel, porque a eficiência do redobramento pode ser baixa e as proteínas redobradas podem ser instáveis devido ao redobramento incompleto, o que torna o processo trabalhoso e pouco rentável.

Apesar das dificuldades encontradas durante a solubilização e redobramento das proteínas dos corpos de inclusão, lectinas de equinodermos que possuem pontes dissulfeto em suas estruturas foram recuperadas biologicamente ativas pelo processo de *refolding*. Na grande

maioria, as lectinas recombinantes expressas na fração insolúvel foram solubilizadas em tampão contendo Ureia 8 M e o redobramento ocorreu por meio de diálises com tampão em gradiente decrescente de ureia (6 M; 4 M; 2 M; 0 M) contendo um sistema redox baseado em glutathiona reduzida e oxidada (Wei *et al.*, 2015; Wang *et al.*, 2018; Zhao *et al.*, 2021; Vasilenko *et al.*, 2012; Hatakeyama *et al.*, 2006).

Para tentar recuperar a proteína rEEL dos corpos de inclusão, as frações contendo a lectina na fração insolúvel foram reunidas e submetidas a lavagens com tampão contendo Tris 50 mM pH 8, NaCl 0,5 M, Ureia 4M e EDTA 1 mM. A fração insolúvel foi solubilizada em tampão desnaturante contendo Ureia 8M, DTT 10 mM e Tris 50 mM pH 8 e o redobramento foi realizado por diálises sucessivas usando o tampão Tris 50 mM, pH 7,5 com NaCl 150 mM e DTT 1 mM, reduzindo progressivamente a concentração de ureia no tampão de redobramento (6 M, 4 M, 2 M, 0 M).

Apesar da solubilização em Ureia 8 M e o redobramento por diálises consecutivas com redução gradual da concentração de ureia seguirem a mesma estratégia descrita com sucesso para outras lectinas de equinodermos, esse procedimento não foi eficiente para a rEEL pois não permitiu recuperar a lectina da fração insolúvel (dados não mostrados).

O problema para rEEL pode estar relacionado ao tampão de redobramento. Agentes redutores fortes, como o DTT, podem ser usados nas etapas iniciais de solubilização para remoção das pontes dissulfeto incorretas. No entanto, no momento do redobramento é necessário um sistema redox equilibrado, ajustando o DTT com um agente oxidante que permita a formação das pontes dissulfetos corretas (Kaur; Kumar; Kaur, 2018). Como descrito para as lectinas de equinodermos citadas anteriormente, combinações de glutathiona reduzida e oxidada em proporções ideais conseguiram manter o dobramento correto, garantindo a funcionabilidade da lectina.

4.3.2 Produção da proteína recombinante rEEL em sistema heterólogo de *E. coli* (França)

4.3.2.1 Subclonagens do gene sintético para os vetores de expressão

O gene sintético da lectina rEEL, com 323 pb incluindo os sítios de restrição, foi inicialmente clonado no vetor *backbone* pEX-A128, delimitado pelos sítios de restrição *NcoI* e *XhoI* e com a sequência gênica otimizada para *Escherichia coli*. O quadro aberto de leitura (ORF) possui 312 pb, codificando uma proteína de 104 aminoácidos. Após a digestão, o inserto foi subclonado nos vetores de expressão pCOLD-TF-3G-Nco, pET39-TEV, pET15b(+) e

pCOLD-HAT-TEV-3G-Nco. A confirmação da subclonagem foi obtida por PCR e digestão enzimática, de acordo com o tamanho esperado para rEEL e os respectivos vetores, visualizado em gel de agarose. Todas as construções apresentaram o inserto no tamanho esperado, e a subclonagem em todos os vetores foi confirmada por sequenciamento (dados não mostrados).

Apesar dos problemas observados na expressão de rEEL em *E. coli* com a formação dos corpos de inclusão, a expressão em sistema procariótico continua sendo o mais atraente devido as principais vantagens de baixo custo e rápido crescimento, além de numerosas ferramentas genéticas para otimização e subclonagens em diferentes vetores e estratégias de expressão (Lobstein *et al.*, 2012). Por esses motivos, decidiu-se por continuar as tentativas de expressão da rEEL em *E. coli*, porém aumentando as chances de sucesso com a otimização de códons raros para esse sistema hospedeiro.

Apesar das inúmeras vantagens da expressão heteróloga em *E. coli*, a produção de proteínas recombinantes neste organismo pode ser prejudicada pelo uso de códons raros. Nesses casos, se o viés de códons do gene a ser expresso for significativamente diferente daquele utilizado pela *E. coli*, as concentrações de RNA transportador (tRNA) da *E. coli* para os códons menos utilizados serão insuficientes para traduzir o RNA de forma eficiente (Rosano; Ceccarelli, 2014).

Isso acontece porque cada aminoácido é codificado por mais de um códon e cada organismo tem sua própria preferência no uso dos diferentes códons disponíveis. As frequências com que os diversos códons são usados variam entre os diferentes organismos, alguns códons que codificam o mesmo aminoácido são usados com mais frequência do que outros, o que reflete na quantidade de tRNA genômico. Em condições heterólogas, se o hospedeiro expressar uma proteína para a qual não possui tRNA abundante devido ao problema de preferência de códons, os níveis de expressão são comprometidos, resultando em nenhuma expressão ou em uma proteína recombinante truncada (Kaur; Kumar; Kaur, 2018; Rosano; Ceccarelli, 2014).

As consequências do uso de códons raros incluem problemas como paralisação da tradução, terminação prematura da tradução, mudança na fase de leitura e incorporação incorreta de aminoácidos. Isso causa redução na quantidade ou qualidade da proteína expressa. Além disso, a competição por tRNA raro também pode afetar negativamente a expressão ou desencadear uma resposta de estresse (Burgess-Brown *et al.*, 2008).

Para contornar esse problema de viés de códons, a estratégia mais comum é a construção de um gene sintético onde a maioria dos códons raros é substituído por aqueles mais

frequentemente usados no hospedeiro. Outra solução é usar uma cepa de *E. coli* geneticamente modificada para conter cópias extras de tRNAs raros para a expressão da proteína exógena. Os códons raros foram definidos para *E. coli* como os códons usados com uma frequência <1% (Kaur; Kumar; Kaur, 2018; Rosano; Ceccarelli, 2014).

Atualmente, existe uma grande disponibilidade comercial de inúmeros vetores de expressão, com diversas origens de replicação, diferentes promotores, regiões de iniciação da tradução, marcadores de resistência a antibióticos, terminadores de transcrição, sítios de restrição, *tags* de purificação e solubilidade, dentre muitas outras características. A seleção do vetor adequado, juntamente com o uso de genes com códons otimizados é, em muitos casos, suficiente para permitir o acúmulo da proteína expressa em um nível elevado. No entanto, achar o vetor com as características ideais para maximizar a produção de determinada proteína é muitas vezes questão de tentativa e erro (Makino; Skretas; Georgiou, 2011).

A escolha dos vetores foi baseada em estratégias pensadas para contornar o problema de formação de corpos de inclusão na expressão de rEEL. O vetor pCOLD-TF-3G-Nco possui o TEE (*Transcriptional Enhancer Element*) e a chaperona molecular *Trigger factor* que podem aumentar a solubilidade da proteína recombinante. O TEE e o *Trigger factor* são expressos, resultando em uma proteína fusionada de maior tamanho, que pode favorecer o dobramento correto da proteína. Além disso, o vetor pCOLD permite a expressão da proteína em temperaturas reduzidas, o que pode melhorar a solubilidade da proteína. O vetor tem *tags* de purificação e sítios de clivagem para diferentes proteases.

No vetor pET39-TEV, a lectina de interesse é expressa fusionada com a proteína DsbA que atua secretando a proteína expressa para o periplasma, ambiente capaz de atuar na formação das pontes dissulfeto, tema que será discutido nas próximas seções. O vetor também possui *tags* de purificação e sítios de clivagem para diferentes proteases.

No vetor pET15b(+), a construção com os sítios de restrição *NcoI* e *XhoI* permite a expressão da proteína sem nenhuma cauda fusionada. Essa estratégia consegue obter uma proteína igual a nativa sem alterações estruturais, permitindo a purificação com o mesmo protocolo da lectina nativa.

No vetor pCOLD-HAT-TEV-3G-Nco, a proteína de interesse é fusionada a uma cauda pequena, composta apenas pela TEE, cauda de purificação com histidina e o sítio de clivagem da protease TEV. Essa estratégia minimiza alterações estruturais da proteína e pode resultar em uma melhor purificação.

4.3.2.2 Expressão da lectina rELEL em sistema heterólogo

Todas as construções foram transformadas na cepa de clonagem *E. coli* NEB5α e na cepa de expressão *E. coli* SHuffle B, após confirmação da subclonagem por sequenciamento. A transformação na cepa de clonagem é para extração do DNA plasmidial e a escolha da cepa de expressão foi baseada na estrutura de rELEL com muitas pontes dissulfeto.

A cepa *E. coli* SHuffle é geneticamente modificada especificamente para a expressão de proteínas recombinantes com ligações dissulfeto. Nessa cepa, a glutaredoxina redutase e a tiorredoxina redutase, enzimas do sistema redox citoplasmático, são deletadas. Além disso, a isomerase de ligação dissulfeto periplasmática DsbC, que corrige pontes dissulfeto, não possui sua sequência sinal e é expressa no citoplasma. Assim, a cepa é capaz de expressar proteínas recombinantes com múltiplas ligações dissulfeto no citoplasma, além de corrigir ligações mal oxidadas e promover o dobramento correto. A produção de diversas proteínas recombinantes contendo ligações dissulfeto foi comprovadamente aumentada na fração solúvel com a utilização dessa cepa (Lobstein *et al.*, 2012).

Para expressão da proteína, a cultura foi crescida a 30 °C, até atingir a densidade ótica de 0,4. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu até atingir DO = 0,8, quando o indutor IPTG foi adicionado na concentração de 0,2 mM e a expressão continuou a 16 °C.

Para o vetor pCOLD-HAT-TEV-3G-Nco, a expressão foi realizada em meio de autoindução LB NYZ, pois a estratégia do mesmo vetor com meio de autoindução, já havia produzido resultados positivos com outras lectinas. As células foram cultivadas a 30 °C até atingir a DO_{600nm} superior a 5,0. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu por 18 horas.

A temperatura de cultivo a 30 °C foi escolhida seguindo as recomendações para a cepa *E. coli* SHuffle (New England Biolabs, 2025). A indução ocorreu a 16 °C, pois temperaturas de indução mais baixas podem ajudar na solubilidade da proteína recombinante (Kaur; Kumar; Kaur, 2018). Além disso, estudos anteriores com lectinas ligantes a ramnose, assim como rELEL, demonstraram que a temperatura de indução de 16 °C produziu ótimos resultados para a cepa SHuffle (Warfel *et al.*, 2023), por isso essa foi a temperatura de indução escolhida para expressar rELEL.

4.3.2.2.1 Expressão da lectina rELEL no vetor pCOLD-TF-TEV-3G-Nco

A construção da proteína recombinante expressa no pCOLD-TF-TEV-3G-Nco origina uma proteína de fusão composta pelo TEE, a cauda de histidina, a chaperana *Trigger factor*, os sítios de clivagem para as proteases HRV 3C e TEV, o sítio de clivagem desativado para trombina, e finalmente, a lectina de interesse (Figura 13).

Figura 13 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-TF-TEV-3G-Nco.

>rELEL expressa em pCOLD-TF-TEV-3G-Nco

```

      10      20      30      40      50      60      70      80
  MNHKVHHHHH HMQVSVETTQ GLGRRVTIT I AADSIETAVK SELVNVAKKV RIDGFRKGKV PMNIVAQRYG ASVRQDVLGD
      90     100     110     120     130     140     150     160
  LMSRNFIDA I I KEKINPAG A PTYVPGEYK I GEDFTYSVEF EVYPEVELQG LEAIEVEKPI VEVTADVDG MLDTLRKQQA
      170     180     190     200     210     220     230     240
  TWKEKDGA V E AEDRVTIDFT G SVDGEEFEG GKASDFVLAM GQGRMIPGFE DGIKGHKAGE EFTIDVTFPE EYHAENLKGK
      250     260     270     280     290     300     310     320
  AAKFAINLKK VEERELPELT A EFIKRFGE DGSVEGLRAE VRKNMERELK SAIRNRVKSQ AIEGLVKAND IDVPAALIDS
      330     340     350     360     370     380     390     400
  EIDVLRQA A Q RFGGNEKQA L ELPRELFEE QAKRRVVVGL LLGEVIRTNE LKADEERVKG LIEEMASAYE DPKEVIEFYS
      410     420     430     440     450     460     470     480
  KNKELMDNMR NVALEEQA V E AVLAKAKVTE K ETTFNELMN QQASAGLEV L FQGFSAGLVP AGSGGENLYF QGGGAMELVS
      490     500     510     520     530     540     550     560
  QLCLKKERV C EGSSLTISC P QKGAGISIA R AIYGRTKTQ V CPSDGATSNV NCKASNALNV VRDLCRGKSS CTVEASNDVF
      570     579
  GDPCMHTYKY LELSYDCSK
  
```

Número de aminoácidos: 579

Peso molecular: 63803,3416

pI: 5,19

TEE

Cauda de histidina

Trigger factor

Sítio da protease HRV 3C

Sítio de trombina desativado

Sítio de clivagem TEV

Lectina

Coefficiente de extinção:

O coeficiente de extinção é: 24870

Em 1mg/ml (15,652 uM) a solução da proteína tem

Abs280nm de: 0,39

Se todas as cisteínas formarem pontes dissulfeto,

o coeficiente de extinção é: 25370

Em 1mg/ml (15,652 uM) a solução de proteína tem

Abs280nm de: 0,40

Fonte: Autora. pI: Ponto isoelétrico. TEE: *Translation Enhancing Element*. HRV 3C: Protease do rinovírus humano. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas a cauda fusionada à rELEL, sem a proteína de interesse, possui 475 aminoácidos com um peso molecular de 52.700,65 Da e com pI de 5. A proteína inteira expressa, com a lectina de interesse, possui 579 aminoácidos com um peso molecular de 63.803,34 Da e com pI de 5,19.

A proteína expressa foi pensada para ser clivada com a protease TEV. Após clivagem, rELEL possui 108 aminoácidos com peso molecular de 11.448,05 Da e com pI de 8,4. A Figura 14 mostra a representação da proteína de interesse rELEL após clivagem com a protease TEV.

Figura 14 – Representação da proteína rELEL obtida após clivagem com a protease TEV.

>rELEL clivada

```

      10      20      30      40      50      60      70      80
GGGAMELVSQ LCLKKERVCE GSSLTISCPQ KGAGISIIARÄ IYGRTKTQVC PSDGATSNVN CKASNALNVV RDLCRGKSSC
      90     100     108
TVEASNDVFG DPCMHTYKYL ELSYDCSK

```

Sítio TEV residual

Número de aminoácidos: 108

Linker residual

Peso molecular: 11448,0533

pI: 8,40

Lectina

Coefficiente de extinção:

O coeficiente de extinção é: 5960

Em 1mg/ml (87,351 uM) a solução da proteína tem Abs280nm de: 0,52

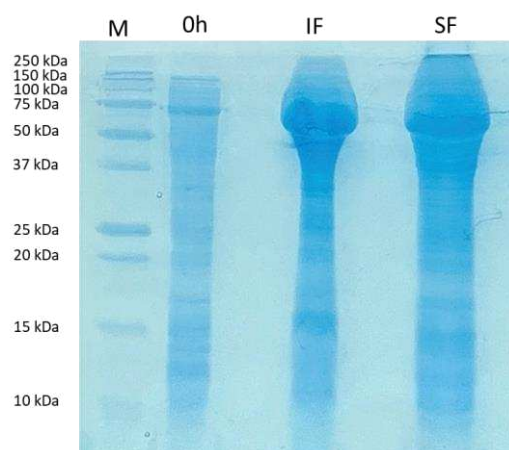
Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 6460

Em 1mg/ml (87,351 uM) a solução de proteína tem Abs280nm de:0,56

Fonte: Autora. pI: Ponto isoelétrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

A expressão de rELELpCOLD-TF-TEV-3G-Nco foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE (Figura 15).

Figura 15 - Análise da expressão de rELELpCOLD-TF-TEV-3G-Nco em SDS-PAGE

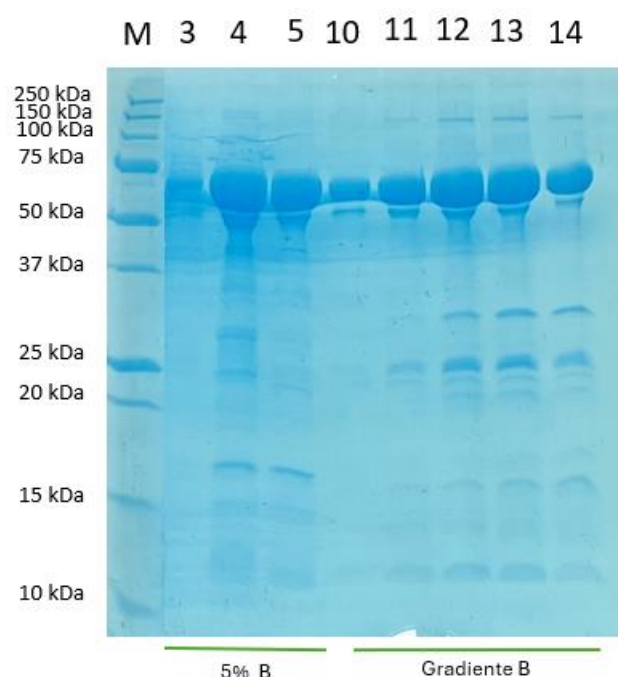


Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. 0h: Alíquota retirada antes da indução. IF: Fração insolúvel. SF: Fração solúvel.

A proteína de interesse rELELpCOLD-TF-TEV-3G-Nco foi expressa tanto na fração insolúvel como na fração solúvel, detectada pela banda visível no gel que não está presente na amostra coletada antes da indução, entre as bandas de 50 e 75 kDa do marcador, corroborando com o tamanho esperado da proteína de aproximadamente 64 kDa.

Como o vetor pCOLD-TF-TEV-3G-Nco possui cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados (IMAC). A coluna foi lavada com tampão contendo imidazol 10 mM e as proteínas retidas a coluna foram eluídas com 5% do tampão com imidazol 500 mM (tampão B), e com um gradiente linear de 5% a 100% do mesmo tampão. As frações foram coletadas e analisadas em SDS-PAGE (Figura 16).

Figura 16 - Análise da purificação de rEELpCOLD-TF-TEV-3G-Nco em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 3 a 5: eluídos com 5% de B. Tubos 10 a 14: eluídos com gradiente de tampão B.

A análise em SDS-PAGE confirmou a ligação e a recuperação da proteína na coluna, observada pela banda de tamanho esperado de 64 kDa nos dois picos. Em ambas as frações eluídas, foi observado contaminantes junto com a proteína de interesse, evidenciado pelas múltiplas bandas presente no gel.

Durante a purificação na cromatografia IMAC, ligações não específicas com a coluna podem ocorrer porque proteínas endógenas do hospedeiro possuem regiões ricas em cisteínas ou histidinas ou regiões carregadas que podem interagir com o níquel imobilizado, além de interações hidrofóbicas, resultando em uma retenção inespecífica e a coeluição da proteína de interesse com contaminantes. Essas ligações podem ser evitadas pelo uso de

imidazol em baixas concentrações no tampão de ligação e de lavagem, que competem seletivamente com interações de baixa afinidade. Além disso, altas forças iônicas, como o uso de tampão com elevada concentração de NaCl, podem reduzir interações inespecíficas (Young; Britton; Robinson, 2012).

A eluição em picos sucessivos ou em gradiente de imidazol pode ajudar na separação dos contaminantes, pois explora as diferenças de afinidade dos contaminantes e da proteína alvo. Os contaminantes geralmente interagem mais fracamente com a coluna e são eluídos com menores concentrações de imidazol, enquanto a proteína de interesse, com a cauda de histidina, se liga com maior afinidade e é eluída com concentrações mais elevadas de imidazol (Bornhorst; Falke, 2000).

A estratégia de eluição inicial com baixa concentração de imidazol foi aplicada na purificação para ajudar na separação dos contaminantes. O primeiro pico foi eluído com 5% do tampão B, que totaliza 35 mM de imidazol, para remover os contaminantes fracamente ligados à coluna. Porém, esse protocolo não foi eficiente, pois a proteína já começou a ser eluída com quantidades baixas de imidazol. Além disso, mesmo em concentrações elevadas de imidazol, a proteína de interesse foi eluída com bastante contaminante.

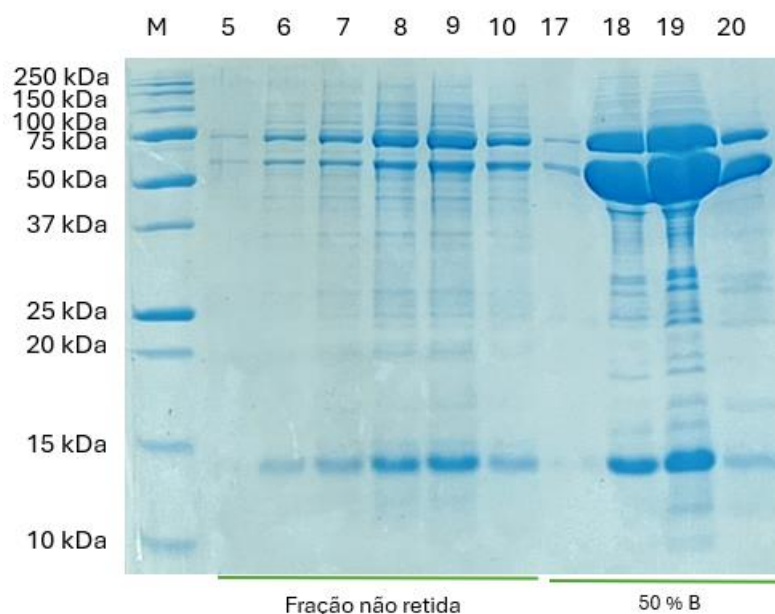
Isso pode acontecer porque algumas proteínas endógenas de *E. coli* tem afinidade real por níquel e se ligam tão fortemente à coluna quanto a cauda de histidina, assim essas proteínas não são eluídas nas lavagens iniciais e só se desligam da coluna em quantidades elevadas de imidazol, junto com a proteína de interesse. Além disso, um contaminante pode não se ligar ao níquel diretamente, mas sim interagir com a própria proteína alvo, formar um complexo estável proteína com proteína, e ser eluído por associação (Bolanos-Garcia; Davies, 2006).

Os tubos coletados referentes a eluição com 5% de B (3,4 e 5) pareciam conter menos contaminantes e por isso foram reunidos e seguiram para digestão com a protease TEV com o objetivo de obter a proteína rEEL clivada.

O tampão foi trocado para Tris 20 mM pH 8, contendo NaCl 100 mM, a proteína foi concentrada para 5 mg/mL e a protease TEV foi adicionada a solução juntamente com 0,5 mM de EDTA. Após a reação de digestão, a concentração de imidazol foi ajustada e a solução foi submetida novamente a cromatografia de afinidade IMAC. A coluna foi lavada com o novo tampão contendo imidazol 10 mM e a fração das proteínas ligadas a coluna, foi eluída com um step de 50% do tampão B. O primeiro pico, das proteínas que não são adsorvidas à matriz, é referente a proteína de interesse clivada, pois perdeu a cauda de histidina que se liga a coluna.

O segundo pico é referente às proteínas não clivadas, que se ligaram a coluna por possuírem a cauda de histidina, além da TEV e da cauda clivada, sem a proteína de interesse. As frações foram coletadas e analisadas em SDS-PAGE (Figura 17).

Figura 17 - Análise da purificação de rEEL clivada com TEV em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 5 a 10: fração lavada da coluna. Tubos 17 a 20: eluídos com 50% de tampão B.

Após a digestão, o esperado é que a maioria das proteínas sejam clivadas, porém o pico retido a coluna teve uma absorbância muito maior que o pico não retido (dados não mostrados), demonstrando que provavelmente a clivagem não havia sido efetiva. A análise em SDS-PAGE confirmou a hipótese, mostrando que havia proteínas clivadas, evidenciada pela presença da banda entre 10 e 15 kDa do marcador, que corrobora com o tamanho esperado para a proteína clivada de aproximadamente 11 kDa. No entanto, ainda é observado a presença de proteínas não clivadas pela banda mais alta próxima a 75 kDa, com o tamanho esperado para rEELpCOLD-TF-TEV-3G-Nco de 64 kDa, demonstrando que a clivagem não foi eficiente.

O vetor pCOLD-TF-TEV-3G-Nco possui duas glicinas inseridas entre o sítio de clivagem TEV e o sítio de restrição *NcoI*. A adição dessas glicinas teve como objetivo criar uma região mais flexível para a protease TEV, melhorando o acesso da enzima ao sítio de clivagem, e consequentemente, aumentando a eficiência da digestão para esse vetor (Waugh, 2011). A escolha do vetor com duas glicinas a mais para rEEL foi baseado na clivagem

ineficiente observada com o vetor pCOLD-TF-TEV-Nco para outras proteínas, porém a mesma ineficiência foi observada para rEELpCOLD-TF-TEV-3G-Nco.

Além disso, ambas as frações possuíam contaminantes e três bandas majoritárias. A banda mais baixa referente a rEEL clivada com o tamanho esperado de 11 kDa, a banda mais próxima a 50 kDa referente a proteína fusionada expressa cortada sem a lectina de interesse, com o tamanho esperado de aproximadamente 53 kDa, e a banda mais próxima a 75 kDa referente as proteínas não clivadas. Isso demonstra que assim como a clivagem, a separação também não foi efetiva.

Como a cromatografia IMAC não foi eficiente em separar as proteínas clivadas, as frações foram submetidas a cromatografia de afinidade com base na especificidade a carboidratos, seguindo o protocolo da lectina nativa (Carneiro *et al*, 2015).

A fração não retida a coluna (tubos coletados 5 a 10) foi submetida a cromatografia em matriz de ramnose e a fração retida a coluna (tubos coletados 17 a 20) foi aplicada em matriz de α -Lactose-Sepharose. As proteínas potencialmente ligadas as colunas foram eluídas com 0,3 M de ramnose, no entanto, nenhuma ligação nas colunas foi observada. Como a retenção na coluna depende da capacidade funcional da lectina, esse dado demonstra que a proteína expressa não estava ativa, provavelmente por dobramento incorreto, assim como observado anteriormente para rEEL expressa no pET32a(+).

Para confirmar a não funcionabilidade da proteína, alíquotas das amostras antes da clivagem e depois da clivagem, a mais de 1 mg/mL, foram submetidas a um teste de atividade hemaglutinante contra eritrócitos de coelho e nenhuma atividade hemaglutinante foi detectada (dados não mostrados).

Conforme mencionado anteriormente, a lectina expressa não funcional foi capaz de se ligar a coluna de níquel imobilizado porque nesse tipo de cromatografia de afinidade, a interação não acontece pelo domínio de reconhecimento a carboidratos da lectina, portanto a ligação a coluna HisTrap não garante o dobramento correto e a atividade funcional da proteína (Li; Su; Janson, 2004; Young; Britton; Robinson, 2012).

Portanto, obter uma proteína recombinante solúvel não é garantia de atividade biológica mantida. Principalmente se a proteína possuir muitas pontes dissulfeto, como é o caso de rEEL, pois a formação correta dessas ligações garante a conformação nativa e funcional da proteína recombinante (Rosano; Ciccarelli, 2014).

Uma lectina encontrada em embriões do ouriço-do-mar *Paracentrotus lividus* foi expressa ativa no vetor pCOLD-TF, onde a proteína foi expressa fusionada com o *trigger factor*.

O plasmídeo foi inserido em *E. coli* BL21 e a expressão foi conduzida com 1 mM de IPTG a 15 °C. A grande diferença para rEEL é que a lectina expressa é uma galectina, não possuindo pontes dissulfeto em sua estrutura (Karakostis *et al.*, 2015).

4.3.2.2.2 Expressão da lectina rEEL no vetor pET39-TEV

A construção da proteína recombinante expressa no vetor pET39-TEV origina uma proteína de fusão composta pela oxidoreductase periplasmática DsbA, a cauda de histidina, a S-tag, os sítios de clivagem para as proteases trombina e TEV e a lectina de interesse (Figura 18).

Figura 18 – Representação esquemática da proteína recombinante expressa no vetor pET39-TEV.

>rEEL expressa em pET39-TEV

```

      10      20      30      40      50      60      70      80
MKKIWLALAG LVLAFSASAA QYEDGKQYTT LEKPVAGAPQ VLEFFSFFCP HCYQFEEVLH ISDNVKKKL P EGVKMTKYHV
  90      100     110     120     130     140     150     160
NFMGGDLGKD LTQAWAVAMA LGVEDKVTVP LFEGVQKTQT IRSASDIRDV FINAGIKGEE YDAAWNSFVV KSLVAQQEKA
 170     180     190     200     210     220     230     240
AADVQLRGVP AMFVNGKYQL NPQGMDTSNM DVFVQYQYADT VKYLSEKKGS TSGSGHHHHH HSAGLVPRGS TAIGMKETA A
 250     260     270     280     290     300     310     320
AKFERQHMD S PDLGTENLYF QGAMELVSQ L CLKKERVCEG SSSLTISCPQK GAGISIA RAI KTQVYGR TCF SDGATSNVNC
 330     340     350     360     367
KASNALNVVR DLCRGKSSCT VEASNDVFGD PCMHYKYLE LSYDCSK

```

DsbA peptideo sinal

DsbA madura

Cauda de histidina

Sítio de trombina

S-tag

Sítio de clivagem TEV

Lectina

Número de aminoácidos: 367

Peso molecular: 40084,6351

pI: 7,24

Coefficiente de extinção:

O coeficiente de extinção é: 35870

Em 1mg/ml (24,947 uM) a solução da proteína tem

Abs280nm de: 0,89

Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 36495

Em 1mg/ml (24,947 uM) a solução de proteína tem

Abs280nm de: 0,91

Fonte: Autora. pI: Ponto isoelétrico. DsbA: Proteína A formadora de pontes dissulfeto. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas a cauda fusionada à rEEL, sem a proteína de interesse, possui 263 aminoácidos com um peso molecular de 28.896,83 Da e com pI de 6,38. A proteína inteira expressa, com a lectina de interesse, possui 367 aminoácidos com um peso molecular de 40.084,63 Da e com pI de 7,24.

A proteína expressa foi pensada para ser clivada com a protease TEV. Após clivagem, rEEL possui 106 aminoácidos com peso molecular de 11.333,95 Da e com pI de 8,4. A Figura 19 mostra a representação da proteína de interesse rEEL após clivagem com a protease TEV.

Figura 19 – Representação da lectina rELEL obtida após clivagem com a protease TEV.

>rELEL clivada

```

      10      20      30      40      50      60      70      80
GAM90ELVSQ100L106 LKKERVCEGS SLTISCPQK106G AGISIA106RAI106Y GRTKTQVCPS DGATSNV106NCK ASNALNV106VRD LCRGKSSCTV
EASNDVFGDP CMHTYKYLE106L SYDCSK

```

Sítio TEV residual

Linker residual

Lectina

Número de aminoácidos: 106

Peso molecular: 11333,9495

pI: 8,40

Coefficiente de extinção:

O coeficiente de extinção é: 5960

Em 1mg/ml (88,230 uM) a solução da proteína tem Abs280nm de: 0,53

Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 6460

Em 1mg/ml (88,230 uM) a solução de proteína tem Abs280nm de: 0,57

Fonte: Autora. pI: Ponto isoelétrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Em *E. coli*, a proteína recombinante é expressa no citoplasma, como esse ambiente é altamente redutor devido a presença de enzimas que reduzem as pontes dissulfeto, a formação dessas ligações ocorre no periplasma bacteriano, localizado entre a membrana citoplasmática e a membrana externa. Assim, proteínas que requerem ligações dissulfeto precisam ser direcionados a esse compartimento para completar seu dobramento (Marco, 2009).

O periplasma é um ambiente capaz de realizar a formação de pontes dissulfeto devido a suas condições oxidantes. Em *E. coli*, a formação dessas ligações ocorre no periplasma pelo grupo de enzimas oxirredutases tiol-dissulfeto conhecidas como proteínas Dsb (*Disulfide Bond proteins*) que compreende DsbABCD e G. A DsbA é altamente potente e oxida resíduos de cisteína à medida que entram no periplasma, por meio de reações de troca tiol-dissulfeto, onde o dissulfeto é transferido para as cisteínas das proteínas. Quando essa troca ocorre, a DsbA fica reduzida e a proteína DsbB é a responsável por regenerar a DsbA e manter em estado oxidado. Essa manutenção ocorre por meio de transferências de elétrons do sítio reduzido de DsbA para as cisteínas de DsbB, que consequentemente passam esses elétrons para quinonas da membrana interna e em seguida para a cadeia transportadora de elétrons da cadeia respiratória bacteriana. Além da via de oxidação, existe uma via de isomerização que revisa e corrige as cisteínas pareadas incorretamente na proteína. O principal componente dessa via é a DsbC (Kaur; Kumar; Kaur, 2018; Stewart; Åslund; Beckwith, 1998; Marco, 2009).

Portanto, como o citoplasma redutor tem um grande impacto no enovelamento correto de proteínas com pontes dissulfeto, uma opção é direcionar a proteína recombinante

expressa para o periplasma. O direcionamento para esse compartimento subcelular ocorre através de peptídeos sinal fundidos à extremidade N-terminal da proteína e a fusão a peptídeos sinal adequados permite o redirecionamento com sucesso para o periplasma. A coexpressão das proteínas DsbA e DsbC ou combinações das proteínas Dsb tem sido empregada com sucesso para o direcionamento e expressão de inúmeras proteínas heterólogas no periplasma bacteriano (Marco, 2009).

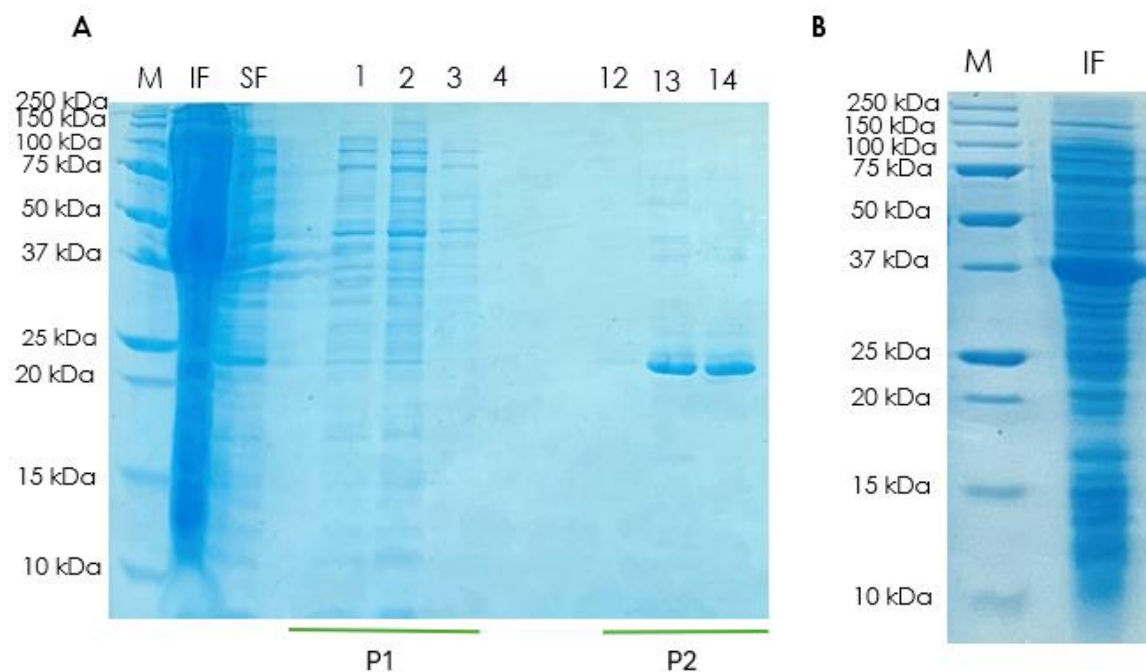
Como o vetor pET39-TEV expressa a proteína DsbA com um peptídeo sinal que direciona a proteína de interesse para o periplasma, a expressão periplasmática de rELELpET39-TEV foi então avaliada experimentalmente. Após o tempo de expressão, as células foram submetidas a uma extração periplasmática com tampão contendo polimixina B sulfato. A fração extraída e a fração insolúvel (fração citoplasmática) foram analisadas por SDS-PAGE (Figura 20).

A fração periplasmática foi diluída 1:1 (v/v) com tampão Tris 50 mM pH 8, contendo NaCl 300 mM para reduzir o EDTA presente no tampão de extração. Como o vetor expressa a cauda de histidina, a fração extraída foi submetida a uma cromatografia de afinidade em coluna cOmpleteTM His-tag, pois são colunas capazes de tolerar pequenas quantidades de EDTA.

A coluna foi equilibrada e as proteínas não ligadas a coluna foram lavadas com o mesmo tampão de diluição. As proteínas retidas a coluna foram eluídas com um gradiente linear de imidazol em tampão Tris 50 mM pH 8, NaCl 300 mM contendo imidazol 500 mM. As frações foram coletadas e analisadas em SDS-PAGE (Figura 20).

A análise em SDS-PAGE mostrou a ligação e a recuperação de proteínas na coluna, observada pela banda evidente entre 20 kDa e 25 kDa do marcador, porém esse tamanho da banda não condiz com o esperado para a proteína expressa rELELpET39-TEV que deveria ter aproximadamente 41 kDa (Figura 20-A). Apesar da saturação dos primeiros poços, a fração solúvel periplasmática (SF) não parece possuir nenhuma banda com o tamanho esperado para a proteína de interesse.

Figura 20 - Análise da purificação periplasmática de rEELpET39-TEV em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. A: Análise das frações de extração e da cromatografia. IF: Fração insolúvel (fração citoplasmática). SF: Fração solúvel extraída do periplasma. Números: tubos coletados. Tubos 1 a 4: referentes ao pico 1 (fração não ligada a coluna). Tubos 12 a 14: referentes ao pico 2 (fração eluída com o gradiente linear). B: Análise da fração citoplasmática.

Por outro lado, analisando a fração insolúvel (Figura 20-B) que corresponde a fração citoplasmática que não teve as células lisadas, é possível observar uma banda próxima a 37 kDa que corrobora com o tamanho esperado de 41 kDa para a proteína de interesse expressa no vetor pET39-TEV.

Com esses resultados, podemos inferir que não houve endereçamento da proteína expressa para o periplasma, pois análises de SDS-PAGE indicam a presença da proteína de interesse apenas na fração citoplasmática. E as bandas visíveis na fração eluída da coluna, com tamanho diferente do esperado para rEELpET39-TEV, são resultantes de interações inespecíficas com a coluna de níquel imobilizado, conforme discutido anteriormente.

A recuperação de proteínas periplasmáticas geralmente é obtida pela ruptura da membrana externa por choque osmótico, com a vantagem de diminuir a contaminação com o material citoplasmático, extraindo frações bem mais puras pois poucas proteínas ficam no periplasma. Essa estratégia já é bastante documentada com resultados satisfatórios e a polimixina B alinhada ao tratamento térmico é uma prática consistente de eficiência comprovada (Martin *et al.*, 2006; Dixon; Chopra, 1986; Wurm *et al.*, 2017).

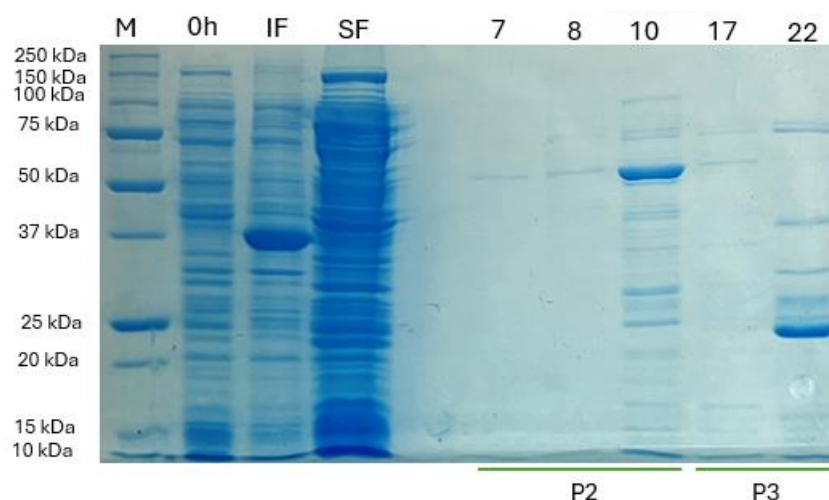
O problema com rELEL não foi a extração periplasmática mas o transporte da proteína expressa do citoplasma para o periplasma. A necessidade de translocar proteínas através da membrana interna é uma etapa delicada, pois apenas duas vias podem direcionar proteínas para alcançar o periplasma. Por isso, proteínas expressas podem se acumular no citoplasma (Martin *et al.*, 2006; Marco, 2009).

Para alcançar o periplasma, as proteínas heterólogas possuem uma sequência sinal N-terminal que as direciona até o *translocon* Sec, um canal condutor de proteínas na membrana citoplasmática. Duas vias podem direcionar proteínas ao *translocon* Sec: a via de direcionamento pós-traducional via SecB (lenta) e a via de direcionamento cotraducional via partícula de reconhecimento de sinal (SRP) (rápida). A sequência do peptídeo sinal está relacionada com a via de direcionamento. A sequência sinal de DsbA guia proteínas para o *translocon* Sec de maneira dependente de SRP. No entanto, estudos anteriores já demonstraram que a saturação da capacidade do translocon Sec é uma das principais desvantagens na produção heteróloga periplasmática, resultando no acúmulo de proteínas recombinantes no citoplasma (Schlegel *et al.*, 2013; Karyolaimos; Gier, 2021).

Portanto, o problema para rELELpET39-TEV pode ter sido essa saturação da capacidade de transporte da proteína expressa no citoplasma para o periplasma.

Como o SDS-PAGE da fração insolúvel indicou a presença da proteína expressa na fração citoplasmática, o *pellet* resultante da extração com polimixina B foi submetido a uma extração citoplasmática usual, onde a fração solúvel foi aplicada em uma HisTrapTM FF. A coluna foi lavada com tampão contendo imidazol 10 mM e as proteínas retidas a coluna foram eluídas com 5% do tampão B, e com um gradiente linear de 5% a 100% do mesmo tampão. As frações foram coletadas e analisadas em SDS-PAGE (Figura 21).

Figura 21 - Análise da purificação citoplasmática de rELELpET39-TEV em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. 0h: alíquota antes da indução. IF: Fração insolúvel. SF: Fração solúvel. Números: tubos coletados. Tubos 7, 8 e 10: referentes ao pico 2 (fração eluída da coluna com 5% de tampão B contendo 500 mM de imidazol). Tubos 17 e 22: referentes ao pico 3 (fração eluída da coluna com gradiente de tampão B contendo 500 mM de imidazol).

A análise em SDS-PAGE mostrou que a proteína de interesse foi expressa na fração insolúvel, detectada pela banda visível no gel que não está presente na amostra coletada antes da indução, próximo a banda de 37 kDa do marcador, corroborando com o tamanho esperado da proteína de aproximadamente 41 kDa. Porém, a mesma banda não foi visualizada na fração solúvel, antes da aplicação da amostra na cromatografia, demonstrando que novamente rELEL foi expressa apenas na fração insolúvel.

A ligação e a recuperação de proteínas na coluna foram observadas pelas bandas evidentes entre 20 kDa e 25 kDa e próxima a 50 kDa do marcador, porém esses tamanhos não condizem com o esperado para a proteína expressa rELELpET39-TEV.

Com esses resultados, podemos inferir que a proteína foi expressa apenas na fração insolúvel. E as bandas visíveis na fração eluída da coluna, com tamanho diferente do esperado, possuindo inclusive, o tamanho observado na extração periplasmática, são novamente resultantes de interações inespecíficas com a coluna de níquel imobilizado.

A relevância da proteína DsbA como proteína fusionada para expressão periplasmática e para o correto enovelamento de proteínas com pontes dissulfeto já foi bastante discutido. Como observado em trabalhos anteriores, a expressão de proteínas com essa

característica foi facilitada e até influenciou positivamente no rendimento da proteína (Marco, 2009; Schlegel *et al.*, 2013; Karyolaimos; Gier, 2021).

O vetor pET39-TEV também expressa uma *tag* fusionada a proteína de interesse chamada, *S-tag*. Esse peptídeo também pode ajudar na solubilização da proteína por ser pequeno, solúvel e interferir minimamente na estrutura da proteína expressa, além de permitir a detecção e purificação da proteína de interesse pela sua sequência (Eraines *et al.*, 2000).

Apesar do uso de todas essas estratégias para rEEL no vetor pET39-TEV, não foi possível expressar a lectina na fração solúvel com o dobramento correto.

4.3.2.2.3 Expressão da lectina rEEL no vetor pET15b(+)

A construção da proteína recombinante expressa no vetor pET15b(+) origina uma proteína sem cauda de fusão, composta apenas pela lectina de interesse (Figura 22).

Figura 22 – Representação esquemática da proteína recombinante no vetor pET15b(+).

>rEEL

```

      10      20      30      40      50      60      70      80
MELVSQLCLK KERVCEGSSL TISCPQKGAG ISIARAIYGR TKTQVCPDGD ATSNVNCKAS NALNVVRDLG RGKSSCTVEA
      90     100     104
SNDVFGDPCM HTYKYLELSY DCSK

```

Lectina

Número de aminoácidos: 104
 Peso molecular: 11205,8188
 pI: 8,40

Coeficiente de extinção:

O coeficiente de extinção é: 5960
 Em mg/ml (89,239 uM) a solução da proteína tem Abs280nm de: 0,53

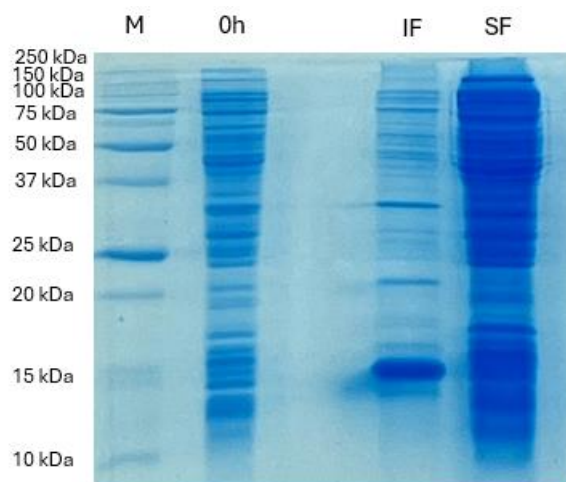
Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 6460
 Em mg/ml (89,239 uM) a solução de proteína tem Abs280nm de: 0,58

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas rEEL possui 104 aminoácidos com um peso molecular de 11.205,82 Da e com pI de 8,4.

A expressão de rEELpET15b(+) foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE (Figura 23).

Figura 23 - Análise da expressão de rELELpET15b(+) em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. 0h: Aliquota retirada antes da indução. IF: Fração insolúvel. SF: Fração solúvel.

A proteína de interesse rELEL foi expressa apenas na fração insolúvel, detectada pela banda visível no gel que não está presente na amostra coletada antes da indução. A banda abaixo de 15 kDa do marcador, corrobora com o tamanho esperado da proteína de aproximadamente 11 kDa. A mesma banda não foi detectada na fração solúvel.

Apesar do indicativa de não expressão na fração solúvel, a amostra seguiu para cromatografia de afinidade, conforme o método usado para purificação da lectina nativa (Carneiro *et al.*, 2015).

A fração solúvel foi aplicada em matriz de ramnose e matriz de α -Lactose-Sepharose, separadamente. As proteínas possivelmente retidas a coluna foram eluídas com 0,3 M de ramnose, mas nenhuma interação com a coluna foi observada, mostrando que de fato não houve expressão na fração solúvel.

Na cromatografia de afinidade, a interação ocorre entre o domínio de reconhecimento a carboidratos (DRC) da lectina e os açúcares imobilizados na matriz cromatográfica. Essa interação depende da integridade funcional do DRC, que envolve o dobramento correto da proteína. Portanto, a ausência de interação nesse tipo de cromatografia indica que o sítio de reconhecimento a carboidratos da proteína recombinante não está funcionalmente ativo, provavelmente devido ao dobramento incorreto no momento da expressão (Li; Su; Janson, 2004; Young; Britton; Robinson, 2012).

A expressão de rEEL na fração insolúvel quando expressa em pET15b(+) era um resultado esperado, visto que a construção não possui proteínas fusionadas à lectina que poderiam ajudar com a solubilidade no ambiente redutor do citoplasma bacteriano. Proteínas ricas em pontes dissulfeto, como é o caso de rEEL, necessitam da aplicação de alguma das estratégias específicas que ajude no dobramento correto da proteína, como já discutido em vários momentos ao longo desse capítulo.

Outra estratégia usada na tentativa de expressão de rEEL na forma solúvel, foi a coexpressão com o sistema CyDisCo (Formação de ligações dissulfeto no citoplasma de *E. coli*). Essa tecnologia consiste em transformar a cepa de expressão com o plasmídeo de interesse e o plasmídeo CyDisCo, facilitando a formação de ligações dissulfeto da proteína recombinante no citoplasma da *E. coli* (Gaciarz *et al.*, 2017).

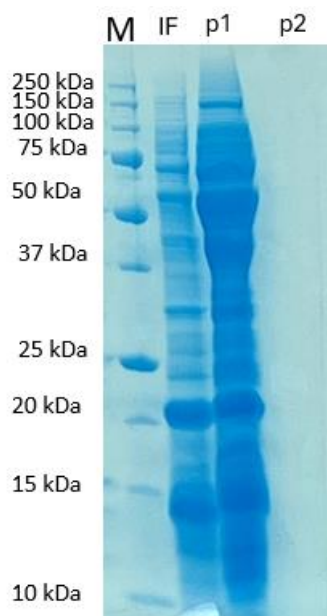
Esse sistema funciona deixando o citoplasma menos redutor, ao coexpressar duas enzimas que catalisam a formação das pontes dissulfeto. A sulfidril oxidase (Erv1p da levedura *Saccharomyces cerevisiae*) que ajuda no dobramento das ligações, e a isomerase de ligação de sulfeto (PDI humana) que ajuda no correto dobramento da proteína e redobrimento em caso de enovelamento incorreto. Essa abordagem permite que proteínas com pontes dissulfeto que expressam na forma de corpos de inclusão, possam se dobrar corretamente sem alterar o ambiente redutor nativo da *E. coli* (Gaciarz *et al.*, 2016, 2017).

A construção rEELpET15b foi cotransformada, na cepa de expressão *E. coli* Shuffle B, com o plasmídeo CyDisCo pMJS205 que contém os genes para Erv1p e PDI sob um promotor induzível Ptac e com um marcador de resistência a cloranfenicol (Gaciarz *et al.*, 2016).

A expressão e a estratégia de purificação da proteína seguiram o mesmo protocolo do vetor pET15b(+). A fração solúvel foi aplicada em matriz de α -Lactose-Sepharose e as proteínas possivelmente retidas a coluna foram eluídas com 0,3 M de ramnose. Novamente, nenhuma interação com a coluna foi observada.

As frações do lisado celular e da cromatografia foram analisadas em SDS-PAGE. Uma banda de tamanho esperado para a rEEL de 11 kDa foi observada na fração insolúvel e nenhuma proteína foi detectada na fração eluída da coluna, mostrando que o sistema CyDisCo não foi capaz de expressar a proteína de forma solúvel (Figura 24).

Figura 24 - Análise da expressão de rEELpET15b(+) com sistema CyDisCo em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. IF: Fração insolúvel. P1: Fração não retida a coluna. P2: Fração retida a coluna, eluída com ramnose 0,3 M.

Vários estudos já comprovaram a eficiência do sistema CyDisCo em expressar proteínas com pontes dissulfeto solúveis e ativas (Gaciarz *et al.*, 2016, 2017). Porém, essa estratégia também não foi eficiente para rEEL.

4.3.2.2.4 Expressão da lectina rEEL no vetor pCOLD-HAT-TEV-3G-Nco

A construção da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco origina uma proteína de fusão composta pelo TEE, a cauda HAT rica em histidina, o sítio de clivagem para a protease TEV e a lectina de interesse (Figura 25).

Figura 25 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco

>rELEL expressa em pCOLD-HAT-TEV-3G-Nco

10 20 30 40 50 60 70 80
 MNHKVKDHLI HNVHKEEHAH AHNKGENLYF QGGGAMELVS QLCLKKERV C EGSSLTISC P QKGAGISIA R AIYGRKTQ V
 90 100 110 120 130 139
 CPDGDGATSNV NCKASNALNV VRDLCRGKSS CTVEASNDVF GDPCMHTYKY LELSYDCSK

TEE	Número de aminoácidos: 139
HAT-Tag	Peso molecular: 15186,2165
	pI: 8,38
Sítio de clivagem TEV	Coefficiente de extinção:
Lectina	O coeficiente de extinção é: 7450
	Em 1mg/ml (65,849 uM) a solução da proteína tem
	Abs280nm de: 0,49
	Se todas as cisteínas formarem pontes dissulfeto,
	o coeficiente de extinção é: 7950
	Em 1mg/ml (65,849 uM) a solução de proteína tem
	Abs280nm de: 0,52

Fonte: Autora. pI: Ponto isoelétrico. TEE: *Translation Enhancing Element*. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas a cauda fusionada à rELEL, sem a proteína de interesse, possui 35 aminoácidos com um peso molecular de 3.998,41 Da e com pI de 7,58. A proteína inteira expressa, com a lectina de interesse, possui 139 aminoácidos com um peso molecular de 15.186,22 Da e com pI de 8,38.

Nesse vetor em estudo, a cauda de histidina convencional (poli-His), geralmente composta por seis resíduos de histidinas consecutivos foi substituída pela cauda HAT-tag (*Histidine Affinity Tag* – Cauda de afinidade a histidina). Ela é um peptídeo curto sintético composto por histidinas e outros aminoácidos espaçadores e neutros que são posicionados para deixar a histidina em um arranjo mais favorável à coordenação com os íons metálicos da cromatografia IMAC. Por isso, essa *tag* apresenta uma afinidade mais específica e uniforme pela matriz metálica, resultando em uma menor ligação inespecífica de proteínas endógenas (Takara, 2013).

A proteína expressa foi pensada para ser clivada com a protease TEV. Após clivagem, rELEL possui 108 aminoácidos com peso molecular de 11.448,05 Da e com pI de 8,4. A Figura 26 mostra a representação da proteína de interesse rELEL após clivagem com a protease TEV.

Figura 26 – Representação da proteína rELEL obtida após clivagem com a protease TEV.

>rELEL clivada

```

      10      20      30      40      50      60      70      80
GGGAMELVSQ LCLKKERVCE GSSLTISCPQ KGAGISIAAR IYGRTKTQVC PSDGATSNVN CKASNALNVV RDLCRGKSSC
      90     100     108
TVEASNDVFG DPCMHTYKYL ELSYDCSK

```

Sítio TEV residual

Linker residual

Lectina

Número de aminoácidos: 108

Peso molecular: 11448,0533

pI: 8,40

Coefficiente de extinção:

O coeficiente de extinção é: 5960

Em 1mg/ml (87,351 uM) a solução da proteína tem Abs280nm de: 0,52

Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 6460

Em 1mg/ml (87,351 uM) a solução de proteína tem Abs280nm de: 0,56

Fonte: Autora. pI: Ponto isoelétrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

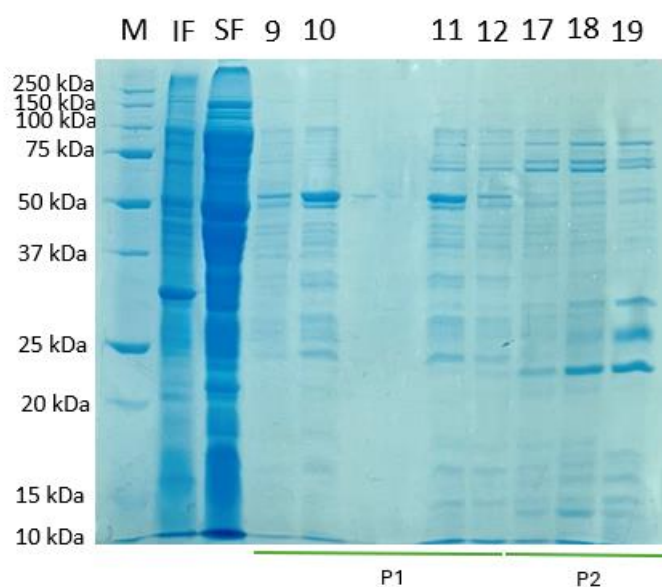
A expressão da lectina de interesse nesse vetor foi conduzida em um meio de autoindução. Esse sistema se baseia na presença simultânea de diferentes fontes de carbonos, geralmente é composto por glicose, glicerol e lactose. Inicialmente, a bactéria consome preferencialmente a glicose, como fonte de carbono e como agente de repressão catabólica do *operon lac*, impedindo a expressão do gene recombinante nas fases iniciais do cultivo. Quando a glicose se esgota, acontece o consumo do glicerol para o crescimento celular e que não exerce repressão no sistema lac. Nesse momento, a lactose é simultaneamente metabolizada e convertida em alolactose, que se liga ao repressor do *operon lac* e ativa a transcrição. Dessa forma, o glicerol sustenta o crescimento celular enquanto a lactose atua como indutor (Studier, 2005; Lobstein *et al.*, 2012).

Portanto, a indução ocorre automaticamente quando a cultura atinge alta densidade celular, resultando em uma expressão controlada, sincronizada e sem a necessidade de adição manual de indutores. Os meios de autoindução geralmente resultam em maiores rendimentos da proteína recombinante porque promove uma separação eficiente entre as fases de crescimento e expressão, permitindo que a expressão só se inicie em altas densidades, o que eleva a produção da proteína exógena. Além disso, a indução não ocorre de forma abrupta como com o IPTG, nesse caso, ela acontece de forma gradual causando um menor estresse metabólico (Studier, 2005).

A expressão de rELELpCOLD-HAT-TEV-3G-Nco foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE (Figura 27).

Como o vetor pCOLD-HAT-TEV-3G-Nco possui cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados. As proteínas retidas a coluna foram eluídas com 5% do tampão B e com um gradiente linear de 5% a 100% do mesmo tampão. As frações foram coletadas e analisadas em SDS-PAGE (Figura 27).

Figura 27 - Análise da purificação de rEELpCOLD-HAT-TEV-3G-Nco em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 9 a 12: eluídos com 5% de B (P1). Tubos 17, 18 e 19: eluídos com gradiente de tampão B (P2). IF: Fração insolúvel. SF: Fração solúvel.

A proteína de interesse expressa possui aproximadamente 15 kDa, no gel de SDS-PAGE não foi detectada nenhuma banda nas frações analisadas com o tamanho esperado para rEELpCOLD-HAT-TEV-3G-Nco, indicando que a proteína não foi expressa em nenhuma condição.

Uma banda na fração insolúvel apareceu entre 25 kDa e 37 kDa, mas não é o tamanho esperado da proteína expressa. Muitas bandas foram visíveis nas frações eluídas da coluna, indicando novamente interações inespecíficas com a coluna de níquel imobilizado.

Estudos anteriores relataram que uma proteína contendo 17 pontes dissulfeto conseguiu ser expressa na fração solúvel por autoindução utilizando a cepa SHuffle B (Lobstein *et al.*, 2012). Para rEELpCOLD-HAT-TEV-3G-Nco, também foi utilizado essa cepa de expressão no meio de autoindução, mas não foi observada a expressão da proteína nem na fração insolúvel.

A produção de proteínas expressas em sistema heterólogo, principalmente aquelas ricas em ligações dissulfeto, tendem a sofrer dobramento incorreto em *E. coli*. Apesar dos inúmeros avanços no campo da expressão de proteínas recombinantes, a produção de proteínas com pontes dissulfeto na forma solúvel e ativa continua sendo um desafio (Nguyen *et al.*, 2021; Ganjave *et al.*, 2025).

Diversas estratégias consideradas eficientes para facilitar a expressão na fração solúvel de proteínas ricas em pontes dissulfeto (Kaur; Kumar; Kaur, 2018; Ganjave *et al.*, 2025) foram testadas com rEEL, incluindo modificação de condições de cultivo, diferentes cepas modificadas geneticamente, diminuição da temperatura de expressão, diminuição da concentração do indutor, diferentes vetores com diferentes promotores, diferentes caudas de fusão, diferentes proteínas fusionadas à lectina, expressão periplasmática, expressão com sistema CyDisCo e nenhuma estratégia foi capaz de expressar a rEEL com seu dobramento correto e funcional.

Em estudos anteriores, os autores relataram que a coexpressão de chaperonas moleculares como o *trigger factor* nas células de SHuffle foi capaz de melhorar a solubilidade de proteínas com pontes dissulfeto (Lobstein *et al.*, 2012). Outros trabalhos relataram que diversas proteínas heterólogas com muitas ligações dissulfeto foram produzidas no citoplasma de células de *E. coli* Origami e SHuffle com altos rendimentos (Makino; Skretas; Georgiou, 2011; Rosano; Ceccarelli, 2014).

No entanto, outros trabalhos mostram que mesmo testando as diferentes cepas geneticamente modificada para expressar com eficiência proteínas contendo pontes dissulfeto, Rosetta-Gami (DE3), Origami e Shuffle T7, não houve expressão das proteínas na fração solúvel. Além disso, no mesmo trabalho, a diminuição da temperatura de expressão também não foi suficiente para aumentar a solubilidade da proteína (Fathi-Roudsari; Akhavian-Tehrani; Maghsoud, 2016; Nguyen *et al.*, 2021).

Portanto, fica claro que alcançar a expressão de proteínas tão desafiadoras como essas, na sua forma solúvel e ativa, envolve muitas tentativas e o sucesso de cada estratégia depende das características individuais de cada proteína.

Para a proteína recombinante rEEL, observou-se que as únicas condições nas quais foi possível obter expressão na fração solúvel, apesar de inativa, envolveram o uso de proteínas fusionadas a lectina, como a tiorredoxina no vetor pET32a(+). Essa solubilidade foi alcançada apenas em combinação com baixas concentrações do indutor e em cepa geneticamente modificada para ter um citoplasma mais oxidante. A coexpressão com a

chaperona *trigger factor* no vetor pCOLD-TF também demonstrou ser crucial para a solubilidade, somada com a diminuição da temperatura de expressão, a cepa geneticamente modificada para expressão de proteínas com ligações dissulfeto, e a otimização de códons. Esses resultados mostram que a presença de proteínas de fusão contribui significativamente para a solubilidade de rEEL, e que fatores como a cepa hospedeira modificada, otimização de códons, a baixa temperatura de indução e a baixa concentração do indutor atuam de forma combinada para favorecer o dobramento de proteínas ricas em ligações dissulfeto.

A grande maioria das lectinas recombinantes de equinodermos que possuem pontes dissulfeto, até em quantidades menores que rEEL, foram expressas apenas na fração insolúvel e foram recuperadas ativas por *refolding* (Wei *et al.*, 2015; Wang *et al.*, 2018; Zhao *et al.*, 2021; Vasilenko *et al.*, 2012; Shimizu *et al.*, 2012; Hatakeyama *et al.*, 2006). E assim como rEEL, os autores relataram que com a lectina MBL-AJ de *Apostichopus japonicus* várias estratégias foram testadas, mas nenhuma foi capaz de alcançar a expressão na fração solúvel (Vasilenko *et al.*, 2012). Para CEL-III de *Cucumaria echinata*, os autores também expressaram a lectina com cauda fusionada de tiorredoxina, e a expressão foi observada apenas na fração insolúvel, sendo necessário a recuperação da proteína funcional por *refolding* (Shimizu; Yonekura; Kouzuma, 2022).

A presença das várias ligações dissulfeto é a principal causa para a dificuldade de dobramento correto e funcional observada para a proteína recombinante rEEL. Provavelmente, assim como observado para as outras lectinas recombinantes de equinodermos, é possível que rEEL consiga ser recuperada de forma solúvel e ativa por meio de estratégias de redobramento a partir dos corpos de inclusão.

Como os procedimentos de solubilização e redobramento são as etapas cruciais, eles exigem otimização específica para cada proteína, e embora tenham muitas desvantagens, como o baixo rendimento, muitas vezes é a única forma de recuperação da proteína ativa e com atividade biológica (Singh *et al.*, 2015).

Apesar de um teste preliminar de *refolding* tenha sido realizado com rEEL, a literatura relata diferentes abordagens alternativas para a solubilização com agentes desnaturantes e para o redobramento, que ainda podem ser avaliadas para a recuperação funcional da proteína. Por exemplo, a solubilização suave apresenta taxas de sucesso superiores às obtidas com a solubilização em altas concentrações de agentes caotrópicos, como a ureia. Esses agentes de solubilização suaves preservam as estruturas secundárias das proteínas,

diminuindo a agregação proteica durante o redobramento e, conseqüentemente, melhorando a recuperação das proteínas (Singh *et al.*, 2015).

Além disso, as proteínas solubilizadas são redobradas pela remoção do agente desnaturante. Essa remoção pode ser feita de diferentes formas, por diluição lenta ou rápida da proteína solubilizada em tampão de renaturação e por diálise da proteína solubilizada na presença de tampão de renaturação com concentrações decrescentes do agente desnaturante. Esses métodos são os mais comuns, principalmente a diálise, que é a estratégia que mais apresenta desvantagens como a agregação proteica. Porém, a baixa concentração da proteína no momento da renaturação e o uso de diferentes métodos de redobramento como a diluição pulsátil e o renovelamento em coluna podem melhorar a recuperação da proteína (Singh *et al.*, 2015).

Inclusive, a lectina do pepino-do-mar *Cucumaria echinata*, CEL-III, com 12 pontes dissulfeto intracadeias, foi recuperada ativa a partir de redobramento pelo método de diluição lenta no tampão de renovelamento (Shimizu *et al.*, 2012). Portanto, estudos futuros ainda podem possibilitar a recuperação da proteína recombinante rEEL com sua atividade biológica preservada.

4.4 Conclusão

A produção heteróloga da lectina rELEL representou um desafio significativo, devido à baixa solubilidade e a ausência de funcionalidade da proteína recombinante mesmo após a utilização de diferentes vetores de expressão, condições experimentais e estratégias de redobramento e cotransformação. Os resultados apresentados mostram que fatores estruturais intrínsecos da proteína podem estar associados à sua complexidade de dobramento e estabilidade, dificultando a obtenção da proteína recombinante de forma solúvel e ativa.

Dessa forma, diante dos desafios observados com as abordagens metodológicas testadas, novas estratégias alternativas podem ser adotadas para expressar essa lectina com sua função biológica preservada, possivelmente com o foco em diferentes metodologias de *refolding*, testando variadas abordagens e soluções de solubilização e redobramento.

5. CAPÍTULO II: EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA LECTINA rADEL PRESENTE NOS OVOS DA LESMA-DO-MAR *Aplysia dactylomela*

5.1 Objetivos

Objetivo geral

Expressar em sistema heterólogo, a lectina presente nos ovos da lesma-do-mar *Aplysia dactylomela*, rADEL (recombinante *Aplysia dactylomela* eggs lectin) e caracterizá-la bioquimicamente e estruturalmente.

Objetivos específicos

- Produzir em sistema heterólogo de *E. coli* a lectina recombinante rADEL na forma solúvel e ativa;
- Produzir em sistema heterólogo de *K. pastoris* a lectina recombinante rADEL na forma solúvel e ativa;
- Confirmar a expressão da lectina em *K. pastoris* por *Immunoblotting*;
- Testar diferentes condições de expressão para obter o maior rendimento da proteína;
- Purificar a lectina recombinante e avaliar a funcionabilidade por atividade hemaglutinante da proteína recombinante;
- Determinar o perfil de inibição por carboidratos da lectina recombinante;
- Obter a estrutura secundária da lectina recombinante por dicroísmo circular;
- Avaliar o comportamento hidrodinâmico da proteína recombinante por meio de Espalhamento dinâmico de luz (*Dynamic Light Scattering* - DLS);
- Obter cristais difratáveis da lectina rADEL para determinar sua estrutura tridimensional por cristalografia de raios X;

5.2 Metodologia

5.2.1 Coleta dos ovos e purificação da lectina nativa

A lectina nativa ADEL, oriunda dos ovos da lesma-do-mar *Aplysia dactylomela*, foi obtida através do protocolo previamente estabelecido por Carneiro e colaboradores (2017b). A coleta e o acesso ao patrimônio genético da espécie foram autorizados e regulados pelos órgãos competentes (SISBIO ID: 33913-12 e SISGEN ID:A1792FE).

Os ovos do invertebrado foram coletados durante a maré de sizígia na zona entre marés da praia do Pacheco, no município de Caucaia, Ceará. O material biológico foi retirado das pedras, lavado e transportado até o laboratório, onde foi imediatamente triturado para extração das proteínas com tampão Tris 20 mM pH 7,6; NaCl 150 mM (TBS), contendo fluoreto de fenilmetilsulfonil (PMSF) 0,1 mM.

O extrato bruto foi submetido a uma cromatografia de afinidade em uma matriz de SepharoseTM 4B previamente equilibrada com o tampão TBS e a lectina foi eluída com 200 mM de α -lactose em TBS. A pureza das frações obtidas foi confirmada por eletroforese (SDS-PAGE) (Laemmli, 1970), e a funcionabilidade da proteína foi avaliada por atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina conforme preferência da lectina. As frações puras e funcionalmente ativas foram reunidas, dialisadas e liofilizadas para uso posterior.

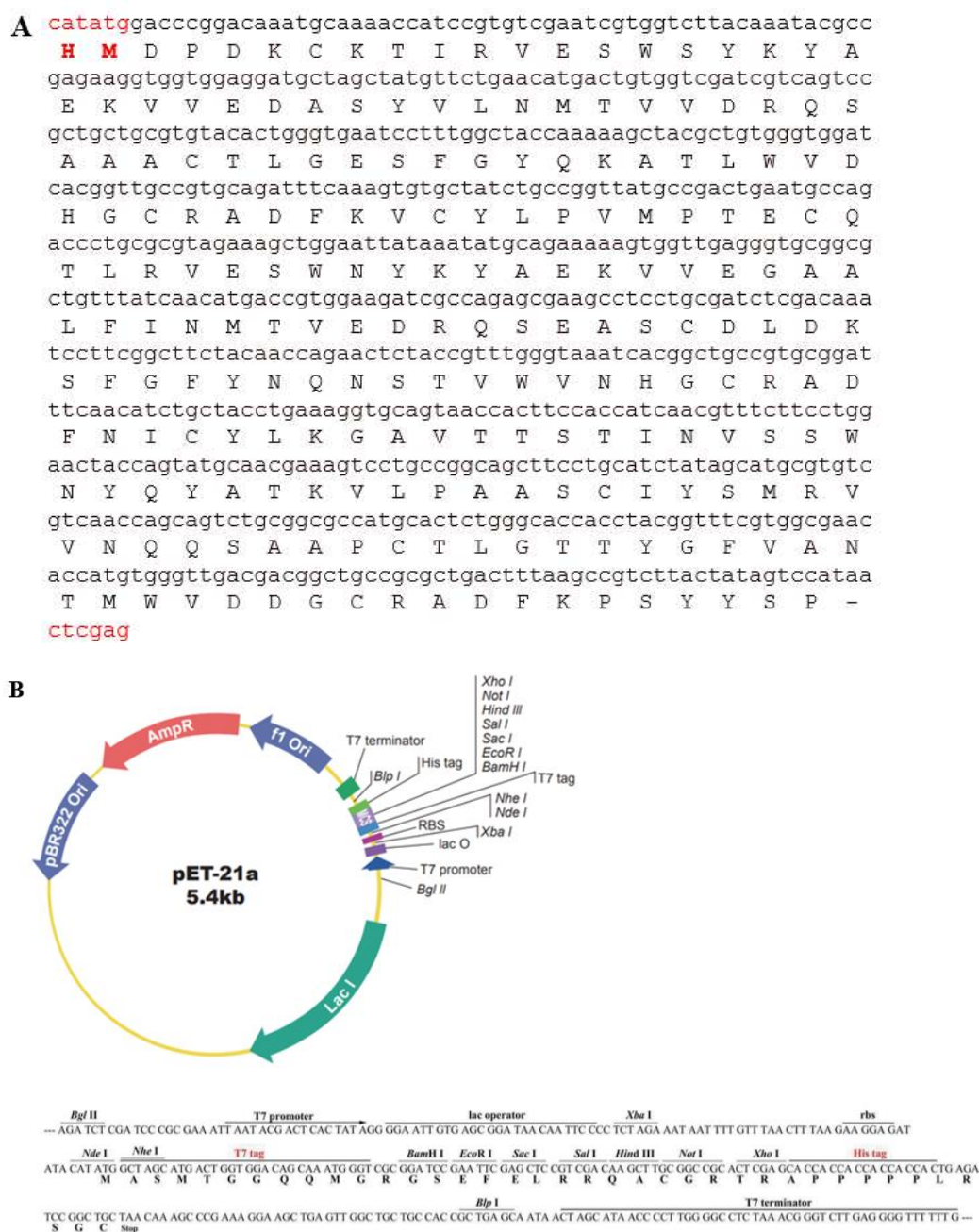
5.2.2 Produção da proteína recombinante *rADEL* em sistema heterólogo de *Escherichia coli* (Brasil)

5.2.2.1 Desenho do gene sintético adquirido no Brasil

A sequência de aminoácidos da lectina nativa (C0HK25.1- UniProt: <https://www.uniprot.org/>) foi utilizada para desenhar o gene sintético (Carneiro *et al.*, 2017b).

O gene sintético foi adquirido pela empresa FastBio (<http://www.fastbio.com.br/>, SP, Brasil) no vetor de expressão pET21a com as regiões de clonagem delimitadas pelas enzimas de restrição *NdeI* e *XhoI*, sem a otimização de códons raros para *Escherichia coli* (Figura 28).

Figura 28 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pET21a.



Fonte: Autora. A – Representação da sequência de nucleotídeos e a respectiva tradução da lectina. A sequência em letras minúsculas representa os nucleotídeos e a sequência em letras maiúsculas representa os aminoácidos traduzidos. As letras em vermelho representam os sítios de clivagem *NdeI* (início) e *XhoI* (final). O símbolo – representa o códon de parada da tradução. B – Representação do vetor de expressão pET21a e seu respectivo mapa. Disponível em: <https://fastbio.com.br/wp-content/uploads/2023/09/GenScript-Lista-e-mapa-de-vetores-bacteria.pdf>.

5.2.2.2 Expressão da lectina rADEL em sistema heterólogo de *E. coli*

O gene sintético foi diluído para uma concentração de $50 \text{ ng} \cdot \mu\text{L}^{-1}$ e inserido, por choque térmico, na cepa de clonagem *Escherichia coli* DH5 α , e nas cepas de expressão: *E. coli* BL21(DE3), *E. coli* Rosetta-gami, *E. coli* Origami B (DE3) plysS e *E. coli* artic Expres, seguindo o protocolo de transformação descrito por Sambrook (2002). A manipulação e uso de organismos geneticamente modificados tem autorização pelo certificado CQB-R007-2016 CTNBio.

As colônias crescidas em LB ágar e selecionadas pelo antibiótico de seleção (ampicilina a $100 \mu\text{g/mL}$) foram inoculadas em meio LB *Broth* contendo o mesmo antibiótico e incubadas a 37°C , *overnight*, sob agitação constante. As células crescidas foram diluídas na proporção de 1:100 em meio LB contendo ampicilina na mesma proporção e a cultura foi crescida a 37°C e induzida de acordo com cada condição testada.

Os testes para avaliar as melhores condições de expressão foram feitos em relação à concentração de IPTG ($0,05 \text{ mM}$; $0,1 \text{ mM}$; $0,5 \text{ mM}$ e $1,0 \text{ mM}$) a 25°C por 20h, a variação de temperaturas de 37°C por 4 h, 30°C por 16h, 25°C por 20h e 18°C por 24 horas e a melhor cepa de expressão. Em todos os experimentos, a indução ocorreu quando a cultura atingiu a $\text{DO}_{600\text{nm}}$ entre 0,4 e 0,5.

Após o tempo de indução em cada teste, as células foram recuperadas por centrifugação a $6000 \times g$, por 5 minutos a 4°C . O sobrenadante foi descartado e o concentrado celular (*pellet*) foi pesado e ressuspenso em tampão Tris 20 mM pH 8,0 com NaCl 150 mM , glicerol 5% e imidazol 20 mM , na proporção de 1g de *pellet* para 10 mL de tampão. As células foram rompidas por sonicação com amplitude de 30% durante 15s com 30s de intervalo. Em seguida, o lisado celular foi centrifugado a $9000 \times g$, por 30 min, a 4°C , e o sobrenadante (fração solúvel) e o precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE.

5.2.2.3 Confirmação da expressão da lectina por espectrometria de massas sequencial

A expressão da lectina recombinante rADEL foi analisada por espectrometria de massas sequencial (MS/MS), com o objetivo de verificar se a proteína produzida correspondia com a estrutura primária da lectina nativa.

As frações insolúveis e solúveis dos testes de expressão foram submetidas a SDS-PAGE 12% (Laemmli, 1970). As bandas referentes às lectinas foram excisadas, reduzidas, alquiladas e, posteriormente, digeridas com a enzima tripsina, de acordo com a metodologia de

Shevchenko e colaboradores (2007). A análise dos peptídeos no equipamento foi realizada conforme descrito na seção 3.2.2.4.

5.2.3 Produção da proteína recombinante rADEL em sistema heterólogo (França)

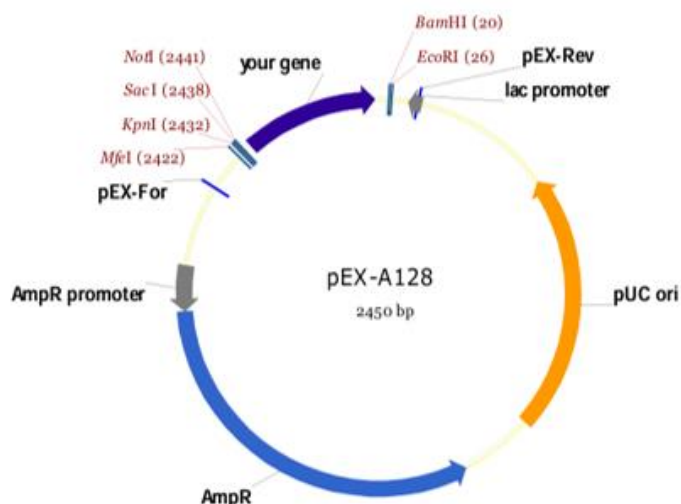
5.2.3.1 Desenho do gene sintético adquirido na França

A sequência de aminoácidos da lectina nativa (C0HK25.1- UniProt: <https://www.uniprot.org/>) foi utilizada para desenhar o gene sintético (Carneiro *et al.*, 2017b).

O gene sintético foi adquirido pela empresa Eurofins Genomics (França) no vetor *backbone* pEX-A128 com a sequência gênica otimizada para *Komagataella pastoris* e delimitada pelas enzimas de restrição NdeI e SalI (Figura 29).

Figura 29 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pEX-A128.

A catatggaccgataagtgcgatacaattagggtggaaagctggagctacaagtacgcg
H M D P D K C D T I R V E S W S Y K Y A
 gagaaggttggtggaagatgcctcctatgtcctcaacatgacggtcgtagatcggcagtc
 E K V V E D A S Y V L N M T V V D R Q S
 gcggctgcgtgtaccctgggggaatcctttggctaccagaaagccaccctgtgggtggat
 A A A C T L G E S F G Y Q K A T L W V D
 catggctgtcgtgctgacttcaaagtttgctatctcccagtaatgccgaccgagtgccaa
 H G C R A D F K V C Y L P V M P T E C Q
 accctccgtgtggagagttggaactacaagtacgcagagaaggtggtggagggggccgca
 T L R V E S W N Y K Y A E K V V E G A A
 ctgttcatcaacatgaccggttgaagatcggcagtcggaggccagttgtgatctggataaa
 L F I N M T V E D R Q S E A S C D L D K
 tcctttggcttctacaatcagaactccacggtgtgggtgaaccacggttgccgtgccgac
 S F G F Y N Q N S T V W V N H G C R A D
 ttcaacatctgttatctgaaaggtgccgtgaccaccagtacaataaacgtgagcagctgg
 F N I C Y L K G A V T T S T I N V S S W
 aactaccaatatgccaccaaggtgcttcccgccgcatcctgcatttactccatgcgagta
 N Y Q Y A T K V L P A A S C I Y S M R V
 gtcaaccaacagtcgcggcccccttgactctgggaaccacttatggttttgtggccaat
 V N Q Q S A A P C T L G T T Y G F V A N
 accatgtgggtggacgatggctgccgtgaggacttcaaaccgagctattattccccataa
 T M W V D D G C R A D F K P S Y Y S P -
 gtcgac

B

MCS of pEX-A128

```

GGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTCTG
GGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTAAGAGAGAAAGG
caattgGGTACCgagctcGCGGCCGCAAGC>your_gene>ACCTGCTT
TTGCTCGCTTGGATCCgaattcAAAGGTGAAATTGTTATCCGCTCACAA
TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTG

```

Start point: gene insertion site

(A of ACCT 3' of "your gene" is bp no. 1)

Multiple Cloning Site: bases 19-30 / 2421-2446

lac Promoter: bases 65-95

pUC origin: bases 419-1007

Ampicillin resistance ORF: bases 1178-2038

pEX-For priming site: bases 2323-2343

pEX-Rev priming site: bases 75-100

Fonte: Autora.

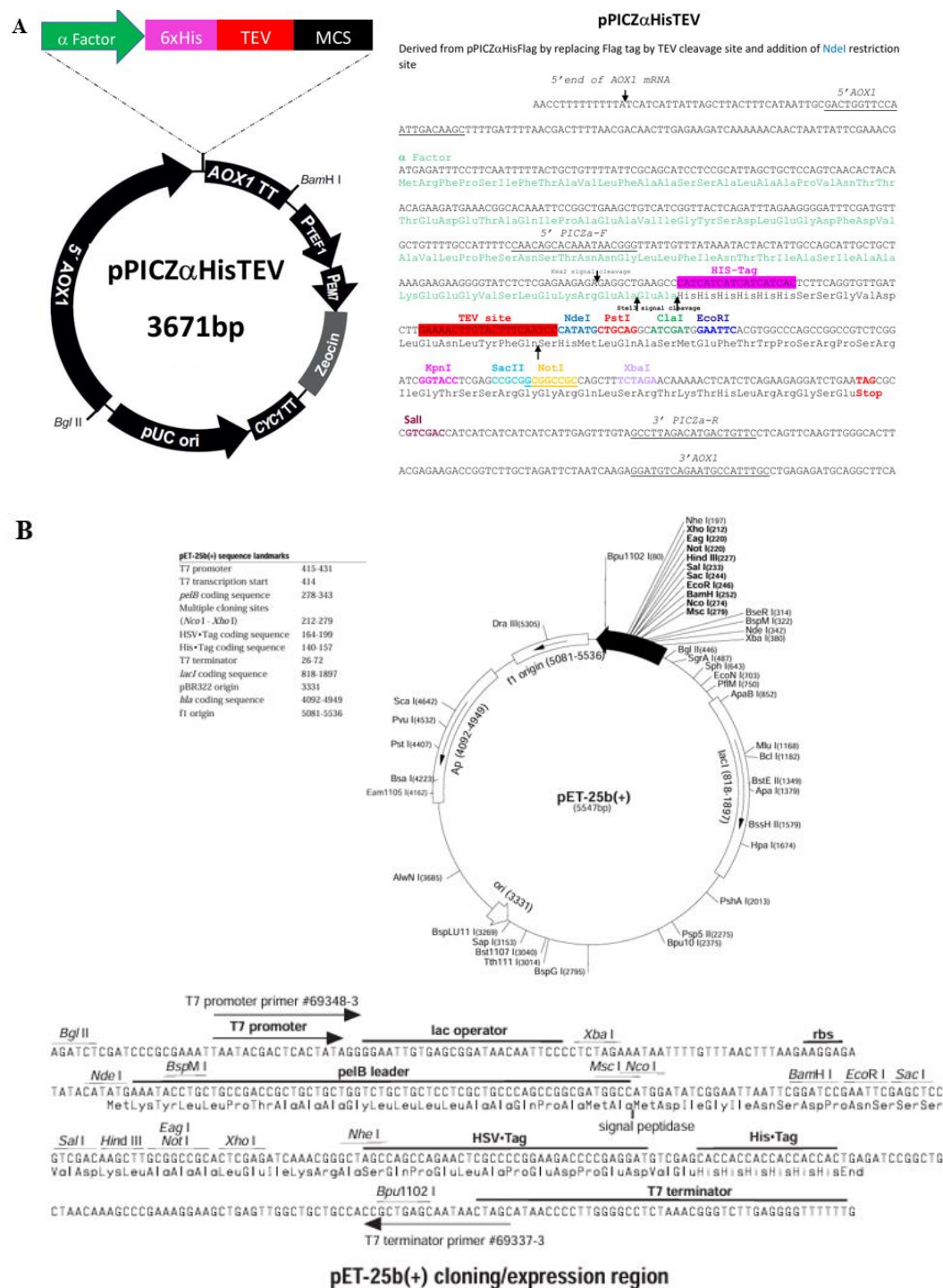
A – Representação da sequência de nucleotídeos e a respectiva tradução da lectina. A sequência em letras minúsculas representa os nucleotídeos e a sequência em letras maiúsculas representa os aminoácidos traduzidos. As letras em vermelho representam os sítios de clivagem *NdeI* (início) e *Sall* (final). O símbolo – representa o códon de parada da tradução. B – Representação do vetor de expressão pEX-A128 e seu respectivo mapa. Disponível em: https://eurofinngenomics.eu/media/1587668/pex-a128_map_seq_v10.pdf.

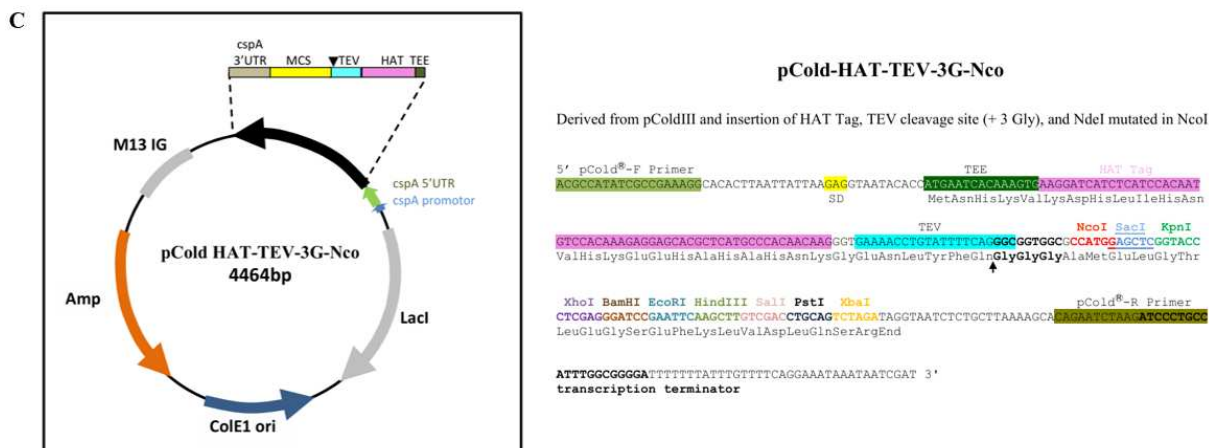
5.2.3.1.1 Subclonagens do gene sintético para os vetores de expressão

O gene sintético inicialmente inserido no vetor *backbone* pEX-A128 foi subclonado nos vetores de expressão pPICZαHisTEV, para expressão na levedura *Komagataella pastoris*, e nos vetores pET25b(+) e pCOLD-HAT-TEV-3G-Nco, para expressão em *E. coli*. O plasmídeo pET25b(+) foi adquirido pela empresa Novagen (França) e os demais foram desenvolvidos pelo grupo GBMS, disponível na coleção do CERMAV- CNRS, Grenoble, França.

Os mapas e as particularidades dos plasmídeos usados para as subclonagens encontram-se na Figura 30 e Tabela 06, respectivamente.

Figura 30 – Representação dos vetores de expressão utilizados para subclonar a sequência gênica.





Fonte: Autora. A – Vetor pPICZαHisTEV. B – Vetor pET25b(+). C – Vetor pCOLD-HAT-TEV-3G-Nco. Os mapas de restrição dos vetores estão disponíveis no banco de dados do CERMAV.

Tabela 06 – Características dos vetores utilizados para a expressão da lectina.

VETOR	ESPECIFICIDADE	PROMOTOR (Proteína)	ANTIBIÓTICO DE SELEÇÃO
pPICZαHisTEV	Derivado do pPICαHisFlag, presença da <i>His-tag</i> , substituindo a <i>Flag tag</i> pela inserção do sítio de clivagem da protease TEV e adição do sítio de restrição <i>NdeI</i> .	AOX	Zeocina
pET25b(+)	Presença de um peptídeo sinal para expressão no periplasma, e das <i>tags His-tag</i> e <i>HSV-tag</i> .	T7 Lac	Ampicilina
pCOLD-HAT-TEV-3G-Nco	Derivado do pCOLDIII contendo TEE, com a inserção da <i>HAT tag</i> , sítio de clivagem TEV, adição de + 3 Gly após o sítio da TEV e a substituição do sítio de restrição <i>NdeI</i> por <i>NcoI</i> .	cspA	Ampicilina

Fonte: Autora. As enzimas de restrição *NdeI* e *SalI* foram empregadas nos vetores pPICZαHisTEV e pET25b(+). A enzimas *NcoI* e *XhoI* foram empregadas no vetor pCOLD-HAT-TEV-3G-Nco. *HSV-tag* é derivada da proteína glicoproteína D do vírus herpes simples (HSV).

O gene sintético adquirido foi diluído para uma concentração de 50 ng.μL⁻¹ e inserido, por choque térmico, na cepa de clonagem *Escherichia coli* NEB5α (Sambrook, 2002).

Após a transformação, as colônias crescidas em LB ágar e selecionadas pelo antibiótico foram inoculadas a 37 °C, *overnight*, em meio LB *Broth* contendo ampicilina (100

µg/mL) para extração do DNA plasmidial. A miniprep foi realizada utilizando o kit de purificação em pequena escala QIAprep® *Spin Miniprep* (Qiagen), seguindo as recomendações do fabricante.

O DNA plasmidial purificado foi quantificado em espectrofotômetro *NanoDrop* (Thermo Fisher Scientific) e avaliado qualitativamente por eletroforese em gel de agarose a 1%, corado com SYBR Safe 1x (Invitrogen).

Para subclonagem nos vetores pPICZαHisTEV e pET25b(+), que possuem os mesmos sítios de restrição, o plasmídeo extraído foi digerido a 37 °C, por 3 horas, usando as enzimas *NdeI* (BioLabs) e *Sall* (BioLabs), de acordo com a configuração do vetor. A reação foi preparada seguindo as recomendações do fabricante (New England BioLabs), utilizando uma proporção de 10 U de enzima/µg de DNA. Após a digestão, os produtos resultantes foram separados por eletroforese em gel de agarose a 1% e o fragmento correspondente ao inserto da proteína foi identificado pelo tamanho em pares de bases (pb) em comparação ao marcador comercial O'GeneRuler 1kb DNA Ladder (*Thermo Scientific*) com intervalo de 250 a 10.000 pb. O fragmento foi purificado usando o kit Monarch® DNA Gel Extraction (New England BioLabs), conforme protocolo disponibilizado pelo fabricante, quantificado em espectrofotômetro *NanoDrop* (Thermo Fisher Scientific).

Os vetores pPICZαHisTEV e pET25b(+) sem inserto e previamente transformados em cepas de *E. coli* NEB5α foram extraídos por miniprep e submetidos ao mesmo protocolo de digestão enzimática com *NdeI* e *Sall*, obtendo extremidades coesas compatíveis com o inserto. Os vetores linearizados foram purificados a partir do gel de agarose a 0,7% usando o mesmo kit Monarch e quantificado em espectrofotômetro *NanoDrop*.

O inserto purificado e o vetor linearizado foram submetidos a uma reação de ligação usando o DNA *Ligation Kit* (Takara), conforme descrito na seção 3.2.5.1.1. Em seguida, 2 µL da reação de ligação foram usados para transformar as células de clonagem *E. coli* NEB5α, por choque térmico (Sambrook, 2002).

Algumas colônias isoladas no meio LB ágar contendo o antibiótico de seleção do vetor (ampicilina 100 µg/mL ou 25 µg/mL zeocina) foram selecionadas para confirmação da transformação por PCR em colônia usando o kit *Taq'Ozyme Purple Mix*, como detalhado na seção 3.2.5.1.1. A T °C de anelamento foi ajustada de acordo com a temperatura de *melting* (Tm) para o par de primer específico do vetor utilizado. Os primers e a temperatura de anelamento usados para cada vetor estão descritos na Tabela 07.

Tabela 07 – Primers e condições de PCR para confirmação de vetores.

<u>Vetor</u>	<u>Primers</u>	<u>T °C de anelamento</u>
pPICZαHisTEV	PICZ α (F)	50 °C
	PICZ α (R)	
pET25b(+)	T7 PRO (F)	50 °C
	T7 TER (R)	
pCOLD-HAT-TEV-3G-Nco	pCOLDTF-F (F)	52 °C
	pCOLD-R(R)	

Fonte: Autora. (F): primer *forward* e (R): primer *reverse*.

Os produtos da PCR foram visualizados por eletroforese em gel de agarose a 1%. Um clone positivo, identificado de acordo com o tamanho esperado do fragmento amplificado, foi selecionado e cultivado em LB ágar e LB *Broth*. A alíquota restante (5 μ L) da suspensão bacteriana em água Mili-Q foi suplementada com 50 μ L de LB *Broth* e plaqueada por estriamento em LB ágar contendo o antibiótico de seleção. O restante da suspensão foi crescido em 5mL de LB *Broth*, a 37 °C, *overnight*, contendo o mesmo antibiótico, e o DNA plasmidial foi extraído por miniprep conforme descrito anteriormente.

Em seguida, o plasmídeo extraído foi quantificado e submetido novamente a reação de digestão a 37 °C, por 1 hora, usando as mesmas endonucleases *NdeI* e *Sall*, para confirmação. Os produtos resultantes da digestão foram analisados em gel de agarose a 1% e a presença do fragmento esperado foi confirmada com base no tamanho em pares de bases. O DNA plasmidial extraído do clone transformado e confirmado por digestão foi enviado para sequenciamento de confirmação pela EUROFINS (França).

Para subclonagem no vetor pCOLD-HAT-TEV-3G-Nco, os sítios de restrição do plasmídeo de rADEL em pEX-A128 foram alterados para *NcoI* e *XhoI* em PCR, a fim de possuírem o mesmo sítio de restrição do vetor de expressão. A reação de PCR foi realizada com o PrimeSTAR 2x Kit (Takara), seguindo as recomendações do fabricante, juntamente com os primers *forward* e *reverse* adquiridos pela Eurofins e específicos para essa reação, na concentração final de 0,2 μ M.

A T °C de anelamento foi ajustada de acordo com a temperatura de *melting* (Tm) para o par de primer específico. Os primers, a Tm e a %GC (guanina/citosina) para cada vetor estão descritos na Tabela 08.

Tabela 08 – Primers para subclonagem no vetor pCOLD-HAT-TEV-3G-Nco.

<u>Primer</u>	<u>%GC</u>	<u>Tm (°C)</u>
ADEL-Nco-Forward ATACCATGGACCCGGATAAGTGCGATACAATTAGG	46%	64,8 °C
ADEL-Xho-Reverse ACACCTCGAGTTATGGGGAATAATAGCTCGG	48%	53,7 °C

Fonte: Autora. A Tm considerada na Tabela é a Tm para homologia total. Tm para o par de primers: $[(TmForward + TmReverse)/2] - 3 = 56\text{ °C}$.

A PCR ocorreu em um termociclador *Eppendorf* programado de acordo com as temperaturas de homologia parcial e total dos primers e com a Tm do par de primers. A reação ocorreu com cinco ciclos de 10 segundos a 98 °C, seguido por 5 segundos a 50 °C e 5 segundos a 72 °C, depois 30 ciclos de 10 segundos a 98 °C, seguido por 5 segundos a 60 °C e 5 segundos a 72 °C. Após a reação de PCR, o termociclador foi programado para incubar as amostras a 16 °C.

Os produtos da PCR foram visualizados por eletroforese em gel de agarose a 1%, e o fragmento correspondente ao inserto da proteína foi identificado pelo tamanho em pares de bases. O fragmento foi purificado usando o kit Monarch[®] DNA Gel Extraction (*New England BioLabs*), conforme protocolo disponibilizado pelo fabricante e quantificado em espectrofotômetro *NanoDrop* (*Thermo Fisher Scientific*).

O fragmento do inserto purificado foi digerido a 37 °C, por 3 horas, usando as enzimas *NcoI* (*BioLabs*) e *XhoI* (*BioLabs*), de acordo com a configuração do vetor. A reação foi preparada seguindo as recomendações do fabricante (*New England BioLabs*), utilizando uma proporção de 10 U de enzima/μg de DNA. Após a digestão, o inserto foi purificado com o DNA *cleanup* (*Qiagen*), seguindo as recomendações do fabricante, para remover as enzimas e o tampão.

O vetor pCOLD-HAT-TEV-3G-Nco foi extraído e submetido ao mesmo protocolo de digestão enzimática com *NcoI* e *XhoI*, como citado anteriormente. O vetor linearizado foi purificado a partir do gel de agarose a 0,7% usando o mesmo kit e quantificado em espectrofotômetro *NanoDrop*.

O inserto purificado e o vetor linearizado foram submetidos a uma reação de ligação usando o DNA *Ligation Kit* (*Takara*), como descrito previamente. Em seguida, 2 μL da reação

de ligação foram usados para transformar as células de clonagem *E. coli* NEB5 α , por choque térmico (Sambrook, 2002).

Algumas colônias isoladas no meio LB ágar contendo o antibiótico de seleção (ampicilina 100 $\mu\text{g/mL}$) foram selecionadas para confirmação da transformação por PCR em colônia usando o kit *Taq'Ozyme Purple Mix*, conforme descrito anteriormente, juntamente com os primers *forward* e *reverse* específicos para o vetor (Tabela 07)

Os produtos da PCR foram visualizados por eletroforese em gel de agarose a 1%. Um clone positivo foi selecionado e cultivado em LB ágar e LB *Broth*, o DNA plasmidial foi extraído por miniprep, quantificado e submetido novamente a reação de digestão com as enzimas *NcoI* e *XhoI*, conforme descrito para os outros vetores. Os produtos resultantes da digestão foram analisados e o DNA plasmidial extraído do clone transformado e confirmado por digestão foi enviado para sequenciamento de confirmação pela EUROFINS (França).

5.2.3.2 Produção da proteína rADEL em sistema heterólogo de *Komagataella pastoris*

5.2.3.2.1 Transformação de *K. pastoris* com o plasmídeo rADELpPICZ α HisTEV

A transformação do vetor em *K. pastoris* e a expressão da lectina no mesmo sistema de expressão seguiu as recomendações do *Pichia Expression Kit: user guide*, Invitrogen (Invitrogen, Thermo Fisher Scientific, 2020; Invitrogen, 2010).

Após a confirmação por sequenciamento, o DNA plasmidial de rADELpPICZ α HisTEV foi extraído de *E. coli* NEB5 α por *midiprep* usando o kit de purificação em média escala QIAprep® *Spin Midiprep* (Qiagen), seguindo as recomendações do fabricante. O DNA plasmidial purificado foi quantificado em *NanoDrop* e linearizado com a enzima *SacI* (*Biolabs*) para uma concentração de 20 μg de inserto e 20 U/ μL de enzima em uma concentração final de 100 μL de reação, por 3 horas a 37 °C, seguindo as orientações do fabricante.

O produto resultante da linearização foi analisado em gel de agarose a 1% e o inserto foi purificado com o DNA *cleanup* (Qiagen), seguindo as recomendações do fabricante, para remover a enzimas e o tampão. O inserto purificado foi quantificado em *NanoDrop* e congelado a - 20 °C.

Para a transformação, a levedura *Komagataella pastoris* GS115 foi crescida em meio YP ágar (1% de ágar de extrato de levedura e 2% de peptona). Uma colônia foi pré

cultivada em 5 mL de meio YPD (1% de extrato de levedura, 2% de peptona e 2% de dextrose), a 30 °C, *overnight*, em agitação constante de 220 rpm.

Um pequeno volume de pré-cultura (20 µL) foi inoculado em 100 mL de meio YPD em *erlenmayer* de 500 mL. As células foram cultivadas a 30 °C, em agitação constante de 220 rpm até atingir a densidade ótica a 600nm (DO_{600nm}) de 1,3 a 1,5. Em seguida, as células foram separadas em dois volumes iguais de cultura (2x40mL) e centrifugadas a 2500 x g, por 5 minutos, a 4 °C, para lavagem do meio. As células foram ressuspensas em 40 mL de água Mili-Q estéril e gelada, e centrifugadas novamente nas mesmas condições. As células foram ressuspensas em 20 mL de água gelada estéril, centrifugadas e ressuspensas em 1,6 mL de Sorbitol 1M estéril gelado. As células foram centrifugadas uma última vez e ressuspensas no volume final de 160 µL de Sorbitol 1M gelado.

As células foram transformadas com o plasmídeo de interesse por eletroporação. Para isso, 80 µL das células foram misturadas com 10 µg de DNA e transferidas para a cubeta de eletroporação. A cubeta foi incubada em gelo por 5 minutos e recebeu um pulso de 1600 volts no eletroporador (*Eppendorf*). Imediatamente depois, 1 mL de sorbitol 1M foi adicionado e a solução incubada por 2 horas a 30 °C, sem agitação em um tubo de 15 mL. Em seguida, as células foram incorporadas em YPDS ágar (1% de ágar de extrato de levedura, 2% de peptona, 2% de dextrose e 1 M de Sorbitol) com 100 µg/mL de zeocina para seleção das células transformadas, e incubadas a 30 °C por três dias.

A confirmação da transformação foi feita por PCR em colônia com o kit *Taq'Ozyme Purple Mix*, conforme descrito anteriormente, juntamente com os primers *forward* e *reverse* específicos para a região do promotor AOX1. Os produtos da PCR foram visualizados por eletroforese em gel de agarose e quatro clones positivos foram selecionados e cultivados a 30 °C em YPDS ágar (1% de ágar de extrato de levedura, 2% de peptona, 2% de dextrose e 1 M de Sorbitol) com 100 µg/mL de zeocina.

5.2.3.2.2 Expressão da lectina rADEL em *K. pastoris*

Primeiramente, um *screening* dos quatro clones positivos foi realizado para avaliar a melhor expressão da proteína. Uma colônia de cada clone, previamente crescido em ágar YPDS com 100 µg/mL de zeocina, foi cultivada em 50 mL de meio GYC (Tabela 09) a 30 °C, *overnight*, sob agitação constante de 220 rpm.

Quando o crescimento celular (DO_{600nm}) atingiu de 2 a 6, o meio de crescimento (GYC) foi substituído pelo meio de indução (MYC). As células foram centrifugadas a 1500 x g, por 5 minutos, a 4 °C e o sobrenadante descartado. O *erlenmayer* foi lavado com água Milli-Q estéril para remoção do glicerol residual e as células foram ressuspensas no meio MYC (Tabela 09). Antes da troca do meio, uma alíquota do *pellet* e do sobrenadante foram coletadas para o controle negativo da expressão e analisadas em SDS-PAGE.

As células foram induzidas por mais 3 dias com a adição de metanol diariamente para a concentração final de 3% (v/v) a 30 °C, 220 rpm. Em todos os dias de indução, alíquotas do sobrenadante e do *pellet* foram retiradas e analisadas em SDS-PAGE.

A composição dos meios de expressão e os volumes de cada componente estão descritos na Tabela 09.

Tabela 09 - Composição dos meios usados na expressão de proteínas em *K. pastoris*.

MEIO	COMPOSIÇÃO	CONCENTRAÇÃO FINAL
YPCA (1,25x)	1,25%: extrato de levedura	1%
	2,5%: peptona	2%
	1,25%: ácido casimínico	3%
GYC	80%: YPCA 1,25x (v/v)	1x
	10%: YNB 10x	1x
	0,2%: Biotina 500x	1x
	10%: Glicerol 10%	1%
MYC	80%: YPCA 1,25X (v/v)	1x
	10%: YNB 10x	1x
	0,2%: Biotina 500x	1x
	3%: Metanol PA	3%
	7% Água	

Fonte: Autora. YPCA = extrato de levedura, peptona e ácido casimínico. YNB = base nitrogenada de levedura. No meio YPCA, as porcentagens referem-se à concentração dos componentes sólidos (% m/v). Nos meios GYC e MYC, as porcentagens referem-se à proporção de volume de cada solução (% v/v) para compor um volume final de 100%. O símbolo “x” indica a concentração dos estoques.

Ao finalizar o tempo de expressão, as células foram recuperadas por centrifugação a 2500 x g, por 10 minutos a 25 °C. O sobrenadante e o concentrado celular (*pellet*) foram separados e congelados a -20 °C.

Após a expressão inicial para *screening*, os clones positivos com melhor rendimento da proteína foram selecionados para expressão em maior escala. Os clones mais rentáveis foram determinados pela comparação da intensidade da banda correspondente a

proteína nos géis de SDS-PAGE e pelo maior pico de absorbância a 280 nm durante a purificação.

A expressão em maior escala foi realizada em 750 mL de meio GYC/MYC em *erlenmeyers* de 3L, seguindo as mesmas proporções da Tabela 09. As induções foram realizadas conforme o protocolo descrito acima, com duração média de 3 dias de expressão. Após indução, os sobrenadantes foram separados por centrifugação e congelados a -20 °C até a etapa de purificação.

5.2.3.2.3 Purificação da lectina recombinante

O vetor pPICZ α contém o peptídeo sinal α -factor que é responsável por secretar para o meio de cultura a proteína expressa em *K. pastoris*. Por esse motivo, após a centrifugação da cultura, o sobrenadante foi submetido a uma precipitação com sulfato de amônio a 70% de saturação, por 16 horas a 25 °C.

Em seguida, o precipitado foi recuperado por centrifugação a 9000 x g, 20 minutos a 4 °C, ressuspenso em 30 mL de tampão Tris 20 mM pH 8, contendo NaCl 500 mM e imidazol 10 mM (Tampão A) e filtrado em 45 μ m.

Como a construção do vetor possui cauda com histidina, a fração solubilizada foi submetida à cromatografia de afinidade a metais imobilizados (IMAC). A fração foi aplicada em uma coluna HisTrapTM FF de 1mL, previamente equilibrada e lavada com o mesmo tampão A. As proteínas mais fracamente ligadas a coluna foram removidas com 5% do mesmo tampão de equilíbrio contendo imidazol 500 mM (tampão B), enquanto as proteínas fortemente ligadas a coluna foram eluídas com um gradiente linear de 5% a 100% do tampão B. A cromatografia foi realizada em um sistema NGC (Bio-Rad) e alíquotas de cada passo de purificação foram coletadas e analisadas por SDS-PAGE.

As frações eluídas da coluna HisTrapTM contendo a proteína de interesse foram identificadas por SDS-PAGE e a funcionabilidade foi confirmada por atividade hemaglutinante contra eritrócitos de coelho a 4% (Sampaio *et al.*, 2002). As frações contendo a proteína pura e funcional foram reunidas para uso posterior.

5.2.3.2.4 Purificação da lectina recombinante a partir do concentrado celular

Para avaliar se a extração citoplasmática era viável, o pellet recuperado após a centrifugação da cultura induzida foi pesado e cada 1g de *pellet* celular foi ressuscitado em 5 mL de tampão A. A enzima denarase foi adicionada na proporção de 1 μ L para cada 25 mL de suspensão celular, seguida de incubação em misturador rotativo por 20 minutos. As células foram lisadas usando um disruptor celular mecânico da *One Shot Cell Disrupter* (Constant System) a 1.9 Kbars.

Em seguida, o lisado celular foi centrifugado a 24.000 x g, por 30 minutos a 10 °C e o sobrenadante obtido filtrado em filtro de 0,45 μ M. Alíquotas do sobrenadante (fração solúvel) e do precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE. A fração solúvel foi aplicada em uma coluna HisTrap™ FF 1mL, e a cromatografia foi conduzida seguindo o mesmo protocolo da purificação do precipitado solubilizado, descrito acima. As frações de cada etapa da purificação foram analisadas por SDS-PAGE e a funcionabilidade foi avaliada por atividade hemaglutinante contra eritrócitos de coelho a 4%.

5.2.3.3 Produção da lectina rADEL em sistema heterólogo de *E. coli*

5.2.3.3.1 Expressão da lectina recombinante

Após a confirmação por sequenciamento, os plasmídeos rADELpET25b e rADELpCOLD-HAT-TEV-3G-Nco foram transformados na cepa de expressão *E. coli* BL21(Star) e *E. coli* Shuffle B, respectivamente, por choque térmico (Sambrook, 2002).

Conforme descrito na seção 3.2.5.2, optou-se pela não realização de novos testes de expressão, adotando as condições padrões recomendadas pelo grupo para os vetores e cepas selecionadas, visando otimização do tempo e dos recursos disponíveis.

Para expressão da proteína, uma colônia de *E. coli* transformada com o plasmídeo de interesse e isolada em LB ágar contendo o antibiótico de seleção, (ampicilina a 100 μ g/mL) foi pré-cultivada em 10 mL de LB *Broth* contendo o mesmo antibiótico a 37 °C, *overnight*, sob agitação constante de 180 rpm.

Todo o volume de pré-cultura foi inoculado em 500 mL de LB *Broth* em *erlenmayer* de 3L contendo o mesmo antibiótico. As células foram cultivadas a 37 °C para *E. coli* BL21 e a 30 °C para *E. coli* Shuffle B, em agitação constante de 180 rpm até atingir a densidade ótica

a 600nm de 0,4. Em seguida, a temperatura foi reduzida para 16 °C em *E. coli* BL21 e para 25 °C em *E. coli* Shuffle B, e o cultivo prosseguiu até atingir DO = 0,8, quando o indutor IPTG (isopropil-beta-D-tiogalactopiranosídeo) foi adicionado ao meio na concentração de 0,2 mM. A expressão continuou na mesma temperatura, 180 *rpm*, por 18 horas.

Após o tempo de expressão, as células foram recuperadas por centrifugação a 5000 x g, por 10 minutos a 25 °C. O sobrenadante foi descartado e o concentrado celular (*pellet*) foi pesado e congelado a -20 °C.

5.2.3.3.2 Purificação da lectina recombinante

A estratégia de purificação da lectina rADEL foi conduzida de acordo com as características da construção inserida em cada vetor de expressão. A Tabela 10 apresenta as diferentes abordagens de purificação empregadas em cada caso.

Tabela 10 – Estratégias de purificação da lectina rADEL em cada vetor.

Vetor (Sítio de restrição)	Cauda/tag	Estratégia	Cromatografia (Matriz)
pET25b(+) (<i>NdeI-XhoI</i>)	Ausente	Purificação seguindo o método da lectina nativa	Afinidade (Sephacrose)
pCOLD-HAT-TEV-3G-Nco (<i>NcoI-XhoI</i>)	<i>HAT-tag</i>	Purificação por afinidade via <i>tag</i> pela presença de histidinas	Afinidade – IMAC (Histrap – Ni ²⁺)

Fonte: Autora. IMAC: Immobilized Metal Affinity Chromatography – Cromatografia de afinidade a metais imobilizados.

Para a construção no vetor pET25b, cada 1g de *pellet* celular foi ressuspensionado em 5 mL de tampão Tris 20 mM pH 7,5 e NaCl 150 mM. Para a construção no vetor pCOLD-HAT-TEV-3G-Nco, o *pellet* celular foi ressuspensionado na mesma proporção em tampão A.

Em ambas as construções, a enzima denarase foi adicionada na proporção previamente descrita, seguida de incubação por 20 minutos. As células foram lisadas em disruptor celular mecânico, e o lisado celular foi centrifugado e filtrado conforme descrito nas seções anteriores. Aliquotas do sobrenadante (fração solúvel) e do precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE.

Como a construção no vetor pET25b não possui cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade, conforme o método usado para purificação da lectina nativa (Carneiro *et al.*, 2017b). A fração solúvel, em tampão Tris 20 mM pH 7,5 e NaCl 150 mM, foi aplicada em uma coluna de Sepharose 6B (Sigma), previamente equilibrada e lavada com o mesmo tampão. As proteínas retidas a coluna foram eluídas com 0,2 M de lactose no tampão de equilíbrio. A cromatografia foi realizada por gravidade a um fluxo de 2 mL/min.

Para a construção no vetor pCOLD-HAT-TEV-3G-Nco, que possui cauda com histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados (IMAC), conforme metodologia descrita anteriormente. A fração foi aplicada em uma coluna HisTrap™ FF de 1mL, e as proteínas retidas a coluna foram eluídas com 5% de tampão B, seguido de um gradiente linear de 5% a 100% do mesmo tampão.

As frações obtidas em cada etapa de purificação, para ambas as construções, foram analisadas por SDS-PAGE.

5.2.4 Avaliação da pureza e funcionabilidade da proteína recombinante (França)

5.2.4.1 Estimativa da massa molecular e avaliação da pureza da lectina

A estimativa da massa molecular e o grau de pureza nas diferentes etapas de expressão e purificação da lectina recombinante foram avaliados por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) (Laemmli, 1970), conforme descrito na seção 3.2.6.

Para as alíquotas do sobrenadante derivado da expressão em *K. pastoris*, as amostras foram preparadas por precipitação com TCA 100% (ácido tricloroacético), na concentração de 1/10 (v/v). A solução foi incubada a 4 °C por 15 minutos e centrifugada a 12.500 x g, 15 minutos. O sobrenadante foi removido, o pellet ressuscitado em 25 µL de tampão de amostra 1x, e a amostra foi aquecida a 100 °C, 5 minutos.

A massa molecular foi estimada em condições redutoras, comparando a migração da proteína com a migração do marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad, com intervalo de 10 kDa a 250 kDa.

5.2.4.2 Atividade hemaglutinante da proteína

A atividade hemaglutinante foi realizada com eritrócitos de coelho não tratado com enzima proteolítica para avaliar a funcionabilidade da proteína expressa, como detalhado na seção 3.2.7.

5.2.5 Immunoblotting para detecção da proteína rADEL expressa (França)

Para detectar a presença da lectina rADEL expressa em *K. pastoris* nas alíquotas de sobrenadante e no concentrado celular, um ensaio de *immunoblotting* foi realizado usando anticorpos anti-His para identificação da cauda de histidina presente na proteína expressa.

As amostras foram separadas por eletroforese em gel de SDS-PAGE 12% com o marcador *Precision Plus Protein Prestained Dual Color Standards* (Bio-Rad), com intervalo de 10 kDa a 250 kDa. Em seguida, as proteínas foram transferidas para uma membrana de nitrocelulose (BioTrace NT, PALL), usando o tampão de transferência Tris 25 mM pH 8,3, Glicina 192 mM, SDS 0,10% (1x) contendo 20% de etanol (v/v). A transferência ocorreu em sistema de blot, Bio-Rad, a 100 v e 50 mA por 1 hora.

O immunoblotting foi realizado usando um sistema de detecção SNAP i.d. 2.0 (Merck), seguindo as recomendações do fabricante. Os sítios inespecíficos foram bloqueados com o tampão TBST (Tris 20 mM pH 7,5, NaCl 150 mM e 0,1% Tween20) acrescido de 1% BSA (p/v) (Albumina Sérica Bovina). Após o bloqueio, a membrana foi incubada por 20 minutos com o tampão TBST acrescido do anticorpo monoclonal anti-poli-histidina-peroxidase, clone HIS-1 (Sigma) na concentração de 1:4.000 (v/v). A membrana foi lavada 4 vezes com tampão TBST, e em seguida incubada com a solução de *blotting* (Immobilon[®] ECL UltraPlus Western HRP Substrate, Sigma) contendo o substrato quimioluminescente para HRP (peroxidase do rábano), para visualização da proteína expressa.

5.2.6 Ensaios de cristalização (França)

Os ensaios de cristalização foram realizados manualmente em placas de cristalização de 48 poços e por triagens robotizadas na plataforma de cristalização do laboratório HTX no EMBL (*European Molecular Biology Laboratory-Grenoble*).

Nas triagens robotizadas, a cristalização foi conduzida pelo método de difusão de vapor com gotas apoiadas (*sitting drops*) de 200 nL em placas de 96 poços, na proporção de 1:1 proteína/reservatório. Após a montagem, as placas foram colocadas no robô de imageamento Vis-UV a 20 °C.

Para a triagem manual, em cada poço foram colocados 200 µL da condição de cristalização e a gota foi composta por 1 µL da solução de proteína + 1 µL da condição do poço. Os poços foram vedados com lamínulas siliconadas e as placas incubadas a 20 °C. Após três dias, as placas foram avaliadas em microscópio, com monitoramentos subsequentes realizadas semanalmente.

Diferentes condições de cristalização foram testadas para a proteína não clivada rADELpPICZαHisTEV, variando a concentração da proteína e os kits de cristalização comerciais. As condições de cristalizações testadas estão listadas na Tabela 11.

Tabela 11 – Condições de cristalizações testadas para rADEL

AMOSTRA	[] (mg/mL)	TAMPÃO INICIAL	LIGANTE	KIT DE CRISTALIZAÇÃO/ MONTAGEM
rADELpPICZαHisTEV	5,0	HEPES 20 mM pH 7,5 NaCl 100 mM	GalNac 10 mM	Morpheus II – Box I e II (Molecular dimensions) / Manual
rADELpPICZαHisTEV	3,5	HEPES 20 mM pH 7,5 NaCl 100 mM	GalNac 10 mM	SG1 Box I e II (Molecular dimensions) / Manual
rADELpPICZαHisTEV	2,4	HEPES 20 mM pH 7,5 NaCl 100 mM	GalNac 10 mM	Triagem robótica

Fonte: Autora. [] mg/mL: concentração da proteína em mg/mL. Tampão inicial: tampão da solução de proteína.

5.2.7 Produção da proteína recombinante rADEL em sistema heterólogo de *K. pastoris* (Brasil)

5.2.7.1 Recuperação do gene sintético francês no Brasil

Na França, a construção gênica da proteína rADEL que apresentou os melhores resultados de expressão e purificação foi a inserida no vetor pPICZ α HisTEV. Por esse motivo, o plasmídeo rADELpPICZ α HisTEV foi trazido para o Brasil em cartões QIAcard FTA (*Flinders Technology Associates*), Sigma.

Para armazenamento, 1 μ g do DNA plasmidial, confirmado por sequenciamento, foi aplicado na superfície do cartão e a região contendo a gota foi demarcada. A recuperação do plasmídeo foi realizada recortando a região da gota, embebedando a membrana em água Mili-Q filtrada e separando o líquido por centrifugação.

5.2.7.2 Transformação de *E. coli* com o gene sintético rADELpPICZ α HisTEV

Após recuperação, o plasmídeo rADELpPICZ α HisTEV foi inserido, por choque térmico, na cepa de clonagem *Escherichia coli* DH5 α (Sambrook, 2002). A manipulação e uso de organismos geneticamente modificados tem autorização pelo certificado CQB-R007-2016 CTNBio.

Para confirmar a transformação na cepa de clonagem, as colônias crescidas em LB ágar e selecionadas pelo antibiótico zeocina (100 μ g/mL), foram inoculadas em meio LB *Broth* contendo o mesmo antibiótico e incubadas a 37 °C, *overnight*.

A confirmação foi feita por PCR em colônia usando a enzima *Taq* DNA Polimerase (Sigma). As reações de PCR foram montadas de acordo com as recomendações do fabricante para o volume final de 25 μ L, contendo tampão 10x (concentração final de 1x), MgCl₂ 25 mM (2,5 mM), dNTP Mix 10 mM (0,4 mM), os primers *forward* e *reverse* específicos para o promotor AOX do vetor pPICZ α HisTEV (5' AOX1 *forward* primer: 5'-GACTGGTTCCAATTGACAAGC-3' e 3' AOX1 *reverse* primer: 5'-GCAAATGGCATTCTGACATCC-3') foram usados na concentração final de 0,4 μ M, *Taq* a 0,02 U/ μ L final e 2 μ L da cultura bacteriana.

As reações ocorreram em um termociclador *Eppendorf* programado para uma desnaturação inicial de 5 min a 94 °C, seguido por 30 ciclos de 30 segundos a 94 °C

(desnaturação), 30 segundos a uma temperatura de anelamento de 50 °C e 60 segundos a 72 °C (extensão). Depois do último ciclo, uma fase de extensão prolongada foi realizada (72 °C por 7 min). Após a reação de PCR, o termociclador foi programado para incubar as amostras a 4 °C. Os produtos da PCR foram visualizados por eletroforese em gel de agarose a 1%, e o clone positivo foi identificado de acordo com o tamanho esperado do fragmento amplificado.

5.2.7.3 Transformação de *K. pastoris* com o gene sintético *rADELpPICZαHisTEV*

Após a confirmação da transformação por PCR, o DNA plasmidial de *rADELpPICZαHisTEV* foi extraído de *E.coli* DH5α por *miniprep* usando o kit de purificação em pequena escala da *GE Healthcare*, seguindo as recomendações do fabricante. O DNA plasmidial purificado foi quantificado em espectrofotômetro e linearizado com a enzima *SacI* (*Thermo Scientific*) para uma concentração de 20 µg de inserto e 20 U/µL de enzima em uma concentração final de 100 µL de reação, por 3 horas a 37 °C, seguindo as orientações do fabricante.

O produto resultante da linearização foi analisado em gel de agarose a 1% e o inserto foi purificado com o DNA *cleanup* (*GE Healthcare*), de acordo com o manual do kit. O inserto purificado foi analisado em gel de agarose e congelado a - 20 °C.

A levedura *Komagataella pastoris* GS115 foi crescida em meio YP ágar e em seguida em meio YPD, a 30 °C, conforme descrito anteriormente. A transformação seguiu exatamente o protocolo detalhado na seção 4.2.3.2.1.

A tentativa de confirmação da transformação foi feita por PCR em colônia. As colônias crescidas em YPDS ágar (1% de ágar de extrato de levedura, 2% de peptona, 2% de dextrose e 1 M de Sorbitol) e selecionadas pelo antibiótico zeocina (200 µg/mL), foram inoculadas em meio YPD contendo o mesmo antibiótico e incubadas a 30 °C, *overnight*. A reação de PCR foi realizada conforme descrito na seção anterior.

Para uma maior seleção, alguns clones selecionados foram replaqueados em YPDS ágar com 200 µg/mL de zeocina e incubados a 30 °C. Uma colônia foi selecionada e cultivada a 30 °C em meio YPD com 200 µg/mL de zeocina, e após o crescimento, essa cultura (colônia 1) foi utilizada para as expressões.

Tentativas de extração do DNA genômico da colônia selecionada para expressão também foram realizadas com o objetivo de confirmação do plasmídeo por PCR. Diferentes extrações foram conduzidas seguindo os protocolos do manual da *Invitrogen* (Invitrogen,

Thermo Fisher Scientific, 2020), de Harju e colaboradores (2004) e de Looke e colaboradores (2011).

5.2.7.4 Expressão da lectina *rADEL* em *K. pastoris*

Para determinar as melhores condições de expressão, diferentes parâmetros foram analisados com base no *Pichia Expression Kit: user guide Invitrogen* (Invitrogen, Thermo Fisher Scientific, 2020).

Os ensaios foram conduzidos avaliando a presença ou a ausência de uma pré cultura, o uso das células para cultura diretamente do estoque congelado ou após replaqueamento em YPD ágar, diferentes tempos de indução (2, 3, 4, 5 e 6 dias) e diferentes densidades celulares (DO_{600nm}) para o início da indução ($DO_{600nm} = 1, 2, 3, 4, 6$ e 8).

As expressões seguiram o protocolo descrito na seção 4.2.3.2.2, com ajustes no volume de cultura e nos parâmetros avaliados. Para os testes da concentração celular e do tempo de indução, a colônia 1 foi previamente crescida em ágar YPD com $100 \mu\text{g/mL}$ de zeocina, e uma colônia isolada foi cultivada em 125 mL de meio GYC em *erlenmayer* de 500 mL , seguindo as mesmas proporções dos meios detalhados na Tabela 09. Quando a cultura atingiu a DO desejada, o meio de crescimento foi substituído pelo meio de indução, e as células foram induzidas por diferentes períodos conforme cada teste.

Para o ensaio com pré cultura, uma colônia previamente crescida em ágar YPD com $100 \mu\text{g/mL}$ de zeocina, foi cultivada em 7 mL de meio GYC em tubo *falcon* de 50 mL , a 30°C , *overnight*. Em seguida, $1,78 \text{ mL}$ da pré cultura foi inoculada em 125 mL de meio GYC em *erlenmayer* de 500 mL . No teste utilizando células diretamente do estoque congelado, as células foram inoculadas na proporção de $1:1000$ em 125 mL de meio GYC em *erlenmayer* de 500 mL . Em ambos os testes, a indução foi realizada de acordo com os melhores parâmetros dos testes iniciais.

Ao finalizar o tempo de expressão, as células foram recuperadas por centrifugação a $2500 \times g$, por 10 minutos a 25°C . O sobrenadante e o concentrado celular (*pellet*) foram separados e analisados por SDS-PAGE. Alíquotas de cada passo de purificação dos diferentes parâmetros avaliados foram coletadas para análise por eletroforese.

O melhor método de expressão foi determinado pela comparação da intensidade da banda correspondente a proteína nos géis de SDS-PAGE, pelo maior pico de absorbância a 280

nm durante a purificação, pelo título de hemaglutinação da fração purificada contra eritrócitos de coelho tratado com tripsina a 3% e pelo tempo de crescimento.

Após estabelecer as melhores condições de expressão, a colônia 1 foi previamente crescida em ágar YPD com 100 µg/mL de zeocina, e uma colônia isolada foi cultivada em 125 mL de meio GYC em *erlenmayer* de 500mL e em 250 mL de meio GYC em *erlenmayer* de 1L, seguindo as mesmas proporções da Tabela 09. A indução foi realizada conforme descrito anteriormente na seção 4.2.3.2.2, ajustando a troca de meio para DO = 1 a 2, e a adição diária de metanol para a concentração final de 3% (v/v) por mais três dias.

Após o período de indução, os sobrenadantes foram separados por centrifugação e congelados a -20 °C até a etapa de purificação.

5.2.7.5 Purificação da lectina recombinante

A purificação da lectina rADEL seguiu a metodologia detalhada anteriormente na seção 4.2.3.2.3. Após a centrifugação da cultura, o sobrenadante foi submetido a uma precipitação com sulfato de amônio a 70%, o precipitado foi ressuspenso em tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 20 mM e centrifugado novamente a 9000 x g, 30 minutos a 4 °C.

O precipitado foi aplicado em uma coluna HisTrap™ FF de 1mL, e as proteínas retidas a coluna foram eluídas com 175 mM ou 270 mM de imidazol em tampão Tris 20 mM pH 8, com NaCl 150 mM e glicerol 5%. A cromatografia foi realizada em um sistema AKTA™ pure (GE).

A pureza das frações contendo a lectina foi avaliada por SDS-PAGE e a funcionabilidade foi confirmada por atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina a 3% (Sampaio *et al.*, 2002). As frações contendo a proteína não clivada e funcional foram reunidas para uso posterior.

5.2.8 Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil)

Os testes de caracterização foram realizados com a proteína não clivada, dialisada em tampão Tris 20 mM pH 7,6, NaCl 150 mM e Glicerol 5%.

5.2.8.1 Estimativa da massa molecular e avaliação da pureza da lectina

A estimativa da massa molecular e o grau de pureza da lectina foram avaliados por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE), corado com Coomassie Brilliant Blue (Laemmli, 1970). O gel de 15% foi submetido a uma corrente elétrica de 25 mA e tensão de 150 V durante 1 h. O grau de pureza da lectina nativa usada nos testes bioquímicos também foi avaliada.

As alíquotas de amostras líquidas foram diluídas em tampão de amostra 5x contendo Tris pH 6,8 (0,06M), glicerol (25%), SDS (2%) e azul de bromofenol (1mg), para concentração final de 1%. As alíquotas de concentrado celular foram ressuspensas no mesmo tampão de amostra 1x.

A massa molecular foi estimada na ausência e na presença de β mercaptoetanol, comparando a migração da proteína com a migração de marcadores moleculares. Os marcadores moleculares usados foram um marcador preparado com as proteínas conhecidas: HSA – 66 kDa (NP_0000468); rSFL - 30 kDa (7UMJ_AA; Chaves *et al.*, 2023) e CCL – 15 kDa (C0HLX7.1; Marques *et al.*, 2018), e o marcador comercial SigmaMarker™ com o intervalo de 6,5 kDa a 66 kDa.

5.2.8.2 Ensaio de atividade hemaglutinante da proteína recombinante

O ensaio de hemaglutinação foi realizado com a proteína recombinante contra eritrócitos de coelho na forma nativa e na forma tratada com a enzima proteolítica tripsina, conforme a preferência da lectina em estudo anteriormente (Carneiro *et al.*, 2017b).

Os eritrócitos de coelho foram lavados, tratados e preparados em uma suspensão a 3%, conforme descrito na seção 3.2.4.

5.2.8.3 Ensaio de inibição da atividade hemaglutinante com a proteína recombinante

A proteína recombinante rADEL foi submetida a um ensaio de inibição da atividade hemaglutinante contra eritrócitos de coelho tratado com tripsina, e a concentração mínima inibitória do açúcar capaz de inibir a aglutinação de eritrócitos também foi determinada.

Para a inibição da atividade hemaglutinante foram utilizados carboidratos e glicoproteínas, seguindo a metodologia descrita por Sampaio *et al.*, (2002). Após determinar os

carboidratos específicos, a concentração mínima inibitória do açúcar capaz de inibir a aglutinação de eritrócitos (CMI – Concentração Mínima Inibitória) foi determinada (Sampaio *et al.*, 2002).

Os carboidratos e glicoproteínas utilizados foram: fenil- β -D-galactopiranosídeo (0,1M), metil- α -D-galactopiranosídeo (0,1M), 4-nitrofenil- α -D-galactose (0,1M), metil- β -D-galactopiranosídeo (0,1M), 2- nitrofenil- β -D-galactose (0,1M), α -metil-manosídeo (0,1M), α -lactose monoidratada (0,1M), L-(-)fucose (0,1M), L-(+)arabinose (0,1M), D-melibiose (0,1M), L-ramnose (0,1M), D-lactulose (0,1M), D-(+)-galactose, D-(+)-ácido galacturônico monoidratado, D-(+)-manose (0,1M), D-(+)-cloridrato de glucosamina (0,1M), mucina do estômago de porco tipo II (1mg/mL), mucina do estômago de porco tipo III (1 mg/mL), metil- β -D-tiogalactosídeo (0,1M), *N*-acetil-D-glucosamina (0,1M), D-(+)-cloridrato de galactosamina (0,1M), D-(+)-glucose (0,1M), D-(+)-rafinose pentaidratada (0,1M), D-(-)-frutose (0,1M), β -lactose (0,1M), D-(-)-ribose (0,1M), sucrose (0,1M), D-(+)-xilose (0,1M), D-(+)-maltose monoidratada (0,1M), fenil- β -D-galactose (0,1M), *N*-acetil-D-galactosamina (0,1M), *N*-acetil-D-manosamina (0,1M), fetuína de soro fetal bovino (1mg/mL), manana de *Saccharomyces cerevisiae* (1mg/mL), mucina submaxilar bovina tipo 1-S (1mg/mL), transferrina humana (1 mg/mL), asialofetuína de soro de bezerro fetal (1mg/mL), tiroglobulina da glândula tireoide do porco (1 mg/mL), ácido tecóico *Staphylococcus aureus* (1 mg/mL), LPS (lipopolissacarídeo) *E. coli* (1 mg/mL) e 4 – nitrofenil- β -D-galactopiranosídeo (0,1M).

5.2.8.4 Dicroísmo circular

A estrutura secundária da lectina recombinante foi analisada por espectroscopia de dicroísmo circular (CD) utilizando um espectropolarímetro Jasco J-815 (Jasco International Co., Tokyo, Japan), equipado com sistema de controle de temperatura Peltier.

As análises foram conduzidas com a proteína a 180 μ g/mL em tampão Tris 20 mM pH 7,6, NaCl 150 mM e glicerol 5% a 20 °C, utilizando uma cubeta de quartzo com caminho óptico de 0,1 cm. Quatro varreduras independentes foram realizadas no intervalo do UV distante (200–250 nm). Os espectros obtidos foram processados e a predição da estrutura secundária foi realizada pela plataforma BestSel (Kardos *et al.*, 2025; Micsonai *et al.* 2022).

5.2.8.5 Espalhamento dinâmico de luz - DLS (*Dynamic Light Scattering*)

O comportamento hidrodinâmico da lectina recombinante foi avaliado pela análise de DLS em duas concentrações proteicas, a fim de identificar os parâmetros que minimizam a formação de agregados.

As amostras foram analisadas em tampão Tris 20 mM pH 7,6, NaCl 150 mM e glicerol 5%, nas concentrações de 0,250 e 0,125 mg/mL. As amostras foram centrifugadas a $10.000 \times g$ por 20 minutos a 4 °C para remoção de agregados e partículas, e em seguida foram transferidas para cubetas descartáveis de poliestireno. As análises de DLS foram realizadas a 20 °C em um equipamento *Zetasizer Advance Ultra* (Malvern Instruments, Worcestershire, UK).

O diâmetro hidrodinâmico da lectina foi calculado a partir das funções de autocorrelação de intensidade, utilizando o modelo de análise fornecido pelo fabricante. Cada condição foi analisada em triplicata, e cada medida representa a média de múltiplas varreduras.

5.3 Resultados e discussão

5.3.1 Produção da proteína recombinante rADEL em sistema heterólogo de *E. coli* (Brasil)

5.3.1.1 Desenho do gene sintético

O gene sintético da lectina rADEL, com 666 pb, incluindo os sítios de restrição, foi inicialmente clonado no vetor pET21a, delimitado pelos sítios de restrição *NdeI* e *XhoI* e sem a otimização de códons raros para *E. coli*. O quadro aberto de leitura (ORF – *Open Reading Frame*) possui 654 pb, codificando uma proteína de 218 aminoácidos. Devido a construção do vetor pET21a e ao gene sintético contendo os sítios de restrição *NdeI* e *XhoI*, com a inserção do códon de parada após a sequência da lectina, a proteína é expressa sem nenhuma cauda fusionada, resultando na expressão apenas da lectina de interesse com peso molecular previsto de 24.546,7 Da. A construção da proteína recombinante expressa está representada na Figura 31.

Figura 31 – Representação esquemática da proteína recombinante expressa rADEL.

>rADEL

```

      10      20      30      40      50      60
MDPDKCKTIR VESWSYKYAE KVEDASYVL NMTVVDRQSA AACTLGESFG YQKATLWVDH

      70      80      90     100     110     120
GCRADFKVCY LPVMPTECQT LRVESWNYKY AEKVVGEAAL FINMTVEDRQ SEASCDLDKS

     130     140     150     160     170     180
FGFYNQNSTV WVNHGCRADF NICYLKGAVT TSTINVSSWN YQYATKVLPA ASCIYSMRVV

     190     200     210     218
NQQSAAPCTL GTTYGFVANT MWVDDGCRAD FKPSYYSP

```

Número de aminoácidos: 218

Peso molecular: 24.546,7043

pI: 5,44

Coeficiente de extinção:

O coeficiente de extinção é: 55350

Em 1mg/ml (40,739 uM) a solução da proteína tem Abs280nm de: 2,25

Se todas as cisteínas formarem pontes dissulfeto, o coeficiente de extinção é: 55975

Em 1mg/ml (40,739 uM) a solução de proteína tem Abs280nm de: 2,28

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

O vetor pET21a é um plasmídeo de expressão usado para produção de proteínas recombinantes em *E. coli*. O diferencial dessa construção é uma menor complexidade, mas sem perder a eficiência pois possui o mesmo promotor T7 do sistema pET. Esse plasmídeo possui uma cauda de histidina no C-terminal que só é expressa se o gene inserido não tiver códon de parada. O pET21a permite expressar proteínas sem adicionar sequências extras no N-terminal e possibilita uma purificação simples por His-tag (Mierendorf *et al.*, 2000; Studier, 1998).

O índice de adaptação de códons (CAI – *Codon Adaptation Index*) da sequência não otimizada foi calculado para *E. coli* pela *GenScript*. O valor foi de 0,83, indicando boa adaptação ao sistema hospedeiro, pois valores acima de 0,8 são considerados bons, e quanto mais perto de 1,0, melhor a expressão naquele sistema. Por esse motivo, a sequência permaneceu não otimizada.

5.3.1.2 Testes de expressão

A cepa *E. coli* DH5 α foi utilizada como cepa de clonagem para replicação do vetor. As cepas de expressão *E. coli* BL21(DE3), *E. coli* Rosetta-gami, *E. coli* Origami B (DE3) plysS e *E. coli* ArcticExpress foram usadas nos testes de expressão da proteína recombinante.

Cada cepa de expressão tem sua individualidade e foram aprimoradas para possuírem diferentes vantagens entre si. Por isso, é interessante a adoção de hospedeiros alternativos a estirpe BL21 para melhorar o desempenho de expressão.

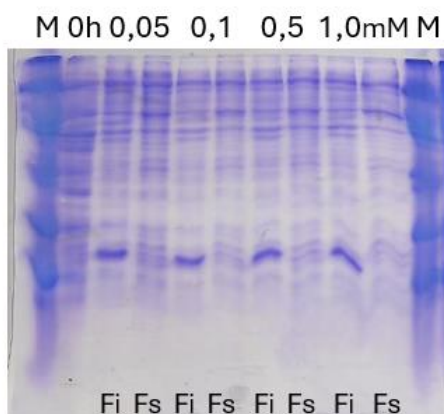
Após o tempo de indução em cada teste, as células foram recuperadas, lisadas e as frações solúveis e insolúveis foram analisadas por SDS-PAGE (Figura 32).

A comparação entre as bandas observadas na fração antes da indução (0h) com as bandas das frações insolúveis e solúveis, revelou que em nenhum dos testes de temperatura (37°C por 4 h, 30°C por 16h, 25 °C por 20h, 18°C por 24 horas) e concentração de IPTG (0,05 mM; 0,1 mM; 0,5 mM e 1,0 mM), foi detectada uma banda que indicasse a expressão de rADEL.

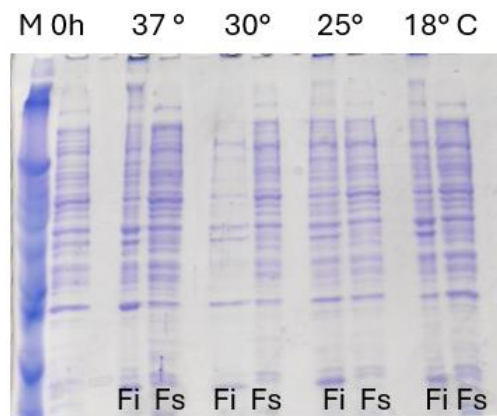
Figura 32 – Testes de expressão da proteína rADEL no vetor pET21a.

***E. coli* BL21 (DE3) rADELpET21a**

A Teste de [IPTG]

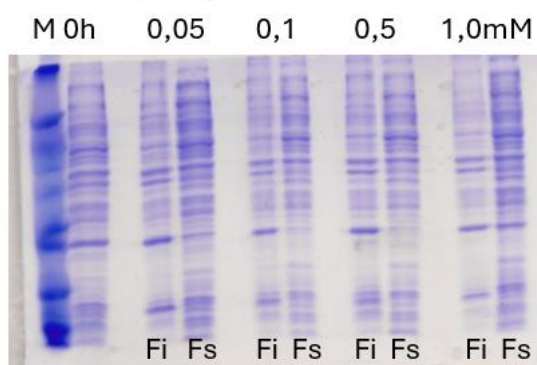


B Teste de Temperatura com 0,05 mM IPTG

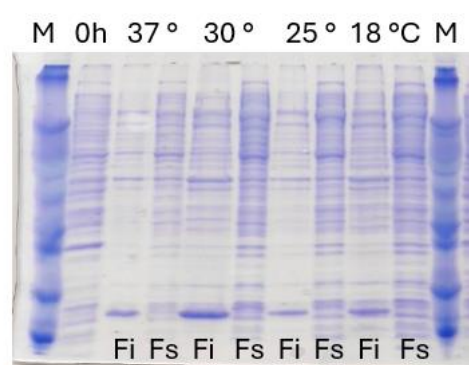


***E. coli* Rosetta-gami rADELpET21a**

C Teste de [IPTG]

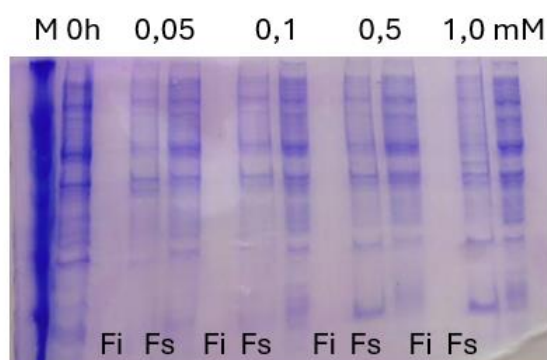


D Teste de Temperatura com 0,05 mM IPTG



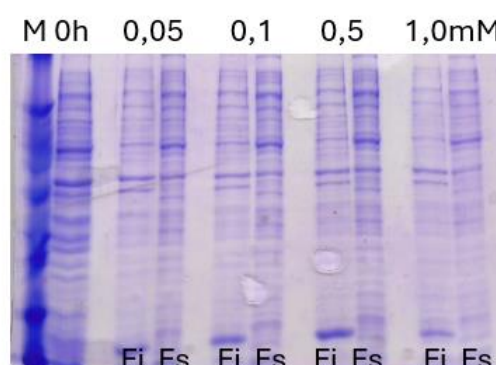
***E. coli* Origami B(DE3)pLysS rADELpET21a**

E Teste de [IPTG]



***E. coli* ArcticExpress rADELpET21a**

F Teste de [IPTG]



Fonte: Autora. Géis de eletroforese SDS-PAGE. M: Marcador molecular. 0h: alíquota do *pellet* celular retirada antes da indução. Fi: Fração insolúvel. Fs: Fração solúvel. **A:** Teste de concentração de IPTG na cepa de *E. coli* BL21(DE3) a 25 °C por 20h. **B:** Teste de temperatura de com 0,05 mM de IPTG na cepa de *E. coli* BL21(DE3). **C:** Teste de concentração de IPTG na cepa de *E. coli* Rosetta-gami a 25 °C por 20h. **D:** Teste de temperatura de com 0,05 mM de IPTG na cepa de *E. coli* Rosetta-gami. **E:** Teste de concentração de IPTG na cepa de *E. coli* Origami(DE3)pLysS a 25 °C por 20h. **F:** Teste de concentração de IPTG na cepa de *E. coli* ArcticExpress a 25 °C por 20h.

A única exceção foi observada no teste de concentração do indutor na cepa *E. coli* BL21, onde uma banda na fração insolúvel, ausente na fração 0h, foi observada em todas as concentrações testadas. As bandas foram digeridas com tripsina e analisadas por espectrometria de massas, porém os peptídeos identificados não corresponderam à sequência de aminoácidos da proteína de interesse. Portanto, a expressão da lectina na fração insolúvel não foi confirmada, indicando que a construção contendo rADEL clonada no vetor pET21, sem cauda de fusão, não resultou na expressão da proteína recombinante.

A proteína de interesse ADEL possui cinco ligações dissulfeto intracadeias e uma intercadeia e é uma proteína glicosilada, com três glicanos ligados à sua estrutura (Carneiro *et al.*, 2017b). Portanto, devido as características intrínsecas da própria proteína, rADEL é uma lectina naturalmente desafiadora para ser expressa em *E. coli*.

As ligações dissulfeto estruturais podem se formar em proteínas no citoplasma quando o gene para a enzima tioredoxina redutase é inativado por mutação, o que acontece com as cepas modificadas para ter um ambiente citoplasmático oxidativo (Singh *et al.*, 2015). A expressão de rADEL foi testada nessas cepas geneticamente modificadas, *E. coli* Rosetta-gami e *E. coli* Origami, porém não foi o suficiente para produzir a proteína nem mesmo na fração insolúvel.

A ausência de expressão da lectina recombinante rADEL em *E. coli* sem uma proteína de solubilidade fusionada era um resultado esperado. Considerando as características estruturais da proteína, que apresenta várias pontes dissulfeto e glicosilação, sua expressão necessita de um ambiente oxidante e vias de processamento pós-traducionais ausentes no citoplasma bacteriano.

Para melhorar o processamento em *E. coli*, algumas medidas poderiam ser adotadas com base nas particularidades estruturais da proteína, como a co-expressão de chaperonas, a adição de uma proteína de fusão para ajudar na solubilidade, a otimização de códons ou a secreção de proteínas no espaço periplasmático (Singh *et al.*, 2015). Alternativamente o uso de sistemas eucariotos capazes de realizar glicosilação poderiam ser explorados para viabilizar a produção recombinante dessa lectina.

5.3.2 Produção da proteína recombinante rADEL em sistema heterólogo (França)

5.3.2.1 Subclonagens do gene sintético para os vetores de expressão

O gene sintético da lectina rADEL, com 666 pb incluindo os sítios de restrição, foi inicialmente clonado no vetor *backbone* pEX-A128, delimitado pelos sítios de restrição *NdeI* e *Sall* e com a sequência gênica otimizada para *Komagataella pastoris*. O quadro aberto de leitura (ORF) possui 654 pb, codificando uma proteína de 218 aminoácidos. Após a digestão, o inserto foi subclonado nos vetores de expressão pPICZ α HisTEV, para expressão na levedura *Komagataella pastoris*, e nos vetores pET25b(+) e pCOLD-HAT-TEV-3G-Nco, para expressão em *E. coli*. A confirmação da subclonagem foi obtida por PCR e digestão enzimática, de acordo com o tamanho esperado para rADEL e os respectivos vetores, visualizado em gel de agarose. Todas as construções apresentaram o inserto no tamanho esperado, e a subclonagem em todos os vetores foi confirmada por sequenciamento (dados não mostrados).

A escolha dos vetores foi baseada em estratégias levando em consideração as características inerentes da proteína, as pontes dissulfeto e a glicosilação. O vetor pPICZ α HISTEV é usado para a expressão de proteínas em leveduras, principalmente em *K. pastoris*. O vetor tem *tags* de histidina para purificação e sítio de clivagem para a protease TEV.

No vetor pET25b(+), a construção com os sítios de restrição *NdeI* e *Sall* permite a expressão da proteína sem nenhuma cauda fusionada. Essa estratégia visa obter uma proteína igual a nativa sem alterações estruturais, permitindo a purificação com o mesmo protocolo da lectina nativa.

No vetor pCOLD-HAT-TEV-3G-Nco, a proteína de interesse é fusionada a uma cauda pequena, composta apenas pelo TEE, cauda de purificação com histidina e o sítio de clivagem da protease TEV. Essa estratégia minimiza alterações estruturais da proteína.

5.3.2.2 Produção da proteína rADEL em sistema heterólogo de *Komagataella pastoris*

Como os padrões de glicosilação podem afetar fortemente a atividade funcional das proteínas, a possibilidade de obter os glicanos semelhantes ao da lectina nativa é importante para que a atividade biológica da proteína seja preservada (Marco, 2009). Por esse motivo, pensando em como a glicosilação pode afetar a funcionabilidade da lectina, e na ausência completa de expressão de rADEL em *E. coli* no primeiro teste, a principal estratégia usada para essa lectina foi a expressão na levedura *K. pastoris*. E para isso, o gene sintético de rADEL foi otimizado para a substituição dos códons raros nesse organismo.

A levedura *Saccharomyces cerevisiae* já foi muito utilizada como plataforma para expressão de proteínas devido às suas múltiplas vantagens, como expressão secretora, presença

de modificações pós-traducionais eucarióticas e glicosilação. No entanto, ela apresenta desvantagens como a hiperglicosilação de proteínas, a secreção de proteínas endógenas dificultando a purificação, o modo fermentativo de respiração, resultando em um nível de crescimento e expressão moderados, e portanto, rendimento menores (Juturu; Wu, 2018).

Por outro lado, a levedura *Komagataella pastoris* ou, anteriormente chamada de *Pichia pastoris*, surgiu como sistema heterólogo de expressão demonstrando muitas vantagens pois usa um modo de respiração aeróbica que lhe permite atingir uma densidade celular mais alta com maiores níveis de expressão, secreta menos proteínas endógenas e principalmente, adiciona uma menor glicosilação, com menos cadeias de manose, o que resulta em uma proteína recombinante mais próxima da nativa (Cereghino; Gregg, 2000; Juturu; Wu, 2018).

Existem diferentes vetores usados para a expressão em *K. pastoris*. A proteína produzida nesse sistema hospedeiro pode ser expressa dentro da célula ou ser secretada para o meio extracelular. O recomendado é que se a proteína de interesse necessite de modificações pós-traducionais, como a glicosilação, ela deve ser secretada para o meio extracelular (Invitrogen, Thermo Fisher Scientific, 2020).

Na expressão heteróloga em *K. pastoris*, a glicosilação de proteínas acontece durante o tráfego intracelular pela via secretória, ou seja, ela precisa ser secretada para que a glicosilação aconteça. Esse processo se inicia no retículo endoplasmático (RE) e é processado no complexo de Golgi. Quando a proteína começa a ser traduzida, se tiver um peptídeo sinal para secreção, a cadeia polipeptídica entra no RE e sofre o dobramento inicial, com a formação das pontes dissulfeto e a N-glicosilação inicial, adicionando o oligossacarídeo de unidade central típico de leveduras ($\text{Man}_{8-9}\text{GlcNAc}_2$), no resíduo de asparagina do motivo de N-glicosilação. Depois a proteína segue para o complexo de Golgi, onde as cadeias dos glicanos são alongadas, inclusive em *K. pastoris* é comum que a glicosilação seja rica em manose, e a O-glicosilação também pode ocorrer. Em seguida, depois de passar pelo RE e pelo complexo de Golgi, a proteína é finalmente secretada para o meio extracelular. Portanto, sem peptídeo sinal, a proteína permanece no citosol e não é glicosilada (Hamilton; Gerngross, 2007).

Dentre as várias opções de vetores para expressão em leveduras, o vetor pPICZ α reúne várias vantagens. Esse vetor é projetado para a clonagem, amplificação e validação do inserto em *E. coli* e para expressão heteróloga em *K. pastoris*. Ele possui um promotor forte e induzível por metanol, o AOX1, com a sequência do fator α que permite o direcionamento da proteína recombinante para a via secretória, onde ocorre o dobramento correto e as modificações pós-traducionais. Além disso, a seleção por zeocina permite uma seleção eficiente

tanto em *E. coli* quanto em *K. pastoris*. Esse vetor é altamente recomendado por possuir a sequência do fator α que há relatos de maior sucesso comparado a outros peptídeos sinais. No sistema pPICZ α , o sinal de secreção fator α ainda é composto pelo sítio de clivagem para a protease Kex2, uma enzima do complexo de Golgi que cliva o peptídeo sinal α , liberando a proteína recombinante madura durante a secreção (Cereghino; Gregg, 2000; Daly; Hearn, 2005).

5.3.2.2.1 Transformação do plasmídeo rADELpPICZ α HisTEV em *K. pastoris*

Após a confirmação da subclonagem por sequenciamento, o DNA plasmidial de rADELpPICZ α HisTEV foi extraído de *E. coli* NEB5 α . O DNA plasmidial purificado foi linearizado com a enzima *SacI* (Biolabs) e o inserto foi purificado.

Na transformação em levedura, a linearização do vetor é um passo essencial para promover a integração genômica e eficiente da construção de expressão, pois diferentemente das bactérias, *K. pastoris* não mantém estáveis os plasmídeos livres dentro das células, o DNA exógeno precisa se integrar ao genoma do organismo para permitir a expressão.

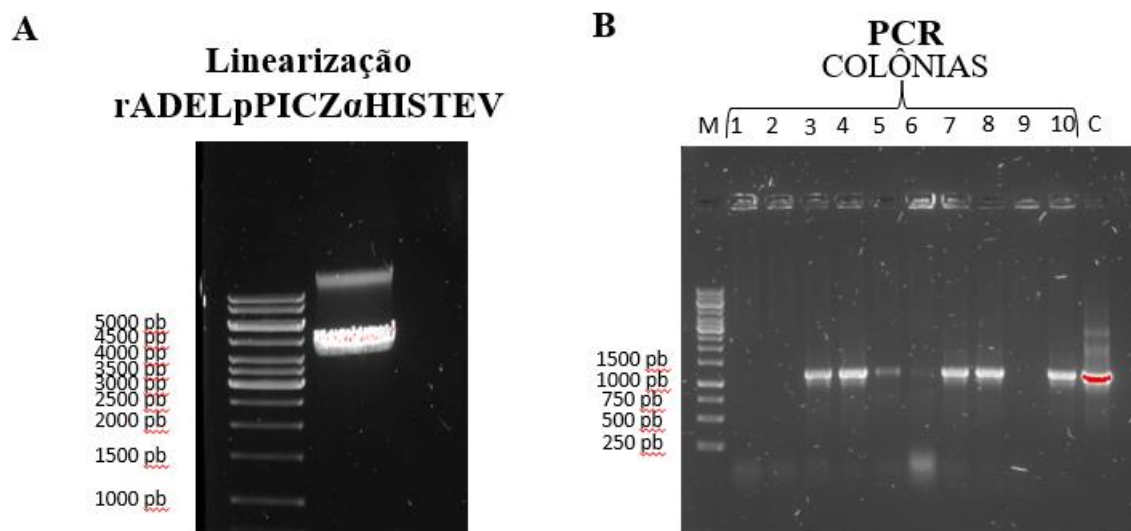
Em *K. pastoris*, os vetores de expressão são integrados ao genoma do organismo visando maximizar a estabilidade das linhagens de expressão. Essa integração pode ocorrer de duas maneiras, uma por *crossing-over* simples e a outra por substituição gênica no locus AOX1. No *crossing-over* simples, o vetor é linearizado dentro do gene marcador (HIS4) ou dentro da região promotora de AOX1 e a extremidade livre estimula eventos de recombinação homóloga que resulta na integração ao genoma, essa é a forma mais comum. Na substituição gênica, o vetor é digerido de modo que a construção de expressão fique flanqueada na região 5' e 3' do gene AOX1, ocasionando a deleção completa do gene que é substituído pelo vetor. Portanto, a linearização do vetor expõe extremidades que favorecem a recombinação homóloga entre as sequências do vetor e do genoma da levedura, promovendo a estabilidade genética do clone transformado e permitindo a expressão eficiente (Juturu; Wu, 2018; Invitrogen, Thermo Fisher Scientific, 2020).

A levedura *Komagataella pastoris* da linhagem GS115 foi transformada com o plasmídeo por eletroporação e a triagem inicial das células transformadas foi realizada com a seleção por zeocina.

A confirmação da transformação foi feita por PCR em colônia e os produtos da PCR foram visualizados por eletroforese em gel de agarose. Os géis de agarose referentes à

linearização do vetor e à confirmação por PCR da transformação estão apresentados na Figura 33.

Figura 33 – Géis de agarose referentes à linearização do vetor e confirmação da transformação em *K. pastoris*.



Fonte: Autora. A – Linearização do vetor rADELpPICZαHISTEV com a enzima *SacI*. B – PCR de confirmação da transformação em *K. pastoris*. M: Marcador comercial O'GeneRuler 1kb DNA Ladder (*Thermo Scientific*). C: Vetor rADELpPICZαHISTEV usado como controle positivo da PCR. Os números 1 a 10 correspondem a colônia selecionada.

A linearização confirmou o plasmídeo do tamanho esperado, rADEL com 666 pb somado ao vetor pPICZαHISTEV de 3.671 pb totalizando 4.337 pb, corroborando com a banda observada próximo a 4.500 pb.

A eletroporação foi eficiente com a PCR confirmando as colônias 3, 4, 5, 7, 8 e 10, com uma banda entre 1500 pb e 1000 pb do marcador, o que corrobora com o tamanho esperado para o fragmento amplificado de 667 pb para o vetor vazio somado aos 666 pb de rADEL, totalizando 1333 pb. O controle positivo usado foi o plasmídeo pPICZαHISTEV, resultando na mesma banda condizente com o tamanho esperado. As colônias 3, 4, 7 e 8 foram selecionadas para os testes de expressão.

5.3.2.2.2 Expressão da lectina rADEL em *K. pastoris*

A expressão da lectina seguiu as recomendações do *Pichia Expression Kit: user guide*, *Invitrogen* para o fenótipo Mut⁺ com a secreção da proteína (*Invitrogen*, *Thermo Fisher*

Scientific, 2020; Invitrogen, 2010). Portanto, testes de expressão avaliando diferentes parâmetros não foram realizados.

Após a confirmação da transformação do vetor em *K. pastoris* por PCR, quatro clones positivos foram selecionados com o antibiótico de seleção zeocina e cultivados em um menor volume para um *screening* com o objetivo de avaliar a melhor expressão da proteína.

A construção de rADEL no pPICZ α HISTEV origina uma proteína fusionada composta apenas pela cauda de histidina, o sítio de clivagem para TEV e a lectina de interesse (Figura 34). O sinal de secreção fator α presente no vetor possui um sítio de clivagem específico para a protease Kex2, resultando na remoção do pró-peptídeo no complexo de Golgi. O vetor ainda possui o sinal de clivagem para a protease Stel3 responsável pela remoção de possíveis resíduos de aminoácidos resultantes da clivagem com Kex2. A ação em conjunto dessas duas proteases garante a obtenção de uma proteína recombinante com o N-terminal corretamente processado, composto apenas pela *tag* de purificação, sítio de clivagem e a proteína.

Apenas a cauda fusionada à rADEL, sem a proteína de interesse, possui 20 aminoácidos com um peso molecular de 2.418,53 Da e com pI de 6,36. A proteína inteira expressa, com a lectina de interesse, possui 238 aminoácidos com um peso molecular de 26.934,14 Da e com pI de 5,67.

A proteína expressa foi idealizada para ser clivada com a protease TEV. Após clivagem, rADEL possui 220 aminoácidos com peso molecular de 24.757,84 Da e com pI de 5,2. A Figura 35 mostra a representação da proteína de interesse rADEL após clivagem com a protease TEV.

Figura 34 – Representação esquemática da proteína recombinante expressa no vetor pPICZαHISTEV

>rADEL expressa em pPICZαHISTEV

```

      10      20      30      40      50
HHHHHHSSGV DLENLYFQSHH MDPDKCDTIRR VESWSYKYAEE KVVEDASYVLL
      60      70      80      90     100
NMTVVDRQSAA AACTLGESFGG YQKATLWVDHH GCRADFKVCYY LPVMPTECQTT
      110     120     130     140     150
LRVESWNYKYY AEKVVEGAALL FINMTVEDRQQ SEASCDLDKSS FGFYNQNSTVV
      160     170     180     190     200
WVNHGCRADFF NICYLKGAVTT TSTINVSSWNN YQYATKVLPAA ASCIYSMRVV
      210     220     230     238
NQQSAAPCTLL GTTYGFVANTT MWVDDGCRADD FKPSYYSPP

```

Cauda de histidina

Número de aminoácidos: 238

Peso molecular: 26934,1356

pI: 5,67

Sítio de clivagem TEV

Lectina

Coefficiente de extinção:

O coeficiente de extinção é: 56840

Em 1mg/ml (37,128 uM) a solução da proteína tem

Abs280nm de: 2,11

Se todas as cisteínas formarem pontes dissulfeto,
o coeficiente de extinção é: 57465

Em 1mg/ml (37,128 uM) a solução de proteína tem

Abs280nm de: 2,13

Fonte: Autora. pI: Ponto isoeletrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Figura 35 – Representação da proteína rADEL obtida após clivagem com a protease TEV.

>rADEL clivada

```

      10      20      30      40      50
SHMDPKCDTT IRVESWSYKYY AEKVVEDASYY VLNMTVVDRQQ SAAACTLGESS
      60      70      80      90     100
FGYQKATLWV DHGCRADFKVV CYLPVMPTECC QTLRVESWNYY KYAEKVVEGAA
      110     120     130     140     150
ALFINMTVEDD RQSEASCDLDD KSFGFYNNQNSS TVVWVNHGCRA DFNICYLKGAA
      160     170     180     190     200
VTTSTINVSSS WNYQYATKVLL PAASCIYSMRR VVNQQSAAPCC TLGTTYGFVAA
      210     220
NTMWVDDGCRR ADFKPSYYSPP

```

Sítio TEV residual

Número de aminoácidos: 220

Peso molecular: 24757,8381

pI: 5,20

Linker residual

Lectina

Coefficiente de extinção:

O coeficiente de extinção é: 55350

Em 1mg/ml (40,391 uM) a solução da proteína tem

Abs280nm de: 2,24

Se todas as cisteínas formarem pontes dissulfeto,
o coeficiente de extinção é: 55975

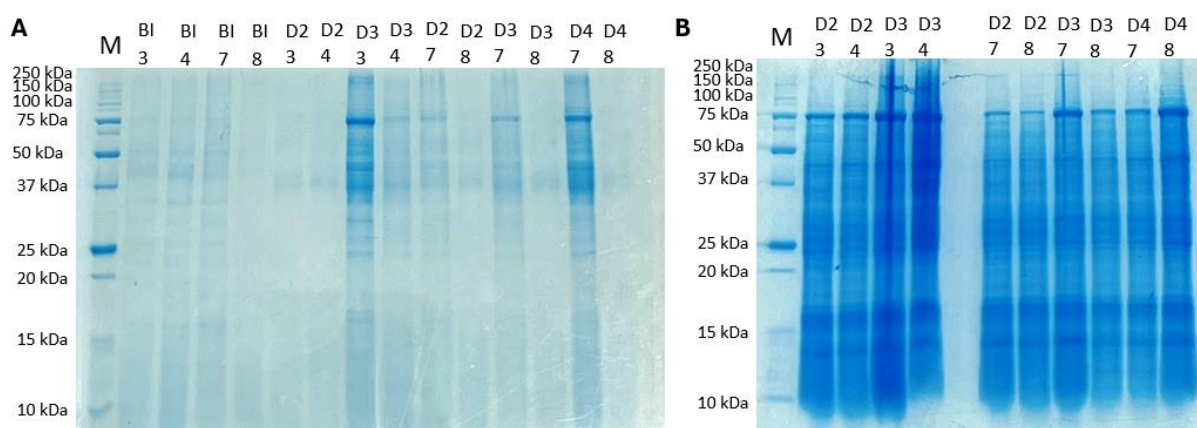
Em 1mg/ml (40,391 uM) a solução de proteína tem

Abs280nm de: 2,26

Fonte: Autora. pI: Ponto isoeletrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

A expressão de rADELpPICZ α HISTEV foi então avaliada experimentalmente. A cultura das quatro colônias iniciou no mesmo dia. Em todos os dias de indução, alíquotas do sobrenadante e do *pellet* foram retiradas e analisadas em SDS-PAGE. Alíquotas retiradas antes da indução foram usadas como o controle negativo de expressão (Figura 36).

Figura 36 - Análise dos sobrenadantes e *pellets* da expressão de rADELpPICZ α HISTEV em SDS-PAGE.

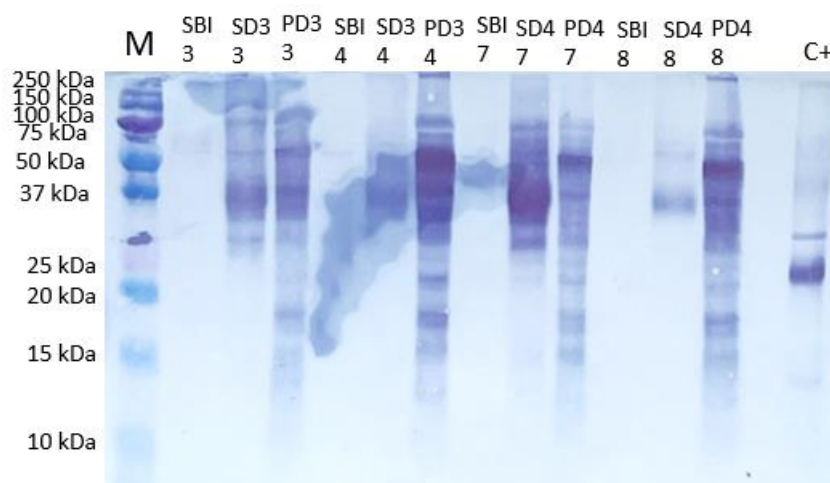


Fonte: Autora. Gel de 12%. A: Análise do sobrenadantes. B: Análise dos *pellets*. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. BI: Alíquota retirada antes da indução. D2: segundo dia de indução, D3: terceiro dia de indução, D4: quarto dia de indução. Números 3,4,7 e 8: colônias selecionadas para o *screening*.

O gel dos sobrenadantes (Figura 36-A) revelou uma banda fraca e difusa na altura do marcador de 37 kDa em todas as colônias e em todos os dias de expressão. Esse tamanho observado é próximo do tamanho esperado para rADELpPICZ α HISTEV de aproximadamente 27 kDa e a mesma banda não está presente nas frações antes da indução de todas as colônias, indicando a presença da proteína sendo expressa e sendo secretada para o meio extracelular. Já para o gel do concentrado celular (Figura 36-B), não foi possível inferir nada por conta da alta concentração de proteínas.

Devido aos resultados pouco conclusivos no SDS-PAGE, uma análise de *immunoblotting* foi realizada com o objetivo de confirmar a presença específica da proteína expressa. A detecção foi realizada usando um anticorpo monoclonal anti-poli-histidina conjugado à peroxidase, que reconhece especificamente a cauda de histidina presente na proteína recombinante expressa. Quando o substrato quimioluminescente é adicionado, a peroxidase catalisa a oxidação do substrato resultando na emissão de luz visível na membrana de nitrocelulose, permitindo a visualização das proteínas com o anticorpo conjugado (Sule, Rivera, Gomes, 2023) (Figura 37).

Figura 37 - Análise de *immunoblotting* dos sobrenadantes e *pellets* da expressão de rADELpPICZ α HISTEV.



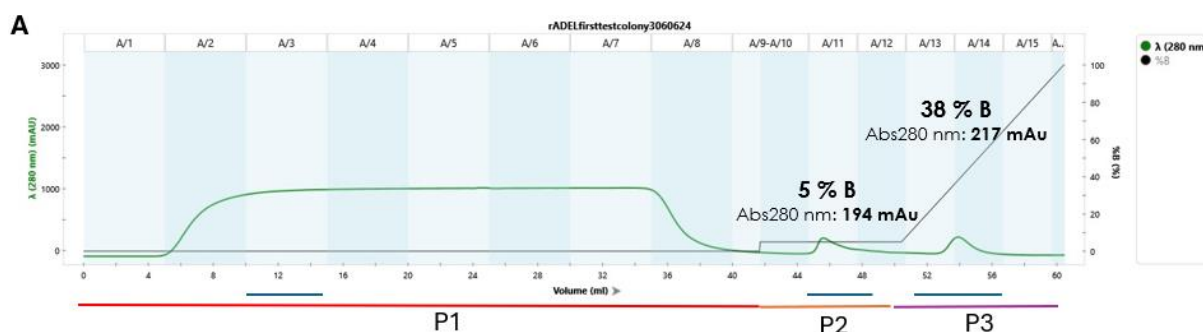
Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Prestained Dual Color Standards*, Bio rad. SBI: Aliquota sobrenadante retirada antes da indução. SD3: sobrenadante terceiro dia de indução, SD4: sobrenadante quarto dia de indução. PD3: *pellet* terceiro dia de indução, PD4: *pellet* quarto dia de indução. Números 3,4,7 e 8: colônias selecionadas para o *screening*. C+: proteína contendo cauda de histidina.

Por meio do *immunoblotting*, foi possível detectar a presença do sinal específico de reação tanto nos sobrenadantes quanto nos *pellets* em todos as colônias e dias de expressão analisados. O sinal observado correspondeu exatamente com a banda fraca e difusa próximo de 37 kDa no SDS-PAGE. As frações antes da indução não apresentaram sinal de reação em nenhuma das colônias avaliadas e o controle positivo reagiu conforme o esperado, mostrando que a reação foi eficiente. Como o tamanho da banda observada é próximo do tamanho esperado para rADELpPICZ α HISTEV, é possível concluir que a lectina de interesse foi expressa, sendo detectada tanto no meio extracelular quanto na fração celular.

Após o último dia de indução (D4), as culturas das quatro colônias foram centrifugadas separadamente e os sobrenadantes e os concentrados celulares foram separados. As frações dos sobrenadantes contendo a lectina recombinante foram precipitadas com sulfato de amônio a 70% e o precipitado foi ressuspenso em tampão A.

Como a proteína expressa possui cauda de histidina, as frações precipitadas e ressolubilizadas de cada colônia foram submetidas à cromatografia de afinidade a metais imobilizados, separadamente. A coluna foi lavada com tampão contendo imidazol 10 mM (P1) e as proteínas retidas a coluna foram eluídas com 5% do tampão com imidazol 500 mM (tampão B) (P2), e com um gradiente linear de 5% a 100% do mesmo tampão (P3) (Figura 38).

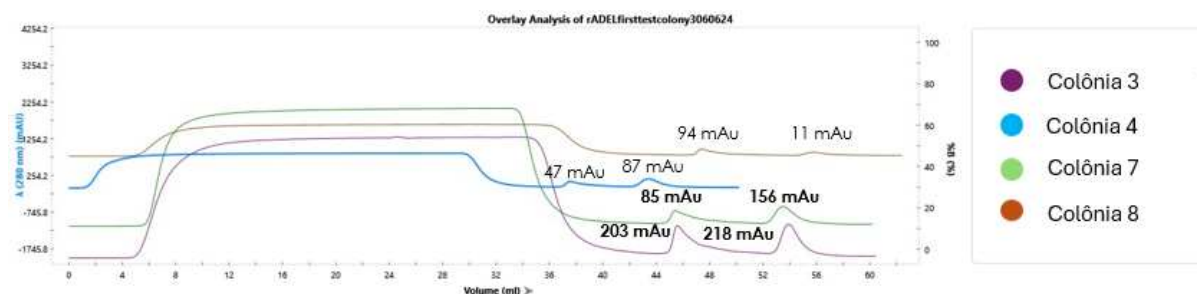
Figura 38 – Cromatograma da purificação de rADELpPICZ α HISTEV colônia 3.



Fonte: Autora. Cromatogramas fornecidos pelo sistema de cromatografia de média pressão NGC (Bio-Rad). A: Cromatograma de purificação da colônia 3. Cromatografia em coluna HisTrapTM FF de 1mL a um fluxo de 1 mL/min. P1: Pico 1 - proteínas lavadas com tampão Tris 20 mM pH 8, NaCl 500 mM e Imidazol 10 mM. P2: Pico 2 - proteínas eluídas com 5% de tampão Tris 20 mM pH 8, NaCl 500 mM e Imidazol 500 mM. Pico 3 - proteínas eluídas com gradiente linear de 5% a 100% de tampão Tris 20 mM pH 8, NaCl 500 mM e Imidazol 500 mM. Eixo inferior: volume em mL. Eixo esquerdo: absorbância a 280 nm em mAu. Eixo direito: tampão de eluição (B) em %. A linha contínua preta indica a % de tampão B (imidazol 500 mM). A linha contínua verde indica a absorbância a 280 nm (mAu). A letra “A” com os números marcados em azul indicam os tubos coletados. As setas indicam as informações de maior absorbância dos picos eluídos. Os traços azuis abaixo do pico indicam os tubos coletados e analisados em SDS-PAGE.

A cromatografia das quatro colônias seguiu o mesmo padrão observados na figura 38. A seleção da colônia com a melhor eficiência de expressão foi baseada na análise conjunta da intensidade das bandas correspondentes à proteína recombinante observadas em SDS-PAGE e na maior absorbância registrada no pico de eluição durante a cromatografia. Para isso, os cromatogramas obtidos das quatro colônias foram sobrepostos, permitindo a comparação direta dos perfis de eluição e a identificação da colônia que apresentou o maior pico da fração ligada a coluna, o que indica uma maior quantidade de proteína eluída, e consequentemente, uma maior quantidade de proteína expressa (Figura 39).

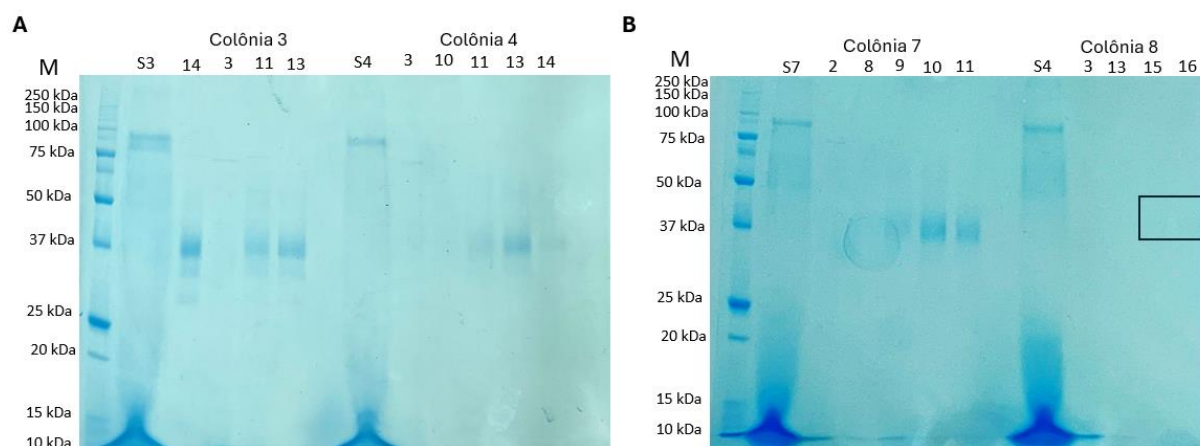
Figura 39 - Sobreposição dos cromatogramas do *screening* de melhor expressão.



Fonte: Autora. Cromatogramas fornecidos pelo sistema de cromatografia de média pressão NGC (Bio-Rad). Eixo inferior: volume em mL. Eixo esquerdo: absorbância a 280 nm em mAu. Eixo direito: tampão de eluição (B) em %. A linha contínua preta indica a % de tampão B (imidazol 500 mM). As linhas contínuas em roxo, azul, verde e laranja indicam os perfis de cromatografias das colônias 3, 4, 7 e 8, respectivamente. Os números em mAu indicam as informações de maior absorbância dos picos eluídos. Os maiores picos estão destacados em negrito.

O primeiro pico recuperado da coluna com 5% de tampão B teve a máxima absorbância(mAu) de 203, 47, 85 e 94 para as colônias 3, 4, 7 e 8, respectivamente. O segundo pico recuperado foi eluído em torno de 26% a 29% de tampão B, tendo a absorbância máxima de 218 mAu, 87 mAu, 156 mAu e 11 mAu para colônias 3, 4, 7 e 8, respectivamente. As máximas absorbâncias do pico foram eluídas em torno 35% de B. As frações foram coletadas e analisadas em SDS-PAGE (Figura 40).

Figura 40 - Análise da purificação de rADELpPIC α HISTEV em SDS-PAGE.



Fonte: Autora. Gel de 12%. A: Análise da purificação das colônias 3 e 4. B: Análise da purificação das colônias 7 e 8. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 2 e 3: fração não ligada a coluna. S: Sobrenadante após precipitação com sulfato de amônio, os números indicam as colônias.

As colônias que apresentaram os maiores picos de absorbância, indicando uma maior quantidade da proteína recombinante eluída, foram as colônias 3 e 7, em concordância com as maiores intensidades das bandas correspondentes observadas em SDS-PAGE. A expressão menos eficiente foi da colônia 8 que apresentou o menor pico de eluição e pouca visualização das bandas correspondentes no gel.

As frações contendo a proteína de interesse foram reunidas e a quantidade de proteína recuperada de cada expressão foi mensurada. Os rendimentos em 50 mL de cultura foram de 0,3 mg, 0,156 mg, 2,1 mg e 0,06 mg para as colônias 3, 4, 7 e 8, respectivamente. Portanto, a colônia 7 apresentou a melhor eficiência de produção da proteína e foi usada para a expressão em maior escala.

A análise em SDS-PAGE também confirma a ligação na coluna e a recuperação da proteína de interesse, observada pela banda difusa próxima a 37 kDa nos dois picos eluídos. Essa banda está ausente no sobrenadante depois da precipitação, demonstrando que esse passo de purificação foi eficiente para recuperar a proteína de interesse. Os tubos 2 e 3 oriundos da

fração que não se ligou a coluna, não apresentaram a proteína de interesse, portanto o sulfato de amônio residual da amostra não atrapalhou a ligação da proteína na coluna.

O tamanho esperado para rADELpPICZ α HISTEV é de aproximadamente 27 kDa, porém ADEL é uma proteína glicosilada. A proteína nativa possui uma massa molecular para a sequência de aminoácidos de 24.450 Da, possuindo quatro sítios de *N*-glicosilação, onde três contêm glicanos ligados: ³⁰NMT, ¹⁰²NMT e ¹²⁶NST (Figura 41) (Carneiro *et al.*, 2017).

Figura 41 – Locais de *N*-glicosilação de ADEL.

>ADEL

DPDKCKTIRV	ESWSYKYAEK	VVEDASYVLN	MTVVDRQSA	ACTLGESFGY
QKATLWVDHG	CRADFKVCYL	PVMPTECQTL	RVESWNYKYA	EKVVEGAALF
INMTVEDRQS	EASCDLDKSF	GFYNQNSTVW	VNHGCRADF	ICYLKGAVTT
STINVSSWN	QYATKVLPA	SCIYSMRVN	QQAAPCTLG	TTYGFVANTM
WVDDGCRADF	KPSYYSP			

Locais de *N*-glicosilação

Motivo não glicosilado

Fonte: Autora. Adaptado de Carneiro *et al.*, 2017. Os números indicam os aminoácidos.

O primeiro sítio contém um oligossacarídeo de 1702,61 Da com uma massa correspondente ao núcleo com o pentassacarídeo Man₃GlcNAc₂ de 892 Da com a adição de cinco resíduos de hexose totalizando 810 Da (162 Da cada hexose). No segundo sítio, a massa molecular do oligossacarídeo varia de 892,36 a 1216,45 Da, correspondendo ao pentassacarídeo nuclear isolado e ao pentassacarídeo acrescido de dois resíduos de hexose totalizando 324 Da, respectivamente. O terceiro sítio apresenta um oligossacarídeo de 1054 Da que consiste no núcleo e um resíduo de hexose, totalizando uma proteína de 28 kDa (24.450 + 1702,61 + 892,36 + 1054 Da) (Carneiro *et al.*, 2017).

A expressão heteróloga em leveduras possui uma característica bastante documentada, a hiperglicosilação. Em todos os eucariotos, a *N*-glicosilação começa no retículo endoplasmático com a transferência de uma unidade do oligossacarídeo precursor Glc₃Man₉GlcNAc₂ para a asparagina presente na sequência do reconhecimento Asn-X-Ser/Thr. Após a remoção das unidades de glicose, e o processamento inicial no retículo endoplasmático,

esse oligossacarídeo é convertido em $\text{Man}_8\text{GlcNAc}_2$. A partir dessa etapa, os padrões de glicosilação de eucariotos basais como *K. pastoris* e outros fungos começam a diferir dos outros eucariotos na maturação dos *N*-glicanos. Em mamíferos, o complexo de Golgi remove resíduos de manose, resultando na formação do pentassacarídeo nuclear $\text{Man}_3\text{GlcNAc}_2$, e a partir dele são sintetizados os tipos de *N*-glicanos, complexos, híbridos e alto teor de manose. No entanto, em *K. pastoris*, o processamento no complexo de Golgi é limitado, com a predominância de *N*-glicanos do tipo alta manose, mantendo o núcleo da *N*-glicosilação em $\text{Man}_{8-9}\text{GlcNAc}_2$, sem a conversão para $\text{Man}_3\text{GlcNAc}_2$. Esse processamento resulta em proteínas expressas glicosiladas com uma maior quantidade de manose (Cereghino; Gregg, 2000).

S. cerevisiae alonga as unidades centrais ligadas a *N* no complexo de Golgi pela adição de cadeias extensas de manose. Essas cadeias geralmente possuem de 50 a 150 resíduos de manose, uma condição denominada hiperglicosilação. Assim, as proteínas heterólogas secretadas apresentam tamanhos heterogêneos da proteína nativa. Tanto *S. cerevisiae* quanto *K. pastoris* apresentam uma maioria de *N*-glicosilação do tipo alto manose. Porém, o comprimento das cadeias de oligossacarídeos adicionadas pós-traducionalmente às proteínas em *K. pastoris* é muito menor do que em *S. cerevisiae*, com uma média de apenas 8 a 14 resíduos de manose por cadeia lateral (Invitrogen, Thermo Fisher Scientific, 2020).

Porém, essa condição de hiperglicosilação pode variar em *K. pastoris* de acordo com a proteína expressa. Algumas proteínas exógenas secretadas em *K. pastoris* parecem ser hiperglicosiladas de forma semelhante às observadas em *S. Cerevisiae*. Outros relatos já demonstram que as proteínas expressas nesse sistema podem não sofrer hiperglicosilação, variando em cada caso (Daly; Hearn, 2005; Juturu; Wu, 2018; Invitrogen, Thermo Fisher Scientific, 2020).

Dessa forma, é realmente esperado que a proteína expressa rADEL, que já possui glicosilação em condições nativas, seja submetida a um grau ainda maior de glicosilação quando expressa em *K. pastoris*, resultando em uma massa estimada maior do que o esperado de aproximadamente 27 kDa.

Carneiro e colaboradores (2017) relataram que os oligossacarídeos ligados à cadeia polipeptídica da ADEL correspondem ao tipo alto-manose. Como naturalmente o processamento no complexo de Golgi é limitado em *K. pastoris* e o núcleo pentassacarídico comum $\text{Man}_3\text{GlcNAc}_2$ é aumentado com resíduos de manose e adicionalmente promovem o alongamento das cadeias laterais com mais resíduos de manose, é compreensível que uma

lectina naturalmente rica em glicosilação do tipo alta manose apresente um aumento ainda maior desse tipo de modificação pós-traducional no sistema heterólogo.

Portanto, mesmo que a proteína recombinante rADEL fosse processada com um padrão de glicosilação exatamente igual a nativa, seria esperado um peso molecular de aproximadamente 31 kDa. No entanto, como o sistema hospedeiro de *K. pastoris* tem a capacidade de aumentar essa glicosilação do tipo alta manose, é aceitável que a proteína expressa apresente um peso molecular aparente significativamente maior, como observado pela presença de uma banda próxima a 37 kDa.

A presença de bandas difusas, um padrão típico observado para proteínas glicosiladas corrobora com a ideia de que o peso molecular observado para rADEL é resultado da hiperglicosilação. Em ambas as frações eluídas, a proteína expressa aparentou pureza, sendo o arrastado observado típico do comportamento de proteínas muito glicosiladas, descrito como *smear* que reflete a heterogeneidade na composição da glicosilação presente na proteína expressa. Isso acontece porque os carboidratos ligados à proteína não se comportam da mesma forma que a cadeia polipeptídica em SDS, resultando em uma migração diferente no gel (Yang *et al.*, 2015).

Para avaliar a funcionabilidade da proteína, as frações eluídas da coluna HisTrapTM contendo a proteína de interesse foram concentradas até 1 mg/mL e submetidas a um teste de atividade hemaglutinante contra eritrócitos de coelho. A atividade hemaglutinante foi detectada, indicando que a proteína de interesse foi expressa ativa e funcional (dados não mostrados). As frações contendo a proteína pura e ativa foram reunidas para uso posterior.

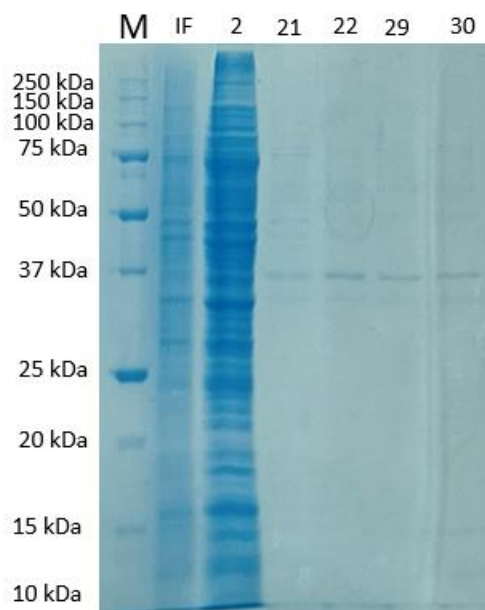
5.3.2.2.3 Purificação da lectina recombinante rADEL a partir do concentrado celular

O *immunoblotting* revelou a presença da proteína recombinante rADEL na fração celular, ou seja, indicando que uma parte da proteína expressa não estava sendo secretada para o meio extracelular. Isso acontece porque a proteína pode iniciar o dobramento no citoplasma, impedindo o reconhecimento da sequência sinal e consequentemente, inibindo a translocação pela via secretória (Damasceno; Huang; Batt, 2012).

Por isso, com o objetivo de avaliar se a extração citoplasmática era viável, e otimizar o rendimento da produção, as células recuperadas após a centrifugação da cultura induzida foram lisadas e alíquotas da fração solúvel e da fração insolúvel foram analisadas por SDS-PAGE.

A fração solúvel foi submetida a uma cromatografia de afinidade a metais imobilizados, utilizando o mesmo protocolo empregado para a purificação da proteína secretada. As frações foram coletadas e analisadas em SDS-PAGE (Figura 42).

Figura 42 - Análise da purificação citoplasmática de rADELpPIC α HISTEV.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. IF: Fração insolúvel. Tubos 2: fração não ligada a coluna. Tubos 21 e 22: fração eluída com 5% de tampão B. Tubos: 29 e 30: Fração eluída com gradiente linear de tampão B.

A análise em SDS-PAGE confirmou a ligação na coluna e a recuperação da proteína de interesse, evidenciada pela presença da banda próxima a 37 kDa nos dois picos de eluição. Essa banda apresentou a mesma mobilidade no gel daquela observada durante a purificação da proteína secretada, indicando que são a mesma proteína. Além disso, essa banda não está visível na fração insolúvel, demonstrando que a proteína expressa estava presente apenas na fração solúvel.

Para avaliar a funcionabilidade da proteína recuperada da extração citoplasmática, as frações eluídas da coluna contendo a proteína de interesse foram submetidas a um teste de atividade hemaglutinante contra eritrócitos de coelho. A atividade hemaglutinante não foi detectada, indicando que a proteína não secretada, mesmo que solúvel, não é funcional (dados não mostrados).

A capacidade de rADEL purificada da fração citoplasmática de se ligar à coluna de níquel imobilizado mesmo não estando funcional, acontece porque a interação nesse tipo de cromatografia de afinidade ocorre exclusivamente entre a sequência de histidina com o níquel carregado na coluna. A interação é independente da capacidade de ligação a carboidratos da lectina. Inclusive, estudos mostram que proteínas recombinantes com cauda de histidina são capazes de se ligar a colunas de níquel imobilizado mesmo com a proteína completamente desnaturada, mostrando que essa interação não significa funcionabilidade (Li; Su; Janson, 2004; Young; Britton; Robinson, 2012).

Como discutido anteriormente, para uma proteína ser glicosilada, ela precisa ser secretada para transitar pela via secretória, pois as etapas da glicosilação acontecem no retículo endoplasmático e no complexo de Golgi. Dessa forma, a proteína recombinante que não foi secretada e permaneceu na célula, possivelmente não passou pela glicosilação. Esse fato corrobora com o perfil da banda observada na purificação citoplasmática que não possuiu o padrão de bandas difusas típico de glicosilação, como foi observada na purificação de rADEL secretada no meio extracelular (Hamilton; Gerngross, 2007).

A ausência da atividade hemaglutinante na proteína purificada da extração citoplasmática reforça a ideia de que a lectina rADEL precisa da glicosilação para manter uma conformação funcional. Esse resultado demonstra que de fato essa modificação pós-traducional é importante para a estabilidade estrutural e atividade biológica de proteínas naturalmente glicosiladas quando expressas de forma recombinante (Marco, 2009).

Além disso, o trânsito pela via secretória também expõe a proteína a um ambiente oxidativo no retículo endoplasmático que permite a formação das pontes dissulfeto da lectina de interesse. Esse ambiente também possui a presença das proteínas sulfeto isomerases que são capazes de corrigir as pontes dissulfeto mal dobradas, todos esses fatores contribuem para o correto dobramento dessa proteína, resultando na funcionabilidade da lectina recombinante (Hamilton; Gerngross, 2007).

Esses resultados também reforçam que obter uma proteína recombinante solúvel não é garantia de atividade biológica mantida. A proteína pode adotar uma conformação estável para ficar solúvel, mas a arquitetura do sítio ativo ainda não é a correta para ter função. No caso de rADEL, os dados indicam que a glicosilação é um requisito essencial para a atividade funcional da lectina (Rosano; Ciccarelli, 2014).

Outros estudos também mostraram que proteínas expressas em *K. pastoris*, podem ser detectadas na célula e serem solúveis no citoplasma. No entanto, a atividade funcional de

proteínas expressas no citoplasma, foi observada em proteínas que não dependiam de mudanças pós-traducionais, como a glicosilação e a formação de pontes dissulfeto. Assim, esses achados corroboram com a ideia de que as proteínas naturalmente glicosiladas e com ligações dissulfeto precisam da passagem pela via secretória para obter uma conformação estrutural correta e manter a função biológica (Zhang; Yang, 2006; Damasceno; Huang; Batt, 2012).

Muitas lectinas são expressas com sucesso em sistema heterólogo de *K. pastoris*. Porém, a maioria dos estudos são com lectinas de origem vegetal ou humana, enquanto relatos para invertebrados marinhos são pouquíssimos na literatura (Oliveira; Teixeira; Domingues, 2013; Martínez-Alarcón; Blanco-Labra, García-Gasca, 2018).

Kuo e colaboradores (2006) relataram a expressão de duas lectinas plasmáticas do caranguejo *Tachypleus tridentatus*, TPL-1 e TPL-2, em *K. pastoris*, onde a lectina foi expressa no vetor pPICZαA e secretada para o meio extracelular. As lectinas purificadas estavam funcionais e apresentaram atividade antibacteriana contra cepas Gram-positivas e Gram-negativas.

As duas lectinas são naturalmente glicosiladas e com pontes dissulfeto, para avaliar se essas modificações pós-traducionais são necessárias para a funcionabilidade das proteínas, os autores avaliaram a influência de uma mutação no sítio de glicosilação, produzindo uma proteína não glicosilada, e a influência de um agente redutor, rompendo as pontes dissulfeto. TPL-1 e TPL-2 perderam a atividade quando não glicosilada e na presença do agente redutor, respectivamente. Esses dados reforçam que as modificações pós-traducionais que esse sistema heterólogo é capaz de realizar são realmente importantes para a manutenção da atividade biológica da proteína recombinante (Kuo *et al.*, 2006).

Uma lectina do tipo C presente no caranguejo *Scylla paramamosain* foi expressa no mesmo sistema utilizando o vetor pPIC9K. A proteína secretada foi purificada e exibiu um amplo espectro de aglutinação bacteriana (Qiu *et al.*, 2021).

Outras proteínas oriundas de invertebrados marinhos foram expressas em *K. pastoris*. Uma enzima presente na *Aplysia californica* também foi expressa no vetor pPICZαA, exibindo atividades enzimáticas comparáveis às da proteína nativa (Munshi; Lee, 1997). Um peptídeo antibacteriano do camarão chinês *Fenneropenaeus chinensis* foi expresso no vetor pPIC9K de forma funcional com atividade antibacteriana e antifúngica (Li *et al.*, 2005). Outro peptídeo antimicrobiano isolado do soro de *Mytilus coruscus* também foi expresso no vetor pPICZαA, exibindo atividade biológica (Meng *et al.*, 2016).

5.3.2.3 Produção da proteína rADEL em sistema heterólogo de *E. coli*

Os plasmídeos rADELpET25b e rADELpCOLD-HAT-TEV-3G-Nco foram transformados na cepa de clonagem *E. coli* NEB5 α e nas cepas de expressão *E. coli* BL21(Star) e *E. coli* Shuffle B, respectivamente, após confirmação da subclonagem por sequenciamento. A transformação na cepa de clonagem objetiva a extração do DNA plasmidial e a escolha da cepa de expressão foi baseada na estrutura de rADEL com muitas pontes dissulfeto.

Apesar da proteína nativa ADEL ser glicosilada, estudos anteriores demonstraram que, em alguns casos, a atividade biológica pode ser mantida mesmo na ausência da glicosilação, desde que se a proteína naturalmente glicosilada adquira o dobramento correto (Kuo *et al.*, 2006). Como *E. coli* é o sistema heterólogo mais simples e mais utilizado, resolvemos avaliar a expressão de rADEL agora na cepa SHuffle B, pois essa cepa apresenta modificações genéticas que favorecem a formação de pontes dissulfeto no citoplasma. Assim, se ocorrer o dobramento adequado da proteína, mesmo na ausência da glicosilação, a atividade biológica talvez seja mantida.

Para expressão da proteína rADEL:pCOLD-HAT-TEV-3G-Nco em *E. coli* SHuffle B, a cultura foi crescida a 30 °C e para rADEL:pET25b(+) em *E. coli* BL21, a cultura foi crescida a 37 °C. Depois de atingir a densidade ótica de 0,4, a temperatura foi reduzida para 16 °C em *E. coli* BL21 e para 25 °C em *E. coli* SHuffle B. O cultivo prosseguiu até atingir DO = 0,8, quando o indutor IPTG foi adicionado na concentração de 0,2 mM e a expressão continuou nas respectivas temperaturas.

A temperatura de 16 °C em *E. coli* BL21 foi usada pois como essa cepa não possui modificações genéticas para obter um citoplasma menos redutor, a estratégia de diminuir a temperatura de indução para ajudar na solubilidade da proteína recombinante foi adotada (Kaur; Kumar; Kaur, 2018).

A temperatura de 25 °C em *E. coli* SHuffle B foi usada pois essa cepa possui um citoplasma menos redutor, que pode favorecer o dobramento correto da proteína, portanto uma nova temperatura de expressão poderia ser testada.

5.3.2.3.1 Expressão da lectina rADEL no vetor pET25b(+)

A construção da proteína recombinante expressa no vetor pET25b(+) origina uma proteína sem cauda de fusão, composta apenas pela lectina de interesse (Figura 43).

Figura 43 – Representação esquemática da proteína recombinante no vetor pET25b(+).

>rADEL expressa em pET25b(+)

```

      10      20      30      40      50
MDPDKCDTIR VESWSYKYAE KVVEDASYVL NMTVVDRQSA AACTLGESFG
      60      70      80      90     100
YQKATLWVDH GCRADFKVCY LPVMPTECQT LRVESWNYKY AEKVVEGAAL
      110     120     130     140     150
FINMTVEDRQ SEASCDLDKS FGFYNQNSTV WVNHGCRADF NICYLKGAVT
      160     170     180     190     200
TSTINVSSWN YQYATKVLPA ASCIYSMRVV NQQAAPCTL GTTYGFVANT
      210     218
MWVDDGCRAD FKPSYYSP

```

Lectina

Número de aminoácidos: 218
 Peso molecular: 24533,6188
 pI: 5,07

Coeficiente de extinção:

O coeficiente de extinção é: 55350
 Em 1mg/ml (40,760 uM) a solução da proteína tem
 Abs280nm de: 2,26

Se todas as cisteínas formarem pontes dissulfeto,
 o coeficiente de extinção é: 55975
 Em 1mg/ml (40,760 uM) a solução de proteína tem
 Abs280nm de: 2,28

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas rADEL possui 218 aminoácidos com um peso molecular de 24.533,62 Da e com pI de 5,07.

A expressão de rADELpET25b(+) foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE.

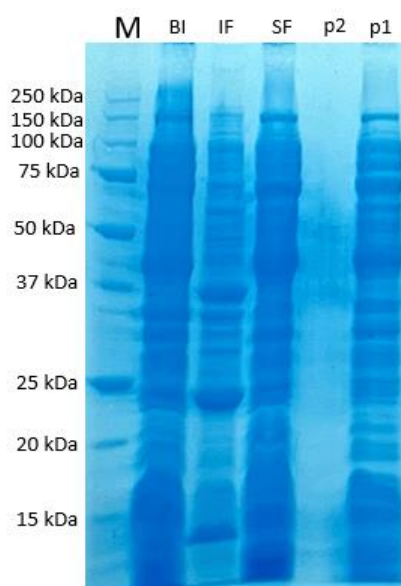
A fração solúvel foi aplicada em matriz de Sepharose 6B e as proteínas possivelmente retidas a coluna foram eluídas com 0,2 M de lactose, conforme o método usado para purificação da lectina nativa (Carneiro *et al.*, 2017b).

Nenhuma interação com a coluna foi observada, demonstrando que mesmo que houvesse expressão na fração solúvel, a proteína não estava ativa pois não houve ligação entre

o sítio de reconhecimento a carboidratos da lectina com os ligantes específicos presente na matriz cromatográfica.

Em seguida, as frações da cromatografia foram avaliadas em SDS-PAGE para confirmar a ausência de ligação na coluna (Figura 44).

Figura 44 - Análise da expressão de rADELpET25b(+) em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Bi: Aliquota retirada antes da indução. IF: Fração insolúvel. SF: Fração solúvel. P1: Fração não ligada à coluna com Tris 20 mM pH 7.5 e NaCl 150 mM. P2: Fração eluída da coluna com 0,2 M de lactose em TBS.

A proteína de interesse rADEL foi expressa apenas na fração insolúvel, detectada pela banda visível no gel próxima a 25 kDa, que não está presente na amostra coletada antes da indução, corroborando com o tamanho esperado para a proteína expressa de aproximadamente 24 kDa. A mesma banda não foi detectada na fração solúvel e nem na fração da cromatografia contendo as proteínas não ligadas a coluna. Como previsto, nenhuma banda foi observada na fração ligada a coluna, confirmando a não interação com a matriz cromatográfica.

A expressão de rADEL na fração insolúvel quando expressa em pET25b(+) é um resultado compreensível, visto que a construção não possui proteínas fusionadas à lectina que poderiam ajudar com a solubilidade no ambiente redutor do citoplasma bacteriano, além disso

a cepa utilizada foi *E. coli* BL21 que não possui as modificações genéticas para expressar no citoplasma proteínas ricas em pontes dissulfeto, como é o caso de rADEL.

Como discutido anteriormente, devido as características intrínsecas da própria proteína, é um desafio expressar rADEL em *E. coli*, sendo necessário adotar estratégias para ajudar na solubilidade, que a construção no vetor pET25b(+) não possui.

Curiosamente, a construção de rADEL no pET21 foi submetido a vários testes de expressão em diferentes cepas, inclusive em *E. coli* modificada geneticamente para ter um citoplasma menos redutor, e não foi observada expressão. No entanto, para a construção de rADEL no pET25b(+), a proteína foi expressa na fração insolúvel em *E. coli* BL21. Além disso, a expressão de rADEL no pET21 foi testada na mesma cepa em temperaturas próximas a 16 °C e com concentrações do indutor mais baixas que 0,2 mM e ainda assim não foi detectada nenhuma expressão.

Embora ambas as construções não contenham cauda de fusão, a diferença entre elas está na otimização da sequência codificante realizada para o sistema de expressão em *K. pastoris*. Apesar da preferência de códons ser diferente para *K. pastoris* e *E. coli*, essa otimização pode ter removido regiões ricas em códons raros para os dois sistemas e reduzido estruturas secundárias no RNAm, melhorando a eficiência da tradução e permitindo a expressão em *E. coli*, ainda que sem resultar no dobramento correto da proteína. Esse resultado indica que a otimização de códons raros realizadas para um sistema heterólogo pode, em alguns casos, favorecer a expressão em outro sistema hospedeiro, mesmo que não resulte na produção da proteína funcional.

5.3.2.3.2 Expressão da lectina rADEL no vetor pCOLD-HAT-TEV-3G-Nco

A construção da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco origina uma proteína de fusão composta pelo TEE, a cauda HAT rica em histidina, o sítio de clivagem para a protease TEV e a lectina de interesse (Figura 45).

Figura 45 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco.

>rADEL em pCOLD-HAT-TEV-3G-Nco

```

      10      20      30      40      50
MNHKVKDHLI HNVHKEEHAH AHNKGENLYF QGGGAMDPDK CDTIRVESWS
      60      70      80      90     100
YKYAEKVVED ASYVLNMTVV DRQSAAACTL GESFGYQKAT LWVDHGCRAD
     110     120     130     140     150
FKVCYLPVMP TECQTLRVES WNYKYAEKVV EGAALFINMT VEDRQSEASC
     160     170     180     190     200
DLDKSFGFYN QNSTVWVNHG CRADFNICYL KGAVTTSTIN VSSWNYQYAT
     210     220     230     240     250
KVLPAASCIY SMRVVNQQSA APCTLGTTYG FVANTMWVDD GCRADFKPSY
     253
YSP

```

Número de aminoácidos: 253

Peso molecular: 28514,0165

pI: 5,98

TEE

HAT-Tag

Sítio de clivagem TEV

Lectina

Coefficiente de extinção:

O coeficiente de extinção é: 56840

Em 1mg/ml (35,070 uM) a solução da proteína tem

Abs280nm de: 1,99

Se todas as cisteínas formarem pontes dissulfeto,
o coeficiente de extinção é: 57465

Em 1mg/ml (35,070 uM) a solução de proteína tem

Abs280nm de: 2,02

Fonte: Autora. pI: Ponto isoelétrico. TEE: *Translation Enhancing Element*. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas a cauda fusionada à rADEL, sem a proteína de interesse, possui 35 aminoácidos com um peso molecular de 3.998,41 Da e com pI de 7,58. A proteína inteira expressa, com a lectina de interesse, possui 253 aminoácidos com um peso molecular de 28.514,02 Da e com pI de 5,98.

A proteína expressa foi projetada para ser clivada com a protease TEV. Após clivagem, rADEL possui 222 aminoácidos com peso molecular de 24.775,85 Da e com pI de 5,07. A Figura 46 mostra a representação da proteína de interesse após clivagem com a protease TEV.

Figura 46 – Representação da proteína rADEL obtida após clivagem com a protease TEV.

>rADEL clivada

```

      10      20      30      40      50
GGGAMDPDKC DTIRVESWSY KYAEKVVEDA SYVLNMTVVD RQSAAACTLG
      60      70      80      90     100
ESFGYQKATL WVDHGCRADE KVCYLPVMPT ECQTLRVESW NYKYAEKVVE
     110     120     130     140     150
GAALFINMTV EDRQSEASCD LDKSFGFYNQ NSTVWVNHGC RADFNICYLK
     160     170     180     190     200
GAVTTSTINV SSWNYQYATK VLPAASCIYS MRVVNQQSAA PCTLGTTYGF
     210     220 222
VANTMWVDDG CRADFKPSY SP

```

Sítio TEV residual

Número de aminoácidos: 222

Linker residual

Peso molecular: 24775,8533

pI: 5,07

Lectina

Coefficiente de extinção:

O coeficiente de extinção é: 55350

Em 1mg/ml (40,362 uM) a solução da proteína tem

Abs280nm de: 2,23

Se todas as cisteínas formarem pontes dissulfeto,
o coeficiente de extinção é: 55975

Em 1mg/ml (40,362 uM) a solução de proteína tem

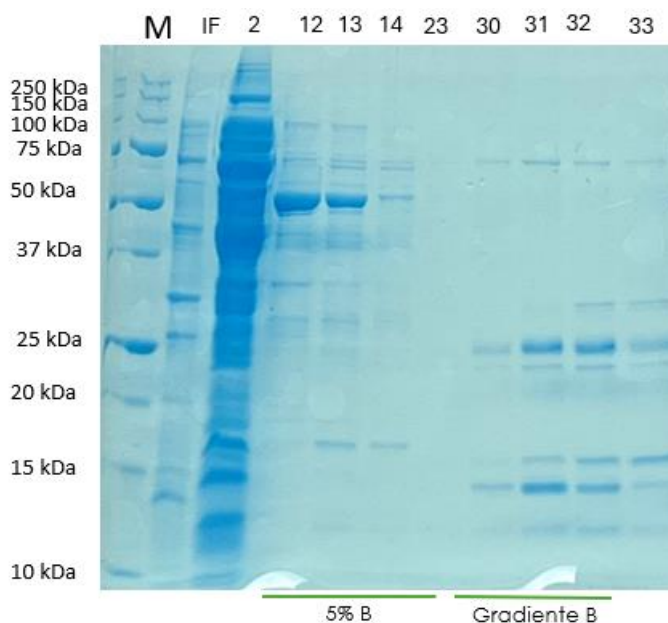
Abs280nm de: 2,26

Fonte: Autora. pI: Ponto isoelétrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

A expressão de rADELpCOLD-HAT-TEV-3G-Nco foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE.

Como o vetor pCOLD-HAT-TEV-3G-Nco possui cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados. As proteínas retidas a coluna foram eluídas com 5% do tampão B e com um gradiente linear de 5% a 100% do mesmo tampão. Os tubos referentes às frações da cromatografia foram coletados e analisados em SDS-PAGE (Figura 47).

Figura 47 - Análise da purificação de rADELpCOLD-HAT-TEV-3G-Nco em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubo 2: Fração não ligada a coluna. Tubos 12 a 23: eluídos com 5% de B. Tubos 30 a 33: eluídos com gradiente de tampão B. IF: Fração insolúvel.

A proteína de interesse expressa possui aproximadamente 29 kDa. No gel de SDS-PAGE, uma banda entre 25 kDa e 37 kDa foi detectada, indicando que houve expressão da proteína na fração insolúvel.

Durante a cromatografia, foi observada a ligação de proteínas na coluna, porém a análise das frações eluídas por eletroforese revelou a presença de muitas bandas de intensidade relevante, com pesos moleculares não condizentes com o esperado para rADEL. Esses resultados indicam a ocorrência de interações inespecíficas de proteínas endógenas com a coluna de níquel imobilizado (Young; Britton; Robinson, 2012).

Em ambos os vetores testados em *E. coli*, a expressão de rADEL ocorreu exclusivamente na fração insolúvel. Nem mesmo a utilização da cepa SHuffle, geneticamente modificada para favorecer a formação de pontes dissulfeto no citoplasma, foi capaz de promover o dobramento correto da proteína. Além disso, a presença do elemento de aumento da tradução (TEE) e o uso de cauda de fusão não foi suficiente para promover a solubilidade da lectina recombinante.

A única variável que pareceu influenciar a expressão, foi a otimização da sequência codificante, mesmo que para *K. pastoris*, pois para a sequência não otimizada não foi detectada

expressão, enquanto para a sequência otimizada houve a expressão na fração insolúvel em ambos os vetores, com ou sem cauda de fusão, e nas cepas *E. coli* BL21 e SHuffle B, que possuem o maquinário completamente diferente.

Apesar da solubilidade ser utilizado como um indicativo de dobramento correto, a análise comparativa das expressões entre os dois sistemas hospedeiros demonstrou que esse critério pode não ser suficiente para garantir a funcionabilidade da proteína recombinante. Na expressão em *K. pastoris*, rADEL foi detectada solúvel na fração celular, porém inativa, mostrando que para essa proteína, a glicosilação é necessária na manutenção da atividade biológica. Esse resultado sugere que, mesmo que rADEL fosse expressa solúvel em *E. coli*, a ausência da glicosilação provavelmente comprometeria sua atividade.

Portanto, os dados observados para rADEL revelam a importância de selecionar um sistema de expressão compatível com as características intrínsecas da proteína e com estratégias de expressão que se aproximem ao máximo das condições da proteína nativa.

No estudo do peptídeo antimicrobiano de *Mytilus coruscus* expresso em *K. pastoris*, os autores também relataram uma primeira tentativa de expressão em *E. coli*. Nesse caso, a proteína possuindo pontes dissulfeto, foi expressa solúvel, mas sem atividade. A funcionabilidade da proteína só foi mantida quando expressa na levedura, demonstrando que o dobramento correto da proteína só foi possível com as modificações pós-traducionais ocorridas em *K. pastoris* e que nem sempre solubilidade é sinônimo de funcionabilidade (Meng *et al.*, 2016).

Lectinas de moluscos aquáticos já foram expressas com sucesso em sistema heterólogo de *E. coli*, no entanto, nesses casos, as proteínas produzidas não eram glicosiladas. O sucesso da expressão com atividade funcional está intimamente relacionado às modificações pós-traducionais necessárias para cada proteína. Lectinas que não possuem pontes dissulfeto em suas estruturas foram expressas de forma solúvel e ativa nesse sistema heterólogo, como a MytiLec isolada do mexilhão mediterrâneo *Mytilus galloprovincialis* (Terada *et al.*, 2016), e a CgFREPI da *Crassostrea gigas* (Yang *et al.*, 2021).

Por outro lado, lectinas que possuem pontes dissulfeto foram expressas predominantemente na fração insolúvel e precisaram de etapas de redobramento para apresentarem funcionabilidade, como a lectina do abalone do Pacífico *Haliotis discus hannai* que possui três pontes dissulfeto (Choi *et al.*, 2022), a lectina do tipo C, AiCTL-7 da vieira *Argopecten irradians* com duas pontes dissulfeto (Kong *et al.*, 2011), a lectina ligante a ramnose da ostra *Pinctada fucata martensii*, possuindo quatro ligações dissulfeto (Zhang *et al.*,

2025) e CgCLec-2 da *Crassostrea gigas* com duas pontes dissulfeto (Li *et al.*, 2015). Quando o refolding não foi necessário, a lectina contendo pontes dissulfeto foi expressa utilizando estratégias que promovem solubilidade, como rCFlec-1 com duas pontes dissulfeto da viera *Chlamys farreri* que foi expressa fusionada a cauda de tiorredoxina (Wang *et al.*, 2007).

Todos esses exemplos reforçam a ideia que para obter uma proteína recombinante funcional, a escolha do sistema hospedeiro precisa ser baseada nas características estruturais e nas modificações pós-traducionais de cada proteína.

5.3.3 Ensaios de cristalização (França)

A lectina recombinante rADEL foi expressa com sucesso utilizando o vetor pPIC Z α HisTEV no sistema heterólogo de *K. pastoris*. Devido a obtenção da proteína recombinante, foram realizados ensaios de cristalização para obter sua estrutura tridimensional. A proteína não clivada rADELpPICZ α HisTEV foi submetida a diferentes condições de cristalização, variando a concentração da proteína (2,4 a 5 mg/mL) e o uso de diferentes kits de cristalização comerciais. A estratégia de cocristalização também foi abordada de acordo com a preferência a carboidratos da lectina (Carneiro *et al.*, 2017b). No entanto, apesar de diferentes condições testadas não foi possível obter cristais da proteína rADEL recombinante.

A dificuldade de cristalização pode estar associada à heterogeneidade estrutural da proteína devido a hiperglicosilação típica da expressão em *K. pastoris*. A presença de glicosilação é um fator que dificulta a cristalização devido a flexibilidade conformacional, que prejudica a organização regular das moléculas, essencial para a nucleação e crescimento ordenado dos cristais (Chang *et al.*, 2007).

5.3.4 Produção da proteína recombinante rADEL em sistema heterólogo de *K. pastoris* (Brasil)

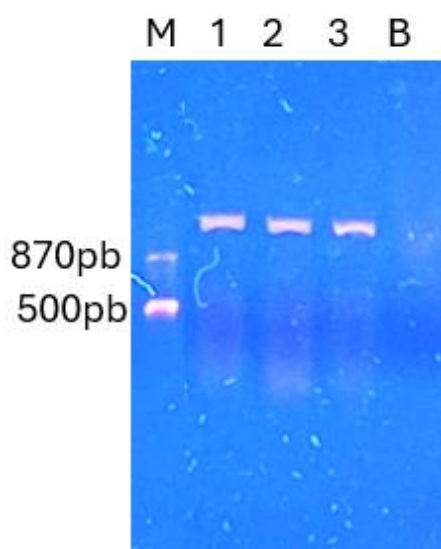
Em todas as estratégias avaliadas, a expressão da lectina rADEL em sistema de *K. pastoris* foi a única capaz de produzir a lectina biologicamente ativa, por esse motivo o plasmídeo rADELpPICZ α HisTEV foi a única construção gênica a ser trazida para o Brasil para dar continuidade aos estudos.

5.3.4.1 Transformação de *E. coli* com o gene sintético

Após a recuperação, o plasmídeo rADELpPICZ α HisTEV foi inserido na cepa de clonagem *Escherichia coli* DH5 α para amplificação e obtenção do DNA plasmidial.

A confirmação da transformação foi feita por PCR usando os *primers* para a região do promotor AOX1, com as colônias selecionadas pelo antibiótico zeocina (100 μ g/mL). Os produtos da reação foram visualizados por eletroforese em gel de agarose (Figura 48).

Figura 48 - Confirmação da transformação de rADELpPICZ α HisTEV em *E. coli*.



Fonte: Autora. Gel de agarose 1%. M: Marcador. Os números indicam as colônias selecionadas. B: Controle negativo da reação, sem o inserto de interesse.

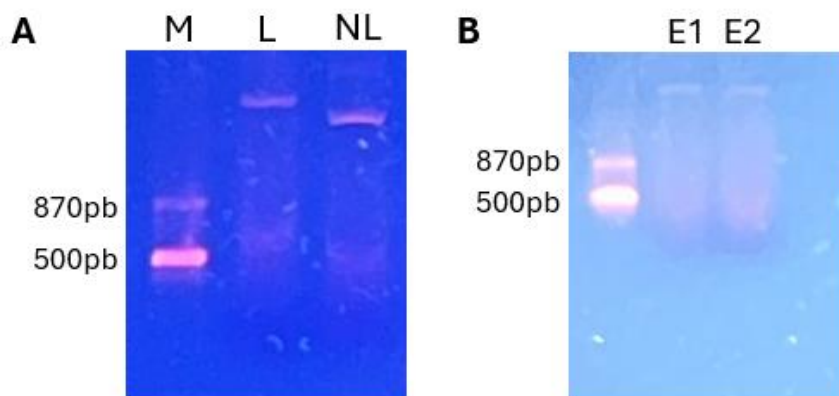
A PCR confirmou a transformação em *E. coli*, com um fragmento amplificado do tamanho esperado de 667 pb para o vetor vazio somado aos 666 pb de rADEL, totalizando 1333 pb, o que corrobora com a banda observada acima de 870 pb. O controle negativo, sem a amostra amplificada, validou a reação.

5.3.4.2 Transformação de *K. pastoris* com o gene sintético

O DNA plasmidial de rADELpPICZ α HisTEV da colônia 1 confirmada por PCR, foi extraído de *E. coli* DH5 α . O DNA plasmidial purificado foi linearizado com a enzima *SacI* e, em seguida, o inserto linearizado foi purificado.

Os géis de agarose referentes à linearização do vetor e a purificação do inserto linearizado estão apresentados na Figura 49.

Figura 49 - Géis de agarose referentes à linearização do vetor e a purificação do inserto



Fonte: Autora. Gel de agarose 1%. A: Gel de agarose referente à linearização do plasmídeo. B: Gel de agarose referente à purificação do plasmídeo linearizado. M: Marcador. L: vetor linearizado. NL: vetor não linearizado. E1: Primeira eluição. E2: Segunda eluição.

A comparação entre o vetor linearizado e o não linearizado em gel de agarose (Figura 49 – A) confirmou a linearização. O DNA linearizado apresentou uma menor mobilidade no gel quando comparado ao plasmídeo não linearizado, migrando mais lentamente. Isso acontece porque plasmídeos circulares (não linearizado) são mais compactos e migram mais rapidamente no gel do que o plasmídeo linear.

A purificação do vetor após a linearização também foi eficiente devido a visualização do plasmídeo no gel de agarose (Figura 49 – B). Para maximizar a recuperação do plasmídeo, durante a purificação foram feitas duas eluições, e as duas amostras foram usadas para a transformação.

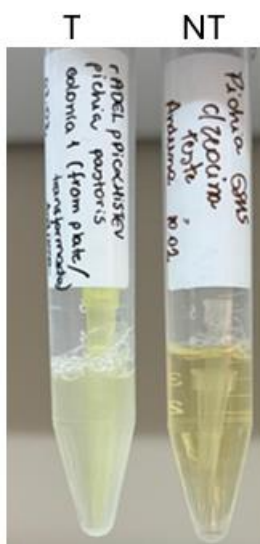
A levedura *Komagataella pastoris* também da linhagem GS115 foi transformada com o plasmídeo por eletroporação e a triagem inicial das células transformadas foi realizada com a seleção por zeocina a 200 µg/mL.

Algumas colônias selecionadas pelo antibiótico zeocina foram crescidas e submetidas a uma tentativa de confirmação da transformação por PCR em colônia. Como controle positivo da reação, foi usado o DNA plasmidial de rADELpPICZαHISTEV proveniente da França, previamente confirmado por sequenciamento e usado na primeira transformação realizada no Brasil, em *E. coli* DH5α, que teve confirmação por PCR. No

entanto, nem o controle positivo e nem as colônias crescidas em zeocina foram amplificadas em PCR.

Como não foi possível confirmar por PCR em colônia, alguns clones foram selecionados e replaqueados em YPDS ágar com 200 µg/mL de zeocina, para uma maior seleção das células transformadas. A colônia 1 foi selecionada e crescida novamente em 200 µg/mL de zeocina, e após o crescimento, essa cultura foi estocada e utilizada para as expressões subsequentes (Figura 50).

Figura 50 - Comparação do crescimento em zeocina das células transformadas e não transformadas.



Fonte: Autora. Crescimento em YPD com 200 µg/mL de zeocina. T: colônia 1 transformada com o plasmídeo rADELpPICαHISTEV, com resistência para zeocina. NT: Células de *K. pastoris* GS115 não transformadas com o plasmídeo, sem resistência para zeocina.

Em muitos trabalhos é relatado a dificuldade da execução de uma PCR em colônia com *K. pastoris* devido a estrutura espessa da parede celular desse organismo, limitando a liberação do DNA genômico em quantidades suficientes para a amplificação. Por isso, muitos protocolos recomendam pré-tratamentos para romper a parede celular com o uso de enzimas ou tratamentos térmicos antes da PCR (Invitrogen, Thermo Fisher Scientific, 2020; Harju *et al.*, 2004; Looke *et al.*, 2011; Linder; Schliwa; Kube-Grandenrath, 1996).

Diferentes tentativas de extração do DNA genômico da colônia 1 selecionada para expressão foram realizadas com o objetivo de confirmação do plasmídeo por PCR. Tratamentos com diferentes detergentes, tampões e choque térmico foram usados, mas não foi possível extrair o DNA genômico da célula. Um método bastante protocolado é o uso de enzimas para romper a parede celular. No entanto, essa abordagem não foi testada no presente trabalho (Invitrogen, Thermo Fisher Scientific, 2020; Harju *et al.*, 2004; Looke *et al.*, 2011; Linder; Schliwa; Kube-Grandenrath, 1996).

Apesar da confirmação da transformação por PCR em colônia não ter sido obtida, esse resultado não gera dúvidas em relação à ocorrência da transformação. A colônia selecionada para expressão apresentou crescimento em concentrações elevadas de zeocina, um critério amplamente utilizado para seleção de clones transformados. O manual do sistema de expressão em *K. pastoris* descreve que transformantes podem ser selecionados com base na hiper-resistência à zeocina, sendo associado à integração estável do vetor no genoma da levedura (Invitrogen, Thermo Fisher Scientific, 2020). Além disso, a colônia foi submetida a sucessivas etapas de replaqueamento e crescimento em quantidades elevadas de zeocina, sob pressão seletiva crescente, favorecendo a eliminação de células não transformadas e a manutenção do fenótipo resistente.

Como a linearização do vetor seguiu o mesmo protocolo realizado na França, é esperado o mesmo fenótipo para a cepa GS115 transformada, o Mut⁺, resultante de uma recombinação do tipo *crossing-over*. Novamente, a confirmação do fenótipo das cepas transformadas não foi realizada, levando em consideração apenas o embasamento teórico.

5.3.4.3 Expressão da lectina rADEL em *K. pastoris*

A recombinação homóloga em *K. pastoris* pode afetar a eficiência da expressão pela ocorrência dos eventos de integração múltipla. Na França, um *screening* inicial foi realizado com quatro clones transformados para avaliar os níveis de expressão.

Entretanto, a seleção por zeocina é dose-dependente, clones que crescem em maiores concentrações do antibiótico tendem a ter mais cópias integradas, e consequentemente, um maior potencial de expressão (Invitrogen, Thermo Fisher Scientific, 2020).

Como a hiper-resistência à zeocina é um indicativo de conter múltiplas cópias do cassete de expressão, não foi realizado um novo *screening* após a nova transformação, pois o clone transformado e selecionado para expressão foi crescido em doses elevadas do antibiótico

e sofreu sucessivas etapas de replaqueamento e crescimento, indicando a presença de múltiplas cópias integradas.

A expressão da lectina seguiu o mesmo protocolo do manual para o fenótipo Mut⁺ com a secreção da proteína (Invitrogen, Thermo Fisher Scientific, 2020; Invitrogen, 2010). Porém, na França, foi observado a diminuição do rendimento da proteína recombinante durante a produção em larga escala, por esse motivo, diferentes parâmetros foram analisados com base no *Pichia Expression Kit* (Invitrogen, Thermo Fisher Scientific, 2020), para avaliar as melhores condições de expressão e otimizar a produção.

A temperatura e a rotação são parâmetros muito bem estabelecidos em inúmeros estudos de produção de proteínas recombinantes em larga escala nesse sistema heterólogo, por isso, essas condições não foram alteradas.

O manual relata a possibilidade de se iniciar as culturas a partir de células frescas plaqueadas, ou a partir de um estoque congelado em glicerol. Esse protocolo também sugere a análise da expressão em dias diferentes: 1 dia, 2 dias, 3 dias e até 4 dias, visando ajustar o melhor tempo de expressão (Invitrogen, Thermo Fisher Scientific, 2020).

Apesar do manual não sugerir mudanças na densidade óptica do momento de indução, esse parâmetro foi bastante alterado na literatura, com diferentes relatos para o momento ideal de troca do meio, por conta disso, essa variável também foi analisada. Além disso, avaliamos se a presença de uma pré-cultura otimizaria a produção.

Portanto, os ensaios foram conduzidos avaliando a presença ou a ausência de uma pré cultura, o uso das células para cultura diretamente do estoque congelado ou após replaqueamento em YPD ágar, os diferentes tempos de indução (2, 3, 4, 5 e 6 dias) e as diferentes densidades celulares (DO_{600nm}) para o início da indução (DO_{600nm} = 1, 2, 3, 4, 6 e 8). As expressões seguiram o mesmo protocolo realizado na França, alterando apenas as variáveis em análise, com o volume de cultura mantido em 25% em relação ao volume total do frasco.

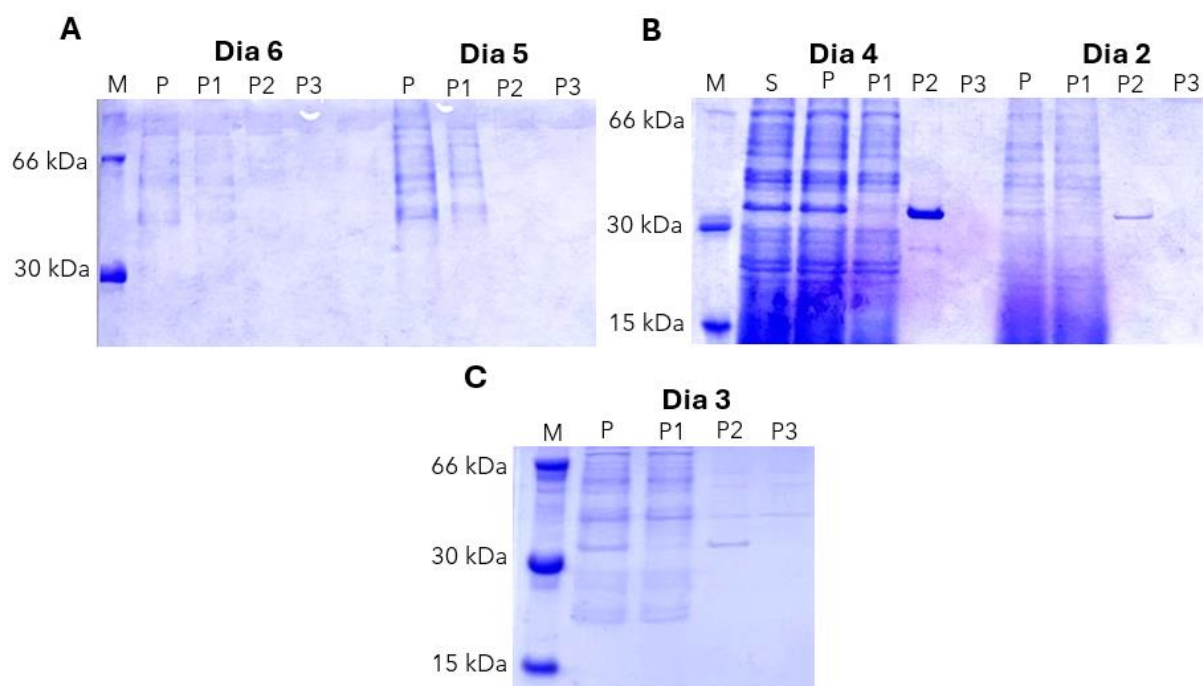
As células que foram cultivadas do estoque congelado em glicerol e as que foram cultivadas inicialmente em uma pré-cultura levaram mais tempo para crescer e atingir a densidade óptica ideal de indução, resultando em estratégias não desejáveis.

Ao finalizar o tempo de expressão dos diferentes ensaios, as células foram recuperadas por centrifugação, o sobrenadante foi submetido a precipitação com sulfato de amônio a 70%, e o precipitado ressuspense em tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 20 mM. A purificação da lectina rADEL seguiu a mesma estratégia de cromatografia de afinidade a metais imobilizados. A coluna foi lavada com tampão contendo

imidazol 20 mM (P1) e as proteínas retidas a coluna foram eluidas com 270 mM de imidazol (P2) e com 500 mM de imidazol no mesmo tampão (P3).

A pureza das frações (Figura 51) e a funcionabilidade da lectina dos diferentes testes foram avaliadas por SDS-PAGE e atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina, respectivamente.

Figura 51 – Avaliação do melhor tempo de expressão para rADEL: Pureza da lectina purificada em SDS-PAGE.



Fonte: Autora. Gel 15%. A: Tempos de indução: 5 e 6 dias. B: Tempos de indução: 4 e 2 dias. C: Tempo de indução: 3 dias. M: Marcador molecular contendo: HSA – 66 kDa (NP_0000468); rSFL - 30 kDa (7UMJ_AA) e CCL – 15 kDa (C0HLX7.1). S: sobrenadante da cultura antes da precipitação. P: Precipitado (AmSO₄ 70%) antes da cromatografia. P1: fração lavada da coluna com tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 20 mM. P2: fração eluída da coluna com tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 270 mM. P3: fração eluída da coluna com tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 500 mM.

A análise dos géis de eletroforese mostrou que o período de indução de 5 e 6 dias não foram eficientes na produção da proteína recombinante rADEL, visto que nada foi observado nas frações eluídas da coluna, correspondente à lectina de interesse (Figura 51 – A). No entanto, as induções por 2, 3 e 4 dias (Figura 51 – B e C) resultaram na produção da lectina, evidenciada pela presença de uma banda acima de 30 kDa, tanto na amostra obtida após precipitação com sulfato de amônio (P) quanto na fração eluída da coluna com 270 mM de imidazol (P2), corroborando com a massa estimada de rADEL próxima a 37 kDa observada nas purificações da França.

Além disso, a análise de SDS-PAGE sugere que a ligação da lectina à coluna e sua posterior recuperação foram eficientes, visto que a lectina não foi detectada nas frações não ligadas a coluna (P1). Da mesma forma, a ausência da proteína na fração eluída com 500 mM de imidazol (P3) indica que a concentração de 270 mM de imidazol foi suficiente para a completa eluição da lectina ligada à matriz cromatográfica.

As frações de rADEL eluídas da coluna (P2) obtidas em todos os períodos de indução foram submetidas ao ensaio de atividade hemaglutinante sob as mesmas condições. Os títulos de hemaglutinação referentes a análise funcional da lectina rADEL, em relação ao período de indução estão resumidos na Tabela 12.

Tabela 12 – Avaliação do melhor tempo de expressão para rADEL: Funcionabilidade da lectina recombinante expressa.

Período de indução (dias)	Maior diluição hemaglutinante (2ⁿ)	Título de hemaglutinação (U.H./mL)
2	2 ⁶	64
3	2 ⁷	128
4	2 ⁹	512
5	2 ⁴	16
6	-	0

Fonte: Autora. A maior diluição hemaglutinante é expressa como 2ⁿ, onde n corresponde ao número de diluições seriadas até o último poço no qual a hemaglutinação é detectada. O título de hemaglutinação é definido como a maior diluição da amostra capaz de promover hemaglutinação visível, expresso em unidade hemaglutinantes por mililitro (U.H./mL).

Os dados de atividade hemaglutinante corroboram com as análises de SDS-PAGE, visto que a banda correspondente à rADEL, com maior intensidade tanto no precipitado quanto na fração purificada, foi observada no quarto dia de indução, seguida pelos dias 3 e 2. Esse mesmo padrão foi observado na atividade hemaglutinante, onde o maior título de hemaglutinação foi detectado no quarto dia, seguido pelos dias 3 e 2, novamente. Apesar de não ter sido detectada nenhuma banda em SDS-PAGE correspondente à lectina no dia 5, a atividade hemaglutinante detectou a presença da proteína recombinante, indicando a expressão da lectina ativa, porém em quantidades baixas que não foram capazes de serem visualizadas em gel. No entanto, no sexto dia de indução não foi detectada a presença da proteína nem em SDS-PAGE e nem pela atividade hemaglutinante, confirmando a ausência da lectina nesse período. Portanto, os resultados indicam que o tempo ótimo de expressão para rADEL é de quatro dias, apresentando a banda mais intensa em SDS-PAGE e o maior título de hemaglutinação.

Através da análise desses resultados podemos entender o perfil temporal de expressão da lectina recombinante rADEL. Nos dias 2 e 3 de indução, a proteína está em fase de produção crescente. No dia 4, a expressão atinge seu ápice de produção, e a partir daí começa a reduzir progressivamente, com níveis baixos de proteína no dia 5 e a ausência total no dia 6. Esses dados validam o tempo sugerido pelo manual, de 2 a 4 dias (Invitrogen, Thermo Fisher Scientific, 2020).

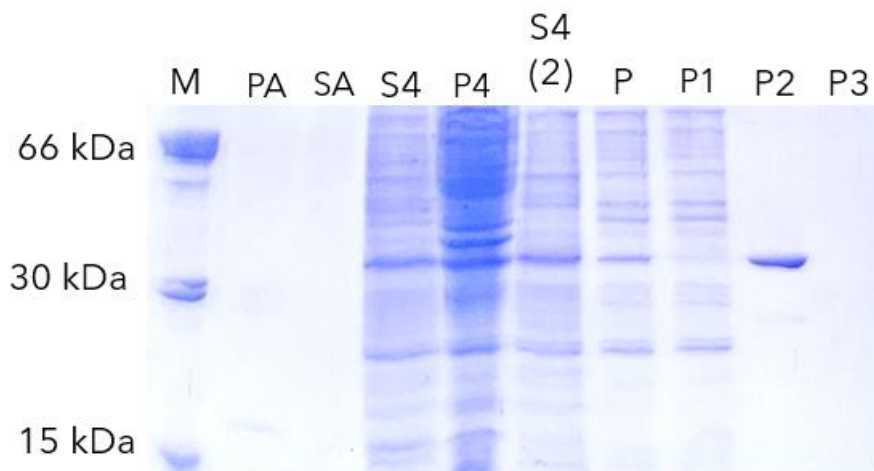
Esse perfil de expressão que atinge seu ápice e começa a reduzir até a ausência total de expressão funcional, como observado no sexto dia, acontece devido às limitações fisiológicas de *K. pastoris*. Induções prolongadas com metanol causam estresse oxidativo pois a degradação desse composto produz formaldeído e peróxido de hidrogênio que podem se acumular nas células gerando toxicidade e diminuindo a viabilidade celular e a capacidade de sintetizar a proteína recombinante ao longo do tempo. Além disso, em cultivos longos ocorre a liberação e ativação de proteases que podem degradar a proteína recombinante expressa, o que explica o motivo de no sexto dia não ser observada mais nenhuma atividade (Daily; Hearn, 2005; Macauley-Patrick *et al.*, 2005).

As diferentes densidades celulares para o início da indução ($DO_{600nm} = 1, 2, 3, 4, 6$ e 8) também foram avaliadas. Os maiores picos de absorbância a 280 nm, correspondentes à eluição da proteína com 270 mM de imidazol, foram observados quando a troca de meio foi realizada em densidades celulares entre 1 e 2, com valores em torno de 200 mAu. Os menores picos de absorbância foram obtidos em induções com densidade celular maior que 6, com valores em torno de 40 mAu (dados não mostrados).

Portanto, dentre os parâmetros avaliados, o melhor método de expressão foi estabelecido com a cultura crescida a partir do replaqueamento e mantida em T °C ambiente, conforme estabelecido anteriormente na França, sem o cultivo em uma pré-cultura, com a indução iniciando na densidade celular entre 1 e 2, e durando 4 dias.

Com as melhores condições determinadas, as expressões em maior escala foram realizadas com a colônia 1, seguindo o protocolo estabelecido. A purificação foi conduzida da mesma forma, com a eluição da proteína a 270 mM de imidazol. Portanto, conseguimos reproduzir com sucesso a expressão e a purificação da lectina recombinante rADEL, mantendo sua funcionabilidade preservada (FIGURA 52).

Figura 52 – Perfil de expressão e purificação da proteína rADEL no Brasil.



Fonte: Autora. Gel 15%. M: Marcador molecular contendo: HSA – 66 kDa (NP_0000468); rSFL – 30 kDa (7UMJ_AA) e CCL – 15 kDa (C0HLX7.1). PA: Pellet antes da indução. SA: Sobrenadante antes da indução. S4: Sobrenadante antes da indução no dia 4. P4: Pellet antes da indução no dia 4. S4(2): Sobrenadante na hora da retirada da cultura no dia 4. P: Precipitado (AmSO_4 70%) antes da cromatografia. P1: fração lavada da coluna com tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 20 mM. P2: fração eluída da coluna com tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 270 mM. P3: fração eluída da coluna com tampão Tris 20 mM pH 8, com NaCl 150 mM, glicerol 5% e imidazol 500 mM.

O perfil de expressão e purificação em SDS-PAGE mostra que, antes da indução, não houve detecção da proteína recombinante sendo expressa nem no *pellet* e nem no sobrenadante. Após a indução, a lectina de interesse passou a ser detectada pela banda acima de 30 kDa, a qual se manteve visível em todas as etapas seguintes do processo até a cromatografia de afinidade. Durante a purificação, a proteína recombinante não foi detectada na fração não retida a coluna (P1), indicando que a ligação à matriz foi eficiente. A eluição de rADEL ocorreu totalmente com 270 mM, pois não foi observada a presença da proteína na fração eluída com 500 mM de imidazol (P3).

A proteína também foi detectada na fração celular, mas conforme comprovado anteriormente, a lectina rADEL que não é secretada não apresenta atividade funcional, pois necessita da glicosilação para se manter biologicamente ativa.

A proteína rADEL purificada no Brasil apresentou um perfil de migração eletroforética semelhante ao observado para a proteína purificada na França, com a massa estimada parecida entre as duas purificações. No entanto, algumas diferenças foram observadas.

Primeiramente, nas purificações realizadas na França, rADEL apresentou bandas difusas, características típicas de glicosilação. Porém, nas purificações realizadas no Brasil, as bandas apareceram mais nítidas e definidas, mesmo apresentando a massa estimada superior ao

valor esperado para rADEL de 27 kDA. Esse aumento de massa, observado nas duas purificações, indica que a proteína permanece hiper glicosilada, apesar da diferença observada no perfil da banda eletroforética.

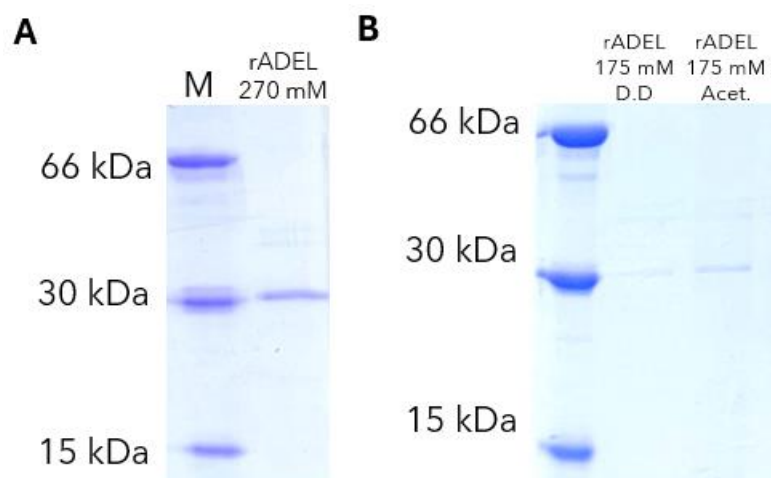
Essa alteração pode ser explicada com a diferença no grau de heterogeneidade da glicosilação. Apesar da maior parte do protocolo de expressão ter sido o mesmo entre as expressões na França e no Brasil, houve algumas mudanças, como o tempo de indução, a densidade celular do início de indução e uma nova transformação. Essas pequenas mudanças podem ter impactado na hiperglicosilação, mudando o padrão da *N*-glicosilação entre as expressões, provavelmente diminuindo o grau de heterogeneidade nos *N*-glicanos, ou seja, pode ter adicionado menos cadeias de manose durante a nova expressão. Essa alteração resultou em uma diferença no perfil observado em SDS-PAGE. Mudanças de migração na presença de *N*-glicosilação são bastantes documentadas na literatura (Scheller *et al.*, 2021; Wilson; Simpson; Cooper-Liddell, 2008).

Outra diferença observada foi em relação a presença de contaminantes. Nas purificações realizadas na França, a proteína foi eluída com um alto grau de pureza. Porém, nas expressões realizadas no Brasil, alguns contaminantes, ainda que em quantidades pequenas, foram detectados na fração eluída com 270 mM de imidazol (Figura 52).

Para tentar retirar esses contaminantes, e obter uma lectina purificada com um maior grau de pureza, foram testadas estratégias de eluição com concentrações crescentes de imidazol. Assim, foram avaliadas eluições com 40 mM, 80 mM, 175 mM e 270 mM de imidazol.

Embora, nas purificações realizadas na França, a lectina tenha sido eluída inicialmente com 35 mM de imidazol, e com o restante sendo recuperado em aproximadamente 170 mM, nas purificações realizadas no Brasil, as concentrações de 40 mM e 80 mM não foram suficientes para remover rADEL da matriz cromatográfica. A eluição com 175 mM permitiu a recuperação da proteína recombinante, porém ainda na presença dos contaminantes observados na eluição com 270 mM (Figura 53). Além disso, quando eluída com 175 mM, a lectina era completamente recuperada nessa concentração, não sendo detectada nenhuma banda referente a proteína na eluição com 270 mM (dados não mostrados).

Figura 53 - Purificação de rADEL com 175 mM e 270 mM de imidazol.



Fonte: Autora. Gel 15%. A: Eluição com 270 mM de imidazol. B: Eluição com 175 mM de imidazol. M: Marcador molecular contendo: HSA – 66 kDa; rSFL – 30 kDa e CCL – 15 kDa. D.D: Diluição direta. Acet.: Precipitação com acetona.

A análise em SDS-PAGE mostrou que o perfil de contaminação permaneceu o mesmo nas duas eluições com 175 mM e 270 mM de imidazol. No entanto, a recuperação da proteína pareceu ser mais eficiente com 270 mM de imidazol. Nessa eluição, a banda correspondente a rADEL foi claramente visualizada quando a amostra foi preparada por diluição direta. Porém, na eluição com 175 mM de imidazol, a banda apresentou menor intensidade, sendo melhor visualizada apenas após precipitação com acetona, e ainda assim com intensidade inferior à observada na eluição com 270 mM.

Essa diferença na presença de contaminantes entre as purificações da França e do Brasil, pode ser devido a alteração das proteínas endógenas secretadas por *K. pastoris*. As pequenas alterações nas condições de cultivo, da mesma forma que podem alterar a glicosilação, também podem alterar o perfil proteico secretado (Radoman *et al.*, 2021; Burgard *et al.*, 2020). Com a mudança nas condições de cultivo, novas proteínas endógenas podem ter sido secretadas junto com a proteína recombinante e interagido fracamente com a coluna durante a purificação.

A necessidade de maiores concentrações de imidazol para a eluição da proteína recombinante, também pode estar indiretamente relacionado com a mudança do perfil de glicosilação. A glicosilação altera o dobramento da proteína e ao alterar a conformação tridimensional, pode afetar a acessibilidade da cauda de histidina, modificando a interação da proteína com a resina cromatográfica.

Portanto, as purificações passaram a ser conduzidas com a eluição das proteínas retidas a coluna realizada com 270 mM de imidazol, na presença dos poucos contaminantes. A funcionabilidade da proteína foi confirmada em cada purificação e as frações contendo a proteína não clivada e ativa foram reunidas para uso posterior.

Mesmo com todos os parâmetros avaliados, a expressão de rADEL apresentou baixo rendimento quando comparado aos valores relatados na literatura (Macauley-Patrick *et al.*, 2005; Juturu; Wu, 2018). Para lectinas, os relatos de rendimento descritos tendem a ser inferiores aos observados para outras proteínas. Por exemplo, TPL-1 e TPL-2 apresentaram rendimentos de 5,2 mg e 0,6 mg por 500 mL de cultura, respectivamente (Kuo *et al.*, 2006). Porém, rADEL apresentou um rendimento médio inferior a 0,2 mg/L, muito abaixo do relatado, indicando limitações do sistema de expressão para essa proteína específica.

Diversas variáveis podem ser ajustadas com o objetivo de melhorar o rendimento da expressão, como selecionar o fenótipo Mut^s, induzir a expressão em uma densidade celular mais alta, escalonar a produção para a fermentação, otimizar os códons raros para *K. pastoris*, alterar o promotor forte AOX1 para um promotor mais fraco AOX2, coexpressar com chaperonas moleculares, expressar com proteínas de fusão, e não selecionar as células com integração múltiplas, pois eventos de integração múltiplas nem sempre leva a melhor expressão da proteína, as vezes pode causar estresse celular e estresse da via secretora. Para rADEL, os códons já foram otimizados e a indução em uma densidade celular alta não foi satisfatória (Macauley-Patrick *et al.*, 2005; Juturu; Wu, 2018).

Além disso, a cepa hospedeira também pode influenciar no nível de transcrição, na eficiência da tradução, na via secretora, na qualidade da proteína e na estabilidade do plasmídeo. Pois um problema muito relatado na expressão de proteínas nesse sistema hospedeiro é a degradação da proteína secretada por proteases, assim cepas deficientes na produção dessas enzimas podem ajudar a melhorar a eficiência de expressão. As proteínas expressas em *K. pastoris*, podem ser potencialmente degradadas no meio de cultura por proteases extracelulares endógenas desse organismo. Essa proteólise pode resultar na redução do rendimento da proteína expressa quando o produto é degradado, na perda da atividade biológica quando o produto é truncado e na contaminação do produto por intermediários de degradação no processamento (Macauley-Patrick *et al.*, 2005).

Algumas estratégias podem ser usadas para controlar essa degradação proteolítica. A utilização de meios tamponados com tampão fosfato é muito relatado, podendo ajustar uma ampla faixa de valores de pH para otimizar a produção da proteína. *K. pastoris* é capaz de

crescer em uma faixa de pH ampla (3,0–7,0) que não afeta significativamente o crescimento. Essa característica permite ajustes de pH para um valor que não seja o ideal para uma protease problemática específica. Diferentes valores de pH foram considerados ótimos para o rendimento de uma proteína recombinante na literatura. Por exemplo, algumas proteínas são suscetíveis a proteases que apresentam atividade ótima em pH neutro. Nesse caso, a expressão utilizando os meios de cultura não tamponados pode ser indicada, pois em meios sem tamponamento, o pH cai para 3 o que inativa muitas proteases de pH neutro. Por outro lado, tamponar o meio pode inativar enzimas que agem em pH baixo, demonstrando a importância da realização dos testes de tamponamento (Invitrogen, Thermo Fisher Scientific, 2020).

A estabilidade da proteína expressa também pode ser melhorada pela adição de suplementos ricos em aminoácidos ao meio de cultura como a peptona. Esses compostos atuam como substratos alternativos e competitivos para proteases, além de reprimir a indução de protease causada pela limitação de nitrogênio (Invitrogen, Thermo Fisher Scientific, 2020; Macauley-Patrick *et al.*, 2005).

Outros parâmetros que podem ser ajustados para inibir a ação de proteases é usar temperaturas de cultivos mais baixas de 30 °C para 23 ou 20 °C, diminuindo a secreção das enzimas. Reduzir a taxa de produção da proteína recombinante, controlando o excesso de metanol também pode diminuir a proteólise. Além disso, o uso de inibidores de proteases no meio de cultura e o uso de linhagens específicas modificadas para serem deficientes nas principais proteases, como as cepas SMD1163, SMD1165 e SMD1168 também podem ajudar a contornar esse problema e melhorar o rendimento da proteína recombinante. O uso dessas cepas com outras estratégias para reduzir a proteólise já foram relatadas como eficientes (Macauley-Patrick *et al.*, 2005). Portanto, a combinação de algumas dessas estratégias de cultivo pode ser eficaz para a obtenção de proteínas recombinantes com a mínima degradação proteolítica.

Depois da análise de diferentes parâmetros, a causa para o baixíssimo rendimento de rADEL pode ser a presença ativa da proteólise. Esse problema pode ser evidenciado pelo perfil temporal de produção da proteína, onde se observa uma redução progressiva da proteína expressa, com detecção da proteína nos dias 2, 3 e 4 de indução, seguidos por uma queda no dia 5 e a ausência total no dia 6. Esse padrão indica que a proteína expressa foi degradada a partir do quinto dia de indução, com a degradação total no sexto dia, demonstrando que o processo de degradação proteolítica pode ser a justificativa para o problema observado.

No caso de rADEL, a lectina pode conter uma sequência de aminoácidos reconhecida por proteases nativas, sendo uma proteína particularmente sensível à degradação. Talvez, realizar a expressão em meios tamponados com tampão fosfato, como sugere o manual da Invitrogen, ajudasse a solucionar esse problema (Invitrogen, Thermo Fisher Scientific, 2020; Macauley-Patrick *et al.*, 2005).

Por isso, é importante a otimização do processo em todos os níveis, não apenas para alcançar o máximo crescimento e produção da proteína recombinante, mas também para obter quantidades máximas da proteína desejada em uma forma biologicamente ativa.

No geral, a otimização da produção de proteínas recombinantes em *K. pastoris* é naturalmente mais complicada e desafiadora devido à complexidade de todo o processo de crescimento, indução e, principalmente, da secreção da proteína pela via secretora. Por isso, muitas vezes, os ajustes de expressão nesse sistema precisam de abordagens combinatórias. Se os fatores limitantes da produção recombinante não forem identificados, a estratégia para otimização não é previsível, nem mesmo para proteínas estruturalmente semelhantes (Juturu, Wu, 2018).

Por exemplo, para as lectinas plasmáticas do caranguejo *Tachypleus tridentatus*, TPL-1 e TPL-2, em *K. pastoris*, a expressão ocorreu em meio tamponado com fosfato, a 30 °C e a indução ocorreu em uma densidade celular de 2 a 6, com 1% metanol a cada 24 horas, durante 9 dias. Enquanto para a lectina do tipo C presente no caranguejo *Scylla paramamosain*, a expressão também foi conduzida em meio tamponado com fosfato, na mesma fase logarítmica e temperatura, mas a indução ocorreu com 0,5% de metanol por apenas 24 hrs (Qiu *et al.*, 2021).

Para a enzima presente na *Aplysia californica*, a expressão atingiu a quantidade de 300 mg/L em uma indução utilizando cepas Mut^S com 0,5% de metanol por dois dias (Munshi; Lee, 1997). Enquanto para o peptídeo antibacteriano do camarão *Fenneropenaeus chinensis*, a indução ocorreu em uma densidade de 6 a 8 em meio tamponado, com a duração de 4 dias a 0,5% de metanol (Li *et al.*, 2005). O outro peptídeo antimicrobiano isolado do soro de *Mytilus coruscus* também foi expresso em meio tamponado, com a indução também na densidade de 6 a 8, porém a 28 °C, com a duração de 4 dias a 1% de metanol e sem a seleção de múltiplas cópias inseridas (Meng *et al.*, 2016). Outras proteínas tiveram a melhor expressão com 1% de metanol, durante dois dias e com a indução começando na densidade de 1. Nesse caso, a influência do pH foi comprovada na ação de proteases (Zhang; Yang, 2006). Enquanto, para outra proteína, os autores testaram diferentes concentrações de metanol e o melhor resultado foi para 3% durante cinco dias de indução (Anggiani; Helianti; Abinawanto, 2018).

Todos esses exemplos demonstram que a estratégia de otimização precisa ser elaborada especificamente para cada proteína e o uso do meio tamponado parece ser um parâmetro importante.

Outro fator que impactou fortemente no rendimento de rADEL foi a ocorrência de precipitação da lectina após a purificação. rADEL apresentou precipitação mesmo na presença de glicerol, agente utilizado para manter a solubilidade de proteínas recombinantes. Essa precipitação não foi observada nas primeiras purificações, na França, passando a acontecer apenas na continuidade das purificações no Brasil, sugerindo a influência das mudanças nas condições de cultivo.

Assim como foi observado para as variações no perfil de purificação, que foram atribuídas às diferenças no padrão de glicosilação devido as variações nas condições de cultivo, é possível que essas mesmas modificações da glicosilação possam contribuir para uma conformação estrutural menos estável, levando a precipitação da proteína. Em *K. pastoris*, pequenas mudanças nas condições de cultivo podem afetar o dobramento, estabilidade, solubilidade e homogeneidade da proteína expressa (Macauley-Patrick *et al.*, 2005; Juturu, Wu, 2018).

Na literatura, relatos sobre a precipitação não intencional de proteínas recombinantes expressas em *K. pastoris* são escassos, mostrando que a precipitação observada pode estar relacionada com características intrínsecas da lectina que serão discutidas mais detalhadamente no próximo capítulo.

5.3.5 Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil)

Devido aos problemas de baixo rendimento, a proteína recombinante rADEL não foi submetida a clivagem da cauda de histidina. Além disso, a cauda presente no vetor de expressão possui um tamanho reduzido, de massa molecular menor que 2,5 kDa, e não possui outras proteínas fusionadas, portanto é pouco provável que a presença dessa cauda influencie de forma significativa as propriedades funcionais da lectina. Dessa forma, os ensaios de caracterização foram realizados com a proteína contendo a cauda e em tampão na presença de glicerol, para ajudar a manter a solubilidade e estabilidade da proteína.

Apesar do uso de proteínas recombinantes com alto grau de pureza ser o ideal para estudos bioquímicos e biológicos, alguns trabalhos relatam que purezas de proteínas recombinantes acima de 90% já são usadas para estudos funcionais posteriores (Jiang *et al.*,

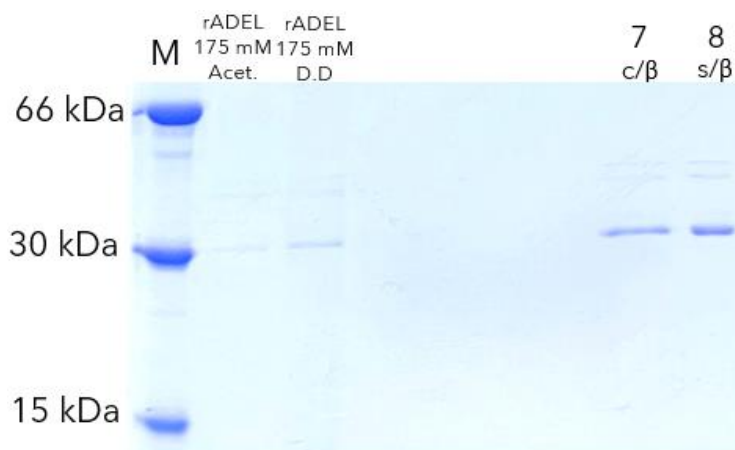
2025). No presente trabalho, apesar da presença de contaminantes residuais, os experimentos subsequentes foram realizados com as proteínas expressas sem a pureza ideal pois os seguintes aspectos foram considerados: o baixo nível de contaminação, a limitação do baixo rendimento, e a reprodutibilidade das tentativas de purificações, onde mesmo com diferentes abordagens, nenhuma foi eficiente para retirar os contaminantes residuais.

Ensaio biológicos e ensaios para avaliar a interação entre a lectina e o ligante estavam planejados para serem executados. Porém, devido ao baixo rendimento de expressão, apenas foi realizado testes que necessitaram de poucas quantidades da lectina para a execução.

5.3.5.1 Estimativa da massa molecular

A massa molecular estimada de rADEL foi de aproximadamente 34 kDa, em condições redutoras e não redutoras (Figura 54 – 7 e 8). Como discutido anteriormente, essa massa aparente maior que a massa esperada de 27 kDa para rADEL, é devido a hiperglicosilação que ocorre no sistema heterólogo de *K. pastoris*.

Figura 54 – Estimativa da massa molecular rADEL



Fonte: Autora. Gel 15%. M: Marcador molecular: HSA – 66 kDa; rSFL – 30 kDa e CCL – 15 kDa. Raias 2 e 3: purificação com 175 mM de imidazol. D.D: Diluição direta. Acet.: Precipitação com acetona. Raias 7 e 8: rADEL purificada na presença e ausência de β mercaptoetanol, respectivamente.

Em SDS-PAGE, a lectina nativa ADEL apresentou bandas de 60 e 29 kDa na ausência e presença de β -mercaptoetanol, respectivamente, pois ADEL é uma proteína dimérica (Carneiro *et al.*, 2017b) formada por duas cadeias polipeptídicas idênticas ligadas por uma ligação dissulfeto. Um comportamento diferente para o observado com rADEL.

Esses resultados sugerem que rADEL foi expressa como um monômero e provavelmente a ligação dissulfeto intermolecular, responsável pela dimerização da proteína nativa, não está sendo formada de forma eficiente no sistema heterólogo ou a hiperglicosilação, que é diferente da forma nativa, impede a dimerização correta da proteína. Apesar de *K. pastoris* permitir a formação de pontes dissulfeto e a glicosilação, a estrutura quaternária nem sempre pode ser reproduzida igual a proteína nativa, principalmente na presença de hiperglicosilação. Isso acontece porque essa adição de glicanos não nativa pode introduzir volumes estéricos adicionais mascarando ou distorcendo as interfaces de interação entre as subunidades. Além de alterar a acessibilidade dos resíduos de cisteínas, impedindo a formação da ponte dissulfeto intermolecular (Khlebodarova *et al.*, 2024; Juturu; Wu, 2018).

5.3.5.2 Ensaio de atividade hemaglutinante da proteína recombinante

A lectina recombinante rADEL apresentou atividade hemaglutinante para os eritrócitos de coelho testados, tanto na forma nativa como na forma tratado com a protease tripsina, assim como observado para a lectina nativa (Carneiro *et al.*, 2017b).

rADEL a uma concentração de 0,26 mg/mL foi capaz de aglutinar eritrócitos nativos com um título de hemaglutinação de 64 U.H./mL e eritrócitos tratados com tripsina em um título de hemaglutinação de 128 U.H./mL. Carneiro e colaboradores (2017b) não relatam o exato título de hemaglutinação, mas mencionam que ADEL demonstrou uma ligeira preferência por eritrócitos de coelho tratados com tripsina, assim como foi observado para a lectina recombinante. Portanto, a lectina possui preferência pelos glicanos que ficam mais acessíveis após o tratamento com a tripsina (Sampaio; Rogers; Barwell, 1998).

A manutenção da atividade hemaglutinante de rADEL, mesmo na presença da hiperglicosilação e da provável ausência de dimerização, que diferem da proteína nativa, indica que a região de reconhecimento a carboidratos permanece funcional (Young; Britton; Robinson, 2012), sugerindo que mesmo com as modificações pós-traducionais de glicosilação e mudanças na estrutura quaternária, o sítio de ligação a carboidrato é funcional por monômero e mantém a funcionabilidade básica da proteína. Portanto, mesmo quando a proteína recombinante não é idêntica a nativa, a atividade biológica pode permanecer preservada.

5.3.5.3 Ensaio de inibição da atividade hemaglutinante

A proteína recombinante rADEL foi submetida a um ensaio de inibição da atividade hemaglutinante para comparar o perfil de reconhecimento a carboidratos da lectina recombinante com a lectina nativa.

O perfil de inibição da atividade hemaglutinante de rADEL e a concentração mínima inibitória do açúcar capaz de inibir a aglutinação de eritrócitos estão mostrados na Tabela 13. Os carboidratos inibidos pela lectina nativa e sua concentração mínima inibitória (Carneiro *et al.*, 2017b) também estão listados na Tabela para facilitar a comparação das duas formas da lectina.

Tabela 13 – Perfil de inibição da atividade hemaglutinante de rADEL

ADEL	CMI	rADEL	CMI
Açúcares		Açúcares	
D-galactose	12.5 mM	D-galactose	25 mM
α -metil-D-galactopiranosídeo	10 mM	α -metil-D-galactopiranosídeo	50 mM
β -metil-D-galactopiranosídeo	5 mM	β -metil-D-galactopiranosídeo	50 mM
D-galactosamina	5 mM	D-galactosamina	50 mM
N-acetil-D-galactosamina	>100 mM	N-acetil-D-galactosamina	50 mM
Fenil- β -D-galactopiranosídeo	6.25 mM	Fenil- β -D-galactopiranosídeo	50 mM
Fenil- β -D-tiogalactopiranosídeo	>50 mM	N.T	N.T
4-nitrofenil- α -D-galactose	10 mM	N.T	N.T
4-nitrofenil- β -D-galactose	5 mM	N.I	N.I
2-nitrofenil- β -D-galactose	10 mM	2-nitrofenil- β -D-galactose	50 mM
O-nitrofenil- β -D-galactose	10 mM	N.T	N.T
D-ácido galacturônico	0.2 μ M	D-ácido galacturônico	3 μ M
L-ramnose	100 mM	N.I	N.I
L-arabinose	100 mM	N.I	N.I
α -lactose (gal β (1 $\rightarrow$ 4) α -glc)	12.5 mM	α -lactose (gal β (1 $\rightarrow$ 4) α -glc)	12,5 mM
β -lactose (gal β (1 $\rightarrow$ 4) β -glc)	6,25 mM	β -lactose (gal β (1 $\rightarrow$ 4) β -glc)	12,5 mM
Lactulose (gal β (1 $\rightarrow$ 4)fru)	6,25 mM	N.I	N.I
Melibiose (gal α (1 $\rightarrow$ 6)glc)	6,25 mM	Melibiose (gal α (1 $\rightarrow$ 6)glc)	6,25 mM
Rafinose (gal α (1 $\rightarrow$ 6)glc α (1 $\rightarrow$ 4)fru)	6,25 mM	Rafinose (gal α (1 $\rightarrow$ 6)glc α (1 $\rightarrow$ 4)fru)	25 mM
Continua			

			Continuação Tabela 13
Glicoproteínas			
PSM tipo III	0,125 mg/mL	PMS tipo III	0,25 mg/mL
Glicolipídeos			
N.T	N.T	Ácido tecóico <i>Staphylococcus aureus</i>	0,5 mg/mL
N.T	N.T	LPS (lipopolissacarídeo) <i>E. coli</i>	0,5 mg/mL

Fonte: Autora. N.I = Não inibiu; N.T = Não testado; PSM: Mucina do estômago de porco. CMI = Concentração mínima inibitória. Adaptada de Carneiro *et al.*, 2017b. Para rADEL, os estoques de açúcares estão em 100 mM e as glicoproteínas e glicolipídeos estão em 1 mg/mL. Para ADEL, os açúcares: α -metil-D-galactopiranosídeo, β -metil-D-galactopiranosídeo, D-galactosamina, 4-nitrofenil- α -D-galactose, 4-nitrofenil- β -D-galactose, 2-nitrofenil- β -D-galactose, O-nitrofenil- β -D-galactose estão em 10 mM de estoque, enquanto o restante está em um estoque de 100 mM, as glicoproteínas também estão em um estoque de 1 mg/mL.

Os açúcares O-nitrofenil- β -D-galactose, 4-nitrofenil- α -D-galactose e Fenil- β -D-tiogalactopiranosídeo não foram testados para rADEL, assim como os glicolipídeos não foram avaliados para ADEL nativa. Portanto, esses carboidratos não serão incluídos na análise comparativa entre as duas formas da lectina.

A lectina recombinante rADEL foi capaz de reconhecer um perfil de carboidratos semelhante ao da lectina nativa, com algumas diferenças. Os açúcares lactulose, ramnose, arabinose e 4-nitrofenil- β -D-galactose não foram capazes de inibir a atividade hemaglutinante de rADEL, portanto, a lectina recombinante não reconheceu esses carboidratos, diferentemente da lectina nativa.

A análise comparativa entre as duas formas da lectina mostrou que, para a maioria dos carboidratos, rADEL foi capaz de reconhecer o ligante, mas apresentou menor afinidade de ligação, demonstrado pelos valores maiores da concentração mínima inibitória quando comparado com ADEL nativa. Esse padrão foi observado para os carboidratos fenil- β -D-galactopiranosídeo, D-galactose, ácido galacturônico, mucina do estômago de porco tipo III, rafinose e β -lactose.

Enquanto para os carboidratos α -metil-D-galactopiranosídeo, 2-nitrofenil- β -D-galactose e N- acetil- D-galactosamina, rADEL demonstrou maior capacidade inibitória da atividade hemaglutinante, sugerindo interações mais fortes em relação à lectina nativa. Já para D-galactosamina, α -lactose, β -metil-D-galactopiranosídeo e melibiose, rADEL interagiu com esses carboidratos na mesma intensidade da lectina nativa. Para α -metil-D-galactopiranosídeo, β -metil-D-galactopiranosídeo e 2-nitrofenil- β -D-galactose a concentração do estoque foi de 10 mM.

As diferenças observadas na intensidade de interação e no perfil de reconhecimento entre a lectina nativa e recombinante é esperado devido a hiperglicosilação naturalmente ocorrida no sistema heterólogo de *K. pastoris*. Apesar da glicosilação não comprometer a integridade do domínio de reconhecimento a carboidratos, pois a lectina recombinante é capaz de reconhecer carboidratos, as cadeias longas de manose adicionadas podem causar um impedimento estérico e alterar o microambiente ao redor do sítio de ligação, funcionando como uma barreira física e dificultando o acesso de determinados ligantes ao domínio de reconhecimento, modificando a orientação da ligação ou afetando interações responsáveis por estabilizar o complexo lectina-carboidrato. Portanto, rADEL pode apresentar diferentes afinidades para alguns carboidratos, ou não reconhecer outros, quando comparada a lectina nativa (Juturu; Wu, 2018; Khlebodarova *et al.*, 2024).

A diminuição de afinidade observada para alguns ligantes pode estar relacionada ao fato de rADEL estar na forma monomérica, contendo apenas um sítio de reconhecimento a carboidratos por molécula, enquanto a lectina nativa é um dímero. A oligomerização de lectinas aumenta a multivalência das interações, que permite a ligação simultânea a vários ligantes, aumentando a intensidade da interação. No entanto, a forma monomérica é capaz de realizar uma única interação por vez, por isso, a intensidade da interação é menor (Feinberg *et al.*, 2001; Leffler *et al.*, 2002).

Assim, a alteração da estrutura quaternária observada em rADEL, provavelmente influenciou a afinidade e a multivalência da interação com alguns carboidratos. Mesmo que rADEL mantenha a funcionabilidade básica do domínio de reconhecimento a carboidratos, a mudança na organização oligomérica causou uma diminuição da eficiência de interação, resultando em uma atividade não idêntica à lectina nativa.

Ao analisar os carboidratos que rADEL não foi capaz de reconhecer, observamos que, na maioria, ramnose, arabinose e 4-nitrofenil- β -D-galactose já apresentavam valores elevados de CMI para ADEL nativa, indicando interações fracas quando comparadas para os outros carboidratos testados. Esse resultado corrobora com a ideia de que a forma monomérica de rADEL apresenta menor eficiência de ligação, pois interações que já eram fracas na proteína dimerica se tornaram ainda mais fracas na proteína monomérica, resultando no não reconhecimento desses ligantes pela lectina recombinante.

Apesar das mudanças, rADEL manteve o padrão principal de especificidade observado para a lectina nativa, preservando a preferência por derivados galactosídicos e

mantendo o ácido galacturônico como o ligante mais potente, mesmo com uma menor intensidade inibitória comparada a ADEL nativa.

A lectina recombinante também foi capaz de reconhecer o ácido tecóico e o LPS presentes na superfície celular das bactérias *S. aureus* e *E. coli*, respectivamente, indicando que provavelmente rADEL pode manter a atividade antibacteriana observada para a lectina nativa (Carneiro *et al.*, 2017b). Apesar de rADEL ser capaz de reconhecer o LPS no ensaio de inibição da atividade hemaglutinante, a lectina nativa não foi capaz de inibir a formação de biofilme por *E. coli* e nem demonstrou ser capaz de aglutinar esse microorganismo.

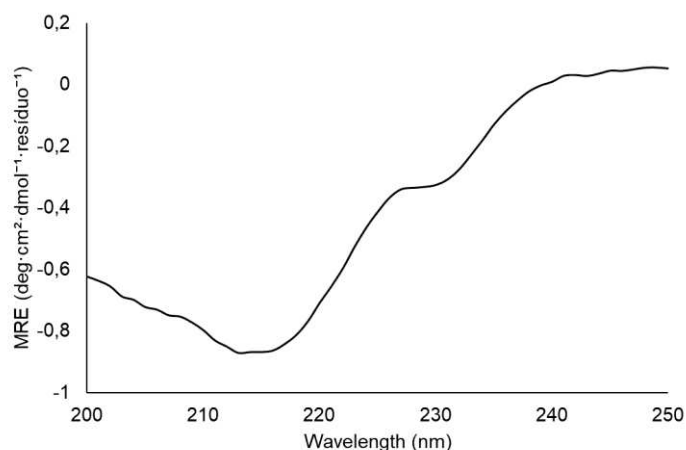
Se ADEL nativa também for capaz de reconhecer o LPS, essa diferença entre o reconhecimento molecular e a atividade biológica pode ocorrer devido à complexidade da organização da superfície celular das bactérias. A lectina pode reconhecer o LPS isolado, mas na membrana externa do microorganismo, outros componentes podem limitar a acessibilidade e a interação necessária para a agregação celular, principalmente devido ao arranjo estrutural mais rígido em bactérias gram-negativas, que pode dificultar o acesso da lectina ao ligante (Nikaido, 2003; Weidenmaier; Peschel, 2008).

5.3.5.4 Dicroísmo circular

O espectro de dicroísmo circular de rADEL apresentou um mínimo bem amplo em aproximadamente 216-218 nm e uma absorção máxima entre 240-250 nm, um padrão característico de proteínas ricas em β -folhas (Greenfield, 2006) (Figura 55).

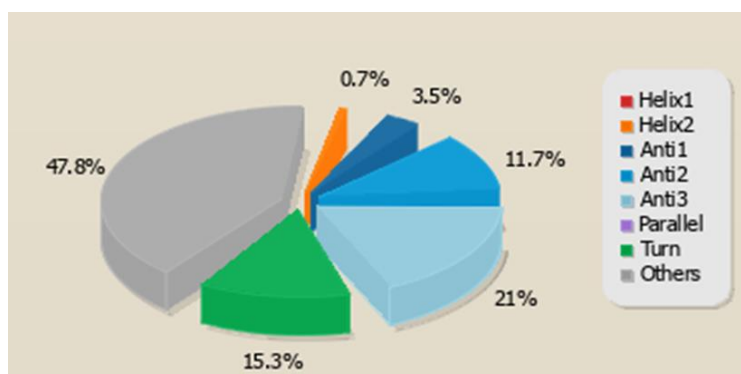
Os espectros obtidos foram processados pela plataforma BestSel (Kardos *et al.*, 2025; Micsonai *et al.*, 2022) e a estrutura secundária da lectina recombinante foi prevista com 0,7% de α -hélice 2 (Helix2), 3,5% de β -folha antiparalela 1 (Anti1), 11,7% de β -folha antiparalela 2 (Anti2), 21% de β -folha antiparalela 3 (Anti3), 15,3% de voltas (Turn) e 47,8% de estrutura desordenada (Others) (Figura 56). Esses dados mostram a predominância das β -folhas, com baixíssimo conteúdo de α -hélice, consistente com o espectro do dicroísmo.

Figura 55 – Espectro de dicroísmo de rADEL.



Fonte: Autora. O eixo X representa o comprimento de onda em nanômetros (nm). O eixo Y representa o MRE (Mean Residue Ellipticity = elipticidade média por resíduo), um parâmetro normalizado que reflete a diferença de absorção da luz circularmente polarizada à direita e à esquerda. A lectina estava em Tampão Tris 20 mM pH 7,6; NaCl 150 mM e Glicerina 5%; pH 7,6 a 0,180 mg/mL.

Figura 56 - Predição da estrutura secundária de rADEL



Fonte: Autora. Gráfico obtido pela plataforma BestSel.

Os espectros para ADEL nativa também exibiram valores de absorção mínima e máxima semelhantes com a lectina recombinante, com a mínima em 215 nm e a máxima em 202 e 229 nm, demonstrando uma predominância também de β -folhas em sua estrutura secundária (Carneiro *et al.*, 2017b).

De acordo com o método de previsão CONTIN que Carneiro e colaboradores (2017b) utilizaram, a estrutura secundária teórica de ADEL consistia em 4% de α -hélice, 44% de folha β , 21% de volta e 32% de espiral, porcentagens semelhantes a estrutura da lectina recombinante. Ao plotar os dados da ADEL nativa no mesmo programa usado para rADEL com o objetivo de facilitar a comparação, temos uma estrutura secundária para a lectina nativa de 5,6 % de β -folha antiparalela 1 (Anti1), 12,2 % de β -folha antiparalela 2 (Anti2), 25,2 % de

β -folha antiparalela 3 (Anti3), 14,5 % de voltas (Turn) e 42,5 % de estrutura desordenada (Others). Com os dados processados no mesmo programa, as estruturas secundárias das lectinas nativas e recombinantes foram ainda mais semelhantes.

Portanto, rADEL foi expressa em *K. pastoris* com um dobramento correto e semelhante ao da lectina nativa.

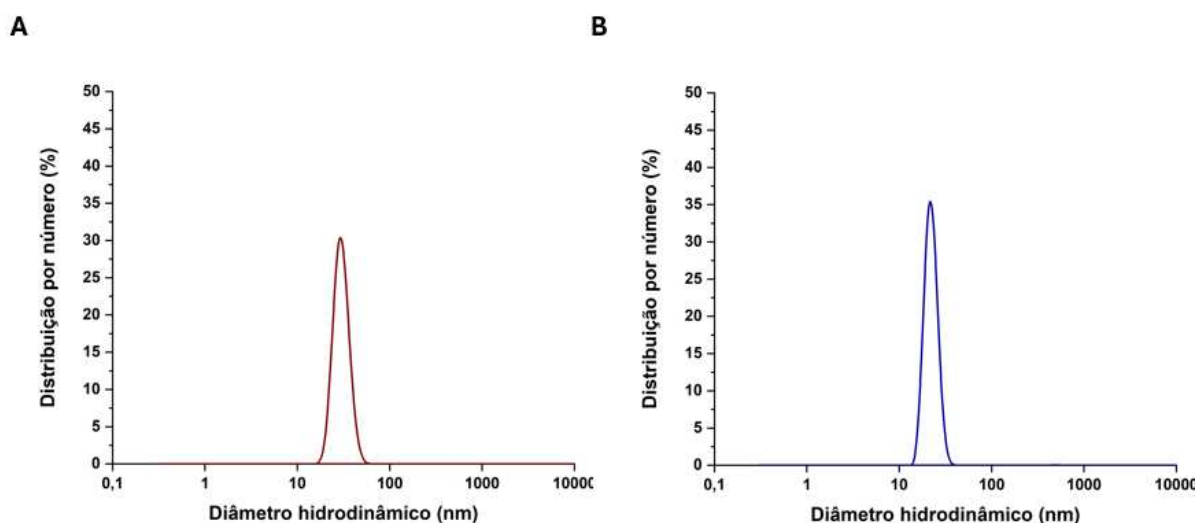
5.3.5.5 Espalhamento dinâmico de luz - DLS (Dynamic Light Scattering)

O comportamento hidrodinâmico da lectina recombinante foi avaliado pela análise de DLS nas concentrações proteicas de 0,250 mg/mL e 0,125 mg/mL, buscando entender em qual concentração a lectina formaria menos agregados.

A análise da proteína por DLS revelou que rADEL apresenta distribuição de tamanhos heterogênea em solução. A 0,250 mg/mL, o diâmetro hidrodinâmico médio foi de aproximadamente 463 nm e o índice de polidispersidade (PDI) foi de 0,43, esses parâmetros indicam uma amostra polidispersa (Erickson, 2009; Maguire *et al.*, 2018). Enquanto a análise da proteína a 0,125 mg/mL apresentou um diâmetro hidrodinâmico médio de aproximadamente 615 nm e o índice de polidispersidade (PDI) foi de 0,5657, esses parâmetros indicam que a amostra mais diluída apresenta partículas maiores em média, com uma maior heterogeneidade.

A distribuição do diâmetro hidrodinâmico de rADEL para 0,250 mg/mL e 0,125 mg/mL obtida por DLS é mostrada na Figura 57 e os parâmetros avaliados nas análises estão resumidos na Tabela 14.

Figura 57 - Distribuição do tamanho hidrodinâmico por números de partículas da lectina rADEL nas duas concentrações.



Fonte: Autora. A: Distribuição do tamanho hidrodinâmico por número de partículas da lectina a 0,250 mg/mL. B: Distribuição do tamanho hidrodinâmico por número de partículas da lectina a 0,125 mg/mL. O eixo x indica o diâmetro hidrodinâmico em nanômetros (nm). O eixo y indica a distribuição por números em porcentagem das partículas (%). A amostra foi analisada em tampão Tris 20 mM pH 7,6, NaCl 150 mM e glicerol 5%, na concentração nas concentrações de 0,250 mg/mL e 0,125 mg/mL. Os dados representam a média de três replicatas independentes indicados pela linha vermelha para 0,250 mg/mL e pela linha azul para 0,125 mg/mL.

Tabela 14 - Parâmetros do espalhamento dinâmico de luz de rADEL a 0,250 mg/mL e 0,125 mg/mL

	Média 0,250 mg/mL	Média 0,125 mg/mL
Diâmetro médio (nm)	463	614,8
Índice de polidispersidade (PI)	0,4267	0,5657
Pico 1: média por intensidade ordenada por área (nm)	275,2	339,5
Pico 1: área por intensidade ordenada por área (%)	85,58	90,65
Pico 2: média por intensidade ordenada por área (nm)	34,21	23,49
Pico 2: área por intensidade ordenada por área (%)	14,42	9,352

Fonte: Autora. Resumo dos dados obtidos por DLS para rADEL a 0,250 mg/mL e 0,125 mg/mL. Valores calculados a partir da distribuição por intensidade, incluindo diâmetro médio (nm), índice de polidispersão (PI) e os principais picos identificados nas distribuições baseadas em intensidade.

Um diâmetro médio de 463 nm e 615 nm é muito maior do que o esperado para uma proteína monomérica ou dimérica, indicando a presença de agregados em solução (Erickson, 2009).

Além disso, a presença de dois picos principais em ambas as concentrações analisadas indica a coexistência de populações de tamanhos distintos. Na concentração de 0,250 mg/mL, foi observado um pico predominante com tamanho médio de 275 nm totalizando aproximadamente 86% da amostra em intensidade, e o outro pico com partículas menores, em torno de 34 nm, representando aproximadamente 14% da amostra. De forma semelhante, para a concentração de 0,125 mg/mL, foi identificado um pico majoritário com tamanho médio de 339 nm, compreendendo aproximadamente 91% da amostra em intensidade, que é acompanhado por um pico minoritário com partículas menores, em torno de 23 nm, totalizando aproximadamente 9% da amostra. Esse resultado corrobora com os valores de PI de 0,4 e 0,5657, pois valores maiores que 0,4 indicam uma amostra altamente dispersa, com a presença de populações de tamanhos distintos (Maguire *et al.*, 2018). Todos esses dados indicam que, nas condições avaliadas, a maior parte de rADEL está agregada em solução, com uma maior agregação a 0,125 mg/mL.

Esse comportamento pode acontecer porque a agregação proteica não depende apenas da concentração, mas também do equilíbrio entre os estados oligoméricos e da estabilidade conformacional, assim como da estabilidade termodinâmica e cinética dos agregados formados. As proteínas podem existir em um equilíbrio entre estado nativo e conformações parcialmente dobradas, essas conformações mal dobradas servem como núcleo para o processo de agregação. Assim, a agregação não possui uma relação simples com a concentração, portanto mesmo após diluição, os agregados podem persistir (Dobson, 2003; Chi *et al.*, 2003).

Além disso, em concentrações mais altas, colisões moleculares frequentes podem favorecer dissociação de agregados instáveis, enquanto em concentrações mais baixas, ocorre menos colisões e os agregados podem se tornar mais estáveis (Chi *et al.*, 2003; Maguire *et al.*, 2018).

A distribuição por números (Figura 57) indica que a maior parte das partículas de rADEL se encontra entre 10 e 100 nm, porém o diâmetro hidrodinâmico médio foi de 463 nm e 615 nm nas concentrações de 0,250 mg/mL e 0,125 mg/mL, respectivamente. Como o diâmetro médio é calculado a partir da distribuição por intensidade, esse comportamento pode acontecer pois no DLS partículas maiores contribuem em maior parte para o sinal de espalhamento, devido a intensidade ser proporcional ao diâmetro elevado a sexta potência (d^6). Assim, a presença de uma população menor de agregados com maior tamanho, é suficiente para

deslocar a média para valores maiores, reforçando a natureza polidispersa da amostra (Maguire *et al.*, 2018).

Esse princípio do DLS explica o deslocamento para a esquerda do gráfico de distribuição por números para a concentração de 0,125 mg/mL, pois esse deslocamento indica a presença, em maior quantidade, de partículas pequenas na amostra, porém a presença de agregados maiores eleva o aumento do diâmetro médio e do PDI. Assim, o gráfico mostra que existe mais partículas pequenas pois é por distribuição de números, mas a maior intensidade (91%) vem das partículas com tamanho médio de 340 nm.

Todas essas análises comprovam que rADEL apresenta tendência a agregação, principalmente na concentração de 0,125 mg/mL. Comportamento observado para outras lectinas recombinantes, conforme discutido anteriormente.

5.4 Conclusão

A lectina recombinante rADEL foi expressa com sucesso em sistema heterólogo de *K. pastoris*, apresentando atividade biológica preservada, porém com rendimento insatisfatório. A caracterização bioquímica e estrutural indicou que a proteína recombinante manteve propriedades compatíveis com a lectina nativa, evidenciando a viabilidade do sistema hospedeiro utilizado. No entanto, as limitações de rendimento observadas demonstram a necessidade de otimização adicional das condições de expressão e purificação. Além disso, a ausência de expressão solúvel e ativa em sistema heterólogo de *E. coli* evidencia a necessidade da escolha do sistema hospedeiro de acordo com as características individuais de cada proteína.

6. CAPÍTULO III: EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOQUÍMICA E ESTRUTURAL DA LECTINA rAmL PRESENTE NA ALGA MARINHA *Amansia multifida*

6.1 Objetivos

Objetivo geral

Expressar em sistema heterólogo a lectina da alga marinha vermelha *Amansia multifida*, rAmL (recombinante *Amansia multifida* lectin), caracterizá-la bioquimicamente e estruturalmente; e avaliar o seu potencial antibacteriano combinado com antibióticos.

Objetivos específicos

- Produzir em sistema heterólogo de *E. coli* a lectina recombinante rAmL de forma solúvel e ativa, testando diferentes condições de expressão para obter o maior rendimento da proteína;
- Purificar a lectina recombinante;
- Comparar a atividade hemaglutinante e o perfil de inibição da lectina recombinante (rAmL) com a de sua forma nativa (AmL);
- Determinar o tampão ideal para manter a estabilidade da proteína recombinante por Ensaio de deslocamento térmico (*Thermal Shift Assay* - TSA);
- Obter a estrutura secundária da lectina recombinante por dicroísmo circular;
- Quantificar a afinidade da lectina recombinante com seu ligante específico por Calorimetria de Titulação Isotérmica (*Isothermal Titration Calorimetry* - ITC);
- Avaliar o comportamento hidrodinâmico da proteína recombinante e comparar o grau de agregação da lectina em diferentes tampões por meio de Espalhamento dinâmico de luz (*Dynamic Light Scattering* - DLS);
- Obter cristais difratáveis da lectina rAmL para determinar sua estrutura tridimensional por cristalografia de raios X;
- Avaliar o efeito antibacteriano da lectina rAmL contra cepas de *Staphylococcus aureus* e *Escherichia coli* em combinação com os antibióticos ampicilina, tetraciclina, estreptomicina e oxacilina;

6.2 Metodologia

6.2.1 Coleta dos exemplares e purificação da lectina nativa

Os exemplares da alga marinha *Amansia multifida* foram coletados durante a maré de sizígia na zona entre marés da praia do Pacheco, no município de Caucaia, litoral oeste do Ceará. O material biológico foi transportado em sacos plásticos até o laboratório, onde foi realizada a limpeza com água destilada, remoção de organismos aderidos e armazenamento a -20 °C para uso posterior. A coleta e o acesso ao patrimônio genético da espécie foram autorizados e regulados pelos órgãos competentes (SISBIO ID: 33913-12 e SISGEN ID:A1792FE).

A lectina presente na alga marinha vermelha *Amansia multifida* (AmL) foi extraída seguindo o protocolo previamente estabelecido (Silva, 2016). As espécimes da alga foram maceradas com nitrogênio líquido para a obtenção de um pó fino. O extrato bruto foi obtido suspendendo o pó em tampão fosfato de sódio 20 mM contendo 150 mM de NaCl, pH 7,0; e mantida sob agitação por 4 horas a 25 °C. Em seguida, o extrato foi filtrado, centrifugado a 9000 x g por 30 minutos (4 °C), e submetido a uma precipitação com 70% de sulfato de amônio. A fração 0/70 foi encaminhada a cromatografia de troca iônica em matriz DEAE-Celulose, e as proteínas não adsorvidas a coluna, contendo a proteína de interesse, foram eluídas com tampão fosfato de sódio 20 mM, pH 7,0.

A pureza das frações obtidas foi confirmada por eletroforese em gel de poliacrilamida/dodecil sulfato de sódio (SDS-PAGE) (Laemmli, 1970), e a funcionabilidade da proteína foi avaliada por atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina conforme descrito por Costa e colaboradores (1999). As frações puras e funcionalmente ativas foram reunidas, dialisadas e liofilizadas para uso posterior.

6.2.2 Produção da proteína recombinante rAmL em sistema heterólogo (França)

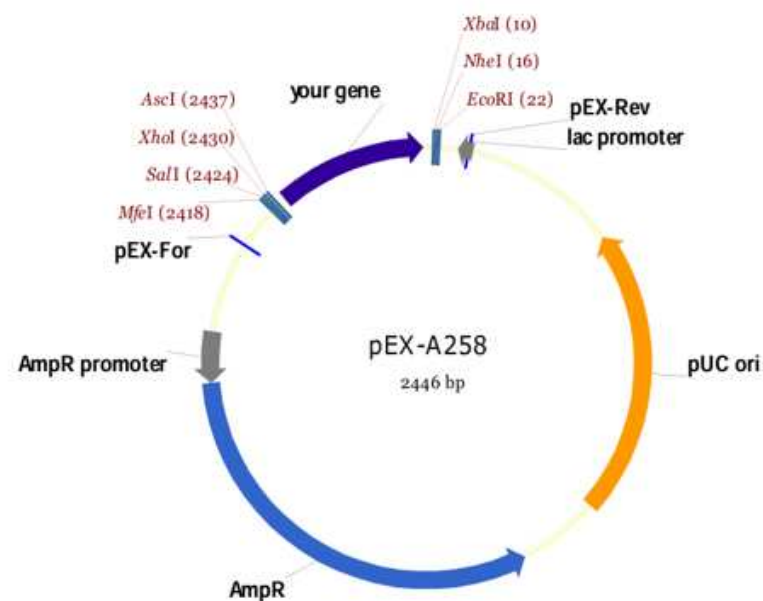
6.2.2.1 Desenho do gene sintético francês

A sequência de aminoácidos da lectina nativa, conforme dados apresentados por Chaves e colaboradores (Chaves *et al.*, 2026), foi utilizada para desenhar o gene sintético.

O gene sintético foi adquirido pela empresa Eurofins Genomics (França) no vetor *backbone* pEX-A258 com a sequência gênica otimizada para os códons de *Escherichia coli* e flanqueadas pelos sítios de restrição para as enzimas *NcoI* e *XhoI* (Figura 58).

Figura 58 – Sequência de nucleotídeos codificantes (minúsculas) e a respectiva tradução em aminoácidos (maiúsculas) da lectina para inserção no vetor pEX-A258.

A ccatgggtcgtcgctacacagtcgagaaccagtggggtgggtcgtcggcaccgtggaatgca
M G R R Y T V E N Q W G G S S A P W N A
gctggcctgtgggttctgggaggacgcccagcagccctgcgttgccgttgatgtgtcc
A G L W V L G G R P K Q P C V A V D V S
tcgacggatggcggtcgcaccttgacgggtacgatgacgtacaaggggagggaccgatt
S T D G G S T L T G T M T Y K G E G P I
ggattcaagggcatccgtaaggacgggtgtaatgtgtatgaggtgaagaatcagtggggg
G F K G I R K D G G N V Y E V K N Q W G
gggtcctctgcgcggtggcatgacgggtggcgagtggggtgatcggttcgcgcacatcgcaa
G S S A P W H D G G E W V I G S R T S Q
tccgtgatcgagctgaagatcaagtcgctcggactcgggtgcgcatttcgatggcacgatg
S V I E L K I K S S D S G A H F D G T M
aagtactccggtgaaggaccgatttgattccaagccacgatgtccgatggtgacacgtat
K Y S G E G P I G F Q A T M S D G D T Y
gacgttgaaaaccagtgaggtagctctgctccgtggcataaggcgggctgtgtggcg
D V E N Q W G G S S A P W H K A G V W A
ctcggcgaccgcgcgttccagaacatcatcatgctcaagggtcacatctcctgacggcgga
L G D R A F Q N I I M L K V T S P D G G
aagacgctaacgggtgacatgcagtagcgaacgcgaggggtcctattggcttccgtggaacc
K T L T G D M Q Y E R E G P I G F R G T
ctttcgcaggcgaataactacgctgtcgaaaaccagtgagggtcacatgcgccttgg
L S Q A N N Y A V E N Q W G G S H A P W
cacaaggctgggaactggattatcggggatcgtgacgggtcagaagggtggttggtgttgc
H K A G N W I I G D R D G Q K V V G V A
atcacgtccagcgataacggcaaacgtttgagggaaccatgacgtactcccgaggggc
I T S S D N G K T F E G T M T Y S R E G
cccattggcttcaagggcattctttcctcctaactcgag
P I G F K G I L S S -

B

MCS of pEX-A258

```

GGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTTCG
GGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTTTGG
caattgGTCGACctcgagGGCGCGCCCGTA>your_gene>GTGGCAGC
TCTAGAgctagcGAATTC TTTGGTGAAATTGTTATCCGCTCACAAATTCC
ACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTG
  
```

Start point: gene insertion site

(G of GTGG 3' of "your gene" is bp no. 1)

Multiple Cloning Site: bases 9-26 / 2417-2442

lac Promoter: bases 61-91

pUC origin: bases 415-1003

Ampicillin resistance ORF: bases 1174-2034

pEX-For priming site: bases 2319-2339

pEX-Rev priming site: bases 71-96

Fonte: Autora.

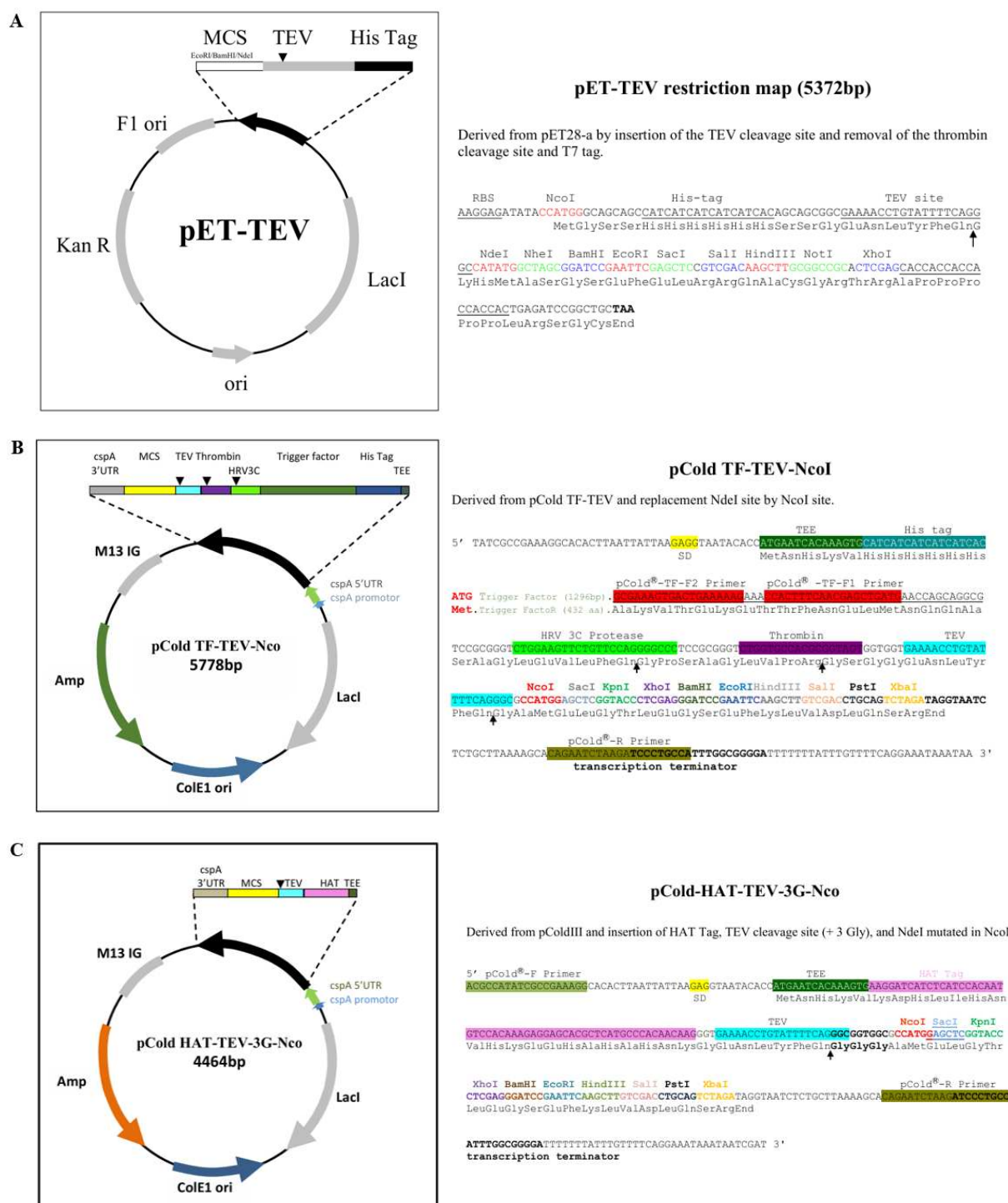
A – Representação da sequência de nucleotídeos e a respectiva tradução da lectina. A sequência em letras minúsculas representa os nucleotídeos e a sequência em letras maiúsculas representa os aminoácidos traduzidos. As letras em vermelho representam os sítios de clivagem *NcoI* (início) e *XhoI* (final). O símbolo – representa o códon de parada da tradução. B – Representação do vetor de expressão pEX-A258 e seu respectivo mapa. Disponível em: https://eurofinngenomics.eu/media/1587669/pex-a258_map_seq_v10.pdf.

6.2.2.1.1 Subclonagens do gene sintético para os vetores de expressão

O gene sintético inicialmente inserido no vetor *backbone* pEX-A258 foi subclonado nos vetores de expressão pET-TEV, pCOLD-TF-TEV-Nco e pCOLD-HAT-TEV-3G-Nco desenvolvidos pelo grupo GBMS (Structural and Molecular GlycoBiology), disponível na coleção do CERMAV- CNRS (Centre de Recherches sur les Macromolécules Végétales - Centre National de la Recherche Scientifique), Grenoble, França.

Os mapas e as particularidades dos plasmídeos usados para as subclonagens encontram-se na Figura 59 e Tabela 15, respectivamente.

Figura 59 – Representação dos vetores de expressão utilizados para subclonar a sequência gênica otimizada.



Fonte: Autora. A – Vetor pET-TEV. B – Vetor pCOLD-TF-TEV-NcoI. C – Vetor pCOLD-HAT-TEV-3G-Nco. Os mapas de restrição dos vetores estão disponíveis no banco de dados do CERMAV.

Tabela 15 – Características dos vetores utilizados para a expressão da lectina

VETOR	ESPECIFICIDADE	PROMOTOR	ANTIBIÓTICO DE SELEÇÃO
pET-TEV	Derivado do pET28a, presença da <i>His-tag</i> , com a inserção do sítio de clivagem da protease TEV e remoção do sítio de clivagem para trombina e da T7 <i>tag</i>	T7 Lac	Canamicina
pCOLD-TF- TEV-Nco	Derivado do pCOLD-TF-TEV.(indução a baixa temperatura com o sítio de clivagem para TEV, trombina e para protease 3C do rinovírus humano. Presença da <i>His-tag</i> , chaperona <i>trigger factor</i> e do TEE (<i>Translation Enhancing Element</i>). Substituição do sítio de restrição <i>NdeI</i> por <i>NcoI</i>	cspA	Ampicilina
pCOLD-HAT- TEV-3G-Nco	Derivado do pCOLDIII contendo TEE, com a inserção da HAT <i>tag</i> , sítio de clivagem TEV, adição de + 3 Gly após o sítio da TEV e a substituição do sítio de restrição <i>NdeI</i> por <i>NcoI</i> .	cspA	Ampicilina

Fonte: Autora. As enzimas de restrição *NcoI* e *XhoI* foram empregadas em todos os três vetores utilizados na subclonagem.

O gene sintético adquirido foi diluído para uma concentração de 50 ng.µL⁻¹ e inserido, por choque térmico, na cepa de clonagem *Escherichia coli* NEB5α, seguindo o protocolo de transformação bacteriana descrito por Sambrook (2002).

A subclonagem seguiu a metodologia detalhada anteriormente na seção 3.2.5.1.1.

Resumidamente, após a transformação, as colônias crescidas em LB ágar e selecionadas pelo antibiótico de seleção foram inoculadas meio LB *Broth* para extração do DNA plasmidial. O DNA plasmidial purificado por *miniprep* foi digerido com as enzimas *NcoI*

e *XhoI*, de acordo com a configuração do vetor. Após a digestão, os produtos resultantes foram separados em gel de agarose e o fragmento correspondente ao inserto da proteína foi purificado.

Os vetores pET-TEV, pCOLD-TF-TEV-Nco e pCOLD-HAT-TEV-3G-Nco foram extraídos por miniprep, submetidos ao mesmo protocolo de digestão enzimática e foram purificados a partir do gel de agarose. O inserto purificado e o vetor linearizado foram submetidos a uma reação de ligação, e 2 µL da reação de ligação foram usados para transformar as células de clonagem *E. coli* NEB5α, por choque térmico.

A confirmação da transformação foi feita por PCR em colônia usando o kit *Taq'Ozyme Purple Mix*. A T °C de anelamento foi ajustada de acordo com a temperatura de *melting* (Tm) para o par de primer específico do vetor utilizado. Os primers e a temperatura de anelamento usados para cada vetor estão descritos na Tabela 16.

Tabela 16 – Primers e condições de PCR para confirmação de vetores.

Vetor	Primers	T °C de anelamento
pET-TEV	T7 PRO (F)	50 °C
	T7 TER (R)	
pCOLD-TF-TEV-Nco	pCOLDTF-F1 (F)	52 °C
	pCOLD-R(R)	
pCOLD-HAT-TEV-3G-Nco	pCOLDTF-F (F)	52 °C
	pCOLD-R(R)	

Fonte: Autora. (F): primer *forward* e (R): primer *reverse*.

Os produtos da PCR foram visualizados em gel de agarose e um clone positivo foi cultivado em LB ágar e LB *Broth*, e o DNA plasmidial foi extraído por *miniprep*. O plasmídeo extraído foi novamente digerido com as enzimas *NcoI* e *XhoI*, para confirmação. O DNA plasmidial extraído do clone transformado e confirmado por digestão foi enviado para sequenciamento de confirmação pela EUROFINS (França).

6.2.2.2 Expressão da lectina *rAmL* em sistema heterólogo a partir da construção gênica francesa

Após a confirmação por sequenciamento, o DNA plasmidial do clone positivo foi transformado, na cepa de expressão *E. coli* BL21(Star), por choque térmico (Sambrook, 2002).

Como mencionado anteriormente, não foi realizado novos testes de expressão, adotando as condições padrões recomendadas pelo grupo para cada construção gênica.

Para expressão da proteína, o mesmo protocolo apresentado na seção 3.2.5.2 foi utilizado. Uma colônia de *E. coli* BL21 (Star) transformada foi pré-cultivada a 37 °C, *overnight*. Todo o volume de pré-cultura foi inoculado em LB *Broth*, e as células foram cultivadas a 37 °C, até atingir a densidade ótica de 0,4. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu até atingir DO = 0,8, quando o indutor IPTG foi adicionado na concentração de 0,2 mM e a expressão continuou a 16 °C, 180 *rpm*, por 18 horas.

Para o vetor pCOLD-HAT-TEV-3G-Nco, a expressão foi realizada em meio de autoindução LB NYZ, seguindo a mesma metodologia descrita na seção na seção 3.2.5.2. As células foram cultivadas a 37 °C, 180 *rpm* até atingir a DO_{600nm} superior a 5,0. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu a 180 *rpm*, por 18 horas.

Após o tempo de expressão para todos os vetores utilizados, as células foram recuperadas por centrifugação a 5000 x *g*, por 10 minutos a 25 °C. O sobrenadante foi descartado e o concentrado celular (*pellet*) foi pesado e congelado a -20 °C.

6.2.2.3 Purificação da lectina recombinante

A estratégia de purificação da lectina rAmL foi conduzida de acordo com as características da construção inserida em cada vetor de expressão. A Tabela 17 apresenta as diferentes abordagens de purificação empregadas em cada caso.

Tabela 17 – Estratégias de purificação da lectina recombinante em cada vetor.

Vetor (Sítio de restrição)	Cauda/tag	Estratégia	Cromatografia (Matriz)
pET-TEV (<i>Nco-XhoI</i>)	Ausente	Purificação seguindo o método da lectina nativa	Troca iônica (DEAE Sepharose)
pCOLD-TF-TEV-Nco (<i>Nco-XhoI</i>)	<i>His-tag</i>	Purificação por afinidade via <i>tag</i> pela sequência de histidinas	Afinidade – IMAC (Hisrap – Ni ²⁺)

Continua

Continuação tabela 17

pCOLD-HAT-TEV- 3G-Nco (<i>Nco-XhoI</i>)	<i>HAT-tag</i>	Purificação por afinidade via <i>tag</i> pela presença de histidinas	Afinidade – IMAC (HisTrap – Ni ²⁺)
---	----------------	--	---

Fonte: Autora. IMAC: Immobilized Metal Affinity Chromatography – Cromatografia de afinidade a metais imobilizados.

Para o vetor pET-TEV, cada 1g de *pellet* celular foi ressuspensionado em 5 mL de tampão Tris 20 mM pH 7 com *vortex*. A enzima denarase foi adicionada, as células foram lisadas com um disruptor celular mecânico, e o lisado celular foi centrifugado, conforme detalhado na seção 3.2.5.3. Aliquotas do sobrenadante (fração solúvel) e do precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE.

Como a construção do vetor pET-TEV não possui cauda de histidina, a fração solúvel foi submetida à cromatografia de troca iônica, conforme o método usado para purificação da lectina nativa (Silva, 2016). A fração solúvel, em tampão Tris 20 mM, pH 7 foi aplicada em uma coluna HiTrapTM DEAE FF de 1mL, previamente equilibrada e lavada com o mesmo tampão. As proteínas retidas a coluna foram eluídas com um gradiente de NaCl 1M em Tris 20 mM, pH 7. A cromatografia foi realizada em um sistema de cromatografia de média pressão NGC (Bio-Rad) a um fluxo de 1 mL/min e alíquotas de cada passo de purificação foram coletadas e analisadas por SDS-PAGE.

Para o vetor pCOLD-TF-TEV-Nco, o *pellet* celular foi ressuspensionado na mesma proporção em tampão Tris 20 mM pH 8, contendo NaCl 500 mM e imidazol 10 mM (Tampão A). A suspensão celular foi submetida ao mesmo procedimento de incubação e lise descrito anteriormente. Já para o vetor pCOLD-HAT-TEV-3G-Nco, o *pellet* foi ressuspensionado no mesmo tampão A com a adição de 5% (v/v) de glicerol. A suspensão celular seguiu o mesmo protocolo de incubação e ruptura celular. Após a lise, alíquotas da fração solúvel e da fração insolúvel obtidas para ambos os vetores foram coletadas por centrifugação e analisadas por SDS-PAGE.

Como a construção dos vetores pCOLD-TF-TEV-Nco e pCOLD-HAT-TEV-3G-Nco possui cauda com histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados (IMAC) de acordo com o protocolo detalhado na seção 3.2.5.3. A fração foi aplicada em uma coluna HisTrapTM FF de 1mL e as proteínas retidas a coluna foram eluídas com 5% do tampão com 500 mM de imidazol (tampão B), e com um gradiente linear de 5% a

100% do tampão B. Para o vetor pCOLD-HAT-TEV-3G-Nco, os tampões foram acrescidos de 5% (v/v) de glicerol.

As frações eluídas da coluna HisTrapTM contendo a proteína de interesse foram identificadas por SDS-PAGE e a funcionabilidade foi confirmada por atividade hemaglutinante contra eritrócitos de coelho a 4% (Sampaio *et al.*, 2002). As frações contendo a proteína não clivada pura e funcional foram reunidas para uso posterior.

6.2.2.4 Digestão da proteína fusionada e repurificação da lectina recombinante

Para remoção da cauda fusionada à lectina, as frações eluídas com 5% de tampão B, foram reunidas e diluídas na proporção de 1:1 com tampão Tris 20 mM pH 8. As frações eluídas com o gradiente de tampão B, tiveram o tampão trocado por centrifugação usando concentradores de 10 kDa com tampão Tris 20 mM pH 8, contendo NaCl 100 mM, conforme detalhado na seção 3.2.5.4.

Após a diluição ou a troca de tampão, a proteína foi concentrada para valores entre 0,2 mg/mL até 2 mg/mL e a protease TEV foi adicionada na razão enzima:substrato de 1:100 juntamente com 0,5 mM de EDTA. A reação de digestão foi incubada a 20 °C por 18 horas.

Após a reação de digestão, a concentração de imidazol foi ajustada e a solução foi reaplicada em coluna HisTrapTM FF de 1mL, previamente equilibrada com tampão Tris 20 mM pH 8, contendo NaCl 100 mM e imidazol 10 mM (tampão C). A fração não retida a coluna foi lavada com o mesmo tampão e a fração das proteínas ligadas a coluna, foi eluída com um step de 50% do tampão B. Todas as frações foram analisadas por SDS-PAGE e a funcionabilidade da proteína após a clivagem foi confirmada por atividade hemaglutinante contra eritrócitos de coelho a 4%.

Para as amostras do vetor pCOLD-HAT-TEV-3G-Nco, tanto as etapas de diluição ou troca de tampão, quanto o tampão C foram acrescidos de 5% (v/v) de glicerol.

Inicialmente, para a completa purificação da proteína clivada, a fração não retida a coluna HisTrapTM foi submetida a cromatografia de exclusão molecular em coluna EnrichTM SEC70 (10 x 300 mm; 24 mL; Bio-Rad) acoplada a um sistema NGC (Bio-Rad). A coluna foi equilibrada e eluída com tampão Hepes 20 mM pH 7,5 com NaCl 150 mM, e a amostra foi aplicada em um volume de 250 µL a 1,5 mg/mL em um fluxo de 1 mL/min.

Porém, após purificação de um novo lote da protease TEV, com uma maior eficiência de clivagem, a etapa de cromatografia de exclusão molecular deixou de ser necessária

e a proteína passou a ser obtida livre de contaminantes após a digestão e recromatografia em coluna de HisTrap™.

A pureza das frações foi avaliada por SDS-PAGE e as frações clivadas, funcionais e livres de contaminantes foram reunidas para uso posterior.

6.2.3 Avaliação da pureza e funcionabilidade da proteína recombinante (França)

6.2.3.1 Estimativa da massa molecular e avaliação da pureza da lectina

A estimativa da massa molecular e o grau de pureza nas diferentes etapas de expressão e purificação da lectina recombinante foram avaliados por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) (Laemmli, 1970), segundo relatado na seção 3.2.6.

A massa molecular foi estimada em condições redutoras, comparando a migração da proteína com a migração do marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad, com intervalo de 10 kDa a 250 kDa.

6.2.3.2 Atividade hemaglutinante da proteína

A atividade hemaglutinante foi realizada com eritrócitos de coelho não tratado com enzima proteolítica para avaliar a funcionabilidade da proteína expressa, como detalhado na seção 3.2.7.

6.2.4 Ensaio de deslocamento térmico (TSA – Thermal Shift Assay) (França)

O ensaio de deslocamento térmico foi realizado com o objetivo de avaliar a estabilidade da proteína e determinar as condições de tampão, pH e concentrações de sais que tornam a lectina mais estável.

O teste foi conduzido com a proteína rAmLpCHT3GNco, não clivada, em Tris 20 mM pH 8, contendo NaCl 100 mM para um *screening* amplo de diferentes condições.

O corante fluorescente SYPRO Orange (Sigma) foi usado na reação como indicador da desnaturação proteica. As reações foram preparadas em placas de 96 poços, onde cada poço possui um volume de 25 µL contendo o corante SYPRO Orange, a proteína e o tampão nas

seguintes concentrações finais: 5x de corante, 0,1 a 0,5 mg/mL de proteína e 50 mM do tampão testado. Um controle negativo foi preparado apenas com o tampão sem proteína e o corante, e o controle positivo continha a proteína no tampão inicial com o corante. A placa foi aquecida de 20 a 100 °C com uma taxa de aquecimento de 1 °C/min, em um equipamento de RT-qPCR (Mini personal Thermal Cycler, Bio-Rad). A intensidade da fluorescência foi mensurada pela determinação da temperatura de melting (T_m) no Biorad CFX Manager *software*.

Os tampões e condições utilizados no teste foram: 1 - Citrato pH 4; 2 - Acetato de sódio pH 4,5; 3 - SPG (Sucrose – Fosfato – Glutamato) pH 4,5; 4 – Acetato de sódio pH 5; 5 - Malonato de sódio pH 5; 6 – Citrato de sódio pH 5,5; 7 – Malonato de sódio pH 5,5; 8 – MES pH 6; 9 – MMT (Malonato; MES; Tris)pH 6; 10 – Bis-Tris pH 6; 11 – MES pH 6,5; 12 – MIB (Malonato ; Imidazol ; Borato) pH 6,5; 13 – HEPES pH 7; 14 – Buffer System 5 pH 7; 15 – SPG pH 7; 16 – HEPES pH 7,5; 17 – Bis-Tris Propano pH 7,5; 18 – Tris pH 8; 19 – Bis-Tris Propano pH 8; 20 – Tris pH 8,5; 21 – Buffer System 6 pH 8,5; 22 – SPG pH 9; 23 – Bicine pH 9; 24 – $MgCl_2$ 20 mM; 25 – $CaCl_2$ 10 Mm; 27 - Água. O número 26 é referente ao controle positivo: a proteína em Tris 20 mM pH 8, contendo NaCl 100 mM.

6.2.5 Ensaios de cristalização (França)

Os ensaios de cristalização foram realizados manualmente utilizando o método de difusão de vapor com gota suspensa em placas de cristalização de 48 poços, conforme descrito anteriormente na seção 4.2.6.

Além disso, triagens robotizadas foram realizadas na plataforma de cristalização do laboratório HTX no EMBL (*European Molecular Biology Laboratory-Grenoble*), também como previamente descrito.

Diferentes condições de cristalização foram testadas para rAmL, variando a amostra (rAmL clivada e rAmLpCHT3GNco não clivada), a concentração da proteína, o tampão inicial da amostra, o ligante para cocristalização e os kits de cristalização comerciais. As condições de cristalizações testadas estão listadas na Tabela 18.

Tabela 18 – Condições de cristalizações testadas para rAmL

AMOSTRA	[]	TAMPÃO INICIAL	LIGANTE	KIT DE CRISTALIZAÇÃO/ MONTAGEM
rAmL	3,49	HEPES 20 mM pH 7,5 NaCl 100 mM Imidazol 10 mM	Man6GlcNac(II) 1 mM	Clear Strategy Screen II (Molecular dimensions) / Manual
rAmL	4,0	Tris 20 mM pH 8 NaCl 100 mM CaCl ₂ 1 mM Glicerol 10%	Man5GlcNac(II) 1 mM	SG1 Box II (Molecular dimensions) / Manual
rAmL pCHT3GNCo	2,78	Tris 20 mM pH 8 NaCl 100 mM CaCl ₂ 1 mM Glicerol 10%	Man6GlcNac(II) 1 mM	SG1 Box I (Molecular dimensions) / Manual
rAmL	0,384	Tris 20 mM pH 8 NaCl 100 mM CaCl ₂ 1 mM Glicerol 10%	Man5GlcNac(II) 1 mM	Triagem robótica
rAmL pCHT3GNCo	1,3	Tris 20 mM pH 8 NaCl 100 mM CaCl ₂ 1 mM Glicerol 10%	Man5GlcNac(II) 1 mM	Triagem robótica

Fonte: Autora. [] mg/mL: concentração da proteína em mg/mL. Tampão inicial: tampão da solução de proteína + ligante.

6.2.6 Produção da proteína recombinante rAmL em sistema heterólogo (Brasil)

6.2.6.1 Desenho do gene sintético adquirido no Brasil

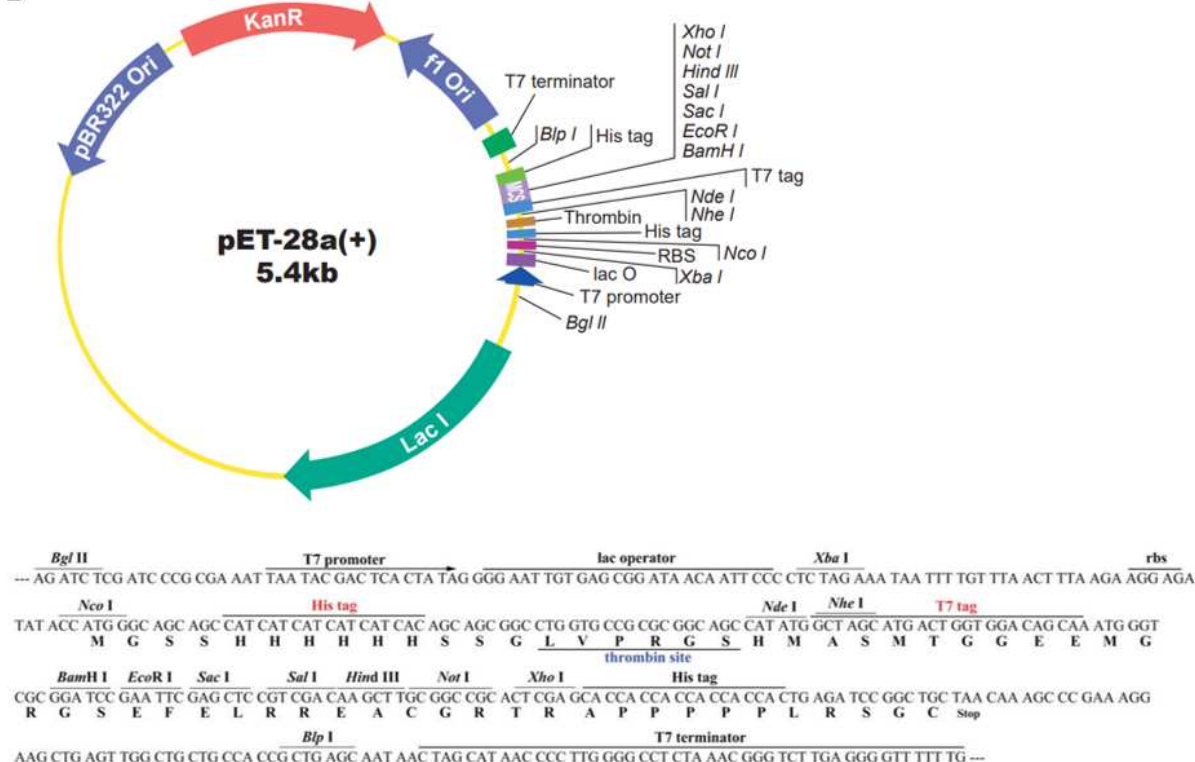
A sequência de aminoácidos da lectina nativa, conforme dados apresentados por Chaves e colaboradores (Chaves *et al.*, 2026), foi utilizada para desenhar o gene sintético.

O gene sintético foi adquirido pela empresa FastBio (<http://www.fastbio.com.br/>, SP, Brasil) no vetor de expressão pET28a(+) com as regiões de clonagem delimitadas pelas enzimas de restrição *NdeI* e *XhoI*, possuindo um sítio de clivagem para a protease TEV e sem a otimização de códons raros em *Escherichia coli* (Figura 60).

Figura 60 – Sequência de nucleotídeos e a respectiva tradução em aminoácidos da lectina para inserção no vetor pET28a(+).

A catatggaaaacctgtattttcaggggtcgctcgctacacagtcgagaaccagtggggtggg
H M E N L Y F Q G R R Y T V E N Q W G G
tcgtcggcaccgtggaatgcagctggcctgtgggttctgggcggacgcccagcagccc
S S A P W N A A G L W V L G G R P K Q P
tgcggttgccgttgatgtgtcctcgacggatggcggtcgaccttgacgggtacgatgacg
C V A V D V S S T D G G S T L T G T M T
tacaagggggagggaccgattggattcaagggcatccgtaaggacgggtgtaattgtgtat
Y K G E G P I G F K G I R K D G G N V Y
gaggtgaagaatcagtggggggggtcctctgcgcggtggcatgacgggtggcgagtgggtg
E V K N Q W G G S S A P W H D G G E W V
atcggttcgcgcacatcgcaatccgtgatcgagctgaagatcaagtcgtcggactcgggt
I G S R T S Q S V I E L K I K S S D S G
gcgcatttcgatggcagatgaagtactccggtgaaggaccgattggattccaagccacg
A H F D G T M K Y S G E G P I G F Q A T
atgtccgatggtgacacgtatgacgttgaaaaccagtgaggaggtagctctgctccgtgg
M S D G D T Y D V E N Q W G G S S A P W
cataagggcggcggtgtgggcgctcggcgaccgcgcgttcagaaacatcatcatgctcaag
H K A G V W A L G D R A F Q N I I M L K
gtcacatctcctgacggcggaagacgctaacgggtgacatgcagtacgaacgcgaggggt
V T S P D G G K T L T G D M Q Y E R E G
cctattggcttcctgtggaacctttcgcaggcgaataactacgctgtcgaaaaccagtggt
P I G F R G T L S Q A N N Y A V E N Q W
ggaggtcacatgcgccttggcacaaggctgggaactggattatcggggatcgtgacgggt
G G S H A P W H K A G N W I I G D R D G
cagaaggtggttggtgttgccatcacgtccagcgataacggcaaaacgtttgagggaaacc
Q K V V G V A I T S S D N G K T F E G T
atgacgtactcccgcgagggccccattggcttcaagggcattctttcctcctaactcgag
M T Y S R E G P I G F K G I L S S -

B



Fonte: Autora. A – Representação da sequência de nucleotídeos e a respectiva tradução da lectina com a inserção da sequência específica de clivagem para a protease TEV. A sequência em letras minúsculas representa os nucleotídeos e a sequência em letras maiúsculas representa os aminoácidos traduzidos. As letras em vermelho e azul representam os sítios de clivagem *NdeI* (início) e *XhoI* (final) e o sítio da protease TEV, respectivamente. O símbolo – representa o códon de parada da tradução. B – Representação do vetor de expressão pET28a(+) e seu respectivo mapa. Disponível em: <https://fastbio.com.br/wp-content/uploads/2023/09/GenScript-Lista-e-mapa-de-vetores-bacteria.pdf>.

6.2.6.2 Recuperação do gene sintético francês no Brasil

Na França, a construção gênica da proteína de interesse que apresentou os melhores resultados de expressão e purificação foi a inserida no vetor pCOLD-HAT-TEV-3G-Nco. Por esse motivo, o plasmídeo rAmLpCHT3GNco foi trazido para o Brasil em cartões QIAcard FTA (*Flinders Technology Associates*), Sigma. O armazenamento e a recuperação do plasmídeo foram realizados conforme descrito na seção 4.2.7.1.

6.2.6.3 Expressão da lectina rAmL em sistema heterólogo a partir das duas construções gênicas

6.2.6.3.1 Transformação bacteriana

O gene sintético não otimizado, adquirido no Brasil, foi diluído para uma concentração de $50 \text{ ng} \cdot \mu\text{L}^{-1}$ e inserido, por choque térmico, na cepa de clonagem *Escherichia coli* DH5 α , e na cepa de expressão *E. coli* BL21(DE3), seguindo o protocolo de transformação descrito por Sambrook (2002). O plasmídeo otimizado adquirido na França, após recuperação, também foi inserido nas mesmas cepas de *E. coli* pelo mesmo método. A manipulação e uso de organismos geneticamente modificados tem autorização pelo certificado CQB-R007-2016 CTNBio.

Para confirmar a transformação na cepa de clonagem e de expressão, as colônias crescidas em LB ágar e selecionadas pelo antibiótico de seleção correspondente ao vetor foram inoculadas em meio LB *Broth* contendo o mesmo antibiótico (ampicilina a $100 \mu\text{g/mL}$ ou Canamicina a $50 \mu\text{g/mL}$) e incubadas a 37°C , *overnight*.

A confirmação foi feita por PCR em colônia usando a enzima *Taq* DNA Polimerase (Sigma). As reações de PCR foram montadas de acordo com as recomendações do fabricante para o volume final de $25 \mu\text{L}$, como detalhado na seção 4.2.7.2.

A $T^\circ\text{C}$ de anelamento foi ajustada de acordo com o vetor utilizado. Os primers e a temperatura de anelamento usados para cada vetor estão descritos na Tabela 19.

Tabela 19 – Primers e condições de PCR para confirmação de vetores.

Vetor	Primers	T $^\circ\text{C}$ de anelamento
pET28a	T7 PRO (F)	50 $^\circ\text{C}$
	T7 TER (R)	
pCOLD-HAT-TEV-3G-Nco	Upstr_VQ (F)	52 $^\circ\text{C}$
	Upstr_VQ (R)	
	Upstr_VE (F)	
	Upstr_VE (R)	

Fonte: Autora. (F): primer *forward* e (R): primer *reverse*. Upstr_VQ e Upstr_VE são primers degenerados para lectinas da família OAAH. **Sequence - Upstr_FW_VQ:** 5' GTI CAG AAT CAR TGG GGI GG 3' (Tm 59.1°C e GC 52.5%). **Sequence - Upstr_FW_VE:** 5' GTI GAR AAT CAR TGG GGI GG 3' (Tm 57.6°C e GC 50%). **Sequence - Upstr_RV_VQ:** 5' CCI CCC CAY TGA TTC TAI AC 3' (Tm 55.9°C e GC 47.5%). **Sequence - Upstr_RV_VE:** 5' CCI CCC CAY TGA TTY TCI AC 3' (Tm 58.1°C e GC 50%).

Os produtos da PCR foram visualizados por eletroforese em gel de agarose a 1%, e o clone positivo foi identificado de acordo com o tamanho esperado do fragmento amplificado.

6.2.6.3.2 Testes de expressão

Os clones da cepa de expressão, previamente transformados e confirmados, foram inoculados em uma pré-cultura de meio LB *Broth* contendo o antibiótico seletivo correspondente (ampicilina a 100 µg/mL ou Canamicina a 50 µg/mL) e incubados a 37° C, *overnight*, sob agitação constante. Para os testes de expressão, as células crescidas foram diluídas na proporção de 1:100 em 20 mL de LB *Broth* contendo o mesmo antibiótico, e a cultura foi crescida e induzida de acordo com cada condição testada.

Os ensaios para determinar as melhores condições de expressão, para ambos os vetores, variaram a concentração de IPTG (0,05 mM; 0,1 mM; 0,2 mM; 0,5 mM e 1,0 mM) e as temperaturas de indução de 37 °C por 4 h, 30 °C por 16h, 25 °C por 20h, 18 °C por 24 horas e pelo método da França de reduzir a T °C antes da indução (descrito na seção 5.2.2.2) nas temperaturas de 25 °C e 16 °C. Em todos os experimentos, a indução ocorreu quando a cultura atingiu a DO_{600nm} entre 0,6 e 0,8.

Após o tempo de indução em cada teste, as células foram recuperadas por centrifugação a 6000 x g, por 5 minutos a 4 °C. O sobrenadante foi descartado e o concentrado celular (*pellet*) foi pesado e ressuspenso em tampão Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 20 mM, na proporção de 1g de *pellet* para 10 mL de tampão. As células foram rompidas por sonicação com amplitude de 30% durante 15s com 30s de intervalo. Em seguida, o lisado celular foi centrifugado a 9000 xg, por 30 min, a 4°C, e o sobrenadante (fração solúvel) e o precipitado (fração insolúvel) foram coletados e analisados por SDS-PAGE.

O melhor método de expressão para cada construção foi determinado pela comparação da intensidade da banda correspondente a proteína nos géis de SDS-PAGE, pelo maior pico de absorbância a 280 nm durante a purificação e pelo título de hemaglutinação da fração purificada contra eritrócitos de coelho tratado com tripsina a 3%.

Após estabelecer o melhor método de expressão para cada vetor, uma cultura de 510 mL de LB *Broth* para cada construção plasmidial foi cultivada com o objetivo de quantificar o rendimento e determinar qual vetor possui a maior rentabilidade de expressão da proteína de interesse. A cultura foi preparada a partir de um pré-inóculo conforme descrito acima, e as células crescidas foram diluídas na proporção de 1:100 em meio LB *Broth* nos volumes de 170 mL (*erlenmayer* de 0,5L) e 340 mL (*erlenmayer* de 1L) contendo o antibiótico na proporção 1:1000. As células foram crescidas a 37 °C até uma absorbância de 0,6 a 0,8 em

600 nm e a indução foi realizada com a concentração do indutor e a temperatura de indução de acordo com o melhor método de expressão.

Após o tempo de indução, o mesmo protocolo para análise da fração solúvel e insolúvel foi reproduzido. As células foram recuperadas por centrifugação e o concentrado celular foi pesado e congelado a -20 °C.

O pellet foi ressuspensão no mesmo tampão de lise (1:10). As células foram rompidas por sonicação e o lisado celular foi centrifugado para separação da fração solúvel e insolúvel.

6.2.6.4 Purificação da proteína rAmL

Como a construção de ambos os vetores possui cauda com histidina, a mesma estratégia de purificação da proteína foi utilizada. A fração solúvel foi submetida à cromatografia de afinidade (IMAC) em uma coluna HisTrap™ FF de 5mL, previamente equilibrada e lavada com o mesmo tampão de lise. As proteínas retidas a coluna foram eluídas com o mesmo tampão de equilíbrio acrescido de imidazol 500 mM. A cromatografia foi realizada em um sistema AKTA™ Pure (GE) ou utilizando uma bomba peristáltica e alíquotas de cada passo de purificação foram coletadas e analisadas por SDS-PAGE.

Para determinar a construção com maior rendimento e a melhor condição de expressão, as purificações iniciais foram realizadas com a eluição da proteína a 270 mM de imidazol. Após a condição de melhor rendimento ser determinada, a purificação passou a ser conduzida com dois steps: 40 mM e 80 mM de imidazol.

A pureza das frações contendo a lectina foi avaliada por SDS-PAGE e a funcionalidade foi confirmada por atividade hemaglutinante contra eritrócitos de coelho tratados com tripsina a 3% (Sampaio *et al.*, 2002). As frações contendo a proteína não clivada pura e funcional foram reunidas para uso posterior.

6.2.6.5 Digestão da proteína fusionada e repurificação da lectina recombinante

Para remoção da cauda fusionada à lectina, as frações eluídas com 270 mM de imidazol de ambos os vetores foram reunidas e dialisadas contra Tris 20 mM pH 8, NaCl 150 mM, glicerol 5% e EDTA 0,5 mM, separadamente. Após a diálise, a protease TEV foi

adicionada à solução na razão enzima:substrato de 1:20. A reação de digestão foi incubada a 20 °C por 18 horas.

Para a repurificação da proteína clivada, a concentração de imidazol na amostra digerida foi ajustada para 20 mM e a solução reaplicada em coluna HisTrap™ FF de 5mL, previamente equilibrada com tampão Tris 20 mM pH 8, contendo NaCl 150 mM, glicerol 5% e imidazol 20 mM. A fração não retida a coluna, contendo a proteína clivada foi lavada com o mesmo tampão e a fração das proteínas ligadas a coluna foi eluída com 270 mM de imidazol. Todas as frações foram analisadas por SDS-PAGE e a funcionabilidade da proteína após a clivagem foi confirmada por atividade hemaglutinante contra eritrócitos de coelho a 3%.

6.2.7 Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil)

A construção plasmidial rAmLpCHT3GNco, que apresentou maior rendimento de expressão, foi usada para as etapas posteriores de caracterização bioquímica e estrutural. Após as melhores condições de expressão serem estabelecidas, a proteína recombinante foi produzida em larga escala e purificada. Os testes de caracterização foram realizados com a proteína não clivada, dialisada em tampão Tris 20 mM pH 8, NaCl 150 mM e Glicerol 5%.

6.2.7.1 Estimativa da massa molecular e avaliação da pureza da lectina

A estimativa da massa molecular e o grau de pureza da lectina foram avaliados por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE), corado com Coomassie Brilliant Blue (Laemmli, 1970), segundo descrito na seção 4.2.8.1. O grau de pureza da lectina nativa usada nos testes bioquímicos também foi avaliada.

A massa molecular foi estimada na ausência e na presença de β mercaptoetanol, comparando a migração da proteína com a migração de marcador molecular: HSA – 66 kDa (NP_0000468, 66kDa); rSFL - 30 kDa (7UMJ_AA; Chaves *et al.*, 2023) e CCL – 15 kDa (C0HLX7.1; Marques *et al.*, 2018), e o marcador comercial SigmaMarker™ com o intervalo de 6,5 kDa a 66 kDa.

6.2.7.2 Ensaio de atividade hemaglutinante da proteína recombinante e nativa

O ensaio de hemaglutinação foi realizado com a proteína recombinante e com a proteína nativa, a fim de comparar a preferência de ambas por eritrócitos de coelho tratados ou não tratados com enzima proteolítica. Os eritrócitos foram utilizados na forma nativa e na forma tratada com a enzima proteolítica tripsina, conforme a preferência da lectina em estudo estabelecida anteriormente (Costa *et al.*, 1999).

A lavagem dos eritrócitos, a preparação da suspensão a 3% e o tratamento com a enzima proteolítica foi conduzido conforme descrito anteriormente na seção 3.2.7.

6.2.7.3 Ensaio de inibição da atividade hemaglutinante com a proteína recombinante e nativa

A proteína recombinante rAmL foi submetida a um ensaio de inibição da atividade hemaglutinante contra eritrócitos de coelho tratado com tripsina para determinar a especificidade a carboidratos da lectina. A proteína nativa também foi submetida ao mesmo ensaio, a fim de comparação da afinidade.

A concentração mínima inibitória do açúcar capaz de inibir a aglutinação de eritrócitos (CMI – Concentração Mínima Inibitória) também foi determinada (Sampaio *et al.*, 2002). Os carboidratos e glicoproteínas utilizados estão listados na seção 4.2.8.3.

6.2.7.4 Dicroísmo circular

A estrutura secundária da lectina recombinante foi analisada por espectroscopia de dicroísmo circular (CD) utilizando um espectropolarímetro Jasco J-815 (Jasco International Co., Tokyo, Japan), equipado com sistema de controle de temperatura Peltier.

As análises foram conduzidas com a proteína a 234 µg/mL em tampão Tris 20 mM pH 8, NaCl 150 mM e glicerol 5% a 20 °C, como descrito na seção 4.2.8.4. Os espectros obtidos foram processados e a predição da estrutura secundária foi realizada pela plataforma BestSel (Kardos *et al.*, 2025; Micsonai *et al.* 2022).

A estabilidade térmica da proteína também foi avaliada por dicroísmo circular (CD) em modo de variação térmica, usando uma cubeta de quartzo com caminho óptico de 0,1 cm. O comprimento de onda foi fixado em 215 nm, região sensível a variações conformacionais de estruturas β-folha, e a variação de temperatura foi conduzida de forma controlada, com o aumento de 3 °C por minuto, em um intervalo de temperatura de 20°C à 100°C. A elipticidade foi registrada continuamente ao longo do aquecimento.

Os dados obtidos foram processados utilizando a plataforma BestSel (Kardos *et al.*, 2025; Micsonai *et al.* 2022). A ferramenta *Thermal Denaturation Analysis* foi empregada para ajustar os dados experimentais a uma curva sigmoidal, e a temperatura média de transição térmica (Temperatura de *melting* - T_m) foi definida como o ponto correspondente a 50 % da variação total de elipticidade entre os estados nativo e desnaturado. A qualidade do ajuste do modelo foi validada com base no coeficiente de determinação (R²) e na distribuição dos resíduos entre valores experimentais e ajustados.

6.2.7.5 Espalhamento dinâmico de luz - DLS (*Dynamic Light Scattering*)

O comportamento hidrodinâmico da lectina recombinante foi avaliado pela análise de DLS em diferentes concentrações proteicas e condições de tampão, com e sem cálcio, a fim de identificar os parâmetros que minimizam a formação de agregados, e consequentemente apresentam menor tendência a precipitação.

As amostras foram analisadas em tampão Tris 20 mM pH 8, NaCl 150 mM e glicerol 5%, com e sem CaCl₂ 20 mM, nas concentrações de 0,5; 0,250 e 0,125 mg/mL, conforme detalhado na seção 4.2.8.5.

O diâmetro hidrodinâmico da lectina foi calculado a partir das funções de autocorrelação de intensidade, utilizando o modelo de análise fornecido pelo fabricante. Cada condição foi analisada em triplicata, e cada medida representa a média de múltiplas varreduras.

6.2.8 Atividade antibacteriana combinada com antibióticos (*Brasil*)

A lectina recombinante não clivada foi submetida a um teste de estabilidade térmica para avaliar se a cauda interferiria na estabilidade da proteína para os testes biológicos executados a 37 °C.

A lectina solúvel em tampão Tris 20 mM pH 8, NaCl 150 mM e glicerol 5%, na concentração de 0,5 mg/mL, foi incubada em termobloco a 37 °C por 1 hora. A atividade hemaglutinante da proteína foi avaliada antes e depois do aquecimento, utilizando eritrócitos de coelho tratados com tripsina a 3% pelo método da dupla diluição seriada (Sampaio *et al.*, 2002).

Além disso, a amostra da lectina usada no teste biológico foi submetida a um ensaio de atividade hemaglutinante para determinar a intensidade da atividade na concentração empregada no ensaio antibacteriano, seguindo o mesmo método descrito acima.

6.2.8.1 *Microorganismos e Condições de Cultivo*

As cepas bacterianas *Staphylococcus aureus* ATCC 25923, *S. aureus* ATCC 700698 e *Escherichia coli* ATCC 11303 foram utilizadas no ensaio. As bactérias foram cultivadas em ágar Mueller-Hinton (MHA; HiMedia, Índia) e incubadas a 37 °C por 24 h. Após o crescimento em meio sólido, as colônias isoladas foram recuperadas, inoculadas em caldo Mueller-Hinton (MHB) e incubadas a 37 °C por 24 h. As células foram coletadas por centrifugação a 9.000 g por 5 min a 4 °C, e foram ressuspensas em MHB. A densidade celular de cada suspensão bacteriana foi ajustada para uma concentração final de 5×10^5 UFC mL⁻¹, com base na densidade óptica a 620 nm.

6.2.8.2 *Avaliação do efeito de rAmLpCHT3GNco combinado aos antibióticos*

A avaliação do efeito da lectina rAmLpCHT3GNco combinada aos antibióticos foi realizada através do método *checkerboard*, descrito no trabalho realizado por Duarte e colaboradores (2023) com adaptações.

O termo *checkerboard* refere-se a múltiplas diluições com dois agentes antimicrobianos em concentrações equivalentes, abaixo ou acima das CIMs (Concentração Inibitória Mínima) para os microrganismos testados. Este método consiste em colunas contendo concentrações diferentes da droga “A” diluídas ao longo do eixo x, e linhas contendo as concentrações da droga “B” ao longo do eixo y (Lorian, 2005).

No presente estudo, foram avaliadas as combinações de rAmLpCHT3GNco com os antibióticos oxacilina, tetraciclina, ampicilina e estreptomicina contra as bactérias cultivadas citadas acima. Para cada combinação testada foram usados 84 poços da microplaca de 96 poços de fundo em "U", correspondentes às linhas A a E. O antibiótico foi diluído no sentido vertical da placa, das linhas A a E, nas concentrações correspondes à CIM: 1/2 x CIM, 1/4 x CIM, 1/8 x CIM e 1/16 x CIM. A lectina rAmLpCHT3GNco foi adicionada na concentração constante de 62,5 µg/mL nas colunas de 1 a 12.

Tanto os antibióticos quanto a lectina foram diluídos separadamente em placas de microtitulação, e logo em seguida foram misturados em uma nova placa, que foi preenchida com 50 µL do antibiótico, 50 µL da lectina e 100 µL do inóculo bacteriano previamente ajustado (5×10^5 UFC/mL). A linha F foi usada como controle da CIM do antibiótico e a linha G como controle do crescimento bacteriano.

Em seguida, as placas foram incubadas a 37°C por 24 horas e para a determinação da nova CIM em combinação com a lectina, a densidade óptica (D.O) foi mensurada no comprimento de onda de 620 nm em leitor de microplacas (Spectramax). Para bactérias que apresentaram valores iguais a 1/2x CIM, consideramos efeito aditivo; valores entre 1/4 e 1/16, efeito sinérgico, valores 1x CIM, não interação, 2x ou 4x CIM, consideramos como efeito antagônico.

6.3 Resultados e discussão

6.3.1 Produção da proteína recombinante rAmL em sistema heterólogo de *E. coli* (França)

6.3.1.1 Subclonagens do gene sintético para os vetores de expressão

O gene sintético da lectina rAmL, com 821 pb incluindo os sítios de restrição, foi inicialmente clonado no vetor *backbone* pEX-A258, delimitado pelos sítios de restrição *NcoI* e *XhoI* e com a sequência gênica otimizada para *Escherichia coli*. O quadro aberto de leitura (*Open Reading Frame* - ORF) possui 810 pb, codificando uma proteína de 270 aminoácidos. Após a digestão, o inserto foi subclonado nos vetores de expressão pCOLD-TF-TEV-Nco, pET-TEV e pCOLD-HAT-TEV-3G-Nco.

A confirmação da subclonagem foi obtida por PCR e digestão enzimática, de acordo com o tamanho esperado para rAmL e os respectivos vetores, visualizado em gel de agarose. Todas as construções apresentaram o inserto no tamanho esperado, e a subclonagem em todos os vetores foi confirmada por sequenciamento (dados não mostrados).

A *E. coli* é o organismo hospedeiro mais utilizado para a produção de proteínas recombinantes heterólogas. Se as proteínas heterólogas não precisam de modificações pós-traducionais complexas, esse é o sistema ideal e preferível a ser usado (Fathi-Roudsari; Akhavian-Tehrani; Maghsoud, 2016; Kaur; Kumar, Kaur, 2018).

Como explicado anteriormente, os aminoácidos são codificados por mais de um códon e cada organismo tem sua preferência de utilização desses diferentes códons, alguns são mais utilizados que outros. Por isso, é importante construir um gene sintético onde os códons raros sejam substituídos por aqueles mais usados pela *E. coli*.

Pelo impacto dos códons raros, o gene sintético de rAmL foi otimizado para *E. coli*, e a escolha dos vetores foi baseada em estratégias para obter um bom rendimento de expressão e de purificação, visto que rAmL não possui modificações pós-traducionais, como pontes dissulfeto, e provavelmente não teria dificuldades de expressão nesse sistema hospedeiro.

O vetor pCOLD-TF-TEV-Nco possui o TEE (*Transcriptional Enhancer Element*) e a chaperona molecular *Trigger factor* que podem aumentar a solubilidade da proteína recombinante. Nesse vetor, o TEE e o *Trigger factor* são expressos antes da proteína de interesse, resultando em uma proteína fusionada de maior tamanho, podendo favorecer o dobramento correto da proteína. Além disso, o vetor pCOLD permite a expressão da proteína

em temperaturas reduzidas, o que pode melhorar a solubilidade da proteína. O vetor tem a cauda de histidina como *tag* de purificação e possui sítios de clivagem para diferentes proteases.

No vetor pET-TEV, a construção com os sítios de restrição *NcoI* e *XhoI* permite a expressão da proteína sem nenhuma cauda fusionada. Essa estratégia permite a obtenção de uma proteína igual a nativa sem alterações estruturais adicionais, permitindo a purificação com o mesmo protocolo da lectina nativa.

No vetor pCOLD-HAT-TEV-3G-Nco, a proteína de interesse é fusionada a uma cauda pequena, composta apenas pelo TEE, cauda de purificação com histidina e o sítio de clivagem da protease TEV. Essa estratégia pode melhorar a purificação e minimiza alterações estruturais da proteína.

6.3.1.2 Expressão da lectina *rAmL* em sistema heterólogo

Todas as construções gênicas foram transformadas na cepa de clonagem *E. coli* NEB5 α e na cepa de expressão *E. coli* BL21 (Star), após confirmação das subclonagens por sequenciamento.

A cepa de clonagem foi utilizada para a propagação do vetor de expressão. A escolha da cepa de expressão foi baseada nos elevados níveis de rendimento observados para *E. coli* BL21 (Star) e na estrutura de *rAmL*, que não possui modificações pós-traducionais, portanto não sendo necessário utilizar cepas geneticamente modificadas para driblar as limitações da expressão em *E. coli*.

Para expressão da proteína, a cultura foi crescida a 37 °C, até atingir a densidade ótica de 0,4. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu até atingir DO = 0,8, quando o indutor IPTG foi adicionado na concentração de 0,2 mM e a expressão continuou a 16 °C.

Para o vetor pCOLD-HAT-TEV-3G-Nco, a expressão foi realizada em meio de autoindução LB NYZ. As células foram cultivadas a 37 °C até atingir a DO_{600nm} superior a 5,0. Em seguida, a temperatura foi reduzida para 16 °C e o cultivo prosseguiu por 18 horas.

A temperatura de cultivo a 37 °C foi escolhida seguindo as recomendações para a cepa *E. coli* BL21 Star, pois é a temperatura ótima de crescimento da estirpe (Tuttle; Trahan; Son, 2021). A temperatura de 16 °C para indução foi usada com o objetivo de ajudar na solubilidade da proteína recombinante.

6.3.1.2.1 Expressão da lectina rAmL no vetor pCOLD-TF-TEV-Nco

A construção da proteína recombinante expressa no vetor pCOLD-TF-TEV-Nco origina uma proteína de fusão composta pelo TEE, a cauda de histidina, a chaperona *Trigger factor*, os sítios de clivagem para as proteases HRV 3C, trombina e TEV, e finalmente, a lectina de interesse (Figura 61).

Figura 61 – Representação esquemática da proteína recombinante expressa no vetor pCOLD-TF-TEV-Nco.

>rAmL expressa em pCOLD-TF-TEV-Nco

```

10      20      30      40      50      60      70      80
MNHKVHHHHH HMQVSVETQ GLGRRVTIT AADSIETAVK SELVNVAKKV RIDGFRKGKV PMNIVAQRYG ASVRQDVLGD
90      100     110     120     130     140     150     160
LMSRNFIDAI IKEKINPAG PTYVPGEYKL GEDFTYSVEF EVYPEVELQG LEAIEVEKPI VEVTADADVDG MLDTLRKQQA
170     180     190     200     210     220     230     240
TWKEKDGA VE AEDRVTIDFT GSVDGEEFEG GKASDFVLAM GQGRMIPGFE DGIKGHKAGE EFTIDVTFPE EYHAENLKGK
250     260     270     280     290     300     310     320
AAKFAINLKK VEERELPELT AEFIKRFGVE DGSVEGLRAE VRKNMERELK SAIRNRVKSQ AIEGLVKAND IDVPAALIDS
330     340     350     360     370     380     390     400
EIDVLRROAA QRFGGNEKQA LELPRELFEE QAKRRVVVGL LLGEVIRTNE LKADEERVKG LIEEMASAYE DPKEVIEFYS
410     420     430     440     450     460     470     480
KNKELMDNMR NVALEEQAVE AVLAKAKVTE KETTFNELMN QQASAGLEVL FQGSPAGLVP RSGGGENLYF QGAMGRRYTV
490     500     510     520     530     540     550     560
ENQWGGSSAP WNAAGLWVLG GRPKQPCVAV DVSSTDGGST LTGTMTYKGE GPIGFKGIRK DGGNVYEVKN QWGGSSAPWH
570     580     590     600     610     620     630     640
DGGEWVIGSR TSQSVIELKI KSSDSGAHFD GTMKYSGEGF IGFQATMSDG DTYDVENQWG GSSAPWHKAG VWALGDRAFQ
650     660     670     680     690     700     710     720
NIIMLKVTSP DGGKTLTGDM QYEREGPIGF RGTLSQANNY AVENQWGGSH APWHKAGNWI IGDRDGQKV VVAITSSDNG
730     740     743
KTFEGTMTYS REGPIGFKGI LSS

```

TEE

Cauda de histidina

Trigger factor

Sítio da protease HRV 3C

Sítio de trombina

Sítio de clivagem TEV

Lectina

Número de aminoácidos: 743

Peso molecular: 81361,3372

pI: 5,22

Coeficiente de extinção:

O coeficiente de extinção é: 96830

Em 1mg/ml (12,291 uM) a solução da proteína tem

Abs280nm de: 1,19

Fonte: Autora. pI: Ponto isoelétrico. TEE: *Translation Enhancing Element*. HRV 3C: Protease do rinovírus humano. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

No vetor pCOLD-TF-TEV-Nco, apenas a cauda fusionada à rAmL, sem a proteína de interesse, possui 473 aminoácidos com um peso molecular de 52.586,54 Da e com pI de 5,04. A proteína inteira expressa, com a lectina de interesse, possui 743 aminoácidos com um peso molecular de 81.361,34 Da e com pI de 5,22.

A proteína expressa foi pensada para ser clivada com a protease TEV. Após clivagem, rAmL possui 272 aminoácidos com peso molecular de 28.920,94 Da e com pI de 6,24. A Figura 62 mostra a representação da proteína de interesse rAmL após clivagem com a protease TEV.

FIGURA 62 – Representação da proteína rAmL obtida após clivagem com a protease TEV.

>rAmL clivada

```

      10      20      30      40      50      60      70      80
GAMGRRYTV E NQWGGSSAPW NAAGLWVLGG RPKQPCVAVD VSSTDGGSTL TGTMTYKGE PIGFKGIRK GGNVYEVKNQ
      90     100     110     120     130     140     150     160
WGGSSAPWHD GGEWVIGSR T SQSVIELKIK SSDSGAHFDG TMKYSGEGPI GFQATMSDGD TYDVENQWGG SSAPWHKAGV
      170     180     190     200     210     220     230     240
WALGDRAFQN IIMLKVTSPD GKTTLTGDMQ YEREGPIGFR GTLSQANNYA VENQWGGSHA PWHKAGNWII GDRDGGQKVVG
      250     260     270     272
VAITSSDNGK TFEGTMTYSR EGPFGKGI L SS

```

Sítio TEV residual

Linker residual

Lectina

Número de aminoácidos: 272

Peso molecular: 28920,9402

pI: 6,24

Coefficiente de extinção:

O coeficiente de extinção é: 77920

Em 1mg/ml (34.577 uM) a solução da proteína tem

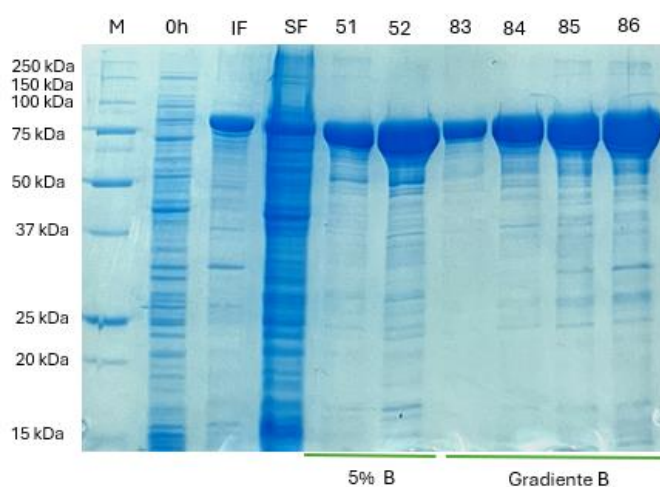
Abs280nm de: 2,69

Fonte: Autora. pI: Ponto isoeletrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

A expressão de rAmLpCOLD-TF-TEV-Nco foi então avaliada experimentalmente. Após o tempo de expressão, a fração solúvel e insolúvel foi analisada por SDS-PAGE.

Como o vetor pCOLD-TF-TEV-Nco possui cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados (IMAC). As proteínas retidas a coluna foram eluídas com 5% do tampão B, e com um gradiente linear de 5% a 100% do mesmo tampão. As frações foram coletadas e analisadas em SDS-PAGE (Figura 63).

Figura 63 - Análise da purificação de rAmLpCOLD-TF-TEV-Nco em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. 0h: Aliquota retirada antes da indução. IF: Fração insolúvel. SF: Fração solúvel. Números: tubos coletados. Tubos 51 e 52: eluídos com 5% de B. Tubos 83 a 86: eluídos com gradiente de tampão B.

A proteína de interesse rAmLpCOLD-TF-TEV-Nco foi expressa tanto na fração insolúvel como na fração solúvel, detectada pela banda visível no gel que não está presente na amostra coletada antes da indução (0h), entre as bandas de 100 e 75 kDa do marcador, corroborando com o tamanho esperado da proteína de aproximadamente 81 kDa.

A análise em SDS-PAGE confirmou a ligação e a recuperação da proteína na coluna, observada pela banda do tamanho esperado de 81 kDa nos dois picos eluídos. Em ambas as frações eluídas, foram observados contaminantes junto com a proteína de interesse, evidenciado pelas múltiplas bandas presente no gel. A intensidade da banda em SDS-PAGE e o elevado pico em absorbância das frações eluídas da coluna, mostram a eficiente expressão da proteína de interesse.

A estratégia de eluição inicial com baixa concentração de imidazol para ajudar na separação dos contaminantes não foi eficiente na purificação de rAmLpCOLD-TF-TEV-Nco, pois a proteína já começou a ser eluída no primeiro pico com 5% do tampão B, que totaliza 35 mM de imidazol. Além disso, mesmo em concentrações elevadas de imidazol, a proteína de interesse foi eluída com bastante contaminante.

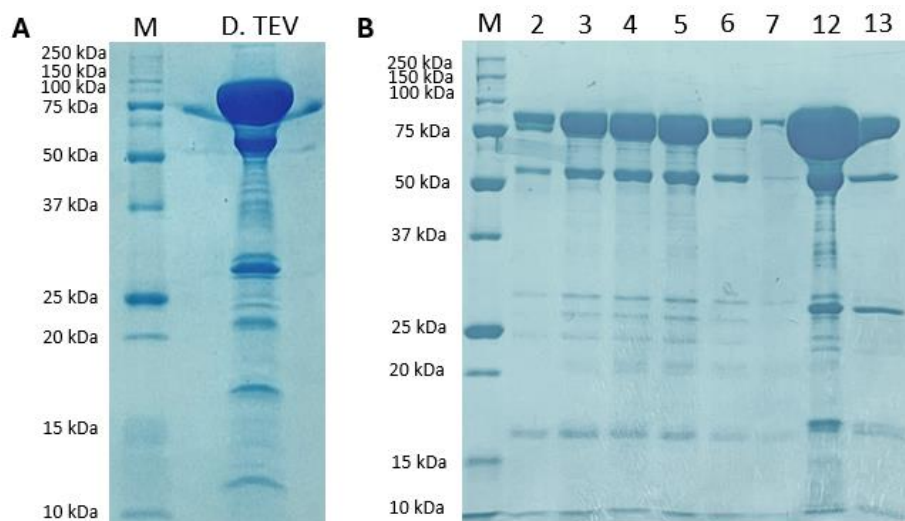
Mesmo na presença dos contaminantes após a cromatografia de afinidade, a proteína rAmLpCOLD-TF-TEV-Nco seguiu para a próxima etapa de purificação, a remoção da cauda fusionada por clivagem enzimática, uma vez que esse procedimento frequentemente contribui para a separação dos contaminantes (Waugh, 2011).

Os tubos coletados referentes a eluição com 5% de B (49 a 70) pareciam conter menos contaminantes e por isso foram reunidos e seguiram para digestão com a protease TEV com o objetivo de obter a proteína rAmL clivada.

Como a fração contendo 5% de tampão B possui baixa concentração de imidazol, uma diluição 1:1 em tampão Tris 20 mM pH 8, foi suficiente para reduzir a concentração de NaCl e de imidazol em quantidades que não interferem na eficiência da clivagem enzimática. A proteína foi concentrada até 2 mg/mL e a protease TEV foi adicionada a solução, juntamente com 0,5 mM de EDTA.

Após a reação de digestão, a concentração de imidazol foi ajustada e a solução foi submetida novamente a cromatografia de afinidade IMAC. A coluna foi lavada com o novo tampão contendo NaCl 100 mM e imidazol 10 mM, e a fração das proteínas não clivadas, ligadas a coluna, foi eluída com um step de 50% do tampão B. As frações foram coletadas e analisadas em SDS-PAGE (Figura 64).

Figura 64 - Análise da purificação de rAmL clivada em SDS-PAGE.



Fonte: Autora. Gel de 12%. A: Análise da solução de proteína após digestão enzimática com TEV e antes da cromatografia. D. TEV: Depois da digestão e antes da cromatografia. B: Análise da clivagem e purificação após a cromatografia. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 2 a 7: fração lavada da coluna. Tubos 12 e 13: eluídos com 50% de tampão B.

Após a digestão, o esperado é que a maioria das proteínas sejam clivadas, porém o pico retido a coluna teve uma absorbância muito maior que o pico não retido (dados não mostrados), demonstrando que provavelmente a clivagem não havia sido efetiva.

A análise em SDS-PAGE da amostra após a clivagem e antes de ser submetida a cromatografia (Figura 64 – A) confirmou a hipótese, mostrando que havia proteínas clivadas, evidenciada pela presença da banda acima de 25 kDa do marcador, que corrobora com o tamanho esperado para a proteína clivada de aproximadamente 29 kDa. No entanto, é observado a presença de proteínas não clivadas pela banda de maior intensidade entre 75 kDa e 100 kDa, com o tamanho esperado para rAmLpCOLD-TF-TEV-Nco de 81 kDa, demonstrando que a clivagem não foi eficiente.

As frações da cromatografia analisadas em SDS-PAGE (Figura 64 – B) mostraram que além da clivagem não ter sido eficiente, a separação também não foi efetiva, pois ambas as frações possuíam contaminantes e três bandas majoritárias. A banda mais baixa referente a rAmL clivada com o tamanho esperado de 29 kDa, a banda mais próxima a 50 kDa referente a proteína fusionada expressa cortada sem a lectina de interesse, com o tamanho esperado de aproximadamente 53 kDa, e a banda de maior intensidade acima de 75 kDa referente as proteínas de interesse não clivadas.

A presença da proteína não clivada na fração lavada com baixa concentração de imidazol sugere redução da interação efetiva com a coluna de níquel. Isso pode acontecer quando a proteína está agregada ou em complexos, reduzindo o acesso da cauda de histidina na coluna (Bio Rad, 2016). De fato, foi observada bastante precipitação após a digestão enzimática, validando a hipótese de agregação.

Como a clivagem não foi efetiva, uma nova tentativa foi realizada trocando o tampão para Tris 20 mM pH 8 contendo NaCl 100 mM, com o objetivo de retirar qualquer imidazol residual, que pudesse atrapalhar a digestão. Além disso, 500 mM de Ureia foi adicionado para tornar o sítio de clivagem mais acessível.

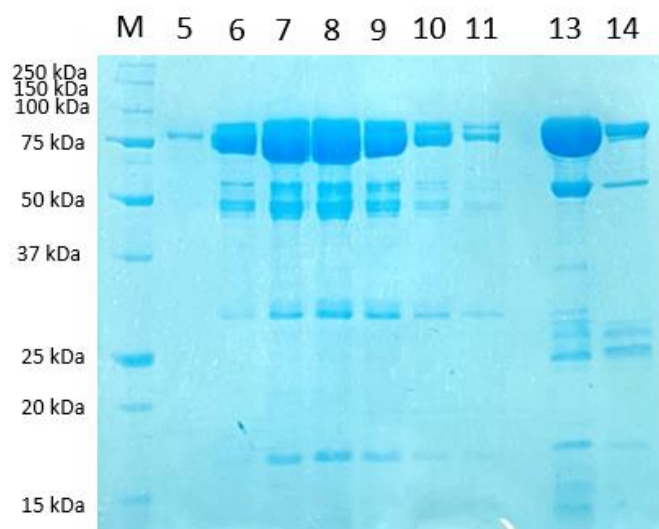
Muitas proteínas fusionadas formam oligômeros compactos, nesses casos, o sítio de clivagem da protease pode ficar estericamente inacessível. A ureia em baixas concentrações, 0,2 a 1M, promove uma desnaturação suave e desagrega complexos, aumentando a flexibilidade da proteína e expondo o sítio para a enzima. Inclusive, a TEV mantém a atividade em baixas concentrações de ureia, sendo uma estratégia viável para aumentar a eficiência da clivagem (Waugh, 2011).

Após a reação de digestão, a concentração de imidazol foi ajustada e a solução foi submetida novamente a mesma cromatografia de afinidade IMAC, para separação das proteínas clivadas e não clivadas (dados não mostrados). Novamente, o segundo pico, referente às proteínas ligadas a coluna, teve a absorbância máxima maior, demonstrando que mais uma vez a clivagem não foi eficiente. As frações foram coletadas e analisadas em SDS-PAGE (Figura 65).

A análise em SDS-PAGE confirmou que as novas estratégias usadas não foram eficientes para uma digestão efetiva. É possível observar a presença de proteínas clivadas, evidenciada pela banda acima de 25 kDa do marcador. No entanto, é observado em maior quantidade, a presença de proteínas não clivadas pela banda de maior intensidade entre 75 kDa e 100 kDa, com o tamanho esperado para rAmLpCOLD-TF-TEV-Nco de 81 kDa, demonstrando que mais uma vez a clivagem não foi eficiente.

Novamente, além da clivagem, a separação também não foi efetiva, pois ambas as frações possuíam contaminantes e com a presença das mesmas bandas majoritárias, a proteína clivada, não clivada e a cauda clivada.

Figura 65 - Análise em SDS-PAGE da purificação de rAmL após segunda clivagem



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 5 a 11: fração lavada da coluna. Tubos 13 e 14: fração eluída com 50% de tampão B.

Apesar da digestão não ter sido efetiva para remover a cauda fusionada à lectina de interesse, para confirmar a funcionabilidade da proteína expressa, alíquotas das amostras antes da clivagem e depois da clivagem, a mais de 1 mg/mL, foram submetidas a um teste de atividade hemaglutinante contra eritrócitos de coelho. A atividade hemaglutinante foi detectada, demonstrando que a lectina foi expressa de forma solúvel e ativa (dados não mostrados).

A adição de uma proteína de fusão ajuda na solubilidade pois além de cobrir parcialmente as regiões hidrofóbicas da proteína expressa, evitando a agregação, no momento do dobramento, a proteína de fusão também carrega a proteína de interesse fusionada, favorecendo o dobramento correto (Costa *et al.*, 2013; Rosano; Ceccarelli, 2014). No entanto, caudas de grande tamanho molecular também apresentam desvantagens.

Essas caudas podem apresentar dificuldades de clivagem enzimática devido ao impedimento estérico, pois como a proteína é muito grande, no momento do dobramento, o sítio de clivagem pode ficar inacessível para a enzima. Além disso, caudas muito grandes fusionadas à proteína de interesse podem apresentar alterações estruturais e funcionais. A proteína de fusão pode introduzir impedimento estérico, modificar a conformação da proteína e afetar sua oligomerização, interferindo no acesso ao sítio ativo de ligação. Embora essas caudas sejam usadas para aumentar a solubilidade, sua permanência pode comprometer a

atividade biológica da proteína alvo, o que torna a remoção dessa cauda uma etapa desejável (Waugh, 2011).

Ainda que a proteína de interesse foi expressa com solubilidade e funcionabilidade, como a clivagem não foi eficiente para separar a rAmL da enorme proteína de fusão presente no vetor pCOLD-TF-TEV-Nco, decidimos utilizar vetores alternativos, sem cauda ou contendo uma cauda de menor peso molecular para minimizar interferências estruturais e permitir uma proximidade maior das propriedades bioquímicas e biológicas com a proteína nativa.

6.3.1.2.2 Expressão da lectina rAmL no vetor pET-TEV

A construção da proteína recombinante expressa no vetor pET-TEV origina uma proteína sem cauda de fusão, composta apenas pela lectina de interesse (Figura 66).

Figura 66 – Representação esquemática da proteína rAmL no vetor pET-TEV.

>rAmL expressa em pET-TEV

```

      10      20      30      40      50      60      70      80
MGRRYTVENQ WGGSSAPWNÄ AGLWVLGGRP KQPCVAVDVS STDGGSTLTG TMTYKGEPTI GFKGIRKDG G NVYEVKNQWG
      90     100     110     120     130     140     150     160
GSSAPWHDGG EWVIGSRISQ SVIELKIKSS DSGAHFDGTM KYSGEPIGF QATMSDGDY DVENQWGGSS APWHKAGVWA
      170     180     190     200     210     220     230     240
LGDRAFQNI MLKVTSPDGG KTLTGDMQYE REGPIGFRGT LSQANNYAVE NQWGGSHAPW HKAGNWIIGD RDGQKVVGVA
      250     260     270
ITSSDNGKTF EGTMTYSREG PIGFKGILSS

```

Lectina

Número de aminoácidos: 270
 Peso molecular: 28792,8095
 pI: 6,24

Coefficiente de extinção:
 O coeficiente de extinção é: 77920
 Em 1mg/ml (34.731 uM) a solução da proteína tem
 Abs280nm de: 2,71

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

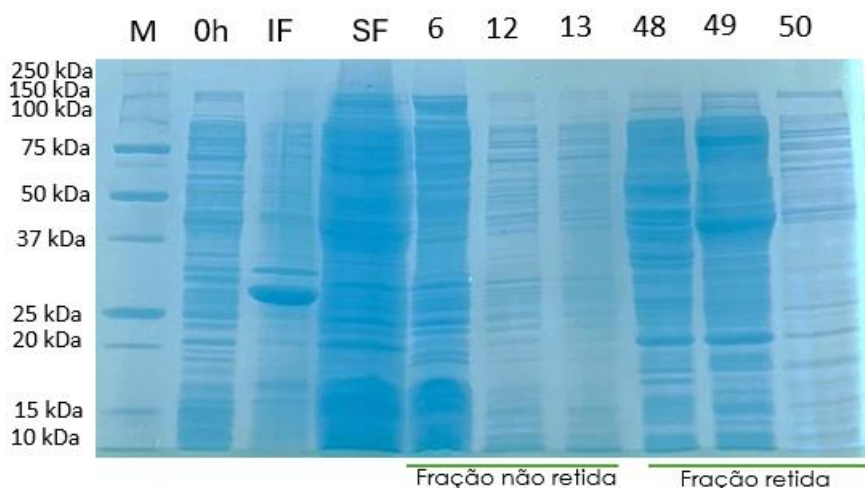
Apenas rAmL possui 270 aminoácidos com um peso molecular de 28.792,81 Da e com pI de 6,24.

A expressão de rAmLpET-TEV foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE.

Como a construção do vetor pET-TEV não possui cauda de histidina, a fração solúvel foi submetida à cromatografia de troca iônica, conforme o método usado para

purificação da lectina nativa (Silva, 2016). A fração solúvel, em tampão Tris 20 mM, pH 7, foi aplicada em uma coluna HiTrap™ DEAE FF de 1mL. A coluna foi lavada com o mesmo tampão sem NaCl e as proteínas retidas a coluna foram eluídas com um gradiente de NaCl 1M em Tris 20 mM, pH 7. As frações foram coletadas e analisadas em SDS-PAGE (Figura 67).

Figura 67 - Análise da purificação de rAmLpET-TEV em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. 0h: Aliquota retirada antes da indução. IF: Fração insolúvel. SF: Fração solúvel. Números: tubos coletados. Tubos 6, 12 e 13: fração não retida a coluna. Tubos 48 a 50: fração retida a coluna.

A proteína de interesse rAmLpET-TEV foi expressa apenas na fração insolúvel, detectada pela banda visível no gel que não está presente na amostra coletada antes da indução (0h). A banda acima de 25 kDa do marcador, corrobora com o tamanho esperado da proteína de aproximadamente 29 kDa. A mesma banda não foi detectada na fração solúvel, e esteve ausente nas frações da cromatografia.

Em ambas as frações obtidas na cromatografia, foi observada a presença de muitas proteínas contaminantes, evidenciada pelas múltiplas bandas presentes no gel e pelos picos de altíssima absorbância durante a purificação. Esse comportamento era esperado, pois a cromatografia de troca iônica não é específica, se baseando exclusivamente em interações eletroestáticas, assim qualquer proteína endógena do hospedeiro com carga líquida negativa pode se ligar a matriz.

Os sistemas de expressão pET são muito populares para produção de proteínas recombinantes em *E. coli* devido à sua alta produtividade, forte controle transcricional e elevada produtividade. Esses vetores contam com uma baixa expressão basal, possuem diferentes

variáveis de marcadores de afinidade, proteínas de fusão e sítios de clivagem (Rosano; Ceccarelli, 2014).

A elevada eficiência desse sistema vem do promotor T7. Esse promotor é muito utilizado para a expressão de proteínas recombinantes, pois é um promotor forte com alta produtividade. Nesse sistema, o gene de interesse é clonado sob um promotor reconhecido pela RNA polimerase do fago T7, induzível pelo sistema *operon lac*. Essa enzima é bastante seletiva e tão ativa que quase todos os recursos da célula são convertidos para a expressão do gene-alvo. Existem relatos na literatura que o produto desejado pode representar mais de 50% da proteína total da célula após algumas horas de indução (Mierendorf *et al.*, 2000).

No entanto, uma alta taxa de produção de proteínas recombinantes induz uma elevada carga metabólica no sistema hospedeiro. Essa carga metabólica é definida como a quantidade de recursos (matéria-prima e energia) que são retirados do metabolismo do hospedeiro para a manutenção e expressão do DNA exógeno. A rápida taxa de síntese proteica somada a uma grande carga metabólica, faz com que o sistema hospedeiro não consiga completar o dobramento correto das proteínas expressas. Assim, sem atingir a conformação correta, as regiões hidrofóbicas dessas proteínas mal enoveladas se associam umas às outras para formar agregados insolúveis de proteínas não nativas, chamados de corpos de inclusão (Singh *et al.*, 2015; Sorensen; Mortensen, 2005).

A formação de corpos de inclusão ocorre como resposta ao acúmulo de proteína desnaturada, e a probabilidade de dobramento incorreto aumenta com o uso rotineiro de promotores fortes e altas concentrações de indutores (Baneyx; Mujacic, 2004; Singh *et al.*, 2015).

Apesar de rAmL não possuir modificações pós-traducionais, o que teoricamente poderia favorecer sua expressão na forma solúvel, outros fatores intrínsecos do sistema podem ter contribuído para a expressão apenas na forma insolúvel. A presença de um promotor forte, como o T7, somada a elevada capacidade de síntese proteica da cepa *E. coli* BL21 Star, pode ter levado a uma alta taxa de expressão, resultando na formação dos corpos de inclusão. Além disso, a construção utilizada não possui proteínas fusionadas à lectina que poderiam ajudar no correto dobramento, e consequentemente na solubilidade da proteína. Todos esses fatores em conjunto podem ter levado rAmL à agregação.

6.3.1.2.3 Expressão da lectina rAmL no vetor pCOLD-HAT-TEV-3G-Nco

A construção da proteína recombinante expressa no vetor pCOLD-HAT-TEV-3G-Nco origina uma proteína de fusão composta pelo TEE, a cauda HAT rica em histidina, o sítio de clivagem para a protease TEV e a lectina de interesse (Figura 68).

Figura 68 – Representação esquemática da proteína recombinante rAmL expressa no vetor pCOLD-HAT-TEV-3G-Nco.

>rAmL expressa em pCOLD-HAT-TEV-3G-Nco

```

      10      20      30      40      50      60      70      80
MNHKVKDHLI HNVHKEEHAH AHNKGENLYF QGGGAMGRRY TVENQWGGSS APWNAAGLWV LGGRPKQPCV AVDVSTDG
  90      100     110     120     130     140     150     160
STLTGTMTYK GEGPIGFKGI RKDGGNVYEV KNQWGGSSAF WHDGGWVIG SRTSQSVIEL KIKSSDSGAH FDGTMKYSGE
 170     180     190     200     210     220     230     240
GPIGFQATMS DGDYTDVENQ WGGSSAPWHK AGVWALGDRA FQNIIMLKVT SPDGGKTLTG DMQYEREGPI GFRGTLSQLN
 250     260     270     280     290     300     305
NYAVENQWGG SHAPWHKAGN WIIGDRDQK VVGVAITSSD NGKTFEGTMT YSREGPIGFK GILSS

```

TEE

HAT-Tag

Sítio de clivagem TEV

Lectina

Número de aminoácidos: 305

Peso molecular: 32773,2072

pI: 6,67

Coefficiente de extinção:

O coeficiente de extinção é: 79410

Em 1mg/ml (30,513 uM) a solução da proteína tem

Abs280nm de: 2,42

Fonte: Autora. pI: Ponto isoelétrico. TEE: *Translation Enhancing Element*. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Apenas a cauda fusionada à rAmL, sem a proteína de interesse, possui 35 aminoácidos com um peso molecular de 3.998,41 Da e com pI de 7,58. A proteína inteira expressa, com a lectina de interesse, possui 305 aminoácidos com um peso molecular de 32.773,21 Da e com pI de 6,67.

A proteína expressa foi pensada para ser clivada com a protease TEV. Após clivagem, rAmL possui 274 aminoácidos com peso molecular de 29.035,04 Da e com pI de 6,24. A Figura 69 mostra a representação da proteína de interesse rAmL após clivagem com a protease TEV.

A expressão de rAmLpCOLD-HAT-TEV-3G-Nco foi então avaliada experimentalmente. Após o tempo de expressão, as células foram recuperadas, lisadas e a fração solúvel e insolúvel foi analisada por SDS-PAGE.

Figura 69 – Representação da proteína rAmL obtida após clivagem com a protease TEV.

>rAmL clivada

```

      10      20      30      40      50      60      70      80
GGGAMGRRYT VENQWGGSSA PWNAAGLWVL GGRPKQPCVA VDVSTDGGG TLTGTMTRYKG EGPIGFKGIR KDGGNVYEVK
      90     100     110     120     130     140     150     160
NQWGGSSAPW HDGGEWVIGS RTSQSVIELK IKSSDSGAHF DGTMKYSGEF PIGFQATMSD GDTYDVENQW GGSSAPWHKA
      170     180     190     200     210     220     230     240
GVWALGDRAF QNIIMLKVTS PDGGKTLTGD MQYEREGPIG FRGTL SQANN YAVENQWGGG HAPWHKAGNW IIGDRDGQKV
      250     260     270     274
VGVAITSSDN GKTFEGMTY SREGPIGFKG ILSS

```

Sítio TEV residual

Linker residual

Lectina

Número de aminoácidos: 274

Peso molecular: 29035,0440

pI: 6,24

Coefficiente de extinção:

O coeficiente de extinção é: 77920

Em 1mg/ml (34.441 uM) a solução da proteína tem

Abs280nm de: 2,68

Fonte: Autora. pI: Ponto isoelétrico. Os dados da cauda expressa estão disponíveis no banco de dados do CERMAV. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

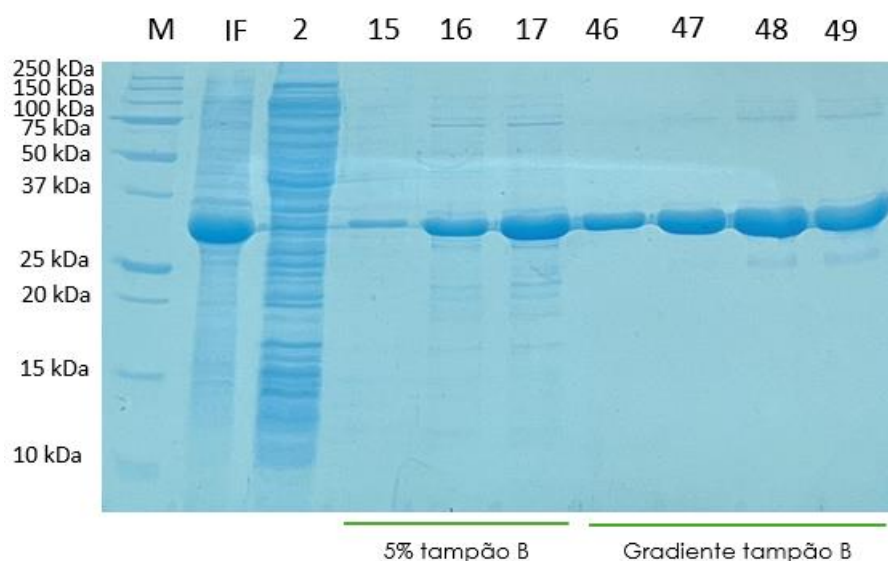
O meio de autoindução foi escolhido para a expressão de rAmLpCOLD-HAT-TEV-3G-Nco, com base nos resultados obtidos anteriormente com o vetor pCOLD-TF-TEV-Nco, no qual a proteína foi expressa de forma solúvel e ativa. Portanto, resolvemos otimizar a produção de rAmL usando um meio de maior eficiência de produção, promovendo maiores rendimentos.

Devido a precipitação observada durante a clivagem de rAmL no vetor pCOLD-TF-TEV-Nco, foi adicionado 5% de glicerol aos tampões de purificação da rAmLpCOLD-HAT-TEV-3G-Nco, com o objetivo de reduzir a agregação e aumentar a estabilidade da proteína expressa em solução.

O glicerol é comumente utilizado em soluções de proteínas recombinantes porque atua como um estabilizante que reduz processos de desnaturação e agregação. Esse composto aumenta a viscosidade do meio e diminui a frequência de colisões intermoleculares, evitando a agregação e mantendo a estabilidade conformacional e a solubilidade (Tripathi, 2016).

Como o vetor pCOLD-HAT-TEV-3G-Nco possui cauda de afinidade a histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados. As proteínas retidas a coluna foram eluídas com 5% do tampão B e com um gradiente linear de 5% a 100% do mesmo. As frações foram coletadas e analisadas em SDS-PAGE (FIGURA 70).

FIGURA 70 - Análise da purificação de rAmLpCOLD-HAT-TEV-3G-Nco em SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. IF: Fração insolúvel. Números: tubos coletados. Tubo 2: fração não ligada a coluna. Tubos 15 a 17: eluídos com 5% de B. Tubos 46 a 49: eluídos com gradiente de tampão B. IF: Fração insolúvel.

A proteína de interesse rAmLpCOLD-HAT-TEV-3G-Nco foi expressa tanto na fração insolúvel como na fração solúvel, detectada pela banda visível no gel entre 25 e 37 kDa do marcador, corroborando com o tamanho esperado da proteína de aproximadamente 33 kDa.

A análise em SDS-PAGE confirmou a ligação e a recuperação da proteína na coluna, observada pela mesma banda com tamanho esperado de 33 kDa nos dois picos eluídos. Em ambas as frações, foi observado a coeluição de contaminantes junto com a proteína de interesse, mesmo após a aplicação de uma etapa inicial de eluição com baixa concentração de imidazol, utilizada para reduzir a presença dos contaminantes. A elevada intensidade da banda correspondente à proteína de interesse em SDS-PAGE e o elevado pico em absorbância das frações eluídas da coluna, mostram a eficiente expressão da proteína recombinante.

A expressão de rAmL observada na fração insolúvel, provavelmente estar relacionada a uma taxa de expressão elevada, que excede a capacidade do maquinário celular, favorecendo a formação de proteínas mal dobradas, o que pode ter limitado a liberação completa da proteína solúvel. Além disso, o método de ruptura celular também pode não estar sendo tão efetivo para liberar toda a proteína do citoplasma.

Com o sucesso da expressão, a proteína rAmLpCOLD-HAT-TEV-3G-Nco seguiu para a etapa de remoção da cauda fusionada por clivagem enzimática. Os tubos referentes a eluição com 5% de tampão B (15 a 20), contendo menos imidazol, foram reunidos e seguiram

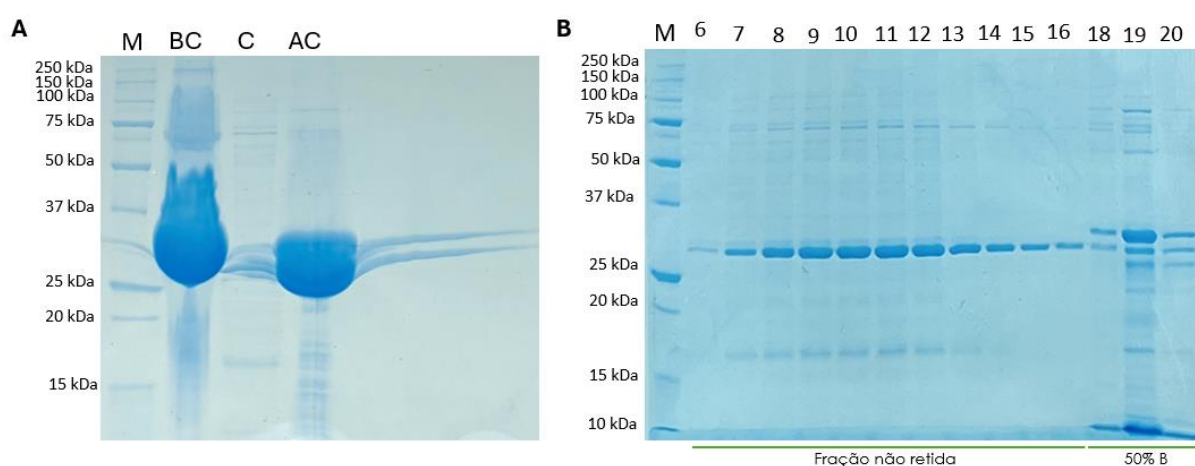
para digestão com a protease TEV, com o objetivo de obter a proteína rAmL clivada e possivelmente livre de contaminantes.

O procedimento foi o mesmo aplicado para a proteína expressa rAmL no vetor pCOLD-TF-TEV-Nco. Após a reação de digestão, a solução foi submetida a cromatografia de afinidade IMAC, onde a coluna foi lavada com o tampão contendo NaCl 100 mM, glicerol 5% e imidazol 10 mM, e a fração das proteínas ligadas a coluna, foi eluída com um step de 50% do tampão B.

O primeiro pico, referente a proteína de interesse clivada, teve a máxima absorbância de 1217 mAu. O segundo pico, referente às proteínas que se ligaram a coluna, foi eluído com o equivalente de 260 mM de imidazol e teve a absorbância máxima de 661 mAu. Esses dados indicam que a clivagem foi eficiente, visto que o maior pico de absorbância foi observado na fração que não se ligou a coluna. As frações foram coletadas e analisadas em SDS-PAGE (Figura 71).

Durante as etapas de diluição e concentração da amostra para a digestão enzimática, e mesmo ao longo de todo o período de clivagem, foi observado a ocorrência de precipitação. Esses precipitados foram analisados em SDS-PAGE para avaliar se a precipitação estava associada à protease TEV ou à própria proteína de interesse (Figura 71).

Figura 71 - Análise da purificação de rAmL clivada em SDS-PAGE.



Fonte: Autora. Gel de 12%. A: Análise do precipitado durante digestão enzimática. BC: Precipitação durante a concentração e diluição, antes de adicionar a TEV. C: Amostra no momento de adicionar a TEV, depois de retirar os precipitados anteriores por centrifugação. AC: Amostra depois do período de incubação com a TEV, antes da cromatografia. B: Análise da clivagem e purificação após a cromatografia. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Números: tubos coletados. Tubos 6 a 16: fração lavada da coluna. Tubos 18 a 20: fração eluída com 50% de tampão B.

A análise em SDS-PAGE do material precipitado (Figura 71 – A) demonstrou que uma fração significativa da proteína rAmL estava agregando e sendo perdida por precipitação,

evidenciada pela presença da banda intensa entre 25 e 37 kDa, correspondendo ao tamanho esperado da proteína. Grande parte dessa perda proteica aconteceu mesmo antes da adição da protease TEV, durante a etapa de concentração da proteína. Por isso, antes da clivagem, a amostra foi centrifugada para remoção do precipitado inicial. Porém, após o período de incubação com a TEV, foi novamente detectada a formação de precipitado contendo a proteína rAmL.

As frações da cromatografia analisadas por SDS-PAGE (Figura 71 – B) mostraram que a clivagem foi eficiente. Na fração não retida a coluna é possível observar a presença de uma banda majoritária entre 25 kDa e 37 kDa, compatível com o tamanho de 29 kDa para a proteína clivada. Na fração correspondente às proteínas retidas a coluna, a clivagem é confirmada pela presença das mesmas bandas observadas na fração não retida, porém com menor intensidade, além da detecção de uma banda adicional de maior intensidade e com massa ligeiramente superior, corroborando com o tamanho da proteína não clivada de 33 kDa. A presença da cauda cortada também é detectada na fração ligada a coluna pela banda abaixo de 10 kDa, compatível com o tamanho da *tag* de aproximadamente 4 kDa. Apesar da presença da proteína não clivada retida a coluna, a maior fração da proteína de interesse foi corretamente clivada.

A eficiência da clivagem depende da acessibilidade do sítio de reconhecimento para a protease. Conforme mencionado anteriormente, caudas de fusão muito extensas podem causar um impedimento estérico, dificultando o acesso da protease ao sítio. Por isso, caudas de fusão menores tendem a facilitar a clivagem ao diminuir a complexidade estrutural ao redor do sítio, aumentando a sua exposição e facilitando a clivagem (Rosano; Ceccarelli, 2014).

Além disso, esse vetor possui a adição de duas glicinas inseridas entre o sítio de clivagem TEV e a proteína de interesse. A adição dessas glicinas teve como objetivo criar uma região mais flexível e espaçadora para a protease TEV, permitindo que o sítio fique mais exposto e melhore o acesso da enzima ao sítio de clivagem, e consequentemente, aumente a eficiência da digestão (Waugh, 2011). A escolha do vetor com uma cauda de fusão pequena somado a adição das duas glicinas, de fato, favoreceu a clivagem, quando comparado aos resultados para rAmL no vetor pCOLD-TF-TEV-Nco.

Para confirmar a funcionabilidade da proteína expressa, alíquotas das amostras de interesse antes da clivagem e depois da clivagem, a 0,5 mg/mL, foram submetidas a um teste de atividade hemaglutinante contra eritrócitos de coelho. A atividade hemaglutinante foi

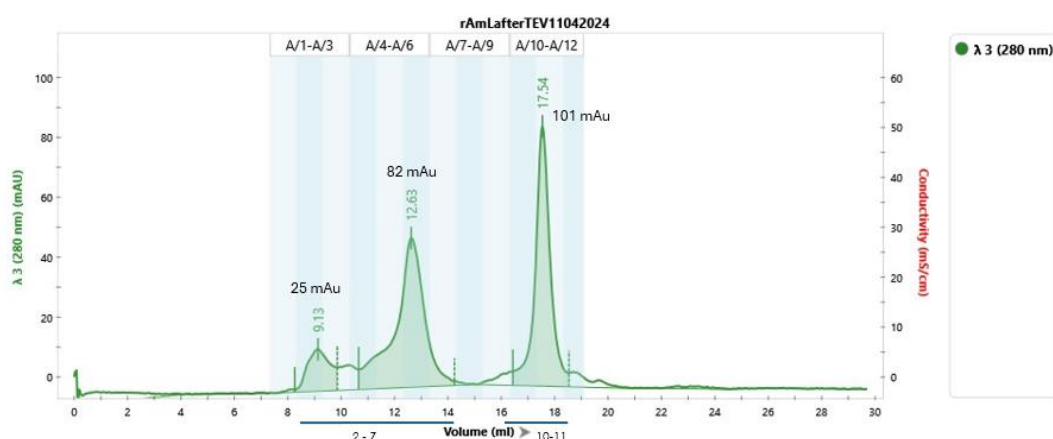
detectada em ambos os casos, demonstrando que a lectina foi expressa de forma solúvel e ativa (dados não mostrados).

Esses resultados mostram que a separação entre as proteínas clivadas e não clivadas também foi efetiva, pois na fração não retida a coluna, apenas a banda referente à rAmL clivada foi detectada. No entanto, a amostra digerida ainda possuía contaminantes.

Após a purificação inicial por IMAC, a amostra pode conter contaminação com as proteínas endógenas que se ligaram inespecificamente a coluna. Quando a clivagem enzimática para retirar a *tag* é realizada, como os contaminantes teoricamente não possuem sítio de clivagem, eles continuam interagindo e se ligando a coluna. Por isso, a digestão enzimática da proteína de interesse pode remover as contaminações. No entanto, a remoção de contaminantes nem sempre é completa porque as proteínas endógenas do hospedeiro podem interagir com a proteína de interesse, formando complexos proteína-proteína e não interagirem com a coluna de níquel. Assim, um novo passo de purificação é necessário para separar esses contaminantes.

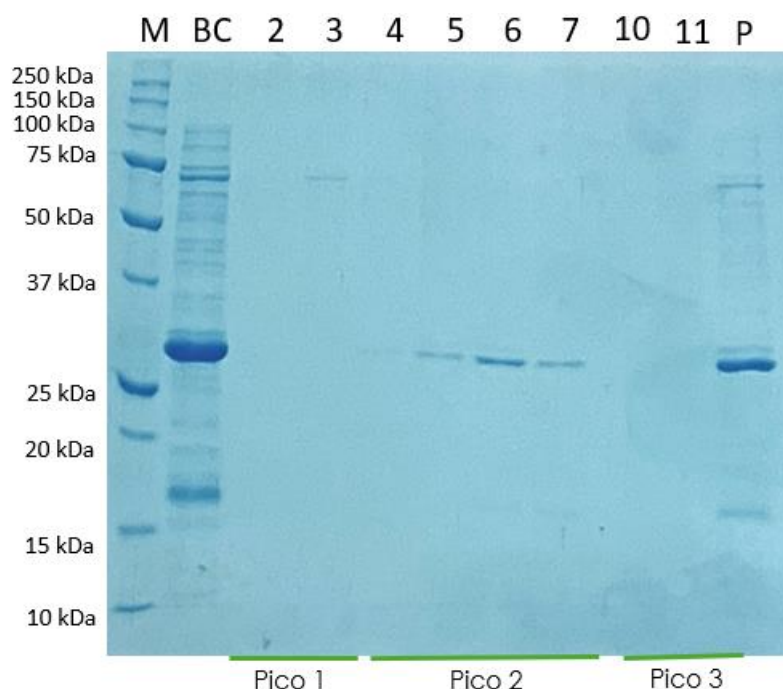
Como as proteínas contaminantes possuíam um tamanho molecular diferente da proteína de interesse (Figura 71), inicialmente, para completar a purificação da proteína clivada, a fração não retida a coluna de HisTrap após a clivagem (tubos 6 a 16) foi submetida a cromatografia de exclusão molecular em coluna Enrich™ SEC70. A coluna foi equilibrada e eluída com tampão Hepes 20 mM pH 7,5 com NaCl 150 mM, e amostra concentrada e aplicada a 1,5 mg/mL (Figura 72). As frações foram avaliadas em SDS-PAGE (Figura 73).

Figura 72 – Cromatograma de exclusão molecular da proteína rAmL clivada.



Fonte: Autora. Cromatograma fornecido pelo sistema de cromatografia de média pressão NGC (Bio-Rad) em coluna Enrich™ SEC70 (10 x 300 mm; 24 mL; Bio-Rad). A coluna foi equilibrada e eluída com tampão Hepes 20 mM pH 7,5 com NaCl 150 mM, e a amostra foi aplicada em um volume de 250 µL a 1,5 mg/mL em um fluxo de 1 mL/min. Eixo esquerdo: absorbância a 280 nm em mAu. A linha contínua verde indica a absorbância a 280 nm (mAu). A letra “A” com os números marcadas em azul indicam os tubos coletados. Os números em preto indicam as informações de maior absorbância dos picos. Os valores em verde indicam o volume de eluição do pico. Os números abaixo do pico indicam os tubos coletados e analisados em SDS-PAGE.

Figura 73 – Análise em SDS-PAGE da purificação de rAmL após cromatografia de gel filtração.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. SB: Amostra antes da cromatografia. P: Precipitado durante a concentração da amostra para aplicar na coluna. Números: tubos coletados. Tubos 2 e 3: pico 1. Tubos: 4 a 7: pico 2. Tubos 10 e 11: pico 3.

Durante a etapa de concentração da amostra para a aplicação na coluna, novamente foi observado a presença de precipitação. A análise do material precipitado em SDS-PAGE mostrou que, novamente, a proteína de interesse estava sendo perdida por agregação, devido a presença da banda intensa entre 25 kDa e 37 kDa, compatível com o tamanho esperado de 29 kDa para rAmL clivada.

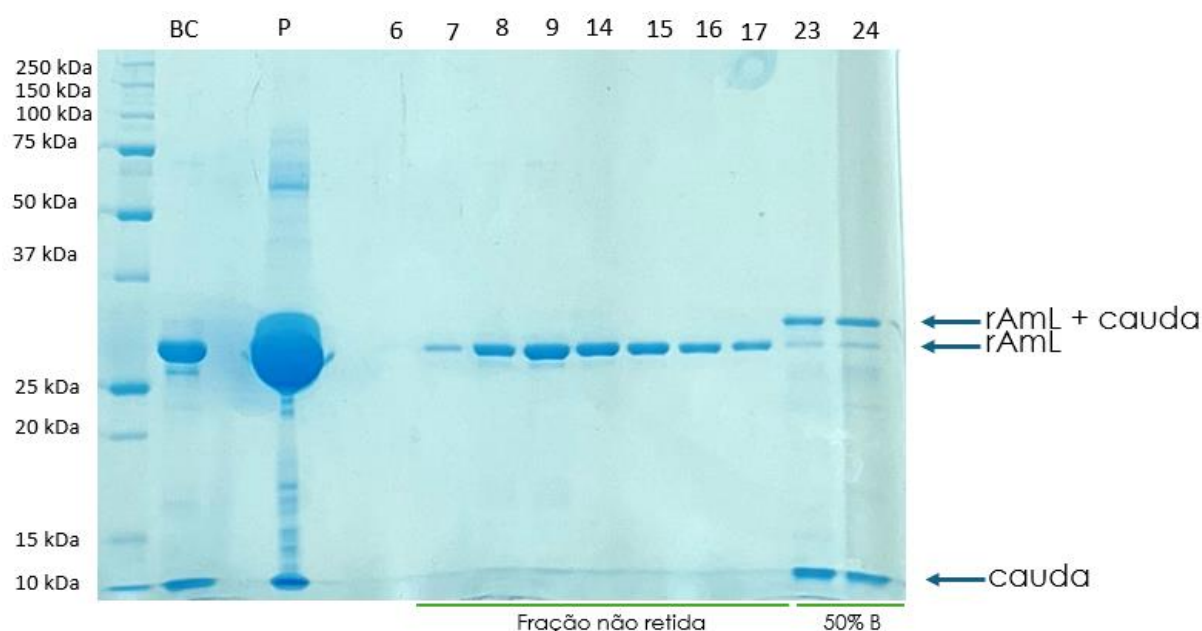
Na cromatografia de exclusão molecular foram identificados três picos, porém a rAmL foi identificada apenas no pico 2, pela presença da banda acima de 25 kDa na fração correspondente. Apesar do pico 3 ter apresentado a maior absorbância, não foi observada banda visível em SDS-PAGE nas frações referentes ao pico 1 e 3. Isso pode acontecer porque o detector UV monitora a absorbância de qualquer composto que absorva na região de 280 nm e não necessariamente apenas proteínas. A presença de ácidos nucleicos na amostra pode ter contribuído para uma absorção nessa região (Janson, 2011).

O objetivo da cromatografia de gel filtração foi alcançado, pois a amostra antes dessa etapa possuía a presença da proteína e dos contaminantes e após a cromatografia, a proteína presente no pico 2 apresentou elevado grau de pureza. A funcionabilidade da proteína

recombinante foi confirmada por atividade hemaglutinante (dados não mostrados). As frações da proteína rAmL pura, clivada e ativa foram reunidas para uso posterior.

Portanto, o protocolo de expressão e purificação da lectina recombinante rAmL estava inicialmente estabelecido. Porém, durante a continuidade dos experimentos de produção da rAmL, foi necessária a purificação de um novo lote da protease TEV otimizada, que apresentou uma maior eficiência de clivagem. Assim, a remoção da cauda fusionada passou a ocorrer de forma mais completa, evitando que contaminantes fossem coeluídos com a proteína clivada. Consequentemente, após a digestão e a recromatografia IMAC, a rAmL foi purificada com alto grau de pureza, sem a necessidade da etapa adicional da cromatografia de exclusão molecular (Figura 74).

Figura 74 – Análise em SDS-PAGE da purificação de rAmL após recromatografia com nova TEV otimizada.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. BC: Amostra antes da cromatografia. P: Precipitado. Números: tubos coletados. Tubos 6 a 17: pico não retido a coluna. Tubos 23 e 24: tubos retidos a coluna.

O novo protocolo resultou em maior rendimento da rAmL ao eliminar a etapa adicional da cromatografia de exclusão molecular. No processo de purificação proteica, cada etapa adicional implica em perdas inevitáveis de material. Por isso, minimizar o número de etapas de purificação é sempre mais vantajoso (Janson, 2011).

Apesar de rAmL ter sido expressa com elevado rendimento, evidenciado pela alta intensidade da banda em SDS-PAGE e pelos picos de absorbância observados durante as etapas

cromatográficas, e de a clivagem enzimática ter ocorrido de forma eficiente, aproximando estruturalmente a proteína recombinante da forma nativa, a rAmL apresentou uma forte tendência a precipitação ao longo de todas as etapas de purificação.

A precipitação foi observada tanto antes quanto depois da clivagem, resultando em perdas significativas de proteína. Por conta dessa agregação, não foi possível determinar com precisão o rendimento final por litro da proteína clivada, visto que grande parte da rAmL era perdida por precipitação, conforme evidenciado na figura 74.

Diversas estratégias foram adotadas para minimizar essa tendência. Por exemplo, a proteína deixou de ser congelada e os experimentos passaram a ser conduzidos preferencialmente em temperatura ambiente, uma vez que o armazenamento na geladeira e os experimentos conduzidos a 4 °C favoreciam a precipitação. Além disso, a concentração da amostra foi evitada sempre que possível, como na etapa da clivagem, sendo mantida em condições diluídas, o que não impactou a eficiência da digestão. Ensaio de deslocamento térmico também foram realizados para identificar os tampões e condições que pudessem dar uma maior estabilidade a proteína, resultados que serão mostrados e discutidos na próxima sessão.

A redução da temperatura pode favorecer a precipitação de proteínas em solução, pois o efeito hidrofóbico responsável pela estabilidade conformacional é dependente da temperatura. A desnaturação pelo frio é causada pela interação específica e fortemente dependente da temperatura entre os grupos apolares das proteínas e a água. Com a diminuição da temperatura, a cadeia polipeptídica, compactada em sua estrutura nativa, se desdobra expondo os grupos apolares internos à água, favorecendo a agregação (Privalov, 1990; Sanfelice; Temussi, 2016). A maior concentração de proteínas também gera uma maior colisão entre as moléculas, aumentando a precipitação. A concentração por ultrafiltração podem gerar concentrações locais elevadas que favorece a nucleação de agregados (Chi *et al.*, 2003; Maguire *et al.*, 2018).

Apesar de adotar todas essas medidas, evitar baixas temperaturas, não concentrar, adição de glicerol e uso de diferentes tampões, a combinação dessas estratégias reduziu apenas parcialmente a perda dessa proteína por agregação. A proteína rAmL continuou apresentando tendência para precipitação, indicando que a instabilidade observada é uma propriedade intrínseca da proteína, provavelmente devido a suas interações intermoleculares que favorece a agregação em diferentes condições.

A análise da sequência primária pelo programa *Protein parameters* revelou que a proteína rAmL possui elevado conteúdo de resíduos hidrofóbicos, superior a 50%. A alta hidrofobicidade é uma das causas reconhecidas de tendência a agregação, isso porque regiões apolares expostas em estados parcialmente desnaturados interagem entre si levando a precipitação. A expressão da proteína recombinante pode não ter levado ao dobramento idêntico ao da proteína nativa, deixando os resíduos hidrofóbicos mais expostos, causando essas interações hidrofóbicas intermoleculares (Chi *et al.*, 2003; Rosano; Ceccarelli, 2014).

As lectinas naturais coexistem biologicamente com variações muito semelhantes da mesma proteína, as isoformas (Costa Filho *et al.*, 2025; Oliveira; Teixeira; Domingues, 2013). Uma possibilidade é que na forma nativa, a coexistência dessas isolectinas contribua com a estabilidade conformacional da proteína em solução, visto que cada isoforma da mesma proteína pode ter propriedades físico-químicas e papel biológico distintos (Oliveira; Teixeira; Domingues, 2013). Assim, como AmL nativa tem isoformas documentadas (Silva, 2016), possivelmente para se tornar estavelmente solúvel, a rAmL recombinante precisaria das isoformas para dar a estabilidade estrutural e consequentemente manter a solubilidade.

Alguns estudos sustentam essa hipótese, Qi e colaboradores (2017) relatam que uma proteína mitocondrial humana expressa na sua forma recombinante apresentou elevada tendência a agregação. Porém, quando as isoformas foram ligadas à essa proteína, elas foram capazes de bloquear as superfícies hidrofóbicas, impedindo a autoassociação dessas regiões, e mantendo a proteína estável e em solução. Portanto, a ausência das variantes estruturais presentes na forma nativa pode contribuir para a precipitação das proteínas recombinantes expressas como uma única isoforma.

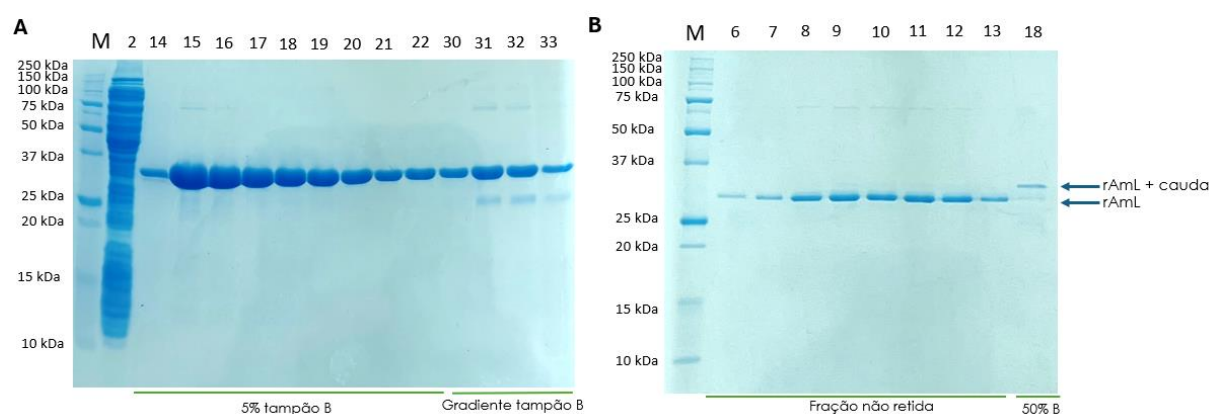
Existem muitos relatos na literatura de proteínas recombinantes com problemas de precipitação, mostrando que a solubilidade pode mudar drasticamente na presença ou na ausência de caudas de fusão (Rosano, Ceccarelli, 2014; Costa *et al.*, 2013). Inclusive, existem estudos de outras lectinas recombinantes de alga marinha que apresentaram alto índice de precipitação, impactando diretamente no rendimento da produção, as lectinas rHML e rBTL de *Hypnea musciformis* e *Alsidium triquetrum*, antigo *Bryothamnion triquetrum*, respectivamente (Nascimento, 2014).

Portanto, o protocolo de expressão heteróloga de uma nova lectina de alga marinha foi estabelecido. Na primeira etapa de cromatografia por afinidade IMAC, antes da clivagem, as frações eluídas apresentavam dois perfis distintos, frações contendo a proteína com alto grau de pureza e frações ainda acompanhadas de contaminantes.

Durante a eluição com baixa concentração de imidazol (5% do tampão B), a proteína passava a ser recuperada de forma progressivamente mais pura. Embora os primeiros tubos pudessem ter contaminantes, à medida que a eluição prosseguia nessa mesma concentração, as frações iam se tornando mais puras. Esse comportamento é explicado pela diferença de afinidade entre a cauda de histidina e as proteínas endógenas contaminantes, que são ligadas mais fracamente à coluna e são removidas nos volumes iniciais.

As frações da primeira cromatografia IMAC que ainda apresentam contaminação eram direcionadas para a etapa de clivagem com a protease TEV para recuperação da proteína clivada. A pureza das frações era sempre confirmada por SDS-PAGE (Figura 75).

Figura 75 – Frações de rAmL não clivadas e clivadas para uso posterior.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. A: Primeira etapa de purificação da cromatografia IMAC. Números: tubos coletados. Tubos 14 a 22: pico eluído com 5% de tampão B. Tubos 30 a 33: pico eluído com gradiente de tampão B. A: Segunda etapa de purificação da recromatografia IMAC. Números: tubos coletados. Tubos 6 a 13: pico não retido a coluna. Tubo 18: pico retido a coluna.

As frações não clivadas e clivadas puras e ativas eram reunidas para uso posterior.

Diversas proteínas de alga marinhas foram expressas na forma recombinante solúvel e biologicamente funcional. As lectinas da alga marinha *Hypnea musciformis* e *Alsidium triquetrum* foram expressas no pET32a-TEV em *E. coli* Rosetta-Gami e *E. coli* BL21(DE3), respectivamente, com as melhores condições de indução para 0,2 mM de IPTG à 16 °C (Nascimento, 2014; Nascimento *et al.*, 2015).

A lectina recombinante de *Solieria filiformis* foi expressa no pET28a em *E. coli* BL21(DE3) com 0.05 mM de IPTG à 25 °C (Chaves *et al.*, 2023). Assim como a lectina da alga vermelha *Kappaphycus alvarezii* que foi expressa no mesmo vetor em *E. coli* SHuffle T7 com 0,1 mM de IPTG à 20 °C (Hirayama *et al.*, 2016).

A griffithsina (GRFT) da alga *Griffithsia sp.* foi expressa no vetor pCold-I também com cauda de histidina, em *E. coli* Rosetta 2. A expressão da proteína foi induzida com 1 mM de IPTG a 22 °C (Ishae *et al.*, 2013). A mesma lectina também foi expressa na forma recombinante no pET28a(+) em meio de autoindução (Giomarelli *et al.*, 2006). Outra lectina de alga vermelha foi expressa na forma ativa e solúvel no vetor pET28a(+) em *E. coli* BL21, a rodobidina da alga *Aglaothamnion callophyllidicola*. A indução ocorreu a 37 °C por 2 horas com 0,5 mM de IPTG (Han *et al.*, 2015).

6.3.2 Ensaio de deslocamento térmico (TSA – Thermal Shift Assay) (França)

Devido aos problemas de precipitação, o ensaio de deslocamento térmico foi realizado com o objetivo de estabelecer as melhores condições de tampão para a lectina expressa manter a estabilidade conformacional e permanecer solúvel.

O ensaio de deslocamento térmico avalia o comportamento estrutural de uma proteína em função da temperatura. A técnica se baseia no monitoramento da fluorescência de um corante hidrofóbico que se liga nas regiões apolares expostas durante a desnaturação térmica da proteína, permitindo avaliar a estabilidade em diferentes condições experimentais. Em sua conformação nativa, os resíduos hidrofóbicos se encontram no interior da estrutura, e não interagem com o corante. À medida que a temperatura aumenta, a proteína começa a desnaturar termicamente, expondo as regiões hidrofóbicas ao corante, resultando em um aumento da fluorescência. A temperatura de *melting* (T_m) é a temperatura onde 50% das moléculas da proteína estão desnaturadas, por isso uma T_m mais alta indica melhor estabilidade, pois é necessária uma maior energia térmica para desdobrar a proteína (Huynh; Partch, 2015).

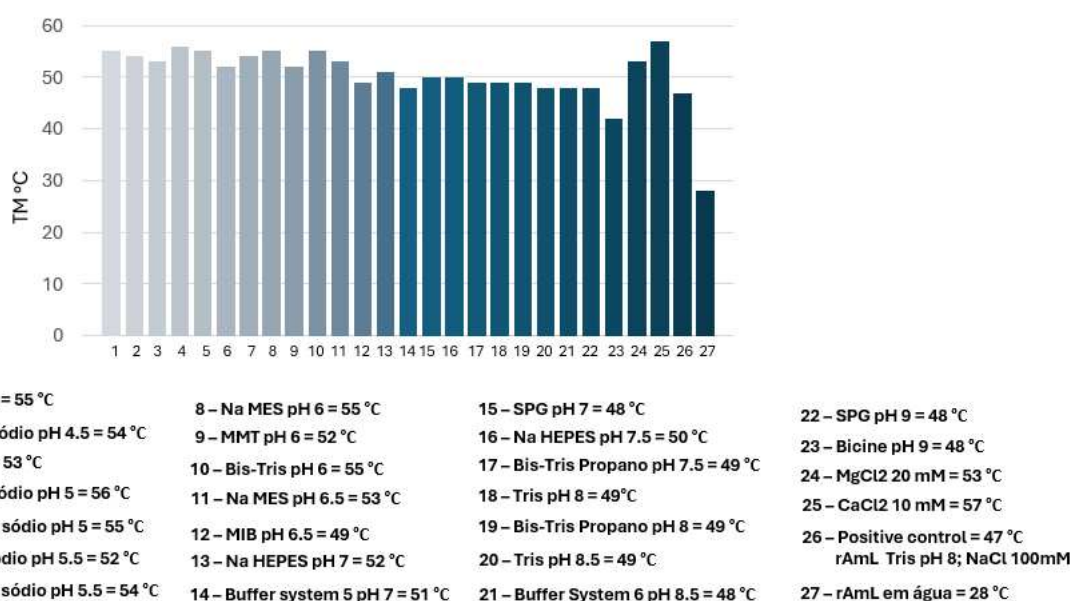
O ensaio foi inicialmente realizado com as duas formas da proteína, clivada e não clivada, em Tris 20 mM pH 8 e NaCl 100 mM para avaliar a presença de possíveis diferenças de estabilidade conformacional devido a presença da cauda de fusão. Porém, esse primeiro experimento foi considerado inválido pois houve contaminação com o corante fluorescente, evidenciado pelo comportamento do controle negativo. Mesmo assim, foi possível observar que as curvas de desnaturação e os valores de T_m foram próximos para as duas formas da proteína, o que não indicou diferenças relevantes de estabilidade térmica.

Com base nesse resultado preliminar e considerando que a cauda presente na construção é pequena, formada apenas pela HAT-tag e pelo sítio de clivagem, o que provavelmente não causaria um grande impacto na estabilidade total da proteína, resolvemos

conduzir o ensaio posterior de TSA apenas com a proteína não clivada visando a otimização do tempo.

O teste foi conduzido com a proteína rAmLpCHT3GNco em Tris 20 mM pH 8, contendo NaCl 100 mM para um *screening* amplo de diferentes condições (Figura 76).

Figura 76 – Estabilidade térmica de rAmL em diferentes condições de tampão determinado por TSA.



Fonte: Autora. O eixo vertical indica a Tm em °C. O eixo horizontal indica as condições experimentais enumeradas (1 a 27). As condições são especificadas abaixo do gráfico junto com a respectiva Tm. A Tm foi determinada a partir do ponto de inflexão das curvas de desnaturação térmica.

De modo geral, é possível observar que as Tm foram maiores em pH ácido a alcalino (pH 4 a 7), com muitas variações de acordo com o agente tamponante. No entanto, a maior Tm foi observada na presença de 10 mM de CaCl₂, indicando que esse sal é capaz de estabilizar a proteína rAmL.

Esse comportamento pode ser explicado pois íons divalentes podem promover estabilização estrutural por reduzir repulsões eletrostáticas entre resíduos carregados negativamente na superfície da proteína, além de estabelecer interações de coordenação transitórias com grupos carboxilato, o que pode diminuir flutuações conformacionais (Chi *et al.*, 2003).

A condição menos estável para a proteína foi em água, com uma Tm de 28 °C, indicando uma perda significativa de estabilidade na ausência de um sistema tamponante e de controle de força iônica. Geralmente em água, as proteínas não apresentam boa estabilidade

pois a ausência de tamponamento e força iônica compromete o equilíbrio de interações eletroestáticas e hidrofóbicas, expondo os resíduos apolares e levando a agregação (Pace; Scholtz; Grimsley, 2014).

Apesar de algumas condições em pH ácido tenham apresentado valores de T_m moderadamente altos, esses tampões não foram priorizados para aplicações posteriores, como cristalização, pois o TSA avalia exclusivamente a estabilidade térmica e não outras variáveis importantes para a cristalização como homogeneidade conformacional e tendência a precipitação de forma ordenada.

Tampões em pH neutro a alcalino são reconhecidos por favorecerem melhor estabilidade e comportamento físico-químico para proteínas em solução. Por isso, os tampões Tris e HEPES, de pH neutro a alcalino, foram escolhidos para os testes de cristalização por combinar boa estabilidade térmica com bons comportamentos físico-químico e serem altamente compatíveis com condições de cristalização de proteínas (Rodriguez-Blanco *et al.*, 2012).

Além disso, foi possível notar que a proteína em Tris 20 mM pH 8, contendo NaCl 100 mM, condições iniciais usadas no teste de TSA, apresentou o maior valor de T_m quando suplementada com 10 mM de CaCl_2 , indicando que o cálcio promove uma maior estabilidade conformacional nessas condições de pH e agente tamponante. Por isso, o cloreto de cálcio passou a ser adicionado nos testes posteriores, e o tampão foi mantido em Tris pH 8 na presença de NaCl.

Paralelamente, outros tampões em pH levemente ácido a neutro, foram testados para avaliar a estabilidade de rAmL e evitar agregação. No entanto, mesmo nas condições que apresentaram T_m mais elevada, a proteína continuou a apresentar precipitação. Da mesma forma, experimentalmente, a adição do CaCl_2 não resultou em uma melhora significativa na solubilidade da proteína, indicando que só o cloreto de cálcio não é suficiente para impedir os processos de agregação, e que a precipitação da rAmL está relacionada com fatores mais complexos, que apenas a desnaturação térmica não foi capaz de prever.

6.3.3 Ensaios de cristalização (França)

A lectina recombinante rAmL foi expressa de forma satisfatória no vetor pCOLD-HAT-TEV-3G-Nco e foi purificada com sucesso na forma clivada e na presença da cauda de fusão. Devido a obtenção da proteína recombinante, foram realizados ensaios de cristalização para determinar sua estrutura tridimensional.

Para os ensaios de cristalização, as duas formas purificadas foram avaliadas, pois a proteína clivada se aproxima da forma nativa, livre de regiões artificiais, e a proteína não clivada pode ser mais propícia a cristalizar, devido a cauda de fusão que pode ajudar na estabilidade estrutural da proteína como um todo.

A proteína não clivada, rAmLpCHT3GNco, e a proteína clivada, rAmL, foram submetidas a diferentes condições de cristalização, variando a concentração da proteína (0,3 a 4 mg/mL), o tampão inicial (em diferentes pHs), o ligante para cocristalização e o uso de diferentes kits de cristalização comerciais.

A cocristalização foi utilizada no ensaio, pois ela pode aumentar a estabilidade conformacional da proteína, ajudando na cristalização (Hassell *et al.*, 2007). A escolha dos ligantes usados no ensaio, Man6GlcNac(II) e Man5GlcNac(II), foi baseado na especificidade da proteína. A lectina AmL é pertencente à família OAAH (Silva, 2016), e esse grupo de lectinas possui preferência para carboidratos *N*-glicanos do tipo alta-manose, com resíduos de manose expostos (Costa Filho *et al.*, 2025).

No entanto, apesar das diferentes condições testadas não foi possível obter cristais da proteína rAmL recombinante em nenhuma das formas purificadas. A cristalização é um processo complexo e a formação de cristais depende do alinhamento de vários fatores como pureza, homogeneidade conformacional, estabilidade, concentração e composição do meio para acontecer. Provavelmente, a altíssima instabilidade de rAmL em solução tenha sido um fator importante para a ausência da formação de cristais (Gao *et al.*, 2017).

6.3.4 Produção da proteína recombinante rAmL em sistema heterólogo (Brasil)

6.3.4.1 Desenho do gene sintético adquirido no Brasil

O gene sintético da lectina rAmL adquirido no Brasil, com 840 pb, incluindo os sítios de restrição, foi inicialmente clonado no vetor de expressão pET28a(+) com as regiões de clonagem delimitadas pelas enzimas de restrição *NdeI* e *XhoI*, possuindo um sítio de clivagem para a protease TEV e sem a otimização de códons raros para *E. coli*.

O quadro aberto de leitura (ORF – *Open Reading Frame*) possui 828 pb, codificando uma proteína de 277 aminoácidos. Devido a construção do vetor pET28a(+), o gene foi expresso em fusão com uma cauda de histidina N-terminal, resultando na expressão da

proteína de fusão His-rAmL com peso molecular previsto de 31.751 Da. A construção da proteína recombinante expressa está representada na Figura 77.

Figura 77 – Representação esquemática da proteína recombinante expressa His-rAmL.

```

>rAmL
      10      20      30      40      50      60
MGSSHHHHHH SSGLVPRGSH MENLYFQGRR YTVENQWGG S SAPWNAAGLW VLGGRRPKQPC
      70      80      90     100     110     120
VAVDVSSTDG GSTLTGTMTY KEGGPIGFKG IRKDGGNVYE VKNQWGGSSA PWHGGEWVI
     130     140     150     160     170     180
GSRTSQSVIE LKIKSSDSGA HFDGTMKYSG EGPIGFQATM SDGDTYDVEN QWGGSSAPWH
     190     200     210     220     230     240
KAGVWALGDR AFQNIIMLKV TSPDGGKTLT GDMQYEREGP IGFRGTLSQA NNYAVENQWG
     250     260     270     280     290     296
GSHAPWHKAG NWIIGDRDGQ KVVGVAITSS DNGKTFEGTM TYSREGPIGF KGILSS

Sítio de restrição NcoI
Cauda de histidina
Sítio clivagem para trombina
Sítio de clivagem TEV

Número de aminoácidos: 296
Peso molecular: 31.750,9947
pI: 6,67

Coeficiente de extinção:
O coeficiente de extinção é: 79410
Em 1mg/ml (31,495 uM), a solução da proteína tem Abs280nm de: 2,50
  
```

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

Para proteínas expressas fusionadas, a remoção da *tag* de fusão pode ser necessária quando se quer obter a proteína de interesse igual a sua forma nativa. Como o objetivo era a clivagem da proteína com a protease TEV, o gene sintético de rAmL foi modificado para conter um sítio de clivagem para a protease, inserido antes da sequência codificadora da lectina, visto que o vetor pET28a(+) não possui a sequência codificadora para essa enzima.

Após clivagem, rAmL possui 269 aminoácidos com peso molecular de 28.661,62 Da e com pI de 6,24. Apenas a cauda expressa possui 27 aminoácidos com peso molecular de 3.107,39 Da e pI de 7,25. A Figura 78 mostra a representação da proteína de interesse rAmL, após remoção da cauda de histidina por clivagem da protease TEV.

Figura 78 – Representação da proteína rAmL obtida após clivagem com a protease TEV.

>rAmL clivada

```

      10      20      30      40      50      60
GRRYTVENQW GGSSAPWNAA GLWVLGGRPK QPCVAVDVSS TDGGSTLTGT MTYKGEGPIG

      70      80      90     100     110     120
FKGIRKDGGN VYEVKNQWGG SSAPWHDGGE WVIGSRTSQS VIELKIKSSD SGAFDGTMK

     130     140     150     160     170     180
YSGEGPIGFQ ATMSDGDITYD VENQWGGSSA PWHKAGVWAL GDRAFQNIIM LKVTSPDGGK

     190     200     210     220     230     240
TLTGDMQYER EGPIGFRGTL SQANNYAVEN QWGGSHAPWH KAGNWIIGDR DGQKVVGVAI

     250     260     269
TSSDNGKTFE GTMTYSREGP IGFKGILSS

```

Número de aminoácidos: 269

Peso molecular: 28.661,6169

pI: 6,24

Coefficiente de extinção:

O coeficiente de extinção é: 77920

Em 1mg/ml (34,890 uM), a solução da proteína tem Abs280nm de: 2,72

Fonte: Autora. pI: Ponto isoelétrico. As informações do peso molecular, pI e coeficiente de extinção foram gerados pelo programa *Protein parameters*, disponível em: <https://protparam.net/index.html>.

O índice de adaptação de códons (CAI – *Codon Adaptation Index*) da sequência não otimizada foi calculado para *E. coli* pela *GenScript*. O valor foi de 0,71. Esse valor não indica uma boa adaptação ao sistema hospedeiro, pois valores acima de 0,8 são considerados bons, e quanto mais perto de 1,0, melhor a expressão naquele sistema.

O vetor pET28a(+) é um plasmídeo de expressão do sistema pET bastante usado para produção de proteínas recombinantes em *E. coli*. Esse vetor contém uma His-tag na extremidade N-terminal e C-terminal, permitindo a escolha da construção para a proteína de interesse e possibilitando a purificação por cromatografia IMAC. Além disso, também possui a T7 tag que permite identificação da construção e o sítio de clivagem para a protease trombina (Novagen, 2005).

Muitas lectinas de algas vermelhas foram expressas de forma eficiente no vetor pET28a(+) (Chaves *et al.*, 2023; Hirayama *et al.*, 2016; Han *et al.*, 2015; Giomarelli *et al.*, 2006).

6.3.4.2 Expressão da lectina rAmL em sistema heterólogo a partir das duas construções gênicas

6.3.4.2.1 Transformação de *E. coli* com os genes sintéticos

Em todas as estratégias avaliadas, a expressão da lectina rAmL no vetor pCOLD-HAT-TEV-3G-Nco foi a construção gênica da proteína de interesse que apresentou os melhores resultados de expressão e purificação. Portanto, o plasmídeo rAmLpCHT3GNco foi a única construção a ser trazida para o Brasil para dar continuidade aos estudos.

Como um gene sintético para rAmL havia sido construído e adquirido no Brasil, os estudos passaram a ser conduzidos com as duas construções gênicas, com o objetivo de comparar o rendimento de produção entre elas, e selecionar a construção mais rentável para os testes físico-químicos e biológicos.

Após a recuperação, os plasmídeos rAmLpCHT3GNco e rAmLpET28a (His-rAmL, adquirido no Brasil) foram inseridos na cepa de clonagem *Escherichia coli* DH5 α para amplificação e obtenção do DNA plasmidial, e na cepa de expressão *E. coli* BL21(DE3).

As duas construções gênicas foram transformadas apenas na cepa de expressão *E. coli* BL21, pois os experimentos realizados previamente na França já haviam sido muito satisfatórios com essa linhagem. Além disso, *E. coli* BL21 é uma das cepas mais utilizadas para expressão heteróloga, apresentando um alto rendimento proteico e alta compatibilidade com vetores do sistema pET (Mital; Christie; Dikicioglu, 2021). Como rAmL não possui modificações pós-traducionais, não houve a necessidade de testar cepas geneticamente modificadas.

6.3.4.2.2 Testes de expressão

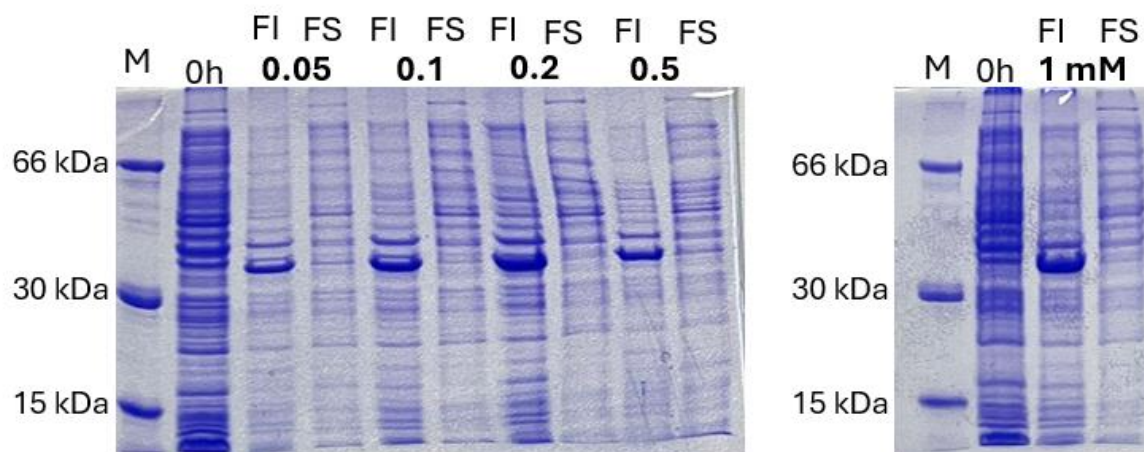
Para otimizar a produção da proteína recombinante, investigamos os três parâmetros mais comumente modificados: temperatura, tempo de indução e concentração do indutor (IPTG). Para ambos os vetores, variaram a concentração de IPTG (0,05 mM; 0,1 mM; 0,2 mM; 0,5 mM e 1,0 mM) e as temperaturas de indução de 37 °C por 4 h, 30 °C por 16h, 25 °C por 20h, 18 °C por 24 horas e pelo método da França de reduzir a T °C antes da indução nas temperaturas de 25 °C e 16 °C.

Após o tempo de indução em cada teste, as células foram recuperadas, lisadas e as frações solúveis e insolúveis foram analisadas por SDS-PAGE para os dois plasmídeos.

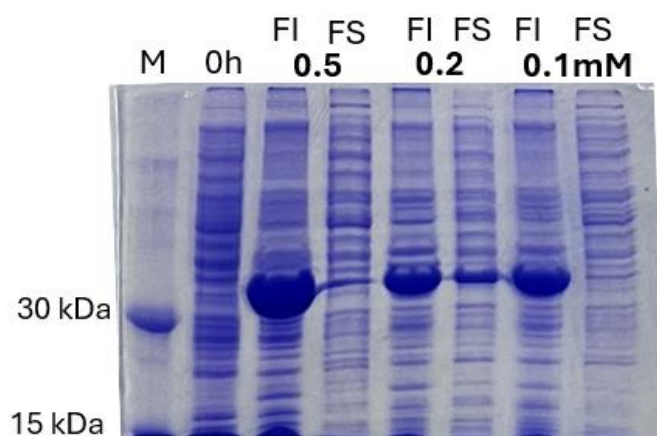
As análises em SDS-PAGE para os testes de expressão da proteína rAmLpCHT3GNco estão apresentados na Figura 79.

Figura 79 – Testes de expressão da proteína rAmL no vetor pCOLD-HAT-TEV-3G-Nco.

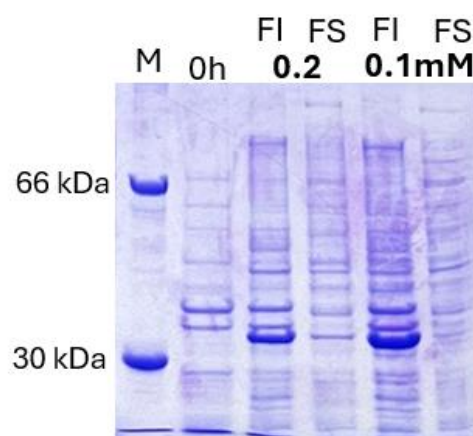
A 37 °C / 4hrs



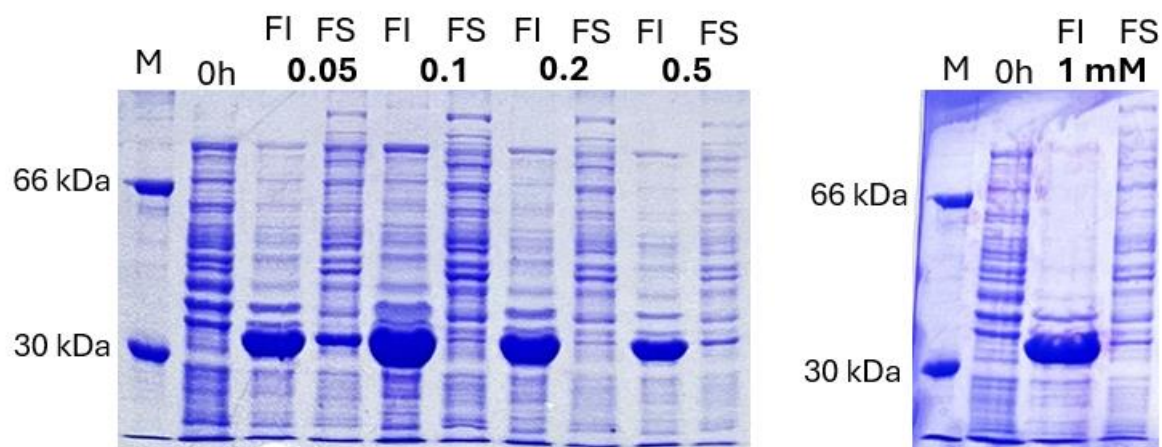
B 30 °C / 16hrs

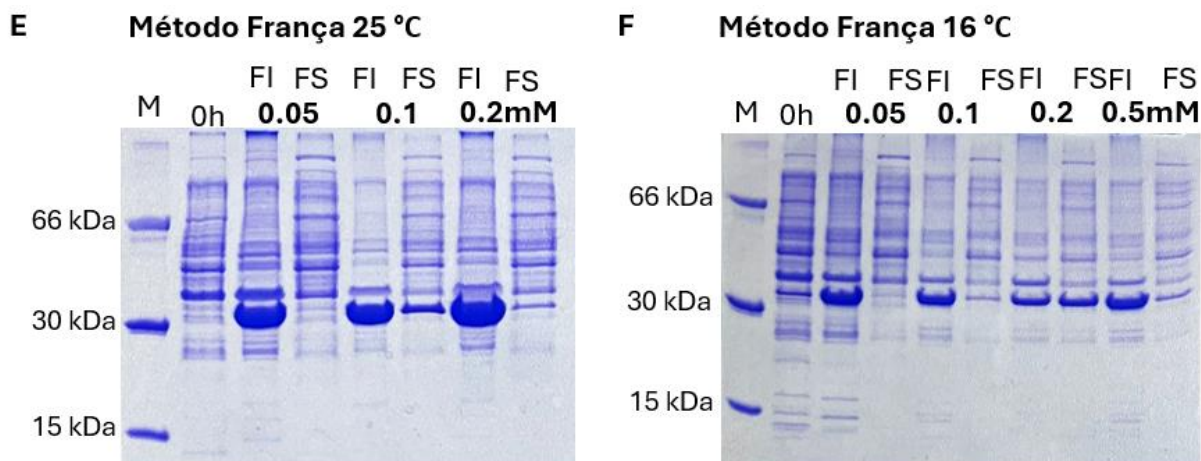


C 18 °C / 24hrs



D 25 °C / 20 hrs





Fonte: Autora. Géis de eletroforese SDS-PAGE. M: Marcador molecular contendo: HSA – 66 kDa; rSFL - 30 kDa; e CCL – 15 kDa. 0h: alíquota do *pellet* celular retirada antes da indução. FI: Fração insolúvel. FS: Fração solúvel. **A:** Teste de concentração de IPTG (0,05; 0,1; 0,2; 0,5 e 1 mM) na temperatura de 37 °C/4hrs. **B:** Teste de concentração de IPTG (0,1; 0,2 e 0,5 mM) na temperatura de 30 °C/16hrs. **C:** Teste de concentração de IPTG (0,1 e 0,2 mM) na temperatura de 18 °C/24hrs. **D:** Teste de concentração de IPTG (0,05; 0,1; 0,2; 0,5 e 1 mM) na temperatura de 25 °C/20hrs. **E:** Teste de concentração de IPTG (0,05, 0,1 e 0,2 mM) na temperatura de 25 °C no método França. **F:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 16 °C no método França.

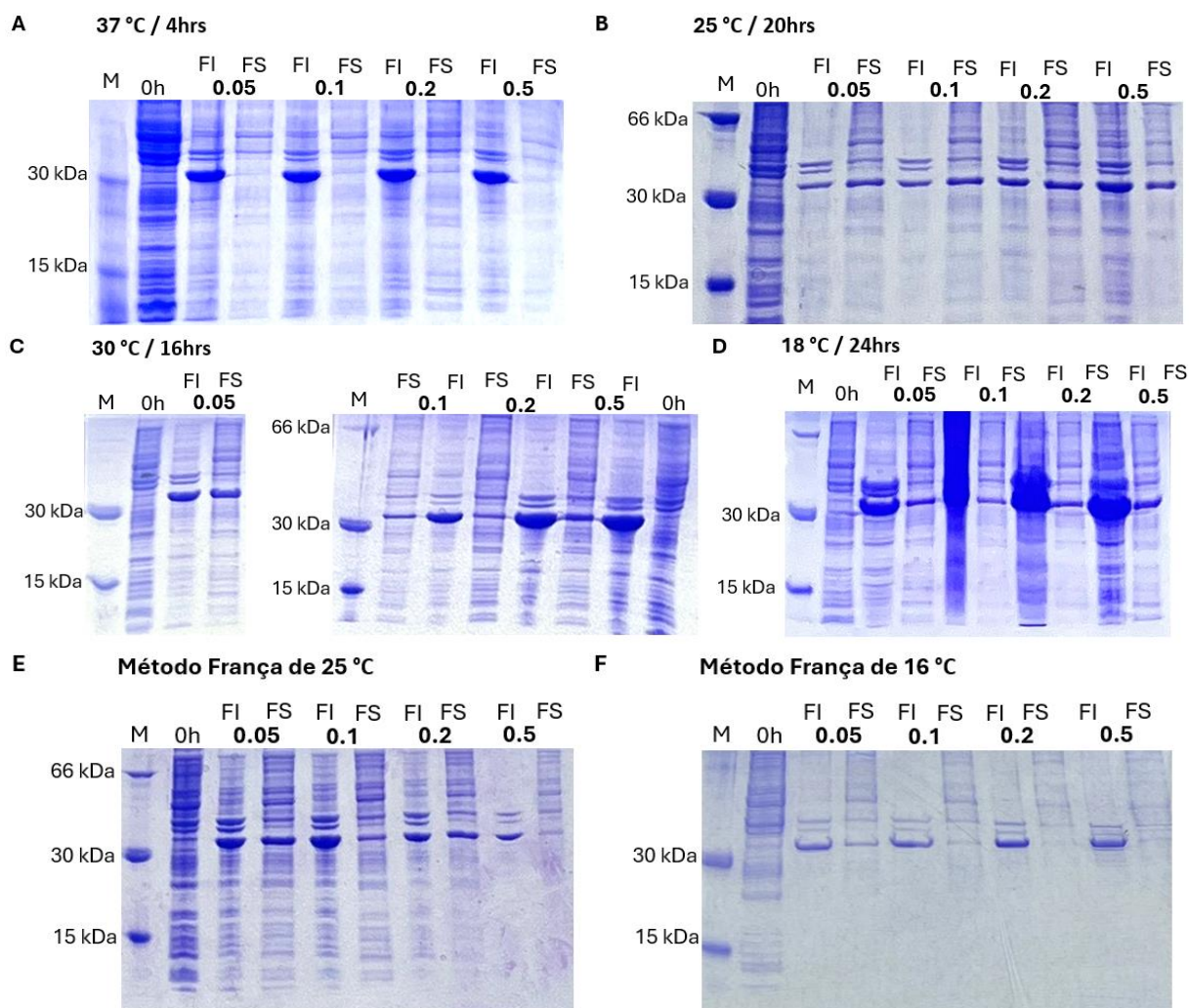
A proteína de interesse rAmLpCHT3GNco foi expressa na fração insolúvel em todas as concentrações de IPTG testadas em todas as temperaturas, detectada pela banda visível no gel acima de 30 kDa, de acordo com o tamanho esperado para a proteína de aproximadamente 33 kDa. Essa banda não está presente na amostra coletada antes da indução. Em 37 °C por 4 h e 18 °C por 24 h não foi observada expressão na fração solúvel.

A lectina foi expressa com solubilidade apenas nas seguintes condições: 0,2 mM de IPTG a 30 °C, 0,05 mM de IPTG e em 0,5 mM com uma intensidade menor a 25 °C, 0,1 mM de IPTG e em 0,2 mM com uma intensidade baixa a 25 °C no método França e 0,2 mM de IPTG e em 0,1 e 0,5 mM com uma intensidade menor a 16 °C no método França.

Na fração solúvel da concentração de 0,5 mM de IPTG a 30 °C aparece uma banda de menor intensidade na altura da proteína expressa, porém aparenta ser uma saturação do gel causada pelo grande volume de expressão na pista da fração insolúvel. Algumas concentrações de IPTG não estão presentes em todos os testes pois não houve necessidade, visto que as expressões de maior intensidade já haviam sido detectadas nas concentrações do indutor avaliadas.

As análises em SDS-PAGE para os testes de expressão da proteína rAmLpET28a(+) estão apresentados na Figura 80.

Figura 80 – Testes de expressão da proteína rAmL no vetor pET28a(+).



Fonte: Autora. Géis de eletroforese SDS-PAGE. M: Marcador molecular contendo: HSA – 66 kDa; rSFL – 30 kDa; e CCL – 15 kDa. 0h: alíquota do *pellet* celular retirada antes da indução. FI: Fração insolúvel. FS: Fração solúvel. **A:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 37 °C/4hrs. **B:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 25 °C/20hrs. **C:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 30 °C/16hrs. **D:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 18 °C/24hrs. **E:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 25 °C no método França. **F:** Teste de concentração de IPTG (0,05; 0,1; 0,2 e 0,5 mM) na temperatura de 16 °C no método França.

A proteína de interesse rAmLpET28a foi expressa na fração insolúvel em todas as concentrações de IPTG testadas em todas as temperaturas, detectada pela banda visível no gel acima de 30 kDa, de acordo com o tamanho esperado para a proteína de aproximadamente 32 kDa. Essa banda não está presente na amostra coletada antes da indução. Em 37 °C por 4h, não foi observada expressão na fração solúvel.

A concentração de 1mM de IPTG não foi usada nos testes de expressão do pET28a(+), pois não havia gerado resultados positivos no vetor pCOLD-HAT-TEV-3G-Nco.

Além disso, usar concentrações mais baixas de indutor é sempre preferível, se a proteína for expressa em menores concentrações, como foi observado.

A lectina foi expressa com solubilidade em todas as concentrações do indutor nas temperaturas de 30 °C, 25 °C, 18 °C e 25 °C no método da França. Em 16 °C no método da França, houve expressão apenas nas concentrações de 0,05 e 0,1 mM de IPTG, em uma intensidade muito baixa, esse comportamento pode ser explicado pelo tipo de vetor. O vetor pET, diferentemente do vetor pCOLD, não possui tolerância a temperaturas muito baixas pois não utiliza o promotor derivado de genes do choque térmico (Baneyx; Mujacic, 2004).

De modo geral, a construção de rAmLpCHT3GNco expressou a proteína em níveis mais elevados, demonstrado pelas bandas de maior intensidade no gel de SDS-PAGE. Porém, a maior parte da proteína foi detectada na fração insolúvel. Muito provavelmente, a taxa de tradução está muito alta, sobrecarregando o maquinário celular da bactéria e levando a agregação da proteína.

Por outro lado, a construção rAmLpET28a exibiu níveis mais baixos de expressão, mas com maior presença da proteína na fração solúvel. Esse resultado pode estar relacionado com a menor taxa de tradução devido à ausência de otimização do gene, reduzindo a pressão no sistema de dobramento da célula e favorecendo a solubilidade.

Curiosamente, a concentração do indutor que resultou na maior intensidade de expressão, variou muito em cada temperatura e construção testada. Entre os parâmetros que mais influenciam na expressão de proteínas recombinantes, a concentração do IPTG é um dos mais difíceis de se estabelecer, porque a expressão gênica está fortemente relacionada a concentração intracelular do IPTG e o modo como essa molécula é absorvida do meio extracelular para a célula. A entrada do IPTG acontece por difusão e depende da permeabilidade da membrana, da atividade metabólica da célula e da temperatura, pois temperaturas diferentes alteram a velocidade do transporte. Portanto, a mesma concentração de IPTG pode gerar resultados diferentes em cada temperatura de indução (Simas; Pessoa Junior; Long, 2023; Kielkopf; Bauer; Urbatsch, 2021).

O melhor método de expressão para cada construção foi determinado pela comparação da intensidade da banda correspondente a proteína nos géis de SDS-PAGE, pelo maior pico de absorbância a 280 nm durante a purificação e pelo título de hemaglutinação da fração purificada contra eritrócitos de coelho tratado com tripsina a 3%.

Para rAmLpCHT3GNco, as frações solúveis de 30 °C a 0,2 mM, 25 °C a 0,05 mM, 25 °C no método França (MF) a 0,1 mM e 16 °C no método França (MF) a 0,2 mM resultaram

na banda mais intensa de proteína expressa em SDS-PAGE. Por isso, foram submetidas a cromatografia de afinidade IMAC, seguindo o mesmo protocolo determinado na França, com a eluição da proteína em um único passo cromatográfico utilizando 270 mM de imidazol. Os maiores picos de absorvância a 280 nm foram para as frações de 30 °C a 0,2 mM, 25 °C a 0,05 mM. A purificação para todas as condições foi confirmada por SDS-PAGE (dados não mostrados).

As frações da proteína eluída da coluna (P2) das quatro condições selecionadas foram submetidas ao ensaio de atividade hemaglutinante sob as mesmas condições. Os títulos de hemaglutinação referentes a análise funcional da lectina estão resumidos na Tabela 20.

Tabela 20 – Avaliação da melhor expressão para rAmLpCHT3GNco: Funcionabilidade da lectina recombinante expressa.

Condição (T °C e [IPTG])	Maior diluição hemaglutinante (2ⁿ)	Título de hemaglutinação (U.H./mL)
30 °C a 0,2 mM	2 ⁴	16
25 °C a 0,05 mM	2 ²	4
25 °C MF a 0,1 mM	2 ¹	2
16 °C MF a 0,2 mM	2 ¹	2

Fonte: Autora. A maior diluição hemaglutinante é expressa como 2ⁿ, onde n corresponde ao número de diluições seriadas até o último poço no qual a hemaglutinação é detectada. O título de hemaglutinação é definido como a maior diluição da amostra capaz de promover hemaglutinação visível, expresso em unidade hemaglutinantes por mililitro (U.H./mL).

O maior título de hemaglutinação foi observado para a fração de 30 °C a 0,2 mM, seguida pela de 25 °C a 0,05 mM, coincidindo com os maiores picos de absorvância durante a purificação da proteína expressa.

Para rAmLpET28a, as frações solúveis de todas as concentrações de IPTG a 30 °C, 25 °C, 18 °C e 25 °C no método França (MF), com exceção da fração de 0,5 mM a 25 °C MF, resultaram em bandas intensas da proteína expressa em SDS-PAGE. Por isso, foram submetidas a mesma cromatografia de afinidade que a construção rAmLpCHT3GNco. Os maiores picos de absorvância a 280 nm foram para as frações de 30 °C a 0,05 mM, 25 °C a 0,05; 0,1 e 0,2 mM, 18 °C a 0,05 e 0,2 mM e 25 °C MF a 0,2 mM. A purificação para todas as condições foi confirmada por SDS-PAGE, e as bandas mais intensas dos picos eluídos coincidiram com os maiores picos de absorvância (dados não mostrados).

As frações da proteína eluída da coluna (P2) das condições que tiveram os maiores picos de absorvância foram submetidas ao ensaio de atividade hemaglutinante sob as mesmas

condições. Os títulos de hemaglutinação referentes a análise funcional da lectina estão resumidos na Tabela 21.

Tabela 21 – Avaliação da melhor expressão para rAmLpET28a: Funcionabilidade da lectina recombinante expressa.

Condição (T °C e [IPTG])	Maior diluição hemaglutinante (2ⁿ)	Título de hemaglutinação (U.H./mL)
30 °C a 0,05 mM	2 ⁹	512
25 °C a 0,05 mM	2 ⁷	128
25 °C a 0,1 mM	2 ⁷	128
25 °C a 0,2 mM	2 ⁷	128
18 °C a 0,05 mM	2 ⁶	64
18 °C a 0,2 mM	2 ⁷	128
25 °C MF a 0,2 mM	2 ⁶	64

Fonte: Autora. A maior diluição hemaglutinante é expressa como 2ⁿ, onde n corresponde ao número de diluições seriadas até o último poço no qual a hemaglutinação é detectada. O título de hemaglutinação é definido como a maior diluição da amostra capaz de promover hemaglutinação visível, expresso em unidade hemaglutinantes por mililitro (U.H./mL).

O maior título de hemaglutinação foi observado para a fração de 30 °C a 0,05 mM. Portanto, as melhores condições de expressão para rAmLpCHT3GNco e rAmLpET28a foram a 30 °C/16 hrs com 0,2 mM de IPTG e 30 °C/16 hrs com 0,05 mM de IPTG, respectivamente.

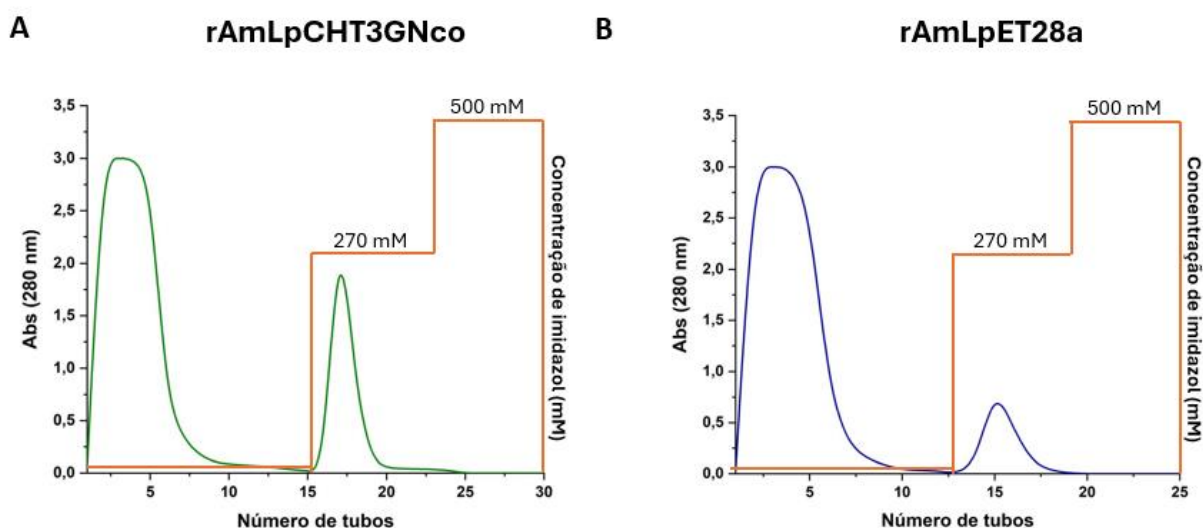
As atividades hemaglutinantes obtidas para as duas construções não podem ser comparadas diretamente, pois durante as purificações de rAmLpCHT3GNco, a coluna cromatográfica apresentou problemas para reter a proteína, sendo substituída apenas nos ensaios com rAmLpET28a. Portanto, parte da proteína presente na fração solúvel de rAmLpCHT3GNco não foi eficientemente retida na coluna, comprometendo a recuperação e a quantificação das amostras. Porém, como todas as quatro condições testadas para rAmLpCHT3GNco foram purificadas utilizando a mesma coluna com desempenho comprometido, esses dados podem ser comparados entre si, permitindo identificar a melhor condição de expressão.

Com o método de expressão mais rentável para cada vetor estabelecido, uma expressão em maior escala foi realizada para cada construção plasmidial com o objetivo de quantificar o rendimento e determinar qual construção possui a melhor produção da proteína de interesse para os testes físico-químicos e biológicos.

Após o tempo de indução na expressão de cada vetor, as células foram recuperadas, lisadas e as frações solúveis e insolúveis foram analisadas por SDS-PAGE.

Como as duas construções possuem cauda de histidina, a fração solúvel foi submetida à cromatografia de afinidade a metais imobilizados (IMAC), separadamente. A coluna foi lavada com tampão contendo imidazol 20 mM (P1) e as proteínas retidas a coluna foram eluídas com 270 mM de imidazol (P2), e com 500 mM de imidazol (P3) (Figura 81).

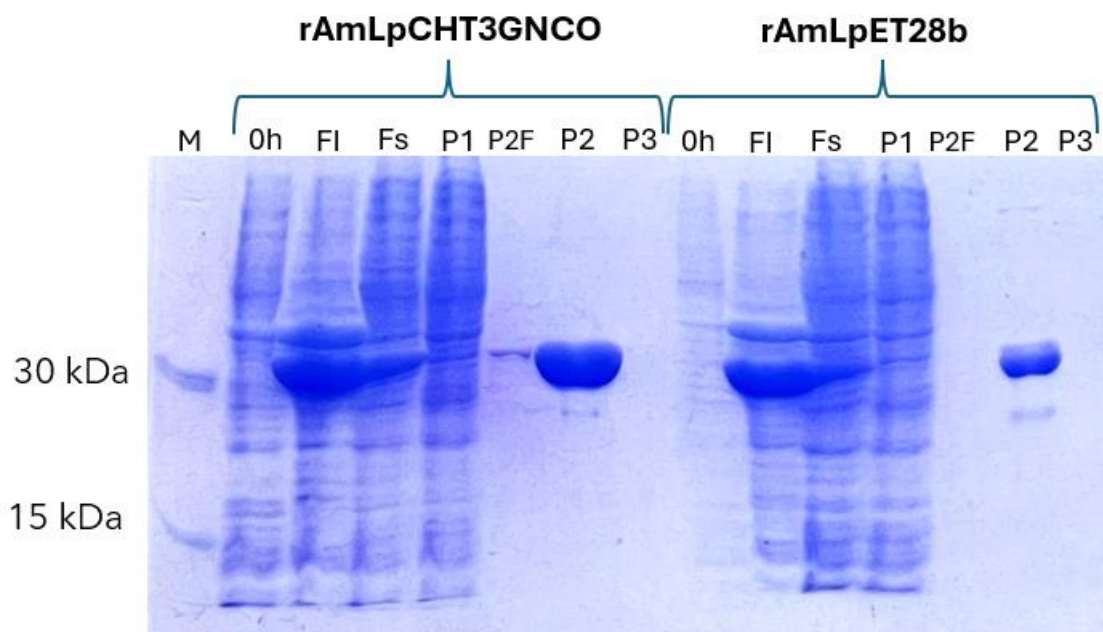
Figura 81 – Cromatograma da purificação da proteína rAmLpCHT3GNco e rAmLpET28a.



Fonte: Autora. Cromatografia em coluna HisTrap™ FF de 5mL a um fluxo de 2,5 mL/min. P1: Pico 1 - proteínas lavadas com tampão Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 20 mM. P2: Pico 2 - proteínas eluídas com Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 270 mM Pico 3 - proteínas eluídas com Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 500 mM. Eixo inferior: número de tubos coletados. Eixo esquerdo: absorbância a 280 nm em Au. Eixo direito: concentração de imidazol em mM. A linha contínua laranja indica a concentração de imidazol 500 mM. A linha contínua verde ou azul indica a absorbância a 280 nm (Au). A: Cromatograma da purificação de rAmLpCHT3GNco. B: Cromatograma de rAmLpET28a.

O primeiro pico recuperado da coluna com 270 mM de imidazol teve a máxima absorbância de 2,583 Au e 0,870 Au para rAmLpCHT3GNco e rAmLpET28a, respectivamente. A eluição com 500 mM de imidazol não resultou em nenhum pico nas duas cromatografias, indicando que toda a proteína havia sido eluída com 270 mM de imidazol. As frações foram coletadas e analisadas em SDS-PAGE (Figura 82).

Figura 82 - Análise da purificação de rAmLpCHT3GNco e rAmLpET28a em SDS-PAGE.



Fonte: Autora. Gel de 15%. M: Marcador molecular contendo: HSA – 66 kDa; rSFL - 30 kDa; e CCL – 15 kDa. 0h: Aliquota retirada antes da indução. FI: Fração insolúvel. Fs: Fração solúvel. P1: Fração lavada da coluna com 20 mM de imidazol. P2: Fração eluída da coluna com 270 mM de imidazol. P2F: Último tubo coletado de absorbância 0,04 Au. P3: Fração eluída da coluna com 500 mM de imidazol. As alíquotas utilizadas nos picos é o tubo de maior absorbância.

A expressão da proteína de interesse para as duas construções seguiu o mesmo padrão observado anteriormente, com a presença da proteína na fração insolúvel e solúvel, evidenciada pela ausência da banda na alíquota antes da indução (0h).

A análise em SDS-PAGE confirmou a ligação e a recuperação da proteína na coluna, observada pela banda acima de 30 kDa, no pico eluído com 270 mM de imidazol, de acordo com o tamanho esperado para ambas as construções. Não foi observada a presença da proteína no P1, indicando que toda a proteína ficou retida na coluna. A presença da lectina no pico eluído com 500 mM de imidazol também não foi detectada, confirmando que toda a proteína havia sido eluída com 270 mM de imidazol.

Comparativamente, o pico de máxima absorbância eluído com 270 mM de imidazol durante a purificação e a intensidade da banda referente ao pico retido analisado em SDS-PAGE, evidenciaram que a construção rAmLpCHT3GNco foi capaz de expressar mais efetivamente a proteína de interesse. A funcionabilidade da proteína em ambas as construções foi confirmada por atividade hemaglutinante (dados não mostrados).

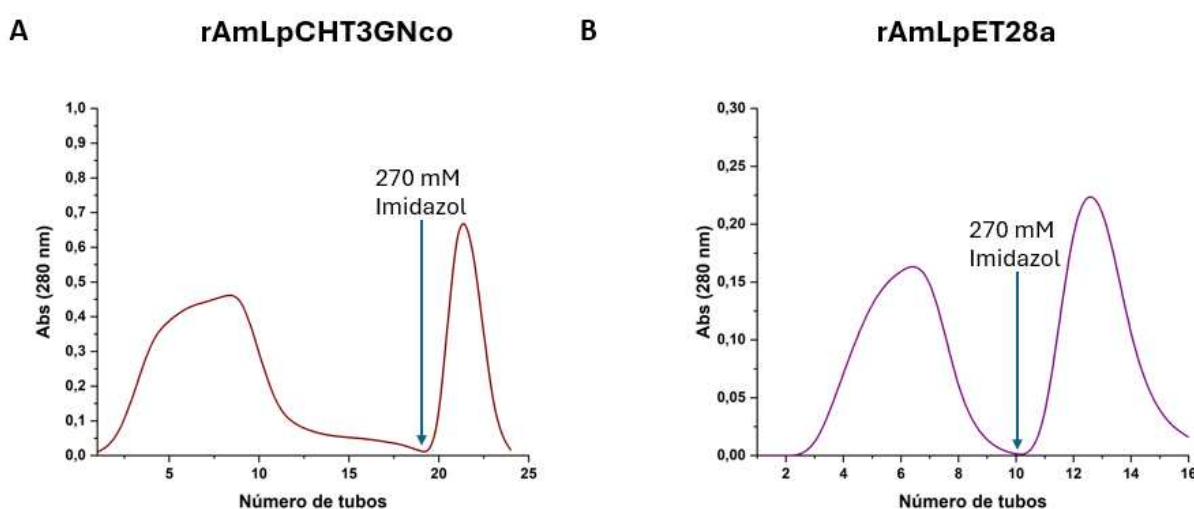
Em ambas as frações eluídas, foi observado contaminantes junto com a proteína de interesse, inclusive, em maior quantidade na construção rAmLpET28a. Conforme discutido anteriormente, durante a purificação na cromatografia IMAC, ligações não específicas com a

coluna podem ocorrer, principalmente quando a eluição acontece em um único passo, com uma concentração elevada de imidazol, que elui simultaneamente todas as proteínas com afinidade baixa, moderada ou alta ao metal imobilizado (YOUNG; BRITTON; ROBINSON, 2012). No entanto, como o objetivo dessa cromatografia era avaliar o rendimento, a eluição em um único passo cromatográfico seria mais otimizada.

Todo o volume dos picos eluidos com 270 mM de imidazol contendo a proteína de interesse em ambas as construções seguiram para a próxima etapa de purificação, a remoção da cauda fusionada com a protease TEV para obter a proteína rAmL clivada. As frações foram dialisadas contra Tris 20 mM pH 8, NaCl 150 mM, glicerol 5% e EDTA 0,5 mM, separadamente, e a protease TEV foi adicionada à solução.

Após a reação de digestão, a concentração de imidazol foi ajustada e a solução foi submetida novamente a cromatografia de afinidade IMAC. A coluna foi lavada com o mesmo tampão contendo imidazol 20 mM (P1), e a fração das proteínas ligadas a coluna, foi eluída com o tampão contendo 270 mM (P2) (Figura 83).

Figura 83 – Cromatograma da purificação da proteína rAmL clivada.

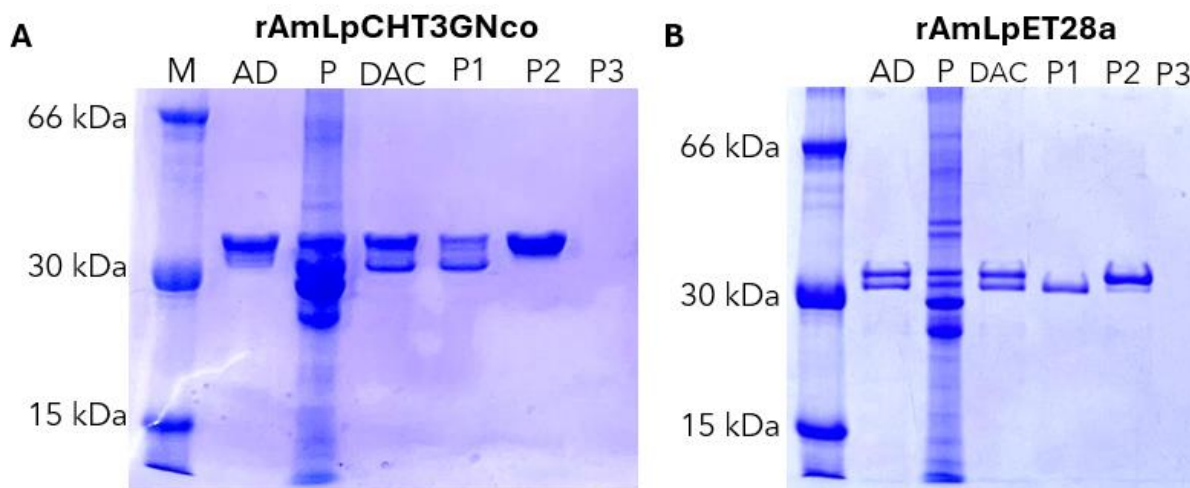


Fonte: Autora. Cromatografia em coluna HisTrap™ FF de 5mL a um fluxo de 2,5 mL/min. P1: Pico 1 - proteínas lavadas com tampão Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 20 mM. P2: Pico 2 - proteínas eluídas com Tris 20 mM pH 8,0 com NaCl 150 mM, glicerol 5% e imidazol 270 mM. Eixo inferior: número de tubos coletados. Eixo esquerdo: absorvância a 280 nm em Au. A linha contínua vinho ou roxo indica a absorvância a 280 nm (Au). A seta indica o momento que o pico das proteínas ligadas a coluna começou a ser eluído. A: Cromatograma da purificação de rAmLpCHT3GNco. B: Cromatograma de rAmLpET28a.

O primeiro pico, das proteínas que não se ligaram a coluna, teve a máxima absorvância de 0,472 Au e 0,169 Au para rAmLpCHT3GNco e rAmLpET28a, respectivamente, enquanto o pico com as proteínas eluídas da coluna teve absorção máxima em 0,777 Au e 0,232

Au para rAmLpCHT3GNco e rAmLpET28a, respectivamente. As frações foram coletadas e analisadas em SDS-PAGE (Figura 84).

Figura 84 - Análise da purificação de rAmL clivada em SDS-PAGE.



Fonte: Autora. Gel de 15%. A: Análise da clivagem e purificação da rAmLpCHT3GNco B: Análise da clivagem e purificação da rAmLpET28a M: M: Marcador molecular contendo: HSA – 66 kDa; rSFL - 30 kDa; e CCL – 15 kDa. AD: Amostra digerida. P: precipitado. DAC: Amostra digerida após centrifugação para retirar o precipitado, antes da cromatografia. P1: Fração lavada da coluna com 20 mM de imidazol. P2: Fração eluída da coluna com 270 mM de imidazol. P3: Fração eluída da coluna com 500 mM de imidazol. As alíquotas utilizadas nos picos é o tubo de maior absorbância.

A análise em SDS-PAGE da amostra rAmLpCHT3GNco (Figura 84 – A) mostrou que a proteína foi clivada, evidenciada pela presença de duas bandas de tamanhos moleculares parecidos, conforme esperado para rAmL de 33 kDa e 29 kDa para a proteína não clivada e clivada, respectivamente. No entanto, ao longo da cromatografia não foi possível separar eficientemente as duas formas, pois as duas bandas permaneceram associadas ao mesmo pico, P1, quando apenas a forma clivada deveria estar presente.

Conforme mencionado anteriormente, a presença da proteína não clivada na fração não retida a coluna, sugere redução da interação efetiva com a coluna de níquel, provavelmente devido a agregação que reduz o acesso da cauda de histidina a matriz (BIO RAD, 2016).

A análise em SDS-PAGE da amostra rAmLpET28a (Figura 84 – B) também mostrou que a proteína foi clivada, igualmente indicada pela presença das duas bandas de tamanhos parecidos. Porém, nesse caso, a separação das duas formas foi eficiente, pois no P1 foi detectada a presença apenas da banda mais baixa, indicando a proteína cortada e no P2 apenas a presença da banda de maior peso molecular, indicando a proteína não clivada.

Adicionalmente, foi observada a presença de precipitação durante a clivagem para as duas construções. Entretanto, para rAmLpCHT3GNco foi constatada a maior perda da

proteína por precipitação, evidenciada pela presença da banda na mesma altura da proteína não cortada no material precipitado. Por outro lado, para rAmLpET28a, a principal banda observada no precipitado corresponde, provavelmente, à protease TEV de aproximadamente 29 kDa, enquanto a banda referente à proteína não clivada foi pouco intensa. Esses resultados indicam que a menor agregação observada em rAmLpET28a permitiu uma interação mais eficiente da proteína não clivada com a coluna de níquel, sendo eluída apenas com 270 mM de imidazol, resultando na separação das duas formas.

A funcionabilidade da proteína em ambas as construções para as frações clivadas (P1) e não clivadas (P2) foi confirmada por atividade hemaglutinante (dados não mostrados), indicando que as proteínas de interesse continuaram ativas após a digestão.

Outras tentativas para a digestão com a TEV e a separação das proteínas clivadas e não clivadas de rAmLpCHT3GNco foram reproduzidas, seguindo exatamente o protocolo eficiente realizado previamente na França. No entanto, nenhuma das tentativas foram efetivas.

Outro parâmetro importante na comparação entre as duas construções gênicas é o rendimento da produção. Todos as variáveis analisadas para determinar qual vetor proporcionou a expressão da proteína mais rentável estão resumidas na Tabela 22.

Tabela 22 - Comparação do rendimento de expressão entre rAmLpCHT3GNco e rAmLpET28a.

	rAmLpET28a	rAmLpCHT3GNco
Volume de cultura	510 mL	510 mL
Melhor condição de expressão	30 °C / 0.05 mM IPTG	30 °C / 0.2 mM IPTG
Rendimento em gramas (g)	3,52 g	4,00 g
Volume de fração solúvel (mL)	35 mL	40 mL
Máxima Abs280nm (Au)	0,870 Au	2,583 Au
Eluição com 270 mM de imidazol		
Volume de pico coletado (mL)	15 mL	28 mL
Coleta até 0,04 Au		
Abs280 nm volume total de pico coletado (Au)	0,372 Au	0,989 Au
Concentração em mg/mL	0,153 mg/mL	0,41 mg/mL
Rendimento da proteína em miligramas (mg)	2,30 mg	11,48 mg
[] x volume		
Abs280 nm antes da digestão (Au)	0,436 Au	0,808 Au
Abs280 nm depois da digestão (Au)	0,356 Au	0,783 Au
Abs280 nm volume total	0,138 Au	0,305 Au
P1(proteínas clivadas) (Au)		(Problema de separação)
		Continua

		Continuação tabela 22
Volume total do P1 (mL)	15 mL	45 mL
Concentração proteínas clivadas	0,050 mg/mL	Não foi possível separar as duas formas
Rendimento de rAmL clivada (mg)	0,75 mg	Não foi possível determinar

Fonte: Autora. A concentração da proteína em mg/mL foi determinada calculando de acordo com o coeficiente de extinção da proteína para cada construção.

O rendimento de proteína expressa em meio litro de cultura foi de 2,30 mg e 11,48 mg para rAmLpET28a e rAmLpCHT3GNco. Apesar da clivagem ter sido eficiente para rAmLpET28a, o rendimento nesse vetor é bem menor quando comparado a outras proteínas recombinantes de alga vermelha (Chaves *et al.*, 2023; Han *et al.*, 2015; Hirayama *et al.*, 2016; Giomarelli *et al.*, 2006).

A diferença entre o rendimento das duas construções está, muito provavelmente, relacionada com a otimização de códons presente no gene sintético da rAmLpCHT3GNco. Muitos estudos já demonstraram que a substituição de códons em um gene heterólogo por códons preferencialmente usados pelo hospedeiro de expressão podem ter um efeito profundo no rendimento de proteínas recombinantes, pois genes altamente expressos tendem a conter códons para os quais a célula possui tRNA abundante (Makino; Skretas; Georgiou, 2011; Kuar; Kumar; Kuar, 2018).

Apesar da clivagem e repurificação da lectina clivada ter sido efetiva apenas na construção rAmLpET28a, o rendimento da produção foi significativamente inferior, principalmente após a recuperação da proteína rAmL digerida. Por isso, decidimos seguir com a construção rAmLpHT3GNco para a realização dos experimentos posteriores que necessitariam de quantidades relevantes da proteína expressa.

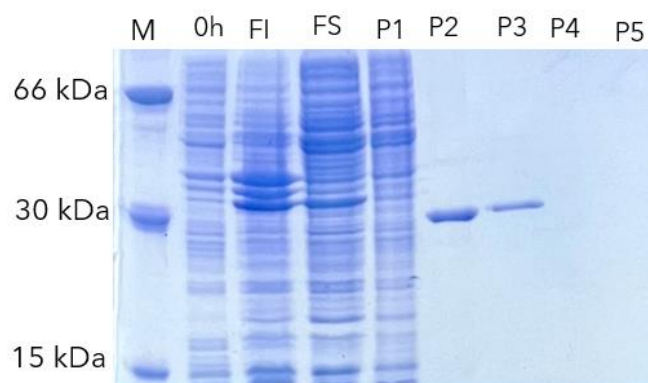
Apesar da clivagem não ter sido eficiente para rAmLpCHT3GNco, o rendimento foi bem superior e muitos estudos bioquímicos e biológicos utilizam proteínas recombinantes com a presença de caudas pequenas sem grandes impactos na atividade funcional (Han *et al.*, 2015; Hirayama *et al.*, 2016; Giomarelli *et al.*, 2006; Gao *et al.*, 2010). Além disso, a cauda do vetor pCOLD-HAT-TEV-3G-Nco é curta, sendo composta apenas pela cauda de afinidade a histidina e o sítio de clivagem, minimizando potenciais interferências estruturais ou funcionais.

Outro motivo para realizar os ensaios posteriores com a construção rAmLpCHT3GNco não clivada é que o problema de precipitação também persistiu nas novas purificações, assim como foi observado anteriormente nas purificações da França. Mesmo a

presença do glicerol em todos os tampões durante as etapas de purificação, não foi suficiente para evitar a agregação. É bastante relatado na literatura que a clivagem da cauda de fusão pode afetar a solubilidade da proteína de interesse, expondo ainda mais as regiões hidrofóbicas (Costa *et al.*, 2013; Rosano; Ceccarelli, 2014). Assim, clivar uma proteína bastante instável pode levar a uma maior precipitação.

Após a construção do melhor rendimento ser determinada para a continuação dos ensaios, a purificação passou a ser conduzida com dois *steps* de 40 e 80 mM de imidazol para retirar as contaminações observada na eluição com 270 mM (Figura 82). Conforme discutido anteriormente, a eluição em picos sucessivos ajuda na separação dos contaminantes por explorar as diferenças de afinidade dos contaminantes e da proteína alvo (Bornhorst; Falke, 2000) (Figura 85).

Figura 85 – Perfil de expressão e purificação da proteína rAmLpCHT3GNco com *steps*.



Fonte: Autora. Gel de 15%. M: Marcador molecular contendo: HSA – 66 kDa; rSFL - 30 kDa; e CCL – 15 kDa. 0h: Alíquota retirada antes da indução. FI: Fração insolúvel. FS: Fração solúvel. P1: Fração lavada da coluna com 20 mM de imidazol. P2: Fração eluída da coluna com 40 mM de imidazol. P3: Fração eluída da coluna com 80 mM de imidazol. P4: Fração eluída da coluna com 270 mM de imidazol. P5: Fração eluída da coluna com 500 mM de imidazol. As alíquotas utilizadas nos picos é o tubo de maior absorbância.

O novo perfil de purificação em SDS-PAGE mostra que a proteína recombinante não foi detectada na fração não retida a coluna (P1), mostrando que a ligação à matriz foi eficiente. A eluição de rAmLpCHT3GNco ocorreu totalmente com 40 mM (P2) e 80 mM (P3) de imidazol, pois não foi observada a presença da proteína na fração eluída com 270 e 500 mM de imidazol (P4 e P5). A lectina foi eluída nos dois *steps* sem a presença de contaminantes, indicando que foram removidos na lavagem com baixa quantidade de imidazol.

Portanto, conseguimos expressar e a purificar com sucesso a lectina recombinante rAmLPCHT3GNco, mantendo sua atividade funcional preservada. Nas etapas posteriores, as purificações passaram a ser realizadas utilizando eluição com os *steps* de imidazol, substituindo o protocolo anteriormente adotado na França, que consistia em uma lavagem inicial com 35 mM de imidazol, seguida de um gradiente linear até 500 mM. Isso porque, no método anterior, a proteína era eluída de forma heterogênea, em frações com e sem contaminação, que precisavam ser separadas e as frações contaminadas seguiam para a clivagem, onde os contaminantes eram removidos. Como a estratégia de produção não inclui mais a etapa de clivagem, a eluição em *steps* mostrou ser mais eficiente, pois permitiu a recuperação da proteína completamente livre de contaminação.

As frações contendo a proteína não clivada pura e funcional foram reunidas para os testes de caracterização bioquímica e os ensaios biológicos.

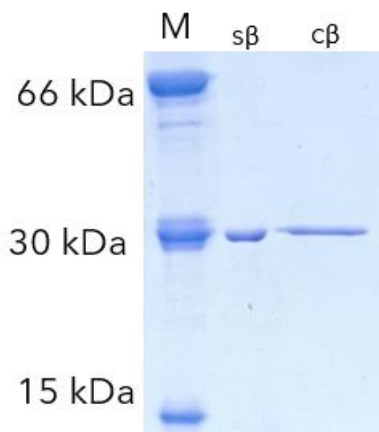
6.3.5 Caracterização bioquímica e estrutural da proteína recombinante purificada (Brasil)

Apesar do problema de tendência a agregação, ao adotar as estratégias de evitar baixas temperaturas, não concentrar e o uso de glicerol, foi possível produzir quantidades suficientes da proteína rAmLpCHT3GNco para os ensaios.

6.3.5.1 Estimativa da massa molecular

A massa molecular estimada de rAmLpCHT3GNco foi de aproximadamente 30 kDa, em condições redutoras (c β) e não redutoras (s β) (Figura 86).

Figura 86 – Estimativa da massa molecular de rAmLpCHT3GNco.



Fonte: Autora. Gel 15%. M: Marcador molecular contendo: HSA – 66 kDa; rSFL - 30 kDa e CCL – 15 kDa. Sβ: na ausência de βmercaptoetanol. Cβ: na presença de βmercaptoetanol.

Nesse gel, a banda referente a proteína ficou exatamente na mesma altura do marcador de 30 kDa. No entanto, ao longo de toda essa seção, os géis mostraram a proteína expressa e pura acima do marcador de 30 kDa, de acordo com o esperado para a proteína rAmLpCHT3GNco com tamanho calculado de 33 kDa.

Esse resultado pode ser devido a uma migração eletroforética ligeiramente diferente nesse gel específico, fazendo com que as bandas coincidissem visualmente com o marcador de 30 kDa, principalmente porque 3 kDa não é uma diferença significativa em gel de poliacrilamida. Em SDS-PAGE, a massa estimada é aparente e pode variar ligeiramente com pequenas diferenças na concentração do gel, na quantidade de acrilamida, na voltagem, no tampão e no tempo de corrida (Idowu *et al.*, 2025).

A lectina recombinante apresentou o mesmo tamanho na ausência e na presença de β-mercaptoetanol, corroborando com os resultados da lectina nativa. AmL é um monômero sem a presença de pontes dissulfeto (Silva, 2016; Costa *et al.*, 1999).

6.3.5.2 Ensaio de atividade hemaglutinante da proteína recombinante e nativa

O ensaio de hemaglutinação foi realizado com a proteína recombinante e com a proteína nativa, a fim de comparar a preferência de ambas por eritrócitos de coelho tratados ou não tratados com a enzima tripsina.

A propriedade hemaglutinante comprova a integridade funcional da proteína recombinante pois demonstra a capacidade de reconhecer seletivamente resíduos específicos de carboidratos na superfície das hemácias, mostrando funcionabilidade pela característica fundamental das lectinas, a capacidade de se ligar a glicanos (O'connor; Monaghan; Cawley, 2017).

A lectina recombinante rAmLpCHT3GNco apresentou atividade hemaglutinante para os eritrócitos de coelho testados, tanto na forma nativa como na forma tratado com a protease tripsina, assim como observado para a lectina nativa AmL. Os títulos de hemaglutinação do ensaio estão resumidos na Tabela 23.

Tabela 23 – Título de hemaglutinação do ensaio para a lectina recombinante e nativa.

	rAmLpCHT3GNco	AmL
Normal	1024 U.H./mL	128 U.H./mL
Tripsina	8192 U.H./mL	256 U.H./mL

Fonte: O título de hemaglutinação é definido como a maior diluição da amostra capaz de promover hemaglutinação visível, expresso em unidade hemaglutinantes por mililitro (U.H./mL). As lectinas estão na concentração de 0,16 mg/mL.

A rAmLpCHT3GNco a uma concentração de 0,16 mg/mL foi capaz de aglutinar eritrócitos nativos com um título de hemaglutinação de 1.024 U.H./mL, 8 vezes a mais que a lectina nativa na mesma concentração. Para eritrócitos tratados com tripsina, o título de hemaglutinação da lectina recombinante foi 8.192 U.H./mL, 32 vezes a mais que a lectina nativa

Costa e colaboradores (1999) não relatam o exato título de hemaglutinação, mas mencionam que AmL demonstrou preferência por eritrócitos de coelho tratados com tripsina, assim como foi observado no presente ensaio para a lectina nativa e recombinante. Esse aumento do título de hemaglutinação observado para os eritrócitos tratados com tripsina ocorre, pois, a enzima remove proteínas da superfície da célula e expõe os resíduos de carboidratos da membrana, deixando os glicanos de preferencias da lectina mais acessíveis para a interação (Sampaio; Rogers; Barwell, 1998).

Outras lectinas recombinantes de algas marinhas vermelhas também apresentaram a mesma preferência por eritrócitos tratados com tripsina observado para a respectiva lectina nativa, como as lectinas recombinantes de *Hypnea musciformis*, *Alsidium triquetrum* (Nascimento, 2014; Nascimento *et al.*, 2015; Nagano *et al.*, 2005). e *Solieria filiformis* (Chaves *et al.*, 2023; Chaves *et al.*, 2018).

A atividade hemaglutinante mais elevada observada para a lectina recombinante pode ser explicada pelas diferenças estruturais e principalmente de homogeneidade em relação à proteína nativa. As lectinas na natureza são encontradas como uma mistura de isoformas com variações estruturais e biológicas. No entanto, a lectina recombinante apresenta uma maior homogeneidade estrutural por possuir apenas uma variação de isoforma, o que pode intensificar a atividade por estar mais pura e não mesclada com outras variações que apresentam atividade de diferentes intensidades (Oliveira; Teixeira; Domingues, 2013; Martínez-Alarcón; Blanco-Labra; García-Gasca, 2018).

Esse mesmo comportamento também foi observado para a lectina recombinante de *Bryopsis plumosa* que apresentou atividade hemaglutinante mais elevada para essa forma do que para a proteína nativa (Hwang *et al.*, 2018).

6.3.5.3 Ensaio de inibição da atividade hemaglutinante com a proteína recombinante e nativa

A proteína recombinante e a proteína nativa foram submetidas a um ensaio de inibição da atividade hemaglutinante para determinar a especificidade a carboidratos das lectinas, a fim de comparação.

O perfil de inibição da atividade hemaglutinante de rAmLpCHT3GNco e AmL, assim como a concentração mínima inibitória do açúcar capaz de inibir a aglutinação de eritrócitos estão mostrados na Tabela 24.

Tabela 24 – Perfil de inibição da atividade hemaglutinante de rAmLpCHT3GNco e AmL.

AmL	CMI	rAmLpCHT3GNco	CMI
<u>Glicoproteínas</u>			
PSM tipo II	0,125 mg/mL	PMS tipo II	62,5 µg/mL
Fetuína	0,125 mg/mL	Fetuína	0,125 mg/mL
Manana	15,62 µg/mL	Manana	15,62 µg/mL
BSM	0,250 mg/mL	BSM	0,5 mg/mL
Asialofetuína	15,62 µg/mL	Asialofetuína	62,5 µg/mL
Tiroglobulina	1,95 µg/mL	Tiroglobulina	7,81 µg/mL
<u>Glicolipídeos</u>			
Ácido tecóico	7,81 µg/mL	Ácido tecóico	62,5 µg/mL
<i>Staphylococcus aureus</i>		<i>Staphylococcus aureus</i>	
LPS	0,125 mg/mL	LPS	0,5 mg/mL
(lipopolissacarídeo) <i>E. coli</i>		(lipopolissacarídeo) <i>E. coli</i>	

Fonte: Autora. PSM: Mucina do estômago de porco. BSM: Mucina submaxilar bovina. CMI = Concentração mínima inibitória. Os estoques de glicoproteínas e glicolipídeos estão em 1 mg/mL.

A lectina recombinante rAmLpCHT3GNco foi capaz de reconhecer um perfil de carboidratos idêntico ao da lectina nativa, com algumas diferenças apenas na intensidade do reconhecimento.

A análise comparativa entre as duas formas da lectina mostrou que, para a maioria dos carboidratos, rAmLpCHT3GNco foi capaz de reconhecer o ligante, mas apresentou menor afinidade de ligação, demonstrado pelos valores maiores da concentração mínima inibitória (CMI) quando comparado com a AmL nativa. Esse padrão foi observado para mucina submaxilar bovina, asialofetuína, tiroglobulina, ácido tecoico de *S. aureus* e lipopolissacarídeo de *E. coli*.

Enquanto para mucina do estômago de porco, rAmLpCHT3GNco demonstrou maior capacidade inibitória da atividade hemaglutinante, sugerindo interações mais fortes em relação à lectina nativa. Já para fetuína e manana, rAmLpCHT3GNco interagiu com essas glicoproteínas na mesma intensidade da lectina nativa. Costa e colaboradores (1999) também relatam a inibição da atividade hemaglutinante da AmL por manana e fetuína.

O mesmo comportamento foi observado para a lectina recombinante de *Alsidium triquetrum* (Nascimento *et al.*, 2015), onde a lectina recombinante apresentou a mesma especificidade que a lectina nativa, porém em menor intensidade.

Os perfis de inibição das KAAs recombinantes de *Kappaphycus alvarezii* também foram semelhantes com os das lectinas nativas. Outro ponto abordado nesse caso, foi a comparação da inibição entre a proteína recombinante clivada e não clivada na presença da cauda de histidina (His-rKAA). A presença da His-tag não resultou em mudanças no perfil de inibição da proteína recombinante (Hirayama *et al.*, 2016). Da mesma forma, a cauda de rAmLpCHT3GNco não pareceu impactar no reconhecimento a carboidrato da proteína, visto que o perfil de inibição foi o mesmo para a lectina nativa e a diminuição da intensidade de reconhecimento também foi observada para a lectina rBTL que estava clivada (Nascimento *et al.*, 2015).

As lectinas de algas vermelhas tendem a ter preferência por glicoproteínas, se ligando preferencialmente a *N*- glicanos. De acordo com essa especificidade, as lectinas de algas podem ser classificadas em lectinas específicas para *N*-glicanos do tipo alto-manose, *N*-glicanos fucosilados e glicanos com resíduos terminais de Galactose e Glicose. Dentre essas, a preferência de ligação mais observada é do tipo alta-manose (Costa Filho *et al.*, 2025).

A análise do perfil de inibição da atividade hemaglutinante indica que tanto a lectina recombinante quanto a lectina nativa, apresentam preferências por glicanos complexos com

maior afinidade por *N*-glicanos com estruturas ricas em alta manose, corroborando com a preferência de ligação das lectinas pertencentes à família OAAH, onde AmL está inserida (Costa Filho *et al.*, 2025; Silva, 2016).

Essa especificidade é evidenciada pela ligação preferencial por manana de levedura e tiroglobulina que apresentaram os CMI's mais baixos. Essas glicoproteínas possuem *N*-glicanos do tipo alta manose expostos. Além disso, a inibição significativa por asialofetuína, em comparação a fetuína, sugere que a remoção de resíduos de ácido siálico favorece o reconhecimento da lectina, possivelmente por aumentar a acessibilidade das estruturas internas.

No entanto, as glicoproteínas associadas à *O*-glicanos, como a mucina submaxilar bovina, apresentou uma inibição mais baixa com menor capacidade de ligação, evidenciada pelo CMI mais alto, mostrando que essas estruturas não são ligantes preferenciais. A fetuína, que possui tanto *N* - glicanos do tipo complexo quanto *O* -glicanos, também foi reconhecida, mostrando que a lectina recombinante e nativa reconhece estruturas *O*- glicanos, de forma moderada.

A única glicoproteína que a lectina recombinante foi capaz de reconhecer com mais intensidade que a nativa foi a mucina do estômago de porco tipo II, que também é uma glicoproteína rica em *O*-glicanos. A interação foi moderada, mas dentre as estruturas ligadas à *O*-glicanos, essa foi a interação mais significativa, sugerindo que a lectina reconhece preferencialmente *O*-glicanos que contém galactose e *N*-acetilgalactosamina pouco sialilados presentes nessa estrutura.

A ausência de inibição para as glicoproteínas transferrina, que contém exclusivamente *N*-glicanos do tipo complexo, e para a mucina do estômago de porco tipo III, que é rica em *O*-glicanos fortemente sialilados, indica que a ligação da lectina é desfavorecida por estruturas sialiladas, indicando que a proteína não possui afinidade por glicanos complexos altamente modificados.

Esses resultados também foram observados para as lectinas de *Eucheuma denticulatum* e *Kappaphycus alvarezii*. Essas proteínas também são específicas para *N*-glicanos de alto manose, com a manana de levedura e a tiroglobulina, que contém *N* -glicanos de alto teor de manose sendo as mais inibitórias. As lectinas KAA recombinantes de *K. alvarezii* também mostraram a mesma especificidade preferencial para manana de levedura e tiroglobulina, e também reconheceram asialofetuína preferencialmente a fetuína. Além disso, a mucina submaxilar bovina também apresentou a menor capacidade de ligação, mostrando que a presença de *O*-glicanos não é preferencial. Tanto as lectinas de *Eucheuma denticulatum* como

de *K. alvarezii* também não reconheceram a transferrina, que possui apenas *N*-glicanos do tipo complexo, e não se ligaram a monossacarídeos (Hung *et al.*, 2015; Hirayama *et al.*, 2016). Portanto, rAmL e AmL possuem o comportamento muito próximo ao dessas lectinas.

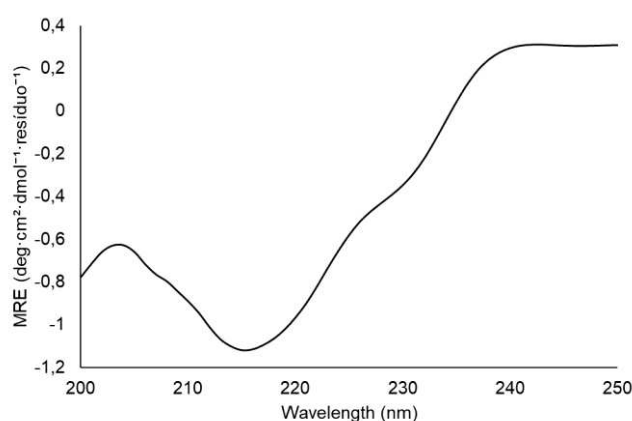
A capacidade de reconhecer o ácido tecóico e o LPS presentes na superfície celular das bactérias *S. aureus* e *E. coli*, respectivamente, indica que a lectina é capaz de se ligar aos glicanos presentes na superfície celular dessas bactérias.

A AML foi capaz de reduzir o número de células viáveis tanto de *S. aureus* quanto de *E. coli*, sendo mais efetiva contra *S. aureus* (Silva, 2016), corroborando com o observado na inibição, pois a lectina tem preferência de ligação para ácido tecóico em comparação ao LPS. Como a lectina recombinante também teve resultados parecidos para o reconhecimento do ácido tecóico e do LPS, provavelmente rAmLpCHT3GNco pode manter a atividade antibacteriana observada para a lectina nativa.

6.3.5.4 Dicroísmo circular

O espectro de dicroísmo circular de rAmLpCHT3GNco apresentou um mínimo em aproximadamente 215-218 nm e uma absorção máxima entre 240-250 nm, um padrão característico de proteínas ricas em β -folhas (Greenfield, 2006) (Figura 87).

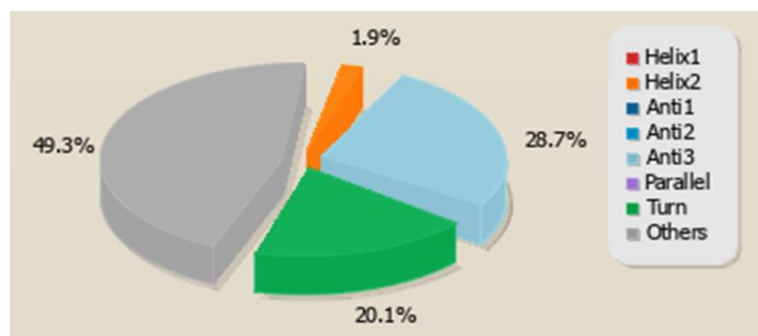
Figura 87 – Espectro de dicroísmo de rAmLpCHT3GNco.



Fonte: Autora. O eixo X representa o comprimento de onda em nanômetros (nm). O eixo Y representa o MRE (*Mean Residue Ellipticity* = elipticidade média por resíduo), um parâmetro normalizado que reflete a diferença de absorção da luz circulamente polarizada à direita e à esquerda. A lectina estava em Tampão Tris 20 mM pH 8; NaCl 150 mM e Glicerina 5%; a 0,234 mg/mL.

Os espectros obtidos foram processados pela plataforma BestSel (Kardos *et al.*, 2025; Micsonai *et al.*, 2022) e a estrutura secundária da lectina recombinante foi prevista com 1,9 % de α -hélice 2 (Helix2), 28,7% de β -folha antiparalela 3 (Anti3), 20,1% de voltas (Turn) e 49,3% de estrutura desordenada (Others) (Figura 88). Esses dados mostram a predominância das β -folhas, com baixíssimo conteúdo de α -hélice, consistente com o espectro do dicroísmo.

Figura 88 - Predição da estrutura secundária de rAmLpCHT3GNco.



Fonte: Autora. Gráfico obtido pela plataforma BestSel.

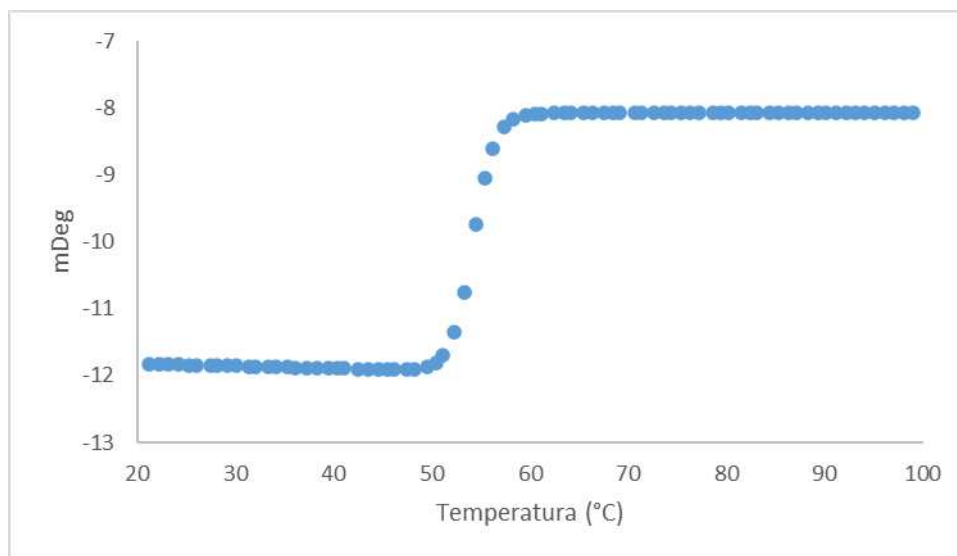
Os espectros para AmL nativa também exibiram valores de absorção mínima e máxima semelhantes com a lectina recombinante, com a mínima em 215 nm e a máxima em 245-250 nm, demonstrando uma predominância também de β -folhas em sua estrutura secundária (Silva, 2016). Mesquita e colaboradores (2021) determinaram por dicroísmo circular a estrutura secundária da lectina nativa. AmL é composta por 4% de α -hélices, 43% de folhas β , 21% de voltas e 32% de conformação desordenada.

rAmLpCHT3GNco apresentou perfis de espectro e proporções de estrutura secundária semelhante a proteína nativa, indicando que a proteína recombinante conserva a estrutura essencial da proteína. Portanto, a lectina recombinante foi expressa com um dobramento correto e semelhante ao da lectina nativa, sem alterações significativas na estrutura global.

A predominância de β -folhas é uma característica de lectinas de algas vermelhas, e é comum em outras lectinas da família OAAH que exibem a estrutura β -barril como observado na estrutura da lectina rSfL-1 (Costa Filho *et al.*, 2025; Chaves *et al.*, 2023).

A estabilidade térmica da proteína recombinante também foi avaliada em um intervalo de temperatura de 20 °C a 100 °C (Figura 89).

Figura 89 – Espectro de dicroísmo de rAmLpCHT3GNco em função da temperatura.



Fonte: Autora. O eixo X representa o incremento da temperatura (°C). O eixo Y representa o mDeg (*Millidegrees* = milidegraus), é a unidade do sinal de elipticidade. A lectina estava em Tampão Tris 20 mM pH 8; NaCl 150 mM e Glicerina 5%; a 0,234 mg/mL. A temperatura média de transição térmica (Temperatura de *melting* - T_m) foi definida como o ponto correspondente a 50 % da variação total de elipticidade entre os estados nativo e desnaturado. A análise da T_m foi em UV distante de 215 nm. A variação de temperatura foi conduzida de forma controlada, com o aumento de 3 °C por minuto, em um intervalo de temperatura de 20°C a 100°C. A ferramenta *Thermal Denaturation Analysis* foi empregada para ajustar os dados experimentais.

O espectro apresentou um platô entre 20 °C e 50 °C, essa constância indica que a proteína está estruturalmente estável nessa faixa de temperatura. A partir de 50 °C até 60 °C, ocorre uma grande variação do sinal, correspondendo ao processo de desnaturação térmica. A partir de 60 °C, a curva se estabiliza indicando que a proteína está completamente desnaturada.

A temperatura de melting, determinada no ponto médio da transição, foi de aproximadamente 54,5 °C, indicando uma estabilidade térmica moderada. Esses dados também estão de acordo com a T_m observada para rAmLpCHT3GNco no teste de deslocamento térmico. A lectina recombinante apresentou uma variação da T_m entre 47 °C a 57 °C de acordo com as condições tamponantes testadas.

Esse comportamento também foi semelhante ao da lectina nativa que demonstrou uma resistência térmica estrutural até 40 °C. A partir dessa temperatura, uma perda gradual de sua estrutura secundária aconteceu com o aumento da temperatura, levando a proteína a um estado mais desdobrado a 85 °C. A temperatura de melting de AmL foi estimada em 42 °C (Mesquita *et al.*, 2021).

Mudanças conformacionais foram observadas entre a estabilidade térmica da proteína recombinante e da proteína nativa. Para a proteína recombinante, a desnaturação ocorre de forma abrupta, entre 50 °C até 60 °C e estabiliza. Enquanto para a proteína nativa, ocorre de forma gradual e só se estabiliza em 85 °C. Apesar das duas lectinas possuírem Tms próximas, a estabilidade térmica foi maior para a proteína recombinante.

Esse comportamento de mudança abrupta do estado estrutural revela a presença de uma amostra estruturalmente mais homogênea, enquanto uma desnaturação gradual sugere a presença de muitos estados conformacionais ou isoformas que se desnaturam progressivamente (Cooper, 2005). Portanto, a presença de isoformas nas formas nativas mantém essa resistência térmica para uma desnaturação rápida, e com a proteína recombinante, que apresenta apenas uma isoforma, a mudança estrutural acontece rapidamente.

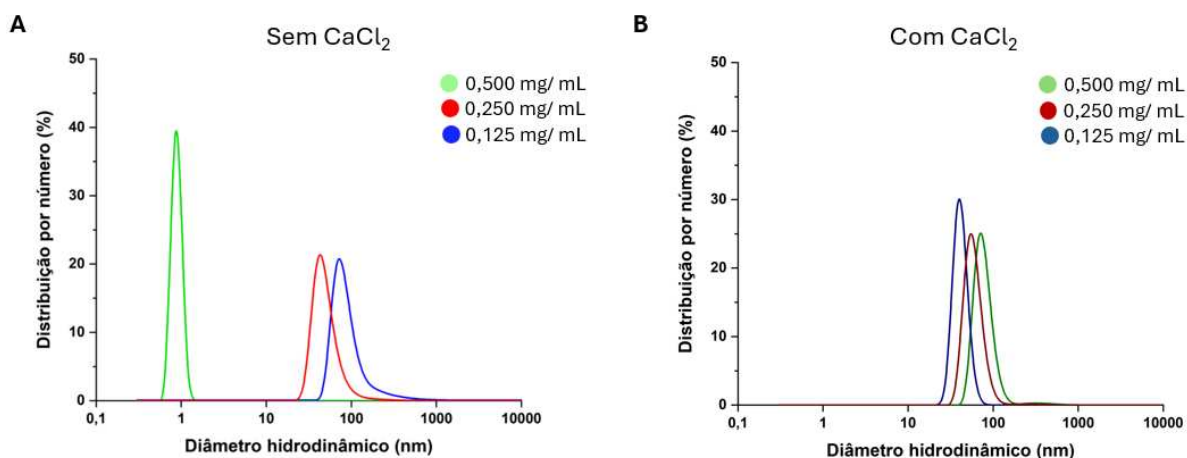
A lectina recombinante rSfL-1, outra lectina da família OAAH, também apresentou estrutura secundária bem próxima da lectina nativa. As temperaturas de melting foram bem semelhantes à Tm de rAmLpCHT3Gnco e AmL, de 53 °C para SfL e 41 °C para rSfL-1. No entanto, nesse caso, a estabilidade térmica da proteína recombinante foi menor que a nativa.

6.3.5.5 Espalhamento dinâmico de luz - DLS (Dynamic Light Scattering)

O comportamento hidrodinâmico da lectina recombinante foi avaliado pela análise de DLS nas concentrações proteicas de 0,5 mg/mL, 0,250 mg/mL e 0,125 mg/mL, em condições de tampão, com e sem cloreto de cálcio, a fim de identificar os parâmetros que minimizam a formação de agregados, e consequentemente apresentam menor tendência a precipitação.

A distribuição do diâmetro hidrodinâmico de rAmLpCHT3GNco para todas as condições analisadas é mostrada na Figura 90 e os parâmetros avaliados nas análises estão resumidos nas Tabelas 25 e 26.

Figura 90 - Distribuição do tamanho hidrodinâmico por números de partículas da lectina rAmLpCHT3GNco.



Fonte: Autora. Distribuição do tamanho hidrodinâmico (nm) por número de partículas da lectina (%) a 0,5 mg/mL (verde), 0,250 mg/mL (vermelho) e 0,125 mg/mL (azul). A: Sem CaCl_2 . B: Com CaCl_2 . O eixo x indica o diâmetro hidrodinâmico em nanômetros (nm). O eixo y indica a distribuição por números em porcentagem das partículas (%). A amostra foi analisada em tampão Tris 20 mM pH 8, NaCl 150 mM e glicerol 5%, csem ou com cloreto de cálcio 20 mM. Os dados representam a média de três replicatas independentes indicados pela linha verde para 0,500 mg/mL, linha vermelha para 0,250 mg/mL e linha azul para 0,125 mg/mL.

Tabela 25 - Parâmetros do espalhamento dinâmico de luz de rAmLpCHT3GNco para as três concentrações sem cloreto de cálcio.

<u>SEM CaCl_2</u>	<u>Média</u> <u>0,500 mg/mL</u>	<u>Média</u> <u>0,250 mg/mL</u>	<u>Média</u> <u>0,125 mg/mL</u>
<u>Diâmetro médio (nm)</u>	1417	296,5	297,4
<u>Índice de polidispersidade (PI)</u>	1	0,528	0,496
<u>Pico 1: média por intensidade ordenada</u> <u>por área (nm)</u>	2318	617	431,3
<u>Pico 1: área por intensidade ordenada</u> <u>por área (%)</u>	92,72	99,33	95,78
<u>Pico 2: média por intensidade ordenada</u> <u>por área (nm)</u>	0,9033	4837	5070
<u>Pico 2: área por intensidade ordenada</u> <u>por área (%)</u>	7,277	0,6683	4219

Fonte: Autora. Resumo dos dados obtidos por DLS para rAmLpCHT3GNco a 0,500 mg/mL, 0,250 mg/mL e 0,125 mg/mL sem cloreto de cálcio. Valores calculados a partir da distribuição por intensidade, incluindo diâmetro médio (nm), índice de polidispersão (PI) e os principais picos identificados nas distribuições baseadas em intensidade.

Tabela 26 - Parâmetros do espalhamento dinâmico de luz de rAmLpCHT3GNco para as três concentrações com cloreto de cálcio.

<u>COM CaCl₂</u>	Média 0,500 mg/mL	Média 0,250 mg/mL	Média 0,125 mg/mL
<u>Diâmetro médio (nm)</u>	311,7	273,2	269
<u>Índice de polidispersidade (PI)</u>	0,4716	0,4786	0,5348
<u>Pico 1: média por intensidade ordenada por área (nm)</u>	520,9	416	326,2
<u>Pico 1: área por intensidade ordenada por área (%)</u>	79,84	84,46	87,27
<u>Pico 2: média por intensidade ordenada por área (nm)</u>	105,5	84,31	48,12
<u>Pico 2: área por intensidade ordenada por área (%)</u>	19,02	15,54	8,698

Fonte: Autora. Resumo dos dados obtidos por DLS para rAmLpCHT3GNco a 0,500 mg/mL, 0,250 mg/mL e 0,125 mg/mL com cloreto de cálcio. Valores calculados a partir da distribuição por intensidade, incluindo diâmetro médio (nm), índice de polidispersão (PI) e os principais picos identificados nas distribuições baseadas em intensidade.

A condição de rAmLpCHT3GNco a 0,500 mg/mL na ausência de cloreto de cálcio apresentou resultados muito discrepantes em relação às outras análises, caracterizado por um diâmetro hidrodinâmico extremamente alto e um índice de polidispersidade máximo ($PDI = 1$), incompatível com o observado para as demais análises. Esse perfil pode sugerir alguma instabilidade durante a medição ou presença de algum agregado macroscópico. Por isso, essa análise não será considerada representativa do comportamento real da proteína em solução, e será desconsiderada na interpretação comparativa dos dados.

De forma geral, as análises mostraram que rAmLpCHT3GNco tem tendência à agregação em solução e com distribuição de tamanhos heterogêneos em solução.

Para as análises sem cloreto de cálcio, o diâmetro hidrodinâmico médio foi de aproximadamente 296 nm e 297 nm para as concentrações de 0,250 mg/mL e 0,125 mg/mL, respectivamente. O índice de polidispersidade (PDI) foi de 0,53 e 0,50 para as concentrações de 0,250 mg/mL e 0,125 mg/mL, respectivamente.

Um diâmetro médio de 300 nm é muito maior do que o esperado para uma proteína monomérica, indicando a presença de agregados em solução (ERICKSON, 2009). E valores de PDI maiores que 0,4 indicam uma amostra altamente dispersa, com a presença de populações de tamanhos distintos (Maguire *et al.*, 2018). A presença de dois picos principais em ambas as concentrações analisadas indica a coexistência de populações de tamanhos distintos, comprovando a natureza heterogênea da amostra (Tabela 25).

A distribuição por números (Figura 90 - A) indica que a maior parte das partículas de rAmLpCHT3GNco se encontra entre 10 e 100 nm, porém o diâmetro hidrodinâmico médio foi de 296 nm e 297 nm nas concentrações de 0,250 mg/mL e 0,125 mg/mL, respectivamente. Como o diâmetro médio é calculado a partir da distribuição por intensidade, a presença de uma população menor de agregados com maior tamanho, é suficiente para deslocar a média para valores maiores, reforçando a natureza polidispersa da amostra (Maguire *et al.*, 2018).

Esse princípio do DLS explica o deslocamento mais para a esquerda do gráfico de distribuição por números para a concentração de 0,250 mg/mL, pois apesar do diâmetro médio e do PDI serem praticamente iguais, esse deslocamento indica a presença, em maior quantidade, de partículas menores na amostra a 0,250 mg/mL, porém a presença de agregados maiores eleva o aumento do diâmetro médio e do PDI.

A análise da proteína com cálcio revelou que rAmLpCHT3GNco também apresenta distribuição de tamanhos heterogênea em solução, com PDI de 0,47 para as concentrações de 0,5 mg/mL e 0,250 mg/mL, e PDI de 0,58 para a concentração de 0,125 mg/mL. Além disso, os diâmetros hidrodinâmicos médios foram de 311,7 nm; 273,2 nm e 269 nm para as concentrações de 0,5 mg/mL, 0,250 mg/mL e 0,125 mg/mL. Novamente, diâmetros médios muito acima do que o esperado, indicando a presença de agregados em solução (Erickson, 2009). A presença de dois picos principais em ambas as concentrações analisadas indica a coexistência de populações de tamanhos distintos, comprovando a natureza heterogênea da amostra (Tabela 26) (Maguire *et al.*, 2018).

Esses parâmetros indicam que as condições testadas não variaram muito, porém a amostra mais diluída apresenta uma menor agregação devido a um menor raio hidrodinâmico. Além disso, a distribuição por números (Figura 90 - B) mostra que quanto maior a concentração, mais o gráfico se desloca para a direita, indicando o aumento do raio hidrodinâmico das partículas. À medida que a concentração aumenta, o raio hidrodinâmico médio também aumenta, indicando maior agregação.

De modo geral, comparativamente, não houve diferenças muito significativas entre as amostras com cloreto de cálcio e sem cloreto de cálcio, tanto o raio hidrodinâmico por distribuição de números (gráficos), como por intensidade (Tabelas), resultou em tamanhos bem próximos. No entanto, na presença do cloreto de cálcio os números do tamanho hidrodinâmico foram menores. Possivelmente, indicando o mesmo comportamento já mencionado no ensaio do deslocamento térmico. Nas análises, o cloreto de cálcio resulta em estabilizações de pequena intensidade, o que não impacta em uma melhora efetiva da tendência a agregação.

Portanto, todas as análises comprovam que rAmLpCHT3GNco apresenta tendência a agregação, com uma discreta estabilização na presença de cloreto de cálcio, onde a proteína permanece mais estável em concentrações menores.

As análises de DLS para as proteínas recombinantes rBTL e rHML, também mostraram formação de agregados com alta dispersidade das amostras, principalmente em concentrações maiores (Nascimento, 2014), mostrando que a formação de agregados não é um problema observado apenas em rAmLpCHT3GNco, mas também outras lectinas recombinantes apresentaram essa desvantagem.

Nesse contexto, as interações entre a lectina recombinante e seu ligante foram planejadas para serem avaliadas por calorimetria de titulação isotérmica (ITC) utilizando um sistema MicroCal PEAQ-ITC (Malvern Instruments). No entanto, a concentração necessária para realizar a análise seria de 100 µg da proteína, o que exigiria uma concentração maior de 3 mg/mL. A execução do experimento foi inviabilizada pois a proteína recombinante apresentou elevada precipitação na concentração necessária, e não foi possível concentrar a proteína além do que 1 mg/mL.

6.3.6 Avaliação da atividade antibacteriana combinada com antibióticos (Brasil)

Antes do teste biológico, a lectina recombinante rAmLpCHT3GNco a 0,5 mg/mL foi submetida a um teste de estabilidade térmica a 37 °C para avaliar se a cauda interferiria na estabilidade térmica da proteína. Após uma hora, a atividade hemaglutinante foi mantida e não foi observada precipitação significativa (dados não mostrados).

Além disso, a amostra da lectina usada no teste biológico, na mesma concentração de 62,4 µg/mL, apresentou um título de hemaglutinação de 2.048 U.H./mL.

A lectina recombinante foi testada juntamente aos antibióticos oxacilina, tetraciclina, ampicilina e estreptomicina para avaliar se era capaz de potencializar o efeito desses fármacos contra as cepas de *S. aureus* ATCC 7006998, *S. aureus* ATCC 25923 e *E. coli* ATCC 11303. Os resultados das análises estão apresentados na Tabela 27.

Tabela 27 - Efeito da lectina rAmLpCHT3GNco combinada aos antibióticos oxacilina, tetraciclina, ampicilina e estreptomicina contra *S. aureus* e *E. coli*.

BACTÉRIAS	ANTIBIÓTICO			EFEITO	ANTIBIÓTICO			EFEITO
	Oxacilina				Tetraciclina			
	CIM1 µg/mL (individual)	CIM2 µg/mL (combinado)	Relação entre CIM		CIM1 µg/mL (individual)	CIM2 µg/mL (combinado)	Relação entre CIM	
<i>S. aureus</i> ATCC 25923	500	250	1/2	Aditivo	0,976	0,244	1/4	Sinergismo
<i>S. aureus</i> ATCC 700698	62,5	31,25	1/2	Aditivo	62,5	15,625	1/4	Sinergismo
<i>E. coli</i> ATCC 11303	125	125	1	Não interação	0,488	0,122	1/4	Sinergismo
	Ampicilina				Estreptomicina			
	CIM1 µg/mL (individual)	CIM2 µg/mL (combinado)	Relação entre CIM		CIM1 µg/mL (individual)	CIM2 µg/mL (combinado)	Relação entre CIM	
<i>S. aureus</i> ATCC 25923	0,976	0,976	1	Não interação	15,625	15,625	1	Não interação
<i>S. aureus</i> ATCC 700698	0,976	0,976	1	Não interação	7,8	3,9	1/2	Aditivo
<i>E. coli</i> ATCC 11303	0,976	>1	>2	Antagônico	15,625	15,625	1	Não interação

Fonte: Autora. Valores da Concentração Inibitória Mínima individual (CIM1), Concentração Inibitória Mínima Combinada (CIM2) dos antibióticos combinados com a lectina rAmLpCHT3GNco, e da relação entre CIM2/CIM1. As concentrações dos antibióticos variaram de 500 a 125 µg/mL para Oxacilina, 62,5 a 0,488 µg/mL para Tetraciclina, de 0,976 µg/mL para Ampicilina, e de 15,625 a 7,8 µg/mL para Estreptomicina. A concentração de rAmLpCHT3GNco se manteve constante (62,5 µg/mL). Valores da relação entre as CIM iguais a 1 foram interpretados como não interação, a 1/2 como efeito aditivo, e valores entre 1/4 a 1/16 foram considerados como efeito sinérgico. Valores igual a >2 foram considerados como efeito antagônico.

A lectina recombinante apresentou efeito de sinergismo associada apenas ao fármaco tetraciclina para as três cepas testadas e com a mesma intensidade, evidenciada pela redução da concentração inibitória mínima de 1/4 em relação ao CIM apenas do antibiótico.

O sinergismo ocorre quando a combinação dos compostos é mais eficaz que a soma dos efeitos individuais, ou seja, a lectina associada a tetraciclina tem efeito mais potente do que apenas o antibiótico. Nesses casos, a lectina pode aumentar a permeabilidade da parede celular, facilitando entrada do fármaco ou pode se ligar ao glicanos na superfície celular da bactéria, alterando a conformação e aumentando a sensibilidade ao antibiótico (European committee for antimicrobial susceptibility testing (EUCAST) of the european society of clinical microbiology and infectious diseases (ESCMID), 2000; Bollenbach, 2015).

Por outro lado, foram observados efeitos aditivos para as duas cepas de *S. aureus* associado ao antibiótico oxacilina e para *S. aureus* ATCC 700698 associada a estreptomicina, evidenciada pela redução da concentração inibitória mínima de apenas 1/2 em relação ao CIM apenas do antibiótico. Um efeito aditivo é a soma dos efeitos individuais, cada composto age

independente, sem cooperação ou interferência entre eles (European committee for antimicrobial susceptibility testing (EUCAST) of the european society of clinical microbiology and infectious diseases (ESCMID), 2000).

Enquanto o efeito de não interação foi observado para *E. coli* associada a oxacilina, para as duas cepas de *S. aureus* com ampicilina e para *E. coli* e *S. aureus* ATCC 25923 com estreptomicina. Essa não interação é devido a não redução da CIM do fármaco sozinho (=1), e significa que a combinação não altera o efeito do antibiótico. O efeito antagônico foi observado apenas em *E. coli* em conjunto a ampicilina, com um CIM maior que o CIM do fármaco. Esse efeito resulta na menor eficiência do que o antibiótico sozinho, indicando que a lectina prejudica a ação do medicamento, possivelmente por bloquear a entrada do antibiótico ou induzir respostas de estresse bacteriano (European committee for antimicrobial susceptibility testing (EUCAST) of the european society of clinical microbiology and infectious diseases (ESCMID), 2000; Bollenbach, 2015).

Os diferentes perfis de interação observados entre a lectina recombinante e os antibióticos testados podem ser explicados pelos diferentes mecanismos de ação desses fármacos. A tetraciclina atua no interior da célula bacteriana, se ligando à subunidade 30s do ribossomo e inibindo a síntese proteica. O efeito sinérgico observado com esse antibiótico sugere que a lectina pode intensificar a ação, ao se ligar nos glicanos da membrana da bactéria e aumentar a permeabilidade da mesma, facilitando a entrada do antibiótico. A estreptomicina também atua na síntese proteica, porém esse fármaco depende de mecanismos ativos de transporte dependentes de gradiente eletroquímico para sua internalização. Assim, a ação da lectina de se ligar na membrana da bactéria não foi suficiente para potencializar significativamente sua atividade, resultando em efeito aditivo ou não interação (Chopra; Roberts, 2001; Nagarajan, 2024).

Esse feito sinérgico associado a tetraciclina contra *S. aureus* e *E. coli*, demonstra a capacidade da lectina de reconhecer os glicanos presente na superfície celular dessas bactérias, o que corrobora com o ensaio de inibição, onde a proteína foi capaz de reconhecer tanto o ácido tecoico, característicos de bactérias gram-positivas, quanto o LPS presente na parede celular das células gram-negativas.

Além disso, rAmLpCHT3GNco apresentou uma maior afinidade nos ensaios de inibição para ácido tecoico, sugerindo uma interação mais forte. Essa melhor interação pode ter resultado em alterações estruturais mais significativas em *S. aureus*, o que favoreceu um efeito

aditivo associado a alguns antibióticos. Em *E. coli*, só foi observado um efeito sinérgico e para os outros fármacos foi efeito de não interação ou antagônico.

A ampicilina e a oxacilina pertencem a classe dos β -lactâmicos que atua na interrupção da síntese da parede celular por meio da inibição das proteínas ligadoras de penicilina (PBPs), resultando no enfraquecimento da parede celular bacteriana (Peechakara; Basit; Gupta, 2018). Os resultados da associação da lectina com esses fármacos promoveram não interação ou antagonismo. Provavelmente, ao reconhecer os glicanos da superfície celular, as lectinas atuam ajudando na entrada dos antibióticos que agem dentro da célula, como a tetraciclina. Enquanto os antibióticos que atuam na parede celular, como a ampicilina e a oxacilina, não se beneficiam da melhora da permeabilidade.

Apesar desses dois antibióticos pertencerem a mesma classe, existem diferenças funcionais entre eles que pode explicar a diferença dos resultados. A ampicilina é altamente sensível a ação da β -lactamase, enzima que inibe a ação do antibiótico, além de depender do crescimento ativo da bactéria para exercer a atividade (Peechakara; Basit; Gupta, 2018). Assim, ao se ligar à membrana das bactérias, a lectina provocou alterações na reorganização da membrana que pode ter comprometido o crescimento bacteriano e consequentemente a eficácia do fármaco, levando ao antagonismo.

Por outro lado, a oxacilina é mais resistente a degradação pela β -lactamase e apresentam maior estabilidade e afinidade pelas PBPs. Essas características deixam esse fármaco menos dependente da taxa de crescimento bacteriano e menos sensível a modificações estruturais da membrana celular causadas pelas lectinas (Bessa *et al.*, 2015). Por isso, a associação da lectina com esse antibiótico não interferiu de modo negativo, causando até um efeito aditivo.

As doenças infecciosas causadas por microorganismos é um grave problema de saúde pública que causa milhares de mortes por ano ao redor do mundo. A descoberta dos antibióticos como medicamentos antimicrobianos foi um marco revolucionário na cura dessas doenças. Durante muitos anos, a monoterapia com antibióticos apresentou sucesso clínico contra as doenças infecciosas. No entanto, o uso desses fármacos apresenta algumas desvantagens, como espectro limitado e probabilidade de resistência adquirida (Santos; Porto; Cavalcanti, 2021; Theuretzbacher; Piddock, 2019).

De fato, o uso indiscriminado de antibióticos levou a disseminação e a aceleração da resistência microbiana, se tornando uma preocupação atual para a farmacoterapia, com cepas multirresistentes se disseminando rapidamente. A resistência antimicrobiana é um grande

desafio que precisa urgentemente de soluções inovadoras. Como uma estratégia que vem sendo explorada, o uso de terapia combinada tem se tornado uma opção viável e eficaz (Napoleão *et al.*, 2025).

A terapia combinada avalia o efeito simultâneo de dois ou mais compostos ativos, visando obter um efeito global maior do que a soma dos seus efeitos individuais. As vantagens dessa abordagem são a expansão do espectro de ação, menor probabilidade de desenvolvimento de resistência, um efeito aditivo ou sinérgico, diminuição dos efeitos secundários e a possibilidade de superar a resistência aos medicamentos (Walvekar; Gannimani; Govender, 2019; Santos; Porto; Cavalcanti, 2021).

Por isso, nos últimos anos, houve uma intensificação de busca por compostos bioativos com potencial terapêutico de combate a essas doenças. Esse crescente interesse, impulsionou a pesquisa voltada para as terapias alternativas baseada nas lectinas (Napoleão *et al.*, 2025).

As lectinas, devido a sua capacidade de se ligar a carboidratos, são capazes de reconhecer os glicanos presentes nas células bacterianas e desencadear a aglutinação ou a inibição do crescimento bacteriano (Elumalai *et al.*, 2019). Essas proteínas agem inibindo o crescimento celular por meio de diferentes mecanismos, como alteração da permeabilidade celular, redução da absorção de nutrientes, formação de poros e consequente extravasamento do conteúdo extracelular ou por meio da interação com receptores de membrana que promovem respostas intracelulares (Mukherjee *et al.*, 2014).

A aplicação desta biomolécula tem vantagens como menor propensão ao desenvolvimento de resistência aos fármacos, devido ao diferente mecanismo de ação das lectinas em comparação com os antibióticos e potencialização da atividade antimicrobiana (Silva *et al.*, 2019). Apesar do potencial biotecnológico das lectinas como molécula antibacteriana e antibiofilme, nos últimos anos, poucos estudos avaliaram o sinergismo entre lectinas e antibióticos convencionais, mostrando que ainda é um viés biotecnológico que pode ser bastante explorado (Santos; Porto; Cavalcanti, 2021).

A ação da lectina AmL nativa também foi avaliada associada aos fármacos tetraciclina e oxacilina, contra as mesmas cepas testadas com a lectina recombinante. Em uma maior concentração, de 125 µg/mL, AmL teve efeito sinérgico apenas com tetraciclina para *S. aureus* ATCC 25923 e feito aditivo com o mesmo antibiótico para *E. coli* e *S. aureus* ATCC 700698 (Menezes, 2024). Não é correto realizar uma comparação direta entre esses resultados, pois as concentrações testadas não são as mesmas devido aos problemas de precipitação em

rAmLpCHT3GNco. Talvez, a lectina recombinante tenha um efeito mais significativo que a nativa, mas seriam necessários testes na mesma concentração para uma comparação conclusiva.

Outras lectinas de algas vermelhas também apresentaram efeito sinérgico associado a tetraciclina contra *S. aureus*, como a SfL, MeL, BSL, BTL, HML e HCL das algas marinhas vermelhas *Solieria filiformis*, *Meristiella echinocarpa*, *Bryothamnium seaforthii*, *Bryothamnium triquetrum*, *Hypnea musciformis*, *Hypnea cervicornis*, respectivamente (Carneiro *et al.*, 2024), demonstrando que a tetraciclina pode ser um fármaco promissor para ser usado em associação com lectinas de organismos marinhos.

Outras lectinas recombinantes também tiveram sua atividade biológica e potencial biotecnológico explorado com a presença das caudas pequenas de histidina em sua estrutura, assim como rAmLpCHT3GNco.

A His-GRFT da alga marinha *Griffithsin*, apresentou a mesma atividade anti-HIV da lectina nativa, com as mesmas características de ligação à gp120, e essencialmente na mesma intensidade da nativa (Giomarelli *et al.*, 2006). A lectina recombinante de *Kappaphycus alvarezzi*, com a His-tag, também apresentou atividade anti-viral, se ligando à gp120 com altas constantes de associação, assim como a lectina nativa (Hirayama *et al.*, 2016). Esses relatos mostram que apenas a cauda de histidina, não tem efeito significativo na atividade biológica da proteína recombinante.

6.4 Conclusão

A lectina rAmL foi expressa com sucesso em sistema heterólogo de *E. coli*, apresentando características bioquímicas e estruturais muito semelhantes a lectina nativa, confirmando a eficiência do sistema hospedeiro utilizado. A caracterização estrutural e funcional da proteína recombinante amplia o conhecimento sobre a lectina para uma futura aplicação biotecnológica. Além disso, a atividade antibacteriana da lectina em associação com antibióticos contra cepas gram-positivas e gram-negativas, destaca o potencial sinérgico dessa molécula, evidenciando sua relevância biotecnológica.

7. CAPÍTULO IV: CARACTERIZAÇÃO ESTRUTURAL DAS LECTINAS ALL E CCL PRESENTES NAS ESPONJAS MARINHAS *Aplysina lactuca* e *Chondrilla caribensis*

7.1 Objetivos

Objetivo geral

Determinar a estrutura tridimensional das lectinas de esponjas marinhas, ALL (*Aplysina lactuca* lectin) e CCL (*Chondrilla caribensis* lectin), e avaliar as interações das mesmas com seus respectivos ligantes.

Objetivos específicos

- Purificar as lectinas nativas ALL e CCL;
- Obter cristais difratáveis das lectinas ALL e CCL;
- Cocrystalizar as proteínas com seus respectivos glicanos alvo;
- Determinar a estrutura tridimensional das lectinas por cristalografia de raios X;
- Avaliar as interações entre as lectinas e seus ligantes;

7.2 Metodologia

7.2.1 Purificação das lectinas nativas

Os exemplares da esponja marinha *Chondrilla caribensis* foram coletados durante a maré de sizígia na zona entre marés da praia do Pedra rachada, no município do Paracuru, litoral oeste do Ceará. Os espécimes da esponja marinha *Aplysina lactuca* foram coletadas em mergulho autônomo no Parque Estadual da Pedra da Risca do Meio, litoral de Fortaleza, Ceará. A coleta e o acesso ao patrimônio genético da espécie foram autorizados e regulados pelos órgãos competentes (SISBIO ID: 33913-12 e SISGEN ID:A1792FE).

A lectina presente na esponja marinha *Chondrilla caribensis*, CCL, foi obtida seguindo o protocolo previamente estabelecido (Marques *et al.*, 2018). As esponjas congeladas foram liofilizadas e trituradas para a obtenção de um pó fino. O pó foi embebido em dez volumes de tampão Tris 50 mM pH 7,6 com NaCl 150 mM (TBS) contendo 0,1M de PMSF (Fluoreto de fenilmetanosulfonil) (1:10 m/v) 0,1 M e mantido sob agitação. O material foi centrifugado e o sobrenadante submetido a uma cromatografia de afinidade em matriz de SepharoseTM 6B previamente equilibrada com TBS. A lectina foi recuperada pela eluição com lactose 0,2 M em TBS. As frações com a proteína purificada foram dialisadas e armazenadas para uso posterior

A lectina ALL, da esponja marinha *Aplysina lactuca*, foi extraída seguindo o protocolo adaptado por Duarte e colaboradores (2023). Pedacos da esponja foram macerados e homogeneizadas em tampão Tris-HCl 50 mM, pH 8,0; contendo NaCl 150 mM e L- cisteína 5 mM, na proporção de 1:2 (p/v). O extrato foi centrifugado e submetido à precipitação com sulfato de amônio a 70% de saturação. Após 4 horas, o precipitado foi recuperado, solubilizado em um pequeno volume do tampão de extração e submetido a cromatografia de afinidade em matriz de SepharoseTM 6B, previamente equilibrada com o mesmo tampão de extração. A lectina foi eluída com glicina 50 mM, pH 2,6; com NaCl 150 mM e L- cisteína 5 mM. As frações contendo a lectina foram reunidas, dialisadas, liofilizadas, e posteriormente submetidas a cromatografia de exclusão molecular em uma coluna Sephacryl S200. A cromatografia foi conduzida com tampão TBS/Cys, e as frações contendo a lectina foram reunidas, dialisadas e armazenadas para uso posterior.

A pureza das lectinas foi confirmada por eletroforese em gel de poliacrilamida/dodecil sulfato de sódio (SDS-PAGE) (Laemmli, 1970), e a funcionabilidade da

proteína foi validada por atividade hemaglutinante (Sampaio *et al.*, 2002). As frações puras e biologicamente ativas foram reunidas, dialisadas e liofilizadas para os ensaios de cristalização. Para a CCL, uma parte das frações foi mantida em solução.

7.2.2 Determinação da estrutura tridimensional das lectinas ALL e CCL por cristalografia de raios X (França)

7.2.2.1 Cristalização

Para ALL, a proteína liofilizada foi inicialmente ressuspensa em 20 mM Hepes pH 7,5 e 100 mM NaCl contendo 1 mM de D-lactose, a uma concentração de 6,8 mg/mL. Duas amostras diferentes de CCL foram submetidas aos testes de cristalização, a proteína liofilizada e a proteína em solução. Inicialmente, a CCL liofilizada foi dissolvida em 20 mM Hepes pH 7,5; 1 mM CaCl₂ e 5 mM D-lactose, a uma concentração de 8,2 mg/mL, ou em 20 mM Tris pH 7,5; 150 mM NaCl e 5 mM D-lactose, a uma concentração de 14,2 mg/mL.

Em seguida, triagens iniciais robotizadas foram realizadas com as duas proteínas na plataforma de cristalização do laboratório HTX no EMBL (*European Molecular Biology Laboratory-Grenoble*), pelo método de difusão de vapor com gotas apoiadas (*sitting drops*) de 200 nL em placas de 96 poços, na proporção de 1:1 proteína/reservatório. Após a montagem, as placas foram colocadas no robô de imageamento Vis-UV a 20 °C.

Posteriormente, triagens adicionais foram realizadas manualmente, utilizando o mesmo método de difusão de vapor com gota suspensa (*hanging drops*) em placas de cristalização de 24 poços. Para a triagem manual da ALL, as amostras foram submetidas a uma cromatografia de exclusão molecular (SEC) na coluna Enrich 70 (10 cm x 300 mm; Bio Rad), em tampão 20 mM Hepes pH 7,5 e 100 mM NaCl, a uma concentração de 7,1 mg/mL. A triagem foi realizada na presença de 1 mM de D-lactose com os kits comerciais BCS *Screen* e SG1 *Screen*, Box I e II da *Molecular Dimensions*. Adicionalmente, para a obtenção de fases suficientemente boas para o refinamento, ALL foi cocristalizada com 10 mM TbXO₄ (*Crystallophore* nº 1, *Polyvalan*, *Molecular Dimensions*), utilizando 25% PEG 8K, 0,2 M AmSO₄ e 0,1 M MES pH 6,5.

Para os ensaios manuais com a CCL liofilizada, a proteína também passou por cromatografia de exclusão molecular (SEC) usando a coluna Tosoh G3000SW (30 cm x 7,5 mm; Tosoh Bioscience) em tampão 20 mM MES pH 6,5, 100 mM NaCl, 0,5 mM CaCl₂, 0,25

mM TCEP e 17 mM trealose. Em seguida, as amostras foram concentradas a 10,9 mg/mL por meio de ultrafiltração em filtro de 3 kDa. A triagem foi realizada na presença de 1 mM de D-lactose com o kit comercial Morpheus II screen, Box I e II da *Molecular Dimensions*.

Para a triagem manual, em cada poço foram colocados 200 μ L da condição de cristalização e a gota foi composta por 1 μ L da solução de proteína + 1 μ L da condição do poço. Os poços foram vedados com lamínulas siliconadas e as placas incubadas a 20 °C. Após três dias, as placas foram avaliadas em microscópio, com monitoramentos subsequentes realizadas semanalmente.

Para os ensaios com a CCL não liofilizada, a proteína foi cocrystalizada com o ligante 3-sialil-lactose (3SL). A proteína foi preparada em tampão 20 mM Tris pH 8,0 com 100 mM NaCl, e concentrada a 11,8 mg/mL. A cocrystalização foi realizada pela diluição da solução estoque de 3SL (100 mM) em uma razão de 1:10 na solução proteica (concentrações finais de 10 mM e 10,6 mg/mL, respectivamente). A triagem manual foi realizada utilizando o mesmo método com o kit comercial Morpheus II screen, Box I e II da *Molecular Dimensions*. As placas foram incubadas a 20 °C e avaliadas diariamente em microscópio.

7.2.2.2 Difração de raios X e resolução da estrutura

Os cristais de ALL obtidos foram transferidos para uma solução crioprotetora composta pela solução do reservatório adicionada de 10% de PSM (*Paratone-Sucrose Mixture*), sendo em seguida montados em *cryoloops* e rapidamente congelados em nitrogênio líquido para difração. Em seguida, os dados de difração foram coletados na linha de luz Proxima 2 do síncrotron SOLEIL (Saint Aubin, França), utilizando um detector Eiger X9M. Os programas XDS (Kabsch, 2010) e XDSme (Legrand, 2019) foram utilizados para o processamento dos dados. Para a resolução da estrutura foi utilizado inicialmente o fluxo de trabalho AlphaFold MR na nuvem do CCP4, empregando a sequência determinada por espectrometria de massas (Duarte *et al.*, 2023).

Os cristais obtidos com o agente de faseamento Terbium (Tb) foram transferidos para uma solução contendo 28% PEG 8K, 0,2 M AmSO₄ e 0,1 M MES pH 6,5 para crioproteção, antes da montagem em *cryoloop* e do congelamento rápido em nitrogênio líquido. O conjunto de dados MAD (Difração anômala de múltiplos comprimentos de onda – *Multi-wavelength Anomalous Diffraction*) foi coletado na linha de luz Proxima 2 do síncrotron Soleil e, após o processamento com XDS (Kabsch, 2010) e XDSme (Legrand, 2019), a determinação

experimental de fases foi realizada utilizando o fluxo de trabalho Crank2 EP na nuvem do CCP4, empregando os conjuntos de dados no pico e no alto remoto (Skubak, 2013; Krissinel, 2022). A revisão da estrutura foi realizada por meio de 10 ciclos de *refmacat* (Yamashita, 2023; Murshudov, 2011), seguida por refinamento por máxima verossimilhança com restrições através do programa *REFMAC 5.8* (Murshudov, 2011), de forma iterativa com a reconstrução manual no programa *Coot* (Emsley, 2010). Cinco por cento das observações foram reservadas para análise de validação cruzada. Os átomos de hidrogênio foram adicionados para cálculos de geometria e de fatores de estrutura. A incorporação do ligante foi realizada após inspeção dos mapas ponderados $2F_o-DF_c$. O modelo final foi validado com o servidor de validação *wwPDB* (<http://wwpdb-validation.wwpdb.org>).

Para CCL, os cristais obtidos de CCL-Lactose e CCL-3SL foram montados diretamente em *cryoloops* e rapidamente congelados em nitrogênio líquido. Os dados de difração foram coletados a 100 K no síncrotron SOLEIL (Saint Aubin, França), nas linhas de luz Proxima 1 e 2, utilizando detectores Eiger 16M e Eiger X9M para os complexos com lactose e 3SL, respectivamente. Os programas *XDS* (Kabsch, 2010) e *XDSme* (Legrand, 2019) foram utilizados para o processamento dos dados, e as etapas subsequentes foram realizadas com o CCP4, versão 8.XX (Agirre, 2023).

A estrutura CCL-lactose foi resolvida por substituição molecular utilizando o PHASER e o modelo da ALL, após processamento com *Chainsaw*, como modelo de busca (Stein, 2008; McCoy, 2007). Para o complexo CCL-3SL, as coordenadas de CCL-lactose foram utilizadas como modelo de busca para substituição molecular no Phaser. Esse modelo foi inicialmente reconstruído utilizando o *ARP/WARP* (Langer, 2008). Em seguida, ambas as estruturas foram refinadas por máxima verossimilhança com restrições utilizando o *REFMAC 5.8* e restrições locais de NCS (Murshudov, 2011), de forma iterativa com a reconstrução manual no *Coot* (Emsley, 2010). Cinco por cento das observações foram reservadas para análise de validação cruzada, e átomos de hidrogênio foram adicionados para cálculos de geometria e de fatores de estrutura. A incorporação do ligante foi realizada após a inspeção dos mapas ponderados $2F_o-DF_c$ depois do primeiro ciclo de refinamento. O modelo foi validado com o servidor de validação do *wwPDB* (<http://wwpdb-validation.wwpdb.org>).

7.2.3 Análise de calorimetria por titulação isotérmica (França)

Para avaliar as interações entre as lectinas ALL e CCL e seus potenciais ligantes alvos usados nos ensaios de cristalização, as análises de calorimetria foram realizadas em um calorímetro MicroCal PEAQ- 31 ITC (Malvern).

As lectinas foram avaliadas em tampão HEPES 20 mM pH 7,5 contendo 100 mM NaCl, na concentração de 100 μ M. Os ligantes foram solubilizados no mesmo tampão na concentração de 10 mM. Os glicanos usados foram D-lactose e *N*-Acetil-D-Lactosamina para as proteínas ALL e CCL, respectivamente.

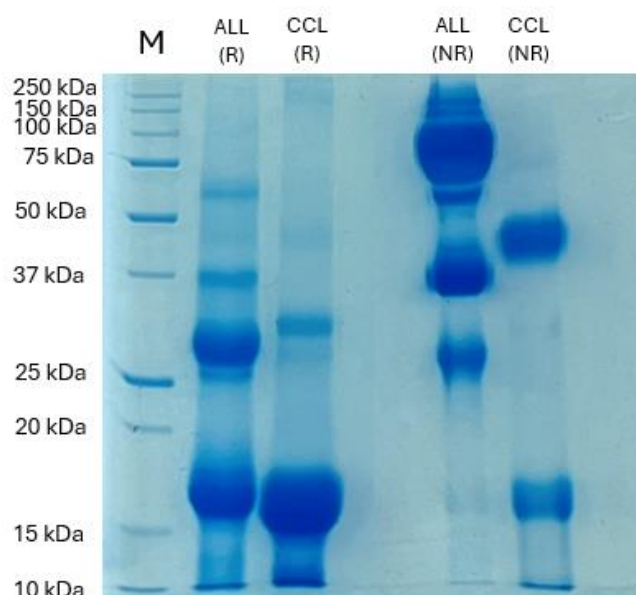
Um total de 20 injeções de 2 μ L de solução do ligante foram adicionadas em intervalos de 120 s, na respectiva solução de proteína, enquanto a solução proteica era agitada a 1000 rpm na célula de 200 μ L. Os experimentos foram conduzidos a 25°C. O controle negativo foi realizado com a injeção de tampão na solução proteica, produzindo variações térmicas de diluição insignificantes. Os dados experimentais dos parâmetros K_d (constante de dissociação), N (número de sítios), ΔH variação de entalpia, $-T\Delta S$ (variação de entropia) e ΔG (energia livre) foram ajustados a uma curva de titulação teórica usando o *software* MicroCal PEAQ-ITC Analysis (Malvern), baseadas no modelo *One Set of Sites*, com ΔH (variação de entalpia), K_a (constante de associação) e n (número de sítios de ligação por monômero) como parâmetros ajustáveis, aplicado a cada titulação seguindo as diretrizes do fabricante. A variação de energia livre (ΔG) e as contribuições de entropia (TDS) foram derivadas da equação: $\Delta G = \Delta H - TDS = 2RT \ln K_a$ (com T sendo a temperatura absoluta e $R = 8,314 \text{ J mol}^{-1} \text{ K}^{-1}$). A constante de afinidade foi determinada pela equação: $K_a = 1/K_d$. Duas titulações independentes foram realizadas para cada açúcar.

7.3 Resultados e discussão

7.3.1 Purificação das lectinas nativas ALL e CCL

Ambas as lectinas nativas foram purificadas e a pureza avaliada por SDS-PAGE em condições redutoras e não redutoras (Figura 91).

Figura 91 – Avaliação da pureza das lectinas ALL e CCL por SDS-PAGE.



Fonte: Autora. Gel de 12%. M: Marcador molecular comercial *Precision Plus Protein Unstained Standards*, Bio rad. Amostra liofilizada ressuspensa em tampão 20 mM Hepes pH 7,5 e 100 mM NaCl. R: Tampão de amostra contendo DTT, condição redutora. NR: Tampão de amostra sem DTT, condição não redutora.

As bandas observadas para ambas as lectinas estão de acordo com as características descritas por Duarte e colaboradores (2023) e Sousa e colaboradores (2021) para ALL e CCL, respectivamente.

Em condições redutoras, duas bandas intensas são visíveis em ALL, referentes ao monômero de 15 kDa e ao dímero de 30 kDa (Duarte *et al.*, 2023). Outras bandas de menor intensidade com maior peso molecular são detectadas, as quais são associadas a oligomerização da lectina. Mesmo na presença do agente redutor, provavelmente o DTT não foi suficiente para evitar a autoassociação pela redução completa das pontes dissulfeto, principalmente se estiverem protegidas no interior da estrutura oligomérica. Para CCL, duas bandas de maior

intensidade são observadas e correspondem ao monômero de 15 kDa e a autoassociação em um dímero de 30 kDa (Sousa *et al.*, 2021; Marques *et al.*, 2018).

Em condições não redutoras, bandas intensas com elevado peso molecular são visíveis em ALL, as quais são referentes às oligomerizações. A menor banda visível apresenta 30 kDa que corresponde ao dímero, formado por uma ponte dissulfeto intercadeia, que em condições redutoras é rompida, conforme o comportamento observado por Duarte e colaboradores (2023). Como CCL, não apresenta pontes dissulfeto em sua estrutura, a mesma banda de 15 kDa é observada nas duas condições. Uma banda de maior peso molecular é observada devido à autoassociação por interações fracas (Sousa *et al.*, 2021; Marques *et al.*, 2021).

As lectinas liofilizadas e ressolubilizadas apresentaram atividade hemaglutinante, confirmando a função biológica mantida (dados não mostrados).

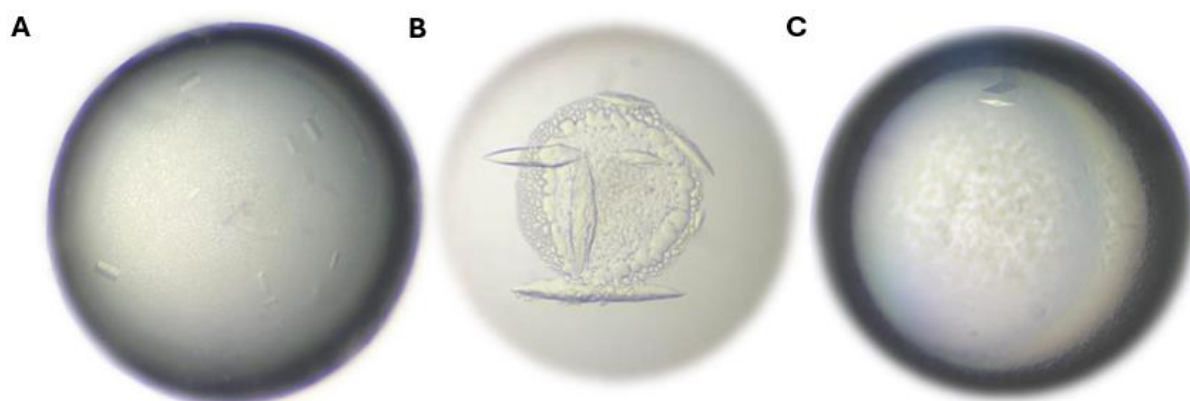
7.3.2 Determinação da estrutura tridimensional da galectina ALL

7.3.2.1 Cristalização da ALL

Na triagem robótica, microcristais foram obtidos em poucos dias com o kit Midas Screen (*Molecular Dimensions*), o que indica que a lectina ALL apresenta propensão à cristalização. Como a triagem robótica permitiu a formação de microcristais, posteriormente as condições foram otimizadas manualmente.

Nas etapas de triagem manual, diferentes morfologias cristalinas foram observadas, incluindo cristais em forma de bastões, losangos espessos e paralelepípedos, obtidos nas condições contendo 0,2 M de sulfato de amônio, 0,1 M MES pH 6,5 e 24–26% de PEG 6K, PEG 8K ou Peg Smear Medium (PSM, *Molecular Dimensions*). Adicionalmente, nessa etapa a utilização de 0,1 M Bis-Tris Propano pH 8,5 favoreceu a formação de cristais bipiramidais alongados, enquanto cristais losangulares foram obtidos em condições contendo 24% PEG 8K, 0,2 M AMSO₄ e 0,1 M MES pH 6,5 (Figura 92).

Figura 92 – Cristais obtidos de ALL.



Fonte: Autora. ALL em 20 mM Hepes pH 7,5 e 100 mM NaCl, a uma concentração de 7,1 mg/mL contendo 1 mM de D-Lactose. A: 24% PEG 8K, 0,2 M AmSO₄ e 0,1 M MES pH 6,5. B: 25% PEG 8K, 0,2 M AmSO₄ e 0,1 M Bis-tris propano pH 8,5. C: 26% PEG 6K, 0,2 M AmSO₄ e 0,1 M MES pH 6,5.

A ocorrência de múltiplas morfologias cristalinas é um fenômeno comum na cristalografia de macromoléculas. Esse comportamento é devido ao polimorfismo cristalino, no qual uma proteína pode se organizar em diferentes arranjos espaciais dependendo das condições de precipitação e do sistema tampão empregado (Cruz-Cabeza; Bernstein, 2014).

Apesar da obtenção de cristais bem definidos nessas condições, as tentativas de resolução estrutural por substituição molecular utilizando o fluxo de trabalho AlphaFold MR na nuvem do CCP4, com base na sequência determinada por espectrometria de massas (Duarte *et al.*, 2023), não resultaram em fases com qualidade suficiente para permitir o refinamento do modelo estrutural. Devido a essa limitação, a estratégia de cocristalização da proteína com 10 mM de TbXo₄ foi adotada, levando à obtenção de cristais também losangulares, os quais produziram dados de difração satisfatórios e permitiram a resolução do problema de fase.

O cristal apresentou uma resolução de 1,85 Å em um sistema ortorrômbico $I2_12_12_1$ com dimensões de célula unitária (Å) de $a = 54,03$; $b = 58,72$; $c = 211,63$, e a presença de um dímero por unidade assimétrica.

As estatísticas de qualidade dos dados de difração utilizados para a resolução da estrutura do complexo ALL-lactose estão resumidas na Tabela 28.

Tabela 28 - Dados de coleta e refinamento do complexo ALL-Lactose.

Complexo	ALL-Lactose
Coleta de dados	
Linha de luz	Soleil PX1
Grupo espacial	$I2_12_12_1$
Parâmetros de célula - a, b, c (Å)	54,03; 58,72; 211,63
a, b, g (°)	90,00; 90,00; 90,00
Protômero na unidade assimétrica	1
Resolução (Å) ^a	45,13–1,85
Número de reflexões	27.924
Completeness (%) ^a	99,97
Refinamento	
Resolução (Å)	1,85
$R_{\text{work}}/R_{\text{free}}$ (%)	17,4 /21,6
R.m.s comprimento de ligações (Å)	0,014
Rmsd ângulo de ligações (°)	1,797
Rmsd Quiral (Å ³)	0,099
Clashscore	1,33
Ramachandran Permitidos	100%
Favorecidos	98,5%
Proibidos	0%

Fonte: Autora. Os valores entre parênteses são referentes à célula de maior resolução.

Na cristalografia de raios X, a determinação estrutural depende do cálculo do mapa de densidade eletrônica, necessitando das amplitudes e das fases dos fatores de estrutura. Se no momento da difração, apenas as intensidades dos feixes difratados serem medidas, sem a informações de fases, não é possível resolver a estrutura. Quando o problema de fases acontece, é necessário a utilização de métodos indiretos para a obtenção do mapa de densidade como a estratégia usado no presente trabalho de cocrystalizar a proteína com metais pesados. Ao introduzir um metal pesado no cristal, ele pode se ligar a sítios específicos da proteína, geralmente associados a resíduos carregados negativamente, gerando diferenças mensuráveis nas intensidades das reflexões difratadas, com elevado poder de espalhamento anômalo de raios X. A identidade das posições permite a estimativa inicial das fases dos fatores da estrutura, possibilitando a construção do mapa de densidade eletrônica (Girard *et al.*, 2003; Chen *et al.*, 2025).

A cocrystalização consiste em adicionar um ligante à proteína no processo de cristalização para obter o complexo proteína-ligante. Essa abordagem foi utilizada nos ensaios,

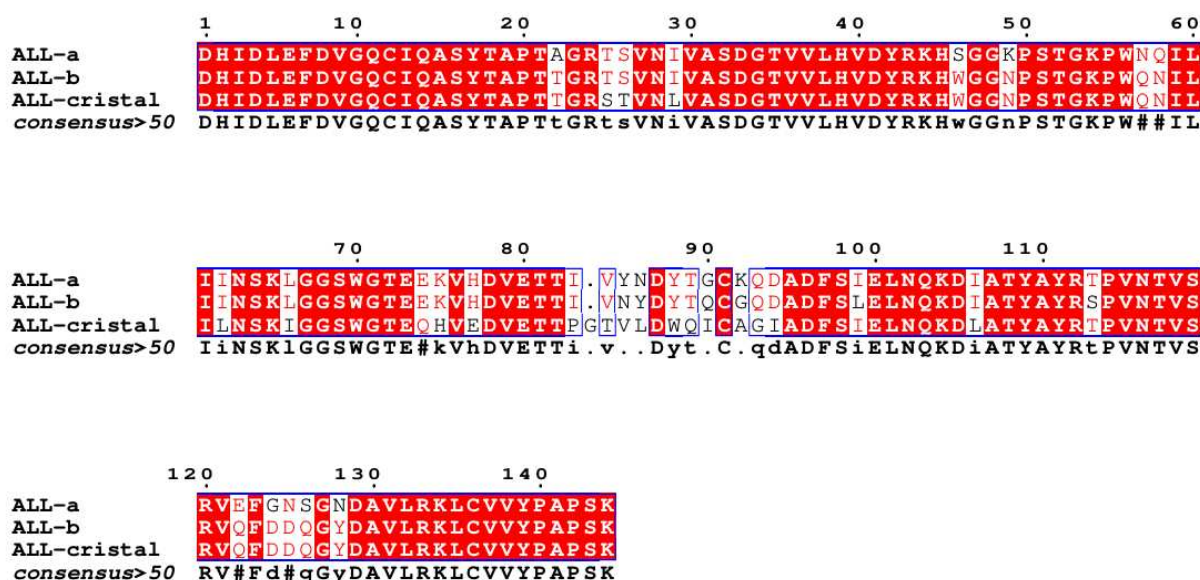
pois ela pode aumentar a estabilidade conformacional da proteína, reduzindo a flexibilidade estrutural e ajudando na formação dos cristais. A técnica, quando bem sucedida, permite a visualização direta do modo de interação da proteína com seu respectivo ligante e os resíduos envolvidos nesse reconhecimento (Hassell *et al.*, 2007).

A formação de cristais para ALL só foi observada na presença da lactose. Tentativas de cristalização na ausência de ligante e na presença de *N*-Acetil-D-Lactosamina não foram bem-sucedidas.

7.3.2.2 Estrutura geral da lectina ALL

A determinação da estrutura tridimensional revelou uma sequência de aminoácidos distinta da observada por espectrometria de massas (Duarte *et al.*, 2023). As estruturas primárias determinadas por espectrometria de massas das isoformas ALL-a e ALL-b e a nova sequência de aminoácidos determinada por cristalografia de raios X estão alinhadas na Figura 93.

Figura 93 – Alinhamento das sequências de ALL determinada por espectrometria de massas e cristalografia.



Fonte: Autora. Os números acima da sequência marcam a contagem de aminoácidos.

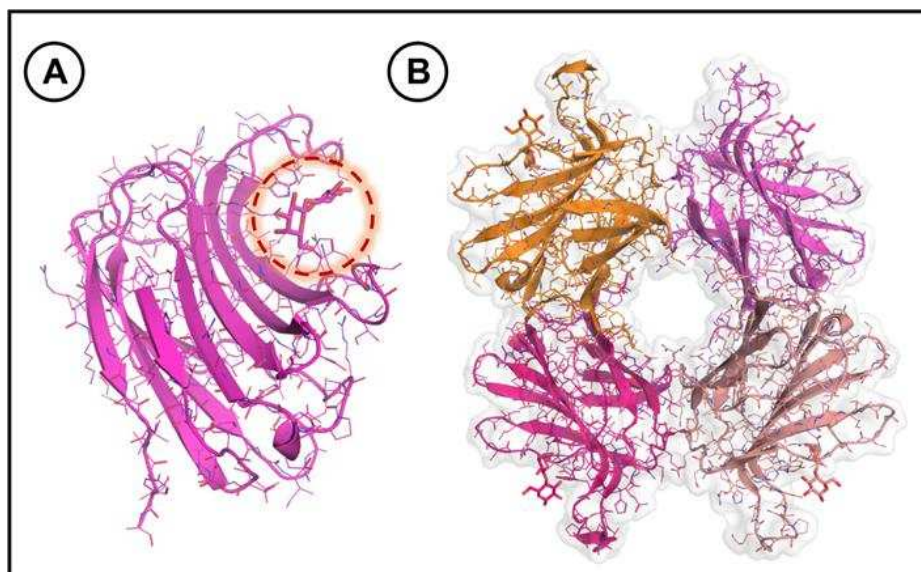
Algumas modificações de aminoácidos são justificadas pelo fato de alguns resíduos apresentarem a mesma massa monoisotópica ou diferença de massa inferior a 1 Da, como a troca de Leu por Ile, Gln por Glu e Asn por Asp. Outras modificações são isobáricas, ou seja,

a soma dos aminoácidos resulta na mesma massa total, como para os pares Ala e Gly e Gln e Lys. A maior substituição foi observada na região entre os aminoácidos 83 até 94, que também corresponde a modificações isobáricas. Portanto, as diferenças entre a estrutura primária determinada por cristalografia e por espectrometria de massas são atribuídas às limitações inerentes à técnica de espectrometria de massas (Habler *et al.*, 2024; Fricker, 2015).

A estrutura primária cristalográfica apresentou 79% de identidade e 91% de similaridade com a isoforma ALL-a, enquanto com a isoforma ALL-b foi 85% de identidade e 93% de similaridade. Assim, os dados indicam que a isoforma cristalizada corresponde, provavelmente à ALL-b.

A estrutura cristalográfica da ALL, composta por 144 aminoácidos, apresenta 100% dos resíduos de aminoácidos localizados em regiões permitidas do gráfico de Ramachandran, sendo 98,5% destes posicionados em regiões favorecidas (Tabela 28). A estrutura do monômero apresenta uma arquitetura global do tipo β -sanduíche, formada por duas folhas β organizadas em arranjo aproximadamente paralelo: uma folha plana constituída por seis fitas β e outra levemente curvada composta por cinco fitas β (Figura 94 - A). As fitas β encontram-se interligadas por alças de comprimentos variáveis, as quais contribuem com resíduos alifáticos e/ou aromáticos responsáveis pela formação do núcleo hidrofóbico da proteína. A ALL apresenta uma ponte dissulfeto intracadeia formada entre os resíduos Cys¹² e Cys⁹² (Figura 95). A estrutura da ALL havia sido predita anteriormente por Duarte e colaboradores (2023), apresentado boa similaridade com a estrutura experimental descrita, atestado por um RMSD de 1,08Å (Carugo; Pongor, 2001).

Figura 94 – Estrutura geral de ALL.

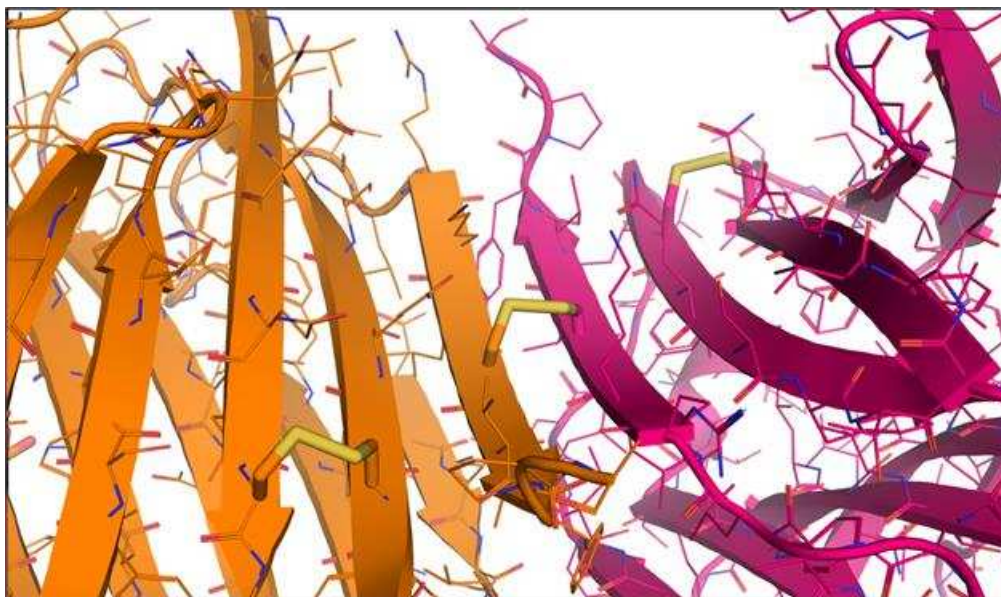


Fonte: Autora. A – Monômero da ALL, com destaque para o sítio de ligação à carboidrato (circundado). B – Forma tetramérica da ALL. A proteína encontra-se representada na forma de *cartoons* e linhas coloridas diferencialmente, estando as cadeias A, B, C e D coloridas em laranja, rosa claro, bege e rosa escuro, respectivamente. Já as moléculas de lactose estão representadas como bastões coloridos da mesma forma.

Na estrutura cristalográfica é possível observar o arranjo tetramérico da lectina, formado por um dímero de dímeros (Figura 94 - B). A interface tetramérica da lectina é formada pelos contatos entre diferentes alças, enquanto a interface dimérica é formada pelo alinhamento das folhas beta planas dos monômeros envolvidos, formando uma folha contínua de 12 fitas, unidas por uma ponte dissulfeto intercadeia entre os resíduos Cys¹³⁷ de cada monômero (Figura 95).

Entretanto, na estrutura cristalográfica é observada na unidade assimétrica um dímero formado pelas subunidades componentes da interface tetramérica. Tal atribuição, se justifica pela ausência nesta interface da ponte dissulfeto característica da interface dimérica da ALL, conforme comprovado experimentalmente por SDS-PAGE em condições redutoras e não redutoras (Figura 91).

Figura 95 - Pontes dissulfeto presentes na estrutura da ALL.



Fonte: Autora. Os diferentes monômeros da ALL estão representados como *cartoons* e linhas coloridos em laranja e rosa, respectivamente. As pontes dissulfeto intracadeias e intercadeia estão representadas como bastões coloridos no mesmo padrão.

A organização do tetrâmero da lectina ALL inicia-se pela formação de um dímero estável, cuja interface é estabilizada por um conjunto de interações entre as cadeias A e D, reforçado por uma rede de pontes aquosas. Nessa interface foram identificadas 16 interações diretas, distribuídas em três contatos hidrofóbicos, onze ligações de hidrogênio e duas interações iônicas, envolvendo principalmente os resíduos Asp¹, His², Glu⁶, Asp⁴, Val¹³⁸, Leu¹³³, Leu¹³⁶, Arg¹³⁴, Tyr¹⁴⁰ e Ile³ (Tabela 29). Os contatos hidrofóbicos, como entre Val¹³⁸-Ile³ e Tyr¹⁴⁰-Leu¹³³, contribuem para o empacotamento da interface, enquanto as múltiplas ligações de hidrogênio, incluindo interações recíprocas entre Asp⁴-Asp⁴, Tyr¹⁴⁰-Arg¹³⁴ e Val¹³⁸-Leu¹³⁶, fornecem especificidade geométrica e estabilidade direcional ao complexo. As interações iônicas entre Glu⁶ e His² indicam ainda a participação de forças eletrostáticas na coesão do dímero. Esse arranjo é adicionalmente estabilizado por diversas pontes aquosas, mediadas sobretudo por His², Glu⁶, Gln¹⁴, Lys¹³⁵, Arg¹³⁴, Tyr¹⁴⁰ e Pro¹⁴¹, configurando uma rede cooperativa de interações indiretas que complementa os contatos diretos e sustenta a organização do dímero.

Tabela 29 - Interações moleculares presentes na interface dimérica da lectina ALL.

Resíduo	Cadeia	Resíduo	Cadeia	Distância (Å)	Tipo de interação
Val138	A	Ile3	D	3,6	Hidrofóbica
Tyr140	A	Leu133	D	3,46	Hidrofóbica
Leu133	A	Tyr140	D	3,46	Hidrofóbica
Tyr140	A	Asp1	D	3,70	Hidrogênio
Asp4	A	Asp4	D	2,88	Hidrogênio
Asp4	A	Asp4	D	2,88	Hidrogênio
Tyr140	A	Arg134	D	2,81	Hidrogênio
Val138	A	Leu136	D	2,89	Hidrogênio
Val138	A	Leu136	D	3,03	Hidrogênio
Val138	A	Leu136	D	3,46	Hidrogênio
Leu136	A	Val138	D	2,89	Hidrogênio
Asp1	A	Tyr140	D	3,7	Hidrogênio
Arg134	A	Tyr140	D	2,81	Hidrogênio
Glu6	A	His2	D	4,68	Iônica
His2	A	Glu6	D	4,68	Iônica

Fonte: Autora. Interações calculadas por meio do servidor PLIP.

A partir dessa unidade dimérica, a associação de dímeros conduz à formação do arranjo tetramérico da ALL, cuja interface é sustentada por um conjunto definido de interações diretas entre as cadeias A e B, complementado por uma rede de pontes aquosas que reforça a associação. Essa interface apresenta 14 interações diretas, distribuídas em quatro contatos hidrofóbicos, oito ligações de hidrogênio e duas interações iônicas, envolvendo predominantemente os resíduos Glu⁸⁰, Thr⁸², Pro⁸³, Val⁸⁶, Asn¹⁰³, Gln¹⁰⁴, Lys¹⁰⁵ e Lys⁴⁴, os quais constituem os principais determinantes estruturais da tetramerização (Tabela 30). Os contatos hidrofóbicos contribuem para o empacotamento da superfície interfacial, enquanto as ligações de hidrogênio e as interações eletrostáticas conferem especificidade geométrica e estabilidade direcional ao complexo. Adicionalmente, a presença de múltiplas pontes aquosas, em especial aquelas mediadas por Glu⁸⁰, Lys⁴⁴, Asn⁵⁸, Thr⁸¹, Thr⁸⁵, Val⁸⁶, Asn¹⁰³ e Gln¹⁰⁴, estabelece uma rede cooperativa de interações indiretas que complementa os contatos diretos e favorece a coesão estrutural do tetrâmero.

Tabela 30 - Interações moleculares presentes na interface tetramérica da lectina ALL.

Resíduo	Cadeia	Resíduo	Cadeia	Distância (Å)	Tipo de interação
Glu80	A	Glu80	B	3,94	Hidrofóbica
Lys105	A	Thr85	B	3,98	Hidrofóbica
Gln104	A	Val86	B	3,68	Hidrofóbica
Val86	A	Gln104	B	3,88	Hidrofóbica
Thr82	A	Glu80	B	2,90	Hidrogênio
Thr82	A	Glu80	B	2,75	Hidrogênio
Glu80	A	Thr82	B	2,81	Hidrogênio
Lys105	A	Pro83	B	2,63	Hidrogênio
Asn103	A	Val86	B	2,89	Hidrogênio
Val86	A	Asn103	B	2,85	Hidrogênio
Pro83	A	Lys105	B	2,86	Hidrogênio
Glu80	A	Lys44	B	4,22	Iônica
Lys44	A	Glu80	B	4,58	Iônica

Fonte: Autora. Interações calculadas por meio do servidor PLIP.

7.3.2.3 Caracterização do sítio de ligação e análise da interação ALL-Lactose

O sítio de ligação a carboidratos da lectina ALL é constituído por um conjunto de resíduos polares e aromáticos que delimitam um bolso de reconhecimento específico, incluindo Arg²⁴, His³⁹, Asp⁴¹, Arg⁴³, Trp⁴⁶, Trp⁵⁶, Ile⁶¹, Asn⁶³, Glu⁷³ e Trp⁷⁰, responsáveis pela conformação estrutural do sítio. A interação com a lactose é predominantemente mediada por uma rede de ligações de hidrogênio estabelecidas com os grupos hidroxila do anel da galactose, com distâncias variando entre aproximadamente 2,6 e 3,1 Å, conferindo alta especificidade ao reconhecimento molecular. Em particular, Arg²⁴ interage com os átomos O³ e O⁴ da galactose, enquanto Arg⁴³ estabelece múltiplas ligações com os átomos O³, O⁴ e O⁵, desempenhando papel central na ancoragem do ligante. His³⁹ contribui por meio de interação com o átomo O⁴, Asn⁶³ interage com o O⁶, e Glu⁷³ estabelece contatos tanto com o O⁶ da galactose quanto com o O³ do resíduo glicose (Tabela 31).

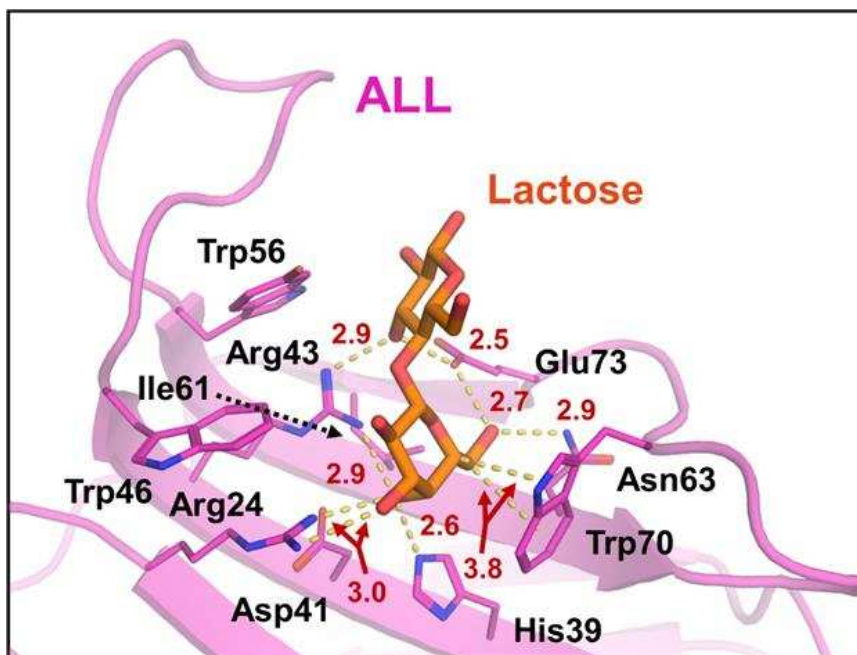
Tabela 31 - Interações moleculares entre ALL e lactose.

ALL		Lactose		Distância (Å)	Tipo de interação
Resíduo	Átomo	Resíduo	Átomo		
Arg24	NH1	Gal	O3	3,06	Hidrogênio
Arg24	NH2	Gal	O3	3,07	Hidrogênio
Arg24	NH1	Gal	O4	2,98	Hidrogênio
His39	NE2	Gal	O4	2,68	Hidrogênio
Arg43	NH1	Glu	O3	2,91	Hidrogênio
Arg43	NH2	Gal	O4	2,98	Hidrogênio
Arg43	NH2	Gal	O5	2,94	Hidrogênio
Asn63	ND2	Gal	O6	2,9	Hidrogênio
Glu73	OE2	Gal	O6	2,78	Hidrogênio
Glu73	OE2	Glu	O3	2,58	Hidrogênio
Trp70	Anel-5	Gal	C5	3,80	H- π
Trp70	Anel-6	Gal	C6	3,87	H- π

Fonte: Autora. Interações calculadas por meio do servidor PLIP.

Adicionalmente, o resíduo aromático Trp⁷⁰ participa do reconhecimento por meio de interações do tipo H- π com os átomos C⁵ e C⁶ do anel de galactose, favorecendo o posicionamento e a estabilização do açúcar no interior do bolso de ligação. Os resíduos Trp⁵⁶, Trp⁴⁶, Asp⁴¹ e Ile⁶¹ complementam esse arranjo por contatos de proximidade e contribuições hidrofóbicas, auxiliando na manutenção da geometria do sítio. Em conjunto, esses dados demonstram que a ligação da lactose à ALL é governada por uma combinação sinérgica de ligações de hidrogênio direcionais, interações aromáticas do tipo H- π e contatos hidrofóbicos estruturais, resultando em um modo de reconhecimento específico e estruturalmente estável do carboidrato (Figura 96).

Figura 96 - Análise das interações presentes no complexo ALL-lactose.



Fonte: Autora. A proteína e os resíduos componentes do sítio de ligação à carboidrato estão representados como *cartoon* e linhas coloridas em rosa, respectivamente, enquanto a molécula de lactose está representada como bastões coloridos em laranja. As interações moleculares estão representadas como linhas tracejadas amarelas e o número indica a distância entre eles em (Å). Algumas interações e resíduos de aminoácidos foram omitidos para propostas de clareza.

7.3.3 Determinação da estrutura tridimensional da galectina CCL

7.3.3.1 Cristalização da CCL

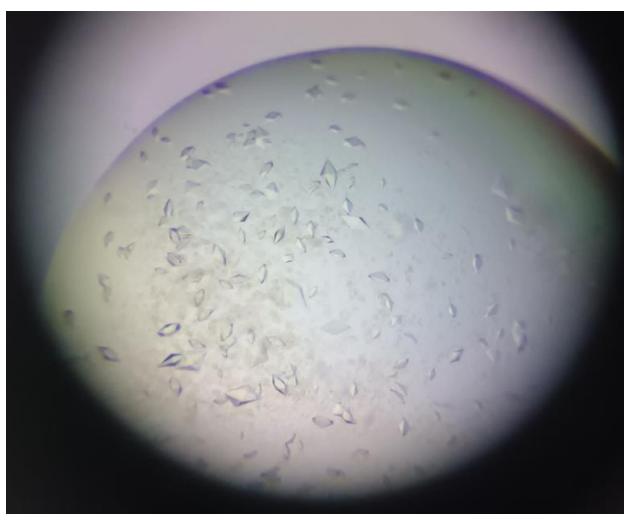
Nas triagens robóticas iniciais de cristalização da CCL liofilizada na presença de D-lactose, não foram obtidos resultados promissores, indicando que as condições avaliadas nessa etapa não foram suficientes para promover a formação de cristais adequados.

Em contraste, nos ensaios manuais subsequentes, conduzidos após a preparação da proteína por cromatografia de exclusão molecular, observou-se a formação de cristais bipiramidais após alguns dias, na condição 2-9 do kit comercial Morpheus II (*Molecular Dimensions*). Essa condição, composta por 2 mM de lantanídeos (0,5 mM de cloreto hexahidratado de ítrio, érbio, térbio e itérbio), 0,1 M do sistema tampão 6 pH 8,5 (Gly-Gly, AMPD) e 50% do precipitante Mix 5 (30% PEG 3K, 40% 1,2,4-butanetriol e 2% NDSB 256), se mostrou favorável à nucleação e ao crescimento cristalino do complexo CCL-lactose. Esses resultados evidenciam a importância da otimização das condições de cristalização e do uso de

sistemas precipitantes e tampões específicos para a obtenção de cristais com morfologia definida.

De forma semelhante, nos ensaios envolvendo a proteína não liofilizada com o ligante 3SL, a formação de cristais bipiramidais foi observada após alguns dias, sob condições correspondentes a 80% da condição 2-10 do kit Morpheus II (*Molecular Dimensions*). Essa condição é composta por 2 mM de lantanídeos (0,5 mM de cloreto hexahidratado de ítrio, érbio, térbio e itérbio), 0,1 M do sistema tampão 6 pH 8,5 (Gly-Gly, AMPD) e 50% do precipitante Mix 7 (20% PEG 4K e 40% hexanetriol) (Figura 97). A obtenção de cristais sob uma formulação semelhante à empregada para o complexo com lactose, diferindo apenas na composição do sistema precipitante, sugere que a presença dos lantanídeos e do sistema tampão alcalino desempenha papel relevante na estabilização do complexo proteína-ligante e na promoção da cristalização.

Figura 97 – Cristais obtidos de CCL-3SL.



Fonte: Autora. CCL não liofilizada em 20 mM Tris pH 8 e 100 mM NaCl, a uma concentração de 10,6 mg/mL contendo 10 mM de 3-sialil-lactose. Condição: 2 mM de lantanídeos, 0,1 M do sistema tampão 6 pH 8,5 e 50% do precipitante Mix 7.

O cristal de CCL-Lactose apresentou uma resolução de 1,55 Å em um sistema monoclinico $I2$ com dimensões de célula unitária (Å) de $a = 36,61$; $b = 36,86$; $c = 94,81$ Å. O cristal de CCL-3SL apresentou uma resolução de 1,95 Å em um sistema tetragonal $P4_122$ com dimensões de célula unitária (Å) de $a = 89,93$; $b = 89,93$; $c = 223,83$. Para os dois cristais, foi observada a presença de um tetrâmero por unidade assimétrica.

As estatísticas de qualidade dos dados de difração obtidos para ambos os complexos (CCL-Lactose e CCL-3SL) estão resumidas na Tabela 32.

Tabela 32 - Dados de coleta e refinamento dos complexos CCL-Lactose e CCL-3SL.

Complexo	CCL-Lactose	CCL-3SL
Coleta de dados		
Linha de luz	Soleil PX2	Soleil PX2
Grupo espacial	<i>I</i> 2	<i>P</i> 4 ₁ 22
Parâmetros de célula - a, b, c (Å)	36,61 ; 36,86 ; 94,81	89,93 ; 89,93 ; 223,83
α, β, γ (°)	90,00 ; 99,17 ; 90,00	90,00 ; 90,00 ; 90,00
Protômeros na unidade assimétrica	1	1
Resolução (Å) ^a	35,68-1,55	50,00–1,95
Número de reflexões	95.115	64.487
Completeness (%) ^a	99,7	99,98
Refinamento		
Resolução (Å)	1,55	1,95
<i>R</i> _{work} / <i>R</i> _{free} (%)	16,3 / 20,4	16,4/18,9
R.m.s comprimento de ligações (Å)	0,012	0,011
Rmsd ângulo de ligações (°)	1,893	1,632
Rmsd Quiral (Å ³)	0,092	0,098
Clashscore		2,48
Ramachandran Permitidos		100%
Favorecidos		96,75%
Proibidos		0%

Fonte: Autora. Os valores entre parênteses são referentes à célula de maior resolução.

A interação com diferentes ligantes pode influenciar diretamente o empacotamento molecular da proteína no cristal, resultando na formação de sistemas cristalinos diferentes. Apesar do dobramento global da proteína ser igual, a presença de ligantes específicos pode induzir alterações conformacionais, especialmente em regiões flexíveis. Essas mudanças afetam as interações responsáveis pelo processo de nucleação e crescimento cristalino, favorecendo diferentes organizações moleculares (Yu; Wang; Peng, 2003).

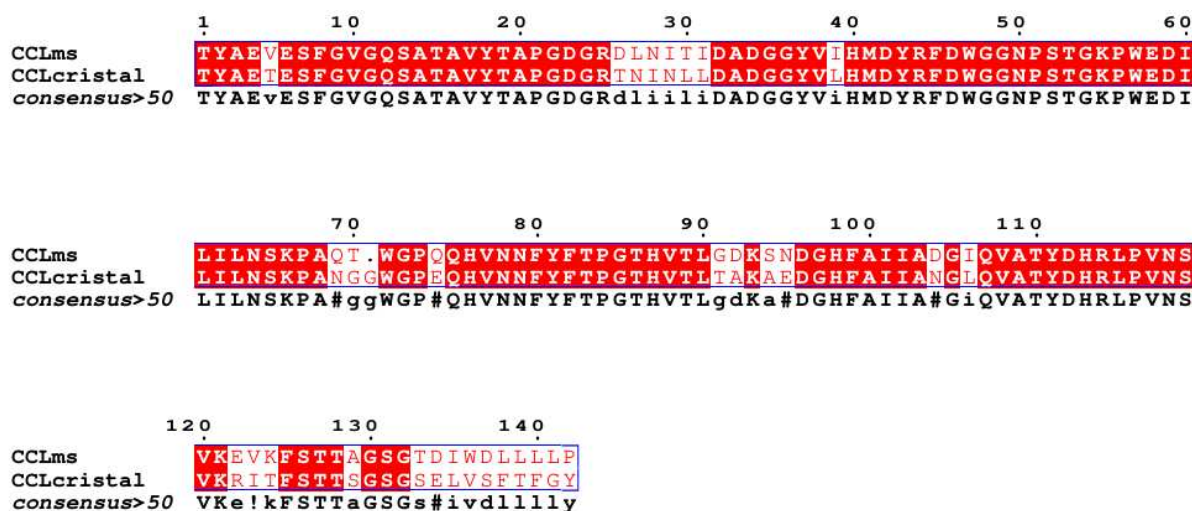
A formação de cristais para CCL não foi observada na ausência de ligante e nem nas tentativas de cocrystalização com *N*-Acetil-D-Lactosamina.

7.3.3.2 Estrutura geral da lectina CCL

A determinação da estrutura tridimensional por cristalografia revelou uma sequência de aminoácidos distinta da observada por espectrometria de massas para CCL (Sousa *et al.*, 2021). A estrutura primária determinada por espectrometria de massas e a nova

sequência de aminoácidos determinada por cristalografia de raios X estão alinhadas na figura 98.

Figura 98 – Alinhamento das sequências de CCL determinada por espectrometria de massas e cristalografia.



Fonte: Autora. Os números acima da sequência marcam a contagem de aminoácidos.

Como observado em ALL, algumas modificações de aminoácidos são justificadas pelas limitações inerentes à técnica de espectrometria de massas, com resíduos que apresentam a mesma massa monoisotópica ou diferença de massa inferior a 1 Da. Outras modificações observadas em CCL são isobáricas, como Gln e Thr por Asn e duas Gly (Habler *et al.*, 2024; Fricker, 2015). A maior concentração de substituições foi observada na região do C-terminal, a partir do aminoácido 133. Nesse caso, especificamente, Sousa e colaboradores (2021) apresentam no mapa peptídico apenas um peptídeo que cobre esse local. Portanto, a baixa sobreposição de fragmentos nessa região impediu a confirmação, justificando as atribuições incorretas nessa região (Sousa *et al.*, 2021).

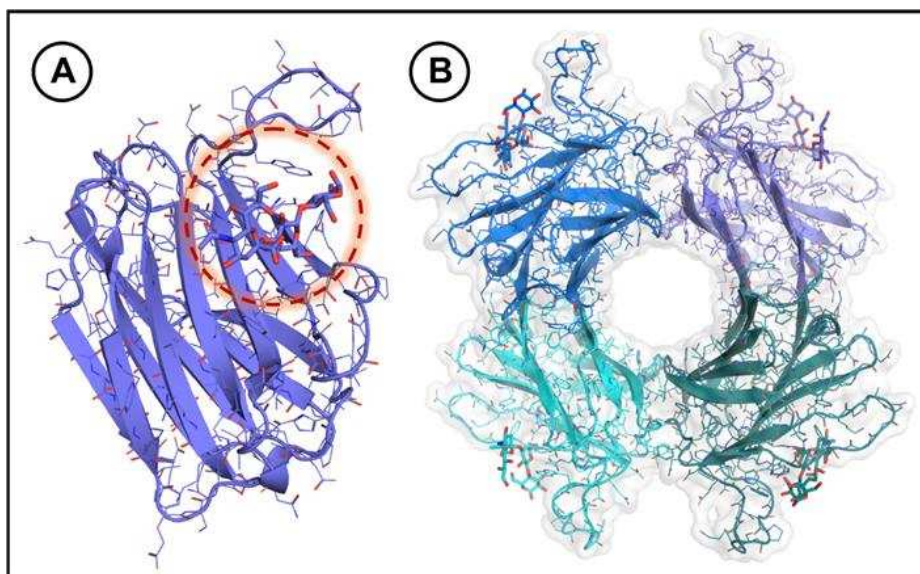
No entanto, para região que compreende os aminoácidos 26 a 31; 94 e 95, os espectros oriundos da espectrometria de massas não encaixam com a sequência observada na estrutura cristalina. Além disso, a massa da sequência determinada pela cristalografia (15.439 Da) também não está de acordo com a massa determinada por espectrometria de massa (15.445 Da) (Marques *et al.*, 2018; Sousa *et al.*, 2021). Talvez essas inconsistências entre as duas estruturas primárias podem ser justificadas pela presença de isoformas nesse organismo já relatada por Marques e colaboradores (2018).

De fato, as lectinas são encontradas na natureza como uma mistura de isolectinas. Proteínas homólogas, estruturalmente semelhantes, mas que diferem em alguns aminoácidos que compõem sua estrutura primária (Van Damme, 2022). Freymann e colaboradores (2012) também relataram discrepâncias entre a estrutura primária determinada por cristalografia da galectina CchG-1 de *Cinachyrella* sp. e as sequências de aminoácidos obtidas por clonagem de cDNA.

Portanto, essas diferenças entre as duas estruturas primárias podem ser atribuídas à presença das isoformas, sugerindo que a estrutura cristalográfica corresponde a uma isolectina diferente das identificadas inicialmente. As duas estruturas primárias de CCL apresentam 78% de identidade e 91% de similaridade.

De forma similar a ALL, a estrutura cristalográfica da CCL também apresentou excelente qualidade, sendo composta por 142 aminoácidos, estando todos localizados em regiões permitidas do diagrama de Ramachandran, com 96,75% dos resíduos dispostos em regiões favorecidas (Tabela 32). O monômero também adota uma conformação do tipo β -sanduíche, formado por duas folhas β aproximadamente paralelas, uma mais plana constituída por seis fitas β e outra levemente curvada composta por cinco fitas β (Figura 99 - A). Essas fitas são conectadas por alças de extensão variável, que contribuem com resíduos alifáticos e aromáticos responsáveis pela organização do núcleo hidrofóbico da proteína. Diferente do que foi descrito para a ALL, a CCL não apresenta pontes dissulfeto em sua estrutura. A estrutura da CCL havia sido predita anteriormente por Sousa e colaboradores (2021), apresentado boa similaridade com a estrutura experimental descrita, atestado por um RMSD de 2,3 Å, levando em consideração as disparidades da sequência de aminoácidos observadas (Figura 98).

Figura 99 - Estrutura geral da lectina CCL.



Fonte: Autora. A – Monômero da CCL, com destaque para o sítio de ligação à carboidrato (circundado). B – Forma tetramérica da CCL. A proteína encontra-se representada na forma de *cartoons* e linhas coloridas diferencialmente, estando as cadeias A, B, C e D coloridas em azul, roxo, verde e ciano, respectivamente. Já as moléculas de 3SL estão representadas como bastões coloridos da mesma forma.

A organização quaternária da CCL é bastante similar ao que foi descrito para a ALL, sendo o seu arranjo quaternário formado por um dímero de dímeros, apresentando assim duas interfaces diméricas e duas tetraméricas. Em vista desta similaridade, a ordem das interfaces seguiu a mesma lógica apresentada para a ALL. A interface dimérica é estabelecida entre as cadeias A e D, apresentando uma maior complexidade interacional. Nessa interface foram identificadas 20 interações diretas, compostas por cinco contatos hidrofóbicos e quinze ligações de hidrogênio, envolvendo predominantemente os resíduos Tyr¹⁴³, Thr¹, Glu⁶, Tyr², Leu¹³⁶, Glu⁴, Lys¹²², Val¹³⁷, Phe¹⁴¹ e Phe¹³⁹ (Tabela 33). Os contatos hidrofóbicos, como Tyr¹⁴³-Thr¹, Tyr¹⁴³-Leu¹³⁶ e as interações recíprocas Glu⁶-Tyr²/Tyr²-Glu⁶, contribuem para o empacotamento apolar da interface, enquanto as numerosas ligações de hidrogênio, incluindo interações recorrentes Glu⁶-Tyr², Glu⁴-Glu⁴, Lys¹²²-Tyr², Tyr¹⁴³-Val¹³⁷ e entre os resíduos Phe¹⁴¹-Phe¹³⁹, fornecem especificidade geométrica e coesão direcional ao complexo. Adicionalmente, a presença de uma ponte aquosa entre os resíduos Glu⁴ de ambas as cadeias, complementa os contatos diretos. Em conjunto, esses dados demonstram que a organização dimérica da CCL resulta de uma integração de interações hidrofóbicas, ligações de hidrogênio extensivas e hidratação interfacial, contribuindo para uma interface estruturalmente bem definida.

Tabela 33 - Interações moleculares presentes na interface dimérica da lectina CCL.

Resíduo	Cadeia	Resíduo	Cadeia	Distância (Å)	Tipo de interação
Tyr143	A	Thr1	D	3,89	Hidrofóbica
Glu6	A	Tyr2	D	3,46	Hidrofóbica
Tyr2	A	Glu6	D	3,39	Hidrofóbica
Tyr2	A	Glu6	D	3,74	Hidrofóbica
Tyr143	A	Leu136	D	3,51	Hidrofóbica
Glu6	A	Thr1	D	3,62	Hidrogênio
Lys122	A	Tyr2	D	3,71	Hidrogênio
Lys122	A	Tyr2	D	3,71	Hidrogênio
Glu6	A	Tyr2	D	3,31	Hidrogênio
Glu6	A	Tyr2	D	3,14	Hidrogênio
Glu4	A	Glu4	D	2,81	Hidrogênio
Glu4	A	Glu4	D	2,79	Hidrogênio
Tyr2	A	Glu6	D	3,12	Hidrogênio
Tyr2	A	Glu6	D	2,83	Hidrogênio
Tyr2	A	Lys122	D	3,52	Hidrogênio
Tyr143	A	Val137	D	2,87	Hidrogênio
Phe141	A	Phe139	D	2,92	Hidrogênio
Phe141	A	Phe139	D	2,98	Hidrogênio
Phe139	A	Phe141	D	2,95	Hidrogênio
Phe139	A	Phe141	D	3,02	Hidrogênio

Fonte: Autora. Interações calculadas por meio do servidor PLIP.

A partir desse dímero base, forma-se o arranjo tetramérico da CCL, cuja interface tetramérica é estabilizada por um conjunto restrito, porém específico, de interações diretas entre as cadeias A e B, complementado por uma rede localizada de pontes aquosas. Nessa interface foram identificadas 11 interações diretas, distribuídas em três contatos hidrofóbicos e oito ligações de hidrogênio, envolvendo principalmente os resíduos Tyr⁸², Leu¹⁰⁷, Thr⁸⁷, Thr⁸⁴, Asn¹⁰⁵ e His⁸⁸ (Tabela 34). Os contatos hidrofóbicos, notadamente entre Tyr⁸²-Tyr⁸² e Leu¹⁰⁷-Thr⁸⁷/Thr⁸⁷-Leu¹⁰⁷, contribuem para o empacotamento apolar da interface, enquanto as ligações de hidrogênio são predominantemente recíprocas entre os resíduos Thr⁸⁴-Tyr⁸² e Asn¹⁰⁵-His⁸⁸, conferindo especificidade geométrica e estabilidade direcional ao tetrâmero. Adicionalmente, a interface é reforçada por pontes aquosas envolvendo os resíduos Thr⁸⁷, His⁸⁸ e Asn¹⁰⁵. Em conjunto, esses elementos indicam que a associação tetramérica da CCL é sustentada também por um arranjo cooperativo de interações hidrofóbicas localizadas, ligações de hidrogênio direcionais e hidratação interfacial, resultando em uma interface compacta e estável.

Tabela 34 - Interações moleculares presentes na interface tetramérica da lectina CCL.

Resíduo	Cadeia	Resíduo	Cadeia	Distância (Å)	Tipo de interação
Tyr82	A	Tyr82	A	3,96	Hidrofóbica
Leu107	A	Thr87	B	3,89	Hidrofóbica
Thr87	A	Leu107	B	3,99	Hidrofóbica
Thr84	A	Tyr82	B	2,88	Hidrogênio
Thr84	A	Tyr82	B	2,90	Hidrogênio
Thr84	A	Tyr82	B	3,23	Hidrogênio
Tyr82	A	Thr84	B	2,86	Hidrogênio
Tyr82	A	Thr84	B	3,21	Hidrogênio
Tyr82	A	Thr84	B	2,94	Hidrogênio
Asn105	A	His88	B	2,83	Hidrogênio
His88	A	Asn105	B	2,86	Hidrogênio

Fonte: Autora. Interações calculadas por meio do servidor PLIP.

7.3.3.3 Caracterização do sítio de ligação e análise da interação CCL-3SL

Assim como para ALL, o sítio de ligação a carboidratos da lectina CCL é delimitado por um conjunto de resíduos com propriedades polares e aromáticas que definem a sua arquitetura, incluindo Arg²⁵, Tyr³⁷, His⁴⁰, Arg⁴⁴, Asn⁶⁴, Trp⁷² e Glu⁷⁵, que fazem contatos diretos com o ligante cocristalizado, bem como os resíduos de Trp⁴⁷, Trp⁵⁷, Asp⁴² e Ile⁶², que na estrutura cristalográfica estão posicionados a aproximadamente 3,2 a 4,0 Å do ligante. Nessa estrutura, a interação com o ligante 3SL ocorre por uma rede de ligações de hidrogênio estabelecidas com grupos funcionais da galactose e do ácido siálico, com distâncias variando entre 2,5 e 3,3 Å, o que pode ser associado à elevada especificidade de reconhecimento (Tabela 35 e Figura 100). Nesse arranjo, Arg²⁵ exerce papel central ao interagir simultaneamente com os átomos O³ e O⁴ da galactose e com o O⁶ do ácido siálico, enquanto Arg⁴⁴ reforça a ancoragem do ligante por meio de contatos múltiplos com O⁴ e O⁵ da galactose e com O³ da unidade de glicose. Contribuições adicionais são fornecidas por His⁴⁰, que interage com O⁴, por Asn⁶⁴, em contato com O⁶, ambos do resíduo de galactose, e por Glu⁷⁵, que estabelece ligações tanto com o O⁶ da galactose quanto com o O³ do resíduo glicose.

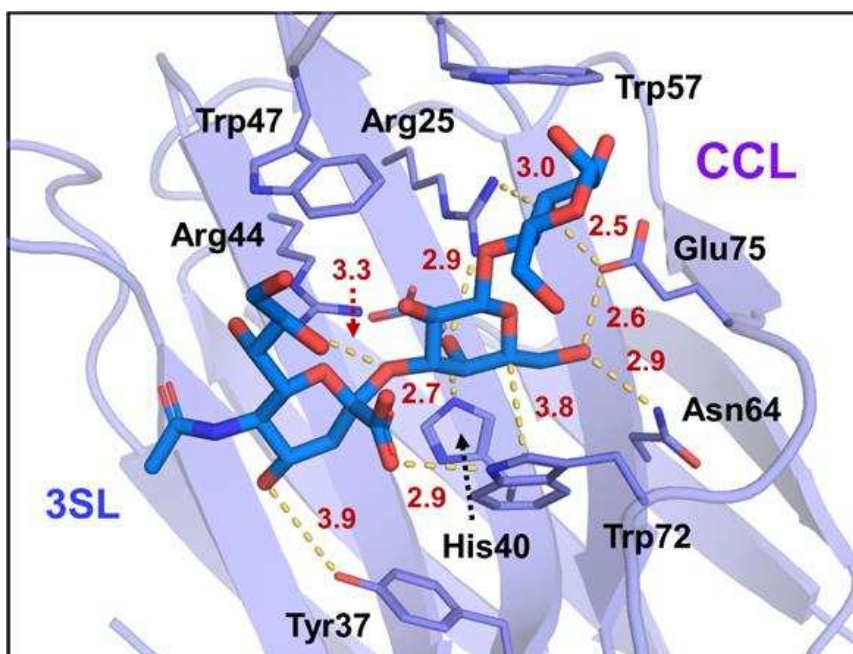
Tabela 35 - Interações moleculares entre CCL e 3SL (3'-sialilactose).

CCL		3SL		Distância (Å)	Tipo de interação
Resíduo	Átomo	Resíduo	Átomo		
Arg25	NH1	Gal	O3	3,14	Hidrogênio
Arg25	NH1	Gal	O4	3,06	Hidrogênio
Arg25	NH2	Gal	O3	3,34	Hidrogênio
Arg25	NH2	Sia	O6	2,98	Hidrogênio
Tyr37	OH	Sia	O4	3,87	Hidrogênio
His40	NE2	Gal	O4	2,73	Hidrogênio
Arg44	NH1	Glu	O3	2,95	Hidrogênio
Arg44	NH2	Gal	O4	2,93	Hidrogênio
Arg44	NH2	Gal	O5	3,00	Hidrogênio
Arg44	NH2	Glu	O3	3,19	Hidrogênio
Asn64	ND2	Gal	O6	2,92	Hidrogênio
Trp72	NE1	Sia	O1A	2,93	Hidrogênio
Glu75	OE2	Gal	O6	2,62	Hidrogênio
Glu75	OE2	Glu	O3	2,54	Hidrogênio
Trp72	Anel-5	Gal	C5	3,84	H- π

Fonte: Autora. Interações calculadas por meio do servidor PLIP.

Além das interações polares, o reconhecimento do 3SL incorpora um componente aromático, onde Trp⁷² participa por meio de uma interação do tipo H- π com o anel galactosídico, favorecendo o alinhamento do ligante no interior da cavidade. Os resíduos Trp⁴⁷, Trp⁵⁷, Asp⁴² e Ile⁶² complementam esse arranjo por contatos de proximidade e contribuições hidrofóbicas que sustentam a geometria do sítio. Em conjunto, esses elementos indicam que a associação entre a CCL e o 3SL resulta da integração de ligações de hidrogênio direcionais, interações aromáticas e efeitos hidrofóbicos estruturais, conferindo ao complexo um modo de ligação seletivo e conformacionalmente estável.

Figura 100 - Análise das interações presentes no complexo CCL-3SL.



Fonte: Autora. A proteína e os resíduos componentes do sítio de ligação à carboidrato estão representados como *cartoon* e linhas coloridas em roxo, respectivamente, enquanto a molécula de 3SL está representada como bastões coloridos em azul. As interações moleculares estão representadas como linhas tracejadas amarelas. Algumas interações e resíduos de aminoácidos foram omitidos para propostas de clareza.

Portanto, no presente trabalho, foram determinadas as estruturas tridimensionais de duas lectinas de esponjas marinhas. Até o momento, apenas duas estruturas tridimensionais de esponjas haviam sido determinadas, sendo apenas uma delas da família das galectinas (Watari *et al.*, 2022; Freymann *et al.*, 2012). Assim, as estruturas apresentadas para ALL e CCL representam, respectivamente, a segunda e a terceira galectina de esponja com a estrutura tridimensional determinada. Esses resultados ampliam significativamente o conhecimento estrutural sobre galectinas de esponjas e fornece novas ideias para a compreensão funcional dessa importante família de lectinas de animais, cujo estudos estruturais ainda são raros.

7.3.3.4 Análise estrutural comparativa entre galectinas de esponjas

Atualmente, algumas galectinas de esponjas marinhas apresentam estruturas tridimensionais descritas, de forma experimental ou teórica, como as apresentadas nesse trabalho, ALL e CCL, e também as galectinas de *Chondrilla australiensis* (hRTL) (Hayashi *et al.*, 2024), *Aiolochoia crassa* (ACrL) (Torres *et al.*, 2025), *Halichondria okadai* (HOL-30)

(Ohwaka *et al.*, 2025) e *Cinachyrella sp.* (CchG-1) (Freyman *et al.*, 2012). Apenas ALL, CCL e CchG-1 possuem estruturas tridimensionais determinadas por métodos experimentais.

A análise dessas estruturas mostra que todas partilham uma arquitetura conservada do domínio de reconhecimento a carboidratos (CRD), caracterizada por um dobramento do tipo β -sanduíche, com duas folhas, sendo uma curvada e responsável pela ligação à carboidratos, e outra plana. Apesar dessa base estrutural comum, essas galectinas diferem na composição de seus sítios de ligação, na presença e conformação de alças adjacentes, e em relação aos seus estados de oligomerização, fatores que em conjunto determinam a especificidade e a afinidade dessas lectinas.

De forma similar, ALL e CCL apresenta uma arquitetura global e estado tetramérico semelhantes entre si e entre outras galectinas. A principal diferença entre elas é a presença de uma ponte dissulfeto intracadeia em ALL.

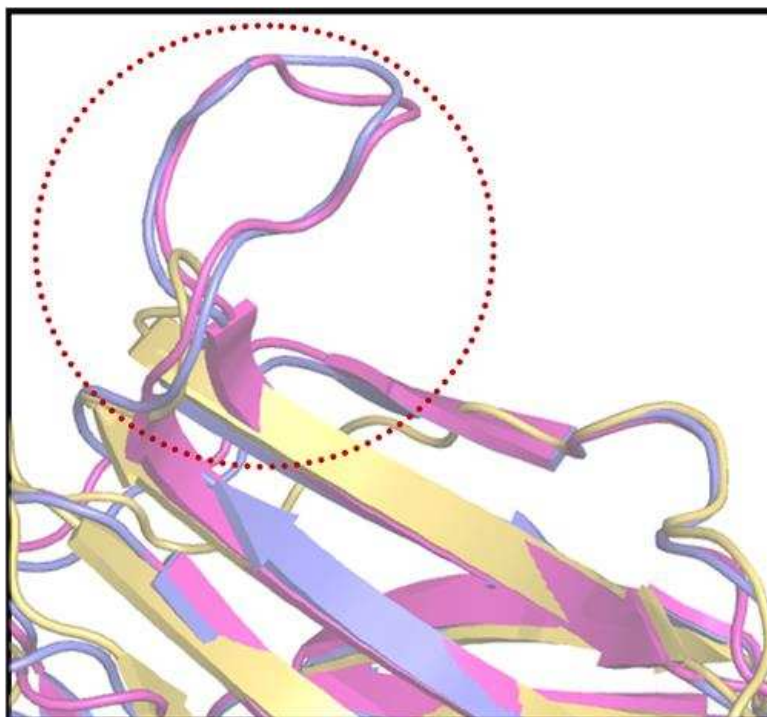
CCL compartilha com hRTL sete resíduos altamente conservados e essenciais à ligação a carboidratos, evidenciando a preservação de um núcleo funcional do CRD entre espécies do gênero *Chondrilla*. A hRTL é descrita como um tetrâmero de subunidades de aproximadamente 15,5 kDa (141 aminoácidos), e assim como CCL, apresenta uma inserção interna (Asn44-Asp59), associada à formação de uma alça específica no sítio de ligação, elemento variante em algumas galectinas de esponjas e que é apontado como potencial modulador de especificidade (Hayashi *et al.*, 2024).

Já a galectina AcrL (*A. crassa*), modelada estruturalmente, é descrita como um dímero, e apresenta seu sítio de ligação subdividido em um “sítio principal”, responsável pelo reconhecimento de galactopiranosídeos, e um “sítio estendido”, que favorece a acomodação de carboidratos mais complexos. O sítio principal é composto pelos resíduos Arg²⁵, His⁴⁰, Asp⁴², Arg⁴⁴, Trp⁴⁷, Ile⁶², Asn⁶⁴ e Trp⁷³, considerados conservados em galectinas, enquanto o sítio estendido inclui resíduos adicionais Lys², Asn²⁹, Asn³⁴, Lys⁶⁶, Tyr¹³⁰ e Val¹³³, descritos como importantes na interação com glicanos sialilados, evidenciando uma expansão funcional do sítio (Torres *et al.*, 2025). Estruturalmente, essa ampliação está associada também a uma inserção interna (Asn⁴⁹-Trp⁵⁷), similar ao observado para CCL e hRTL, produzindo uma alça mais longa no sítio de ligação, apontado como um determinante direto da capacidade de reconhecer carboidratos de maior complexidade.

Por sua vez, a galectina de *Cinachyrella sp.* (CchG-1), também apresenta uma composição de aminoácidos em seu sítio de ligação (Arg²⁸, Tyr³², Arg⁴⁷, Arg⁵¹, Asn⁶⁰, Glu⁵³, Trp⁶⁸ e Glu⁷¹) similar aos das demais galectinas descritas, bem como um arranjo funcional

tetramérico (Freyman *et al.*, 2012). Entretanto, a alça destacada nas galectinas acima apresenta-se reduzida em CchG-1, dotada de apenas seis aminoácidos, enquanto em CCL, hRTL e ACrL essa alça apresenta 16 resíduos, e em ALL 11 resíduos. Essa redução pode estar associada a uma menor capacidade de reconhecimento de ligantes complexos (Figura 101).

Figura 101 – Alça componente do sítio de ligação a carboidratos.



Fonte: Autora. As estruturas referentes às galectinas CCL, ALL e CchG-1 estão representadas na forma de *cartoons* coloridos em roxo, rosa e amarelo, respectivamente. A alça componente do sítio de ligação está destaca por um círculo vermelho tracejado.

Em contraste com as galectinas descritas acima, a HOL-30 (*H. okadae*) possui domínios repetidos, contendo dois CRDs por cadeia polipeptídica de aproximadamente 30 kDa e formando em solução um complexo de aproximadamente 60 kDa, compatível com uma organização dimérica. Embora os dois CRDs apresentem estrutura tridimensional semelhante, possuem baixa identidade de sequência, apresentando resíduos moderadamente conservados (Ohwaka *et al.*, 2025). Um detalhe importante é a ausência em HOL-30 da “alça estendida” observada em ALL, CCL, hRTL e ACrL, sugerindo um modo de reconhecimento a ligantes galactosídicos menos complexos, o que está de acordo com sua afinidade por estruturas do tipo LacNAc e moléculas relacionadas.

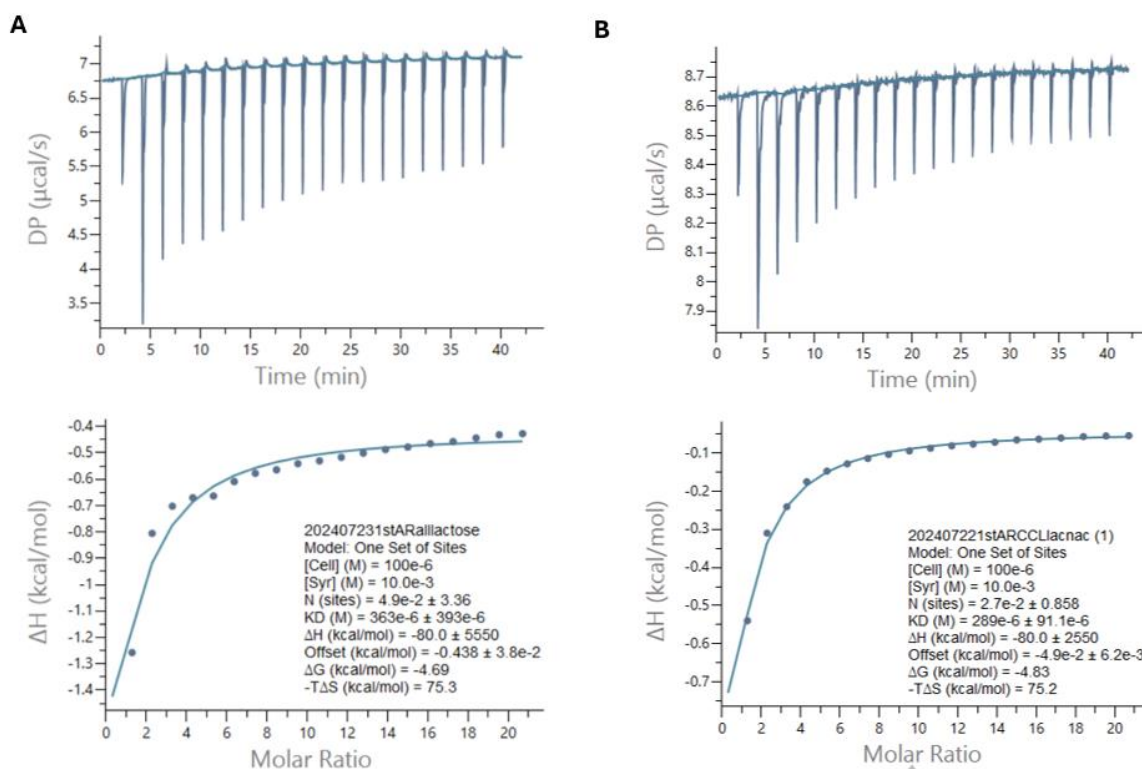
Portanto, essa comparação estrutural evidencia um núcleo conservado de resíduos do CRD preservado entre galectinas protótipo (ALL, CCL, hRTL, CchG-1 e ACrL) e de

domínios repetidos (HOL-30). A variação de especificidade surge como resultado da composição específica de resíduos periféricos do sítio principal e da presença, ausência ou extensão da alça adjacente. Essas diferenças estruturais explicam a diversidade de perfis de reconhecimento a carboidratos observada entre galectinas de esponjas.

7.3.4 Análise de calorimetria por titulação isotérmica

A interação entre a ALL com a D-lactose e a CCL com *N*-Acetil-D-Lactosamina (LacNac) foi avaliada por calorimetria de titulação isotérmica (Figura 102), com o objetivo de analisar a interação proteína-ligante em solução e caracterizar os parâmetros termodinâmicos envolvidos, fornecendo informações para a escolha desses ligantes nos ensaios de cocristalização. Os parâmetros de interação estão resumidos na Tabela 36.

Figura 102 – Análise da interação ALL-Lactose e CCL-LacNac por calorimetria de titulação isotérmica.



Fonte: Autora. A: Interação entre ALL e lactose. B: Interação entre CCL e *N*-Acetil-D-lactosamina. O painel superior apresenta o termograma correspondente aos picos de fluxo de calor (DP, $\mu\text{cal/s}$) registrado após injeções sucessivas de lactose na solução contendo a proteína. O painel inferior mostra a isoterma de ligação, obtida pela integração dos picos do termograma, expressa em variação de entalpia ΔH (kcal/mol) em função da razão molar ligante/proteína. A curva contínua representa o ajuste dos dados experimentais ao modelo de um único conjunto de sítios de ligação, a partir do qual foram determinados os parâmetros da interação.

Tabela 36 - Parâmetros da ligação entre a lectina e o ligante.

	N	K _D (μ M)	Δ H (kcal/mol)	Δ G (kcal/mol)	-T Δ S (kcal/mol)	K _a (M ⁻¹)
ALL-lactose	0,49	363	-80	-4,69	75,3	2,75 x 10 ³
CCL-LacNac	0,27	289	-80	-4,83	75,2	3,46 x 10 ³

Fonte: Autora. N: número de sítios aparente por unidade de proteína. K_D: Constante de dissociação. Δ H: Variação de entalpia. Δ G: Variação de energia livre de Gibbs. -T Δ S: Contribuição entrópica. K_a: Constante de associação.

A interação das duas lectinas e seus respectivos ligantes foram caracterizadas por um processo exotérmico e espontâneo, evidenciado pela Δ H e Δ G negativas, indicando que nas condições experimentais avaliadas, a proteína e o ligante tendem a se associar espontaneamente em solução (Kabiri; Unsworth, 2014). A contribuição entrópica positiva, indica que ela é desfavorável, com a perda de liberdade conformacional durante a formação do complexo (Freire, 2008). Um comportamento característico entre lectina e carboidrato, onde a proteína sofre conformações para interagir com o ligante de forma específica.

A constante de dissociação (K_D) para ALL-Lactose foi de 363 μ M, evidenciando uma afinidade moderada entre a lectina e o carboidrato. O número aparente de sítios de ligação estimado é de 0,49 sítio de ligação por monômero. Enquanto para CCL a constante de dissociação foi de 289 μ M, também com uma afinidade moderada, e o número aparente de sítios de ligação estimado é de 0,27 sítio de ligação por monômero.

Considerando que ambas as galectinas apresentam uma organização tetramérica, o valor de N = 0,49 observado para ALL é compatível com a existência de aproximadamente dois sítios funcionais por tetrâmero. Enquanto para CCL, o N = 0,27 sugere a presença de um sítio funcional por tetrâmero. Esse comportamento indica que nem todas as subunidades do oligômero contribuem igualmente para a ligação, possivelmente por acessibilidade diferencial dos domínios de reconhecimento a carboidratos, devido a oligomerização (Campoy; Freire, 2005).

Os experimentos de calorimetria de titulação isotérmica demonstraram que ambas as lectinas avaliadas, CCL e ALL, são capazes de reconhecer seus respectivos glicanos alvo em solução na faixa de micromolar. Esses resultados estão de acordo com o observada para galectinas, onde o domínio de reconhecimento a carboidratos apresenta afinidade variável de acordo com a complexidade do ligante na faixa de micromolar e milimolar (Cummings *et al.*, 2022). A lectina CCL apresentou interação mais significativa com N-acetil-D-lactosamina, visto que apresentou uma constante de dissociação menor (Kabiri; Unsworth, 2014).

Valores de afinidade nessa faixa são compatíveis com aqueles descritos para galectinas humanas avaliadas por calorimetria de titulação isotérmica. No estudo conduzido por Sindrewicz e colaboradores (2019), a galectina-3 apresentou K_D de aproximadamente 24 μM para LacNAc e cerca de 93 μM para lactose, enquanto a interação com galactose ocorreu na faixa milimolar (3 mM), evidenciando maior afinidade por dissacarídeos contendo galactose terminal quando comparados a monossacarídeos. De forma semelhante, a galectina-2 apresentou interação para LacNAc com K_D em torno de 654 μM , e com lactose de aproximadamente 1,3 mM, reforçando que valores na faixa micromolar a milimolar são característicos das interações galectina-carboidrato quando avaliadas com ligantes simples em solução (Sindrewicz *et al.*, 2019).

Em outros estudos envolvendo galectina-3 e variantes estruturais desse domínio, García Caballero e colaboradores (2020) demonstraram que a interação com LacNAc ocorre com valores de K_D também situados na faixa de micromolar, independentemente do arranjo estrutural da proteína, indicando que o domínio de reconhecimento a carboidratos mantém afinidade semelhante mesmo diante de diferentes apresentações oligoméricas.

Estudos com lectinas de esponjas marinhas demonstram que afinidades mais elevadas são observadas quando os ligantes apresentam caráter multivalente. No trabalho de Andrade e colaboradores (2025), a interação da lectina AfiL com a mucina gástrica suína apresentou constante de dissociação de aproximadamente 1,71 μM , valor significativamente inferior àqueles observados para dissacarídeos isolados, evidenciando o efeito da multivalência na estabilização do complexo lectina-carboidrato. No mesmo estudo, a interação da AfiL com lactose apresentou K_D na faixa milimolar, consideravelmente superior ao observado para a mucina, reforçando que carboidratos simples promovem interações de menor afinidade quando comparados a glicoconjugados multivalentes (Andrade *et al.*, 2025).

Portanto, os valores de K_D obtidos para CCL-LacNAc e ALL-lactose estão em concordância com o observado para as galectinas humanas e outras lectinas de esponja marinha, corroborando que a afinidade na faixa de micromolar é uma característica intrínseca das interações entre lectinas e carboidratos.

7.4 Conclusão

A determinação das estruturas tridimensionais das lectinas ALL e CCL por cristalografia de raios X permitiu a elucidação detalhada de suas organizações estruturais e dos mecanismos envolvidos no reconhecimento a carboidratos. As análises de cocristalização e calorimetria de titulação isotérmica confirmaram a especificidade das interações lectina-carboidrato, fornecendo evidências experimentais importantes fundamentais para a compreensão funcional dessas proteínas. Portanto, esse capítulo contribui significativamente para a ampliação do conhecimento estrutural das lectinas de organismos marinhos, cujo dados ainda são escassos na literatura.

8. CONCLUSÃO DA TESE

O presente trabalho contribuiu de forma significativa para ampliar o conhecimento estrutural e funcional de lectinas de organismos marinhos, um grupo ainda pouco explorado nesse contexto. As inúmeras abordagens adotadas, desde os desafios de produção recombinante até a elucidação estrutural e o potencial biotecnológico dessas biomoléculas, permitiram não apenas a caracterização individual de diferentes lectinas, mas também a compreensão dos fatores que condicionam sua expressão, funcionalidade e interação com ligantes.

A determinação das estruturas tridimensionais das galectinas ALL e CCL representa um avanço muito relevante, visto que estruturas determinadas de esponjas marinhas são raras. Além disso, a caracterização da interação lectina-carboidrato forneceu informações fundamentais para entender os mecanismos de reconhecimento molecular dessas proteínas e ampliar os estudos fisiológicos e biotecnológicos.

Adicionalmente, os estudos de expressão heteróloga permitiram avaliar diferentes sistemas de produção recombinante, evidenciando tanto o potencial quanto os desafios envolvidos para obter uma lectina recombinante ativa. Os resultados evidenciaram que a obtenção de lectinas recombinantes funcionalmente ativas permanece um desafio dependente da natureza estrutural da proteína.

A ausência de solubilidade e funcionalidade da rEEL, mesmo após diversas estratégias de expressão, destaca a complexidade associada ao dobramento correto de proteínas ricas em pontes dissulfeto. As dificuldades enfrentadas nesse trabalho para expressar a lectina de equinodermo é condizente com o relatado na literatura, reforçando que essas moléculas, de fato, apresentam características estruturais complexas que impõem desafios à sua expressão heteróloga.

No entanto, os estudos também demonstram avanços significativos pois conseguimos expressar uma lectina de alga marinha em *E. coli* e uma lectina de invertebrado marinho em sistema de *K. pastoris*, abordagem pouco descrita na literatura para lectinas oriundas desses organismos. rADEL demonstrou que sistemas eucarióticos, como *K. pastoris*, podem favorecer a obtenção de proteínas biologicamente ativas. Esses resultados reforçam que a escolha do sistema de expressão depende das características estruturais da lectina-alvo.

Por outro lado, a expressão bem-sucedida da rAmL em *E. coli*, com manutenção de características bioquímicas e funcionais semelhantes à proteína nativa, demonstra que sistemas procarióticos podem ser altamente eficientes para determinadas lectinas. Além disso, a

atividade antibacteriana observada em associação com antibióticos evidencia o potencial dessas proteínas no contexto do combate à resistência bacteriana.

Portanto, os resultados apresentados demonstram que as dificuldades encontradas na expressão de algumas lectinas e o sucesso em outros casos, evidenciam que não existe uma abordagem universal, mas a necessidade de estratégias de acordo com a natureza de cada proteína. Ao mesmo tempo, os avanços estruturais e funcionais obtidos reforçam a importância desses estudos para orientar aplicações futuras, fornecendo uma base importante para estudos de exploração do potencial biotecnológico das lectinas marinhas, principalmente em aplicações na área da saúde.

REFERÊNCIAS

- ABREU, T. M. *et al.* Antinociceptive and anti-inflammatory activities of the lectin from marine red alga *Solieria filiformis*. **Planta medica**, [s.l.]. v. 82, n. 07, p. 596-605, 2016.
- AGIRRE, J. *et al.* The CCP4 suite: integrative software for macromolecular crystallography. **Biological Crystallography**, [s.l.]. v. 79, n. 6, p. 449-461, 2023.
- AHMAD, I. *et al.* Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. **Protein expression and purification**, [s.l.]. v. 144, p. 12-18, 2018.
- AHMED, M. K. *et al.* An Update of Lectins from Marine Organisms: Characterization, Extraction Methodology, and Potential Biofunctional Applications. **Marine Drugs**, [s.l.]. v. 20, n. 7, p. 430, 2022.
- AINES, R. T. *et al.* The S⁺ tag fusion system for protein purification. **Methods in enzymology**, [s.l.]. v. 326, p. 362–376, 2000.
- ALÉSSIO, F. M. **Portal de Zoologia de Pernambuco**, [s.l.]. 2023. Disponível em: www.portal.zoo.bio.br. Acesso em: 23 Maio 2023.
- ANAM, C. *et al.* Cytotoxicity of crude lectins from red macroalgae from the southern coast of Java island, Gunung Kidul Regency, Yogyakarta, Indonesia. In: **IOP Conference Series: Materials Science and Engineering**. IOP Publishing, p. 012017, 2017.
- ANDRADE, F. R. N. *et al.* Afil, a Lectin from *Aplysina fistularis*, Exhibits Antibiofilm and Synergistic Antibacterial Activity Against Resistant Bacteria. **Microorganisms**, [s.l.]. v. 13, n. 6, p. 1349, 2025.
- ANGGIANI, M.; HELIANTI, I.; ABINAWANTO, A. Optimization of methanol induction for expression of synthetic gene *Thermomyces lanuginosus* lipase in *Pichia pastoris*. In: **AIP Conference Proceedings**, [s.l.]. AIP Publishing LLC, p. 020157, 2018.
- BANEYX, F.; MUJACIC, M. Recombinant protein folding and misfolding in *Escherichia coli*. **Nature biotechnology**, [s.l.]. v. 22, n. 11, p. 1399–1408, 2004.
- BARTON, C. *et al.* Pharmacokinetics of the antiviral lectin griffithsin administered by different routes indicates multiple potential uses. **Viruses**, [s.l.]. v. 8, n. 12, p. 331, 2016.
- BEKIARI, M. Marine bioprospecting: understanding the activity and some challenges related to environmental protection, scientific research, ethics, and the law. In: **Blue Planet Law: The Ecology of our Economic and Technological World**, [s.l.]. Cham: Springer International Publishing, 2023.
- BESSA, L. J. *et al.* Synergistic effects between thioxanthenes and oxacillin against methicillin-resistant *Staphylococcus aureus*. **Microbial drug resistance**, [s.l.]. v. 21, n. 4, p. 404–415, 2015.
- BIO RAD. **Proteus IMAC handbook**: mini and mid spin columns. 4.ed. Inglaterra: 2016. 58p.
- BOLANOS-GARCIA, V. M.; DAVIES, O. R. Structural analysis and classification of native proteins from *Escherichia coli* commonly co-purified by immobilised metal affinity

chromatography. **Biochimica et biophysica acta – general subjects**, [s.l.]. v. 1760, n. 9, p. 1304–1313, 2006.

BOLLENBACH, T. Antimicrobial interactions: mechanisms and implications for drug discovery and resistance evolution. **Current opinion in microbiology**, [s.l.]. v. 27, p. 1-9, 2015.

BOONSRI, N. *et al.* Protein extract from red seaweed *Gracilaria fisheri* prevents acute hepatopancreatic necrosis disease (AHPND) infection in shrimp. **Journal of Applied Phycology**, [s.l.]. v. 29, n. 3, p. 1597-1608, 2017.

BORNHORST, J. A.; FALKE, J. J. Purification of proteins using polyhistidine affinity tags. **Methods in enzymology**, [s.l.]. v. 326, p. 245–254, 2000.

BOYD, W. C.; SHAPLEIGH, E. Specific precipitating activity of plant agglutinins (lectins). **Science**, [s.l.]. v. 119, p. 419, 1954.

BROOKS, S. A. Lectin histochemistry: historical perspectives, state of the art, and the future. **Histochemistry of single molecules**, [s.l.]. p. 93-107, 2017.

BURGARD, J. *et al.* The secretome of *Pichia pastoris* in fed-batch cultivations is largely independent of the carbon source but changes quantitatively over cultivation time. **Microbial biotechnology**, [s.l.]. v. 13, n. 2, p. 479–494, 2020.

BURGESS-BROWN, N. A. *et al.* Codon optimization can improve expression of human genes in *Escherichia coli*: a multi-gene study. **Protein expression and purification**, [s.l.]. v. 59, n. 1, p. 94–102, 2008.

CAMPOY, A. V.; FREIRE, E. ITC in the post-genomic era...? Priceless. **Biophysical chemistry**, [s.l.]. v. 115, n. 2-3, p. 115-124, 2005.

CARNEIRO, R. F. *et al.* Isolation, biochemical characterization and antibiofilm effect of a lectin from the marine sponge *Aplysina lactuca*. **International Journal of Biological Macromolecules**, [s.l.]. v. 99, p. 213–222, 2017a.

CARNEIRO, R. F. *et al.* Toward enhanced antibiotic efficacy: exploring the synergistic potential of marine-derived lectins against human pathogenic bacteria. **Anais da Academia Brasileira de Ciências**, [s.l.]. v. 96, n. 4, p. e20240072, 2024.

CARNEIRO, R. F. *et al.* L-Rhamnose-binding lectin from eggs of the *Echinometra lucunter*: Amino acid sequence and molecular modeling. **International journal of biological macromolecules**, [s.l.]. v. 78, p. 180-188, 2015.

CARNEIRO, R. F. *et al.* Purification, biochemical characterization, and amino acid sequence of a novel type of lectin from *Aplysia dactylomela* eggs with antibacterial/antibiofilm potential. **Marine Biotechnology**, [s.l.]. v. 19, n. 1, p. 49-64, 2017b.

CARUGO, O.; PONGOR, S. A normalized root-mean-square distance for comparing protein three-dimensional structures. **Protein science**, [s.l.]. v. 10, n. 7, p. 1470-1473, 2001.

CATANZARO, E. *et al.* Antitumor potential of marine and freshwater lectins. **Marine drugs**, [s.l.]. v. 18, n. 1, p. 11, 2020.

CEREGHINO, J. L.; CREGG, J. M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. **FEMS microbiology reviews**, [s.l.]. v. 24, n. 1, p. 45-66, 2000.

CHANG, V. T. *et al.* Glycoprotein structural genomics: solving the glycosylation problem. **Structure**, [s.l.]. v. 15, n. 3, p. 267–273, 2007.

CHAVES, R. P. *et al.* Structural characterization of two isolectins from the marine red alga *Solieria filiformis* (Kützinger) PW Gabrielson and their anticancer effect on MCF-7 breast cancer cells. **International journal of biological macromolecules**, [s.l.]. v. 107, p. 1320–1329, 2018.

CHAVES, R. P. *et al.* Structural study and antimicrobial and wound healing effects of lectin from *Solieria filiformis* (Kützinger) PW Gabrielson. **Biochimie**, [s.l.]. v. 214, p. 61–76, 2023.

CHAVES, R. P. *et al.* Structural characterization and bacterial agglutination of a new OAAH-type lectin from the marine red alga *Amansia multifida* JV Lamouroux. **Journal of Applied Phycology**, [s.l.]. p. 1-14, 2026.

CHBEL, A. *et al.* Marine biomolecules: a promising approach in therapy and biotechnology. **European Journal of Biological Research**, [s.l.]. v. 11, n. 1, p. 122-133, 2021.

CHEN, K. *et al.* Anomalous-dispersion phase matching in a hypergap crystal. **Physical review A**, [s.l.]. v. 112, n. 4, p. 043513, 2025.

CHEUNG, R. C. F. *et al.* Marine lectins and their medicinal applications. **Applied microbiology and biotechnology**, [s.l.]. v. 99, n. 9, p. 3755-3773, 2015.

CHI, E. Y. *et al.* Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. **Pharmaceutical research**, [s.l.]. v. 20, n. 9, p. 1325–1336, 2003.

CHOI, M. J. *et al.* Characterization of a C-type lectin domain-containing protein with antibacterial activity from Pacific abalone (*Haliotis discus hannai*). **International journal of molecular sciences**, [s.l.]. v. 23, n. 2, p. 698, 2022.

CHOPRA, A. K. *et al.* Improved synthesis of *Salmonella typhimurium* enterotoxin using gene fusion expression systems. **Gene**, [s.l.]. v. 144, n. 1, p. 81–85, 1994.

CHOPRA, I.; ROBERTS, M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. **Microbiology and molecular biology reviews**, [s.l.]. v. 65, n. 2, p. 232-260, 2001.

CONINCK, T.; VANHAEREN, H.; VAN DAMME, E. J. M. The Achilles Heel of Protein Biochemistry: Insolubility of Recombinant Proteins—A Case Study About Producing a Rice Enzyme. **International Journal of Molecular Sciences**, [s.l.]. v. 26, n. 18, p. 8974, 2025.

COOPER, A. Heat capacity effects in protein folding and ligand binding: a re-evaluation of the role of water in biomolecular thermodynamics. **Biophysical chemistry**, [s.l.]. v. 115, n. 2-3, p. 89-97, 2005.

COSTA FILHO, M. F. *et al.* Lectins of algal origin: 25 years of progress in structural insights and biotechnological applications. **Journal of Applied Phycology**, [s.l.]. v. 37, n. 5, p. 3151-3168, 2025.

- COSTA, F. H. F. *et al.* Purification and characterization of a lectin from the red marine alga *Amansia multifida*. **Physiology and Molecular Biology of Plants**, [s.l.]. v. 5, p.53-61, 1999.
- COSTA, S. J. *et al.* The novel Fh8 and H fusion partners for soluble protein expression in *Escherichia coli*: a comparison with the traditional gene fusion technology. **Applied microbiology and biotechnology**, [s.l.]. v. 97, n. 15, p. 6779–6791, 2013.
- CRUZ-CABEZA, A. J.; BERNSTEIN, J. Conformational polymorphism. **Chemical reviews**, [s.l.]. v. 114, n. 4, p. 2170-2191, 2014.
- CUMMINGS, R. D. *et al.* Galectins. **Essentials of Glycobiology**. Cold. Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022.
- DALY, R.; HEARN, M. T. W. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. **Journal of molecular recognition**, [s.l.]. v. 18, n. 2, p. 119–138, 2005.
- DAMASCENO, L. M.; HUANG, C. J.; BATT, C. A. Protein secretion in *Pichia pastoris* and advances in protein production. **Applied microbiology and biotechnology**, [s.l.]. v. 93, n. 1, p. 31–39, 2012.
- DATTA, D.; TALAPATRA, S. N.; SWARNAKAR, S. An overview of lectins from freshwater and marine macroinvertebrates. **World Scientific News**, [s.l.]. n. 46, p. 77-87, 2016.
- DEB, S.; MANFIELD, I.; CARPENTER, B. A novel method for the purification of recombinant subunit I of the Dolichos biflorus seed lectin. **Molecular biotechnology**, [s.l.]. v. 8, n. 1, p. 1-6, 1997.
- DIXON, R. A.; CHOPRA, I. Leakage of periplasmic proteins from *Escherichia coli* mediated by polymyxin B nonapeptide. **Antimicrobial agents and chemotherapy**, [s.l.]. v. 29, n. 5, p. 781–788, 1986.
- DOBSON, C. M. Protein folding and misfolding. **Nature**, [s.l.]. v. 426, n. 6968, p. 884–890, 2003.
- DRESCH, R. R. *et al.* Biological activities of ACL-I and physicochemical properties of ACL-II, lectins isolated from the marine sponge *Axinella corrugata*. **Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology**, [s.l.]. v. 161, n. 4, p. 365-370, 2012.
- DRESCH, R. R. *et al.* ACL-I, a lectin from the marine sponge *Axinella corrugata*: Isolation, characterization and chemotactic activity. **Comparative 110 Biochemistry Physiology – Part C: Toxicology and Pharmacology**, [s.l.]. v. 148, p. 23–30, 2008.
- DUARTE, J. A. *et al.* Structural characterization of a galectin from the marine sponge *Aplysina lactuca* (ALL) with synergistic effects when associated with antibiotics against bacteria. **Biochimie**, [s.l.]. v. 214, p. 165-175, 2023.
- EL-MARADNY, Y. A. *et al.* Lectins purified from medicinal and edible mushrooms: Insights into their antiviral activity against pathogenic viruses. **International journal of biological macromolecules**, [s.l.]. v. 179, p. 239-258, 2021.

ELUMALAI, P. *et al.* The Role of Lectins in Finfish: A Review. **Fisheries Science & Aquaculture**, [s.l.]. v. 27, n. 2, p. 152-169, 2019.

EMSLEY, P. *et al.* Features and development of Coot. **Biological crystallography**, [s.l.]. v. 66, n. 4, p. 486-501, 2010.

ERICKSON, H. P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. **Biological procedures online**, [s.l.]. v. 11, n. 1, p. 32, 2009.

EUROPEAN COMMITTEE FOR ANTIMICROBIAL SUSCEPTIBILITY TESTING (EUCAST) OF THE EUROPEAN SOCIETY OF CLINICAL MICROBIOLOGY AND INFECTIOUS DISEASES (ESCMID). Terminology relating to methods for the determination of susceptibility of bacteria to antimicrobial agents. **Clinical microbiology and infection**, [s.l.]. v. 6, n. 9, p. 503-508, 2000.

FATHI-ROUDSARI, M.; AKHAVIAN-TEHRANI, A.; MAGHSOUDI, N. Comparison of three *Escherichia coli* strains in recombinant production of reteplase. **Avicenna journal of medical biotechnology**, [s.l.]. v. 8, n. 1, p. 16, 2016.

FEINBERG, H.; MITCHELL, D. A.; DRICKAMER, K.; WEIS, W.I. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. **Science**, [s.l.]. v. 294, n. 5549, p. 2163–2166, 2001.

FREIRE, E. Do enthalpy and entropy distinguish first in class from best in class?. **Drug discovery today**, [s.l.]. v. 13, n. 19-20, p. 869-874, 2008.

FREYMANN, D. M. *et al.* Structure of a tetrameric galectin from *Cinachyrella* sp. (ball sponge). **Acta Crystallography**, [s.l.]. v. 68, p.1163-1674, 2012.

FRICKER, L. D. Limitations of mass spectrometry-based peptidomic approaches. **Journal of the American Society for Mass Spectrometry**, [s.l.]. v. 26, n. 12, p. 1981–1991, 2015.

FUJIMOTO, Z.; TATENO, H.; HIRABAYASHI, J. Lectin structures: classification based on the 3-D structures. **Lectins**, [s.l.]. p. 579-606, 2014.

GABIUS, H. J. *et al.* What is the sugar code?. **ChemBioChem**, [s.l.]. v. 23, n. 13, p. e202100327, 2022.

GĄCIARZ, A. *et al.* Efficient soluble expression of disulfide bonded proteins in the cytoplasm of *Escherichia coli* in fed-batch fermentations on chemically defined minimal media. **Microbial cell factories**, [s.l.]. v. 16, n. 1, p. 108, 2017.

GACIARZ, A. *et al.* Systematic screening of soluble expression of antibody fragments in the cytoplasm of *Escherichia coli*. **Microbial cell factories**, [s.l.]. v. 15, n. 1, p. 22, 2016.

GALLOWAY, C. A.; SOWDEN, M. P.; SMITH, H. C. Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. **Biotechniques**, [s.l.]. v. 34, n. 3, p. 524-530, 2003.

GANJAVE, S. D. *et al.* Effect of thioredoxin tag, oxidizing environment, and temperature on the global metabolome of *Escherichia coli* strains. **Applied biochemistry and biotechnology**, [s.l.]. p. 1–18, 2025.

- GAO, X. *et al.* Soluble cytoplasmic expression, rapid purification, and characterization of cyanovirin-N as a His-SUMO fusion. **Applied microbiology and biotechnology**, [s.l.]. v. 85, n. 4, p. 1051–1060, 2010.
- GAO, Z. *et al.* Recent developments in the crystallization process: toward the pharmaceutical industry. **Engineering**, [s.l.]. v. 3, n. 3, p. 343–353, 2017.
- GARCÍA CABALLERO, G. *et al.* Influence of protein (human galectin-3) design on aspects of lectin activity. **Histochemistry and cell biology**, [s.l.]. v. 154, n. 2, p. 135–153, 2020.
- GARDÈRES, J. *et al.* Porifera Lectins: Diversity, physiological roles and biotechnological potential. **Marine drugs**, [s.l.]. v.13, p.5059–5101, 2015.
- GHAZARIAN, H.; IDONI, B.; OPPENHEIMER, S. B. A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. **Acta histochemica**, [s.l.]. v. 113, n. 3, p. 236–247, 2011.
- GIOMARELLI, B. *et al.* Recombinant production of anti-HIV protein, griffithsin, by auto-induction in a fermentor culture. **Protein expression and purification**, [s.l.]. v. 47, n. 1, p. 194–202, 2006.
- GIRARD, E. *et al.* High-phasing-power lanthanide derivatives: taking advantage of ytterbium and lutetium for optimized anomalous diffraction experiments using synchrotron radiation. **Biological crystallography**, [s.l.]. v. 59, n. 10, p. 1877–1880, 2003.
- GORAKSHAKAR, A. C.; GHOSH, K. Use of lectins in immunohematology. **Asian journal of transfusion science**, [s.l.]. v. 10, n. 1, p. 12, 2016.
- GREENFIELD, N. J. Using circular dichroism spectra to estimate protein secondary structure. **Nature protocols**, [s.l.]. v. 1, n. 6, p. 2876–2890, 2006.
- GUIRY, M. D.; GUIRY, G. M. AlgaeBase. **World-wide electronic publication, National University of Ireland**, Galway, 2025. Disponível em: <http://www.algaebase.org>. Acesso em: 10.jan.26.
- HABLER, K. *et al.* Understanding isotopes, isomers, and isobars in mass spectrometry. **Journal of mass spectrometry and advances in the clinical lab**, [s.l.]. v. 33, p. 49–54, 2024.
- HAMILTON, S. R.; GERNGROSS, T. U. Glycosylation engineering in yeast: the advent of fully humanized yeast. **Current opinion in biotechnology**, [s.l.]. v. 18, n. 5, p. 387–392, 2007.
- HAN, J. W. *et al.* Functional recombinants designed from a fetuin/asialofetuin-specific marine algal lectin, rhodobindin. **Marine drugs**, [s.l.]. v. 13, n. 4, p. 2183–2195, 2015.
- HARJU, S.; FEDOSYUK, H.; PETERSON, K. R. Rapid isolation of yeast genomic DNA: Bust n'Grab. **BMC biotechnology**, [s.l.]. v. 4, n. 1, p. 8, 2004.
- HASSELL, A. M.; *et al.* Crystallization of protein–ligand complexes. **Biological crystallography**, [s.l.]. v. 63, n. 1, p. 72–79, 2007.
- HATAKEYAMA, T. *et al.* Characterization of a recombinant C-type lectin, rCEL-IV, expressed in *Escherichia coli* cells using a synthetic gene. **Biochimica et Biophysica Acta (BBA)-General Subjects**, [s.l.]. v. 1760, n. 3, p. 318–325, 2006.

HATAKEYAMA, T *et al.* Characterization of a recombinant C-type lectin, rCEL-IV, expressed in *Escherichia coli* cells using a synthetic gene. **Biochimica et biophysica acta – general subjects**, [s.l.]. v. 1760, n. 3, p. 318–325, 2006.

HAYASHI, R. *et al.* Novel galectins purified from the sponge *Chondrilla australiensis*: unique structural features and cytotoxic effects on colorectal cancer cells mediated by TF-antigen binding. **Marine Drugs**, [s.l.]. v. 22, n. 9, p. 400, 2024.

HIRAYAMA, M. *et al.* High-mannose specific lectin and its recombinants from a carrageenophyta *Kappaphycus alvarezii* represent a potent anti-HIV activity through high-affinity binding to the viral envelope glycoprotein gp120. **Marine biotechnology**, [s.l.]. v. 18, n. 1, p. 144–160, 2016.

HUNG, L. D. *et al.* Purification, primary structure, and biological activity of the high-mannose N-glycan-specific lectin from cultivated *Eucheuma denticulatum*. **Journal of applied phycology**, [s.l.]. v. 27, n. 4, p. 1657–1669, 2015.

HUYNH, K.; PARTCH, C. L. Analysis of protein stability and ligand interactions by thermal shift assay. **Current protocols in protein science**, [s.l.]. v. 79, n. 1, p. 28.9. 1-28.9. 14, 2015.

HWANG, H. J.; *et al.* Induction of recombinant lectin expression by an artificially constructed tandem repeat structure: a case study using *Bryopsis plumosa* mannose-binding lectin. **Biomolecules**, [s.l.]. v. 8, n. 4, p. 146, 2018.

IDOWU, O. S. *et al.* Optimizing SDS-PAGE for accurate protein characterization in nutritional research and food quality assessment. **International journal of innovative science and research technology**, [s.l.]. v. 10, n. 1, 2025.

INVITROGEN, THERMO FISHER SCIENTIFIC. **BL21 Star™(DE3) One Shot® BL21 Star™(DE3)pLysS One Shot® Chemically Competent Cells**. 1.ed. EUA: 2010. 24p.

INVITROGEN, THERMO FISHER SCIENTIFIC. **Pichia Expression Kit user guide**: For expression of recombinant proteins in *Pichia pastoris*. 2.ed. EUA: 2020. 111p.

INVITROGEN. **pGAPZ A, B, and C pGAPZα A, B, and C**: Pichia expression vectors for constitutive expression and purification of recombinant proteins. EUA: 2010. 44p.

JANDÚ, J. J. B. *et al.* Targeting the immune system with plant lectins to combat microbial infections. **Frontiers in pharmacology**, [s.l.]. v. 8, p. 671, 2017.

JANSON, J. C. Protein purification: principles, high resolution methods, and applications. New York: John Wiley & Sons, 2011.

JIANG, Y. *et al.* Protective immunity induced by Tp0136 epitope vaccines with mRNA LNP or protein delivery. **npj Vaccines**, [s.l.]. 2025.

JUAN, L. L. *et al.* Pharmaceutical applications of lectins. **Journal of Drug Delivery Science and Technology**, [s.l.]. v. 42, p. 126-133, 2017.

JUTURU, V.; WU, J. C. Heterologous protein expression in *Pichia pastoris*: latest research progress and applications. **ChemBioChem**, [s.l.]. v. 19, n. 1, p. 7-21, 2018.

KABIRI, M.; UNSWORTH, L. D. Application of isothermal titration calorimetry for characterizing thermodynamic parameters of biomolecular interactions: peptide self-assembly

and protein adsorption case studies. **Biomacromolecules**, [s.l.]. v. 15, n. 10, p. 3463-3473, 2014.

KABSCH, W. XDS, **Acta Crystallogr D Biol Crystallogr**, [s.l.]. v. 66, p. 125-132, 2010.

KARAKOSTIS, K.; *et al.* Heterologous expression of newly identified galectin-8 from sea urchin embryos produces recombinant protein with lactose binding specificity and anti-adhesive activity. **Scientific reports**, [s.l.]. v. 5, n. 1, p. 17665, 2015.

KARYOLAIMOS, A.; DE GIER, J. W. Strategies to enhance periplasmic recombinant protein production yields in *Escherichia coli*. **Frontiers in bioengineering and biotechnology**, [s.l.]. v. 9, p. 797334, 2021.

KAUR, J.; KUMAR, A.; KAUR, J. Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. **International journal of biological macromolecules**, [s.l.]. v. 106, p. 803-822, 2018.

KHLEBODAROVA, T. M. *et al.* *Komagataella phaffii* as a platform for heterologous expression of enzymes used for industry. **Microorganisms**, [s.l.]. v. 12, n. 2, p. 346, 2024.

KIELKOPF, C. L.; BAUER, W.; URBATSCH, I. L. Expression of cloned genes in *E. coli* using IPTG-inducible promoters. **Cold Spring Harbor Protocols**, [s.l.]. v. 2021, n. 2, p. pdb.prot102137, 2021.

KILPATRIC, D. C. Animal lectins: a historical introduction and overview. **Biochimica et Biophysica Acta**, [s.l.]. v. 1572, p. 187-197, 2002.

KILPATRICK, David C. **Handbook of animal lectins, properties and biomedical applications**. Eddinburg: British library, p.468, 2000.

KIM, G. H.; KLOTCHKOVA, T. A.; WEST, J. A. From protoplasm to swarmer: regeneration of protoplasts from disintegrated cells of the multicellular marine green alga *Microdictyon umbilicatum* (Chlorophyta). **Journal Phycology**, [s.l.]. v. 38, p. 174–84, 2002.

KLOCHKOVA, T. A. *et al.* Experimental hybridization between some marine coenocytic green algae using protoplasts extruded in vitro. **Algae**, [s.l.]. v. 20, n. 3, p. 239-249, 2005.

KONG, P. *et al.* A novel C-type lectin from bay scallop *Argopecten irradians* (AiCTL-7) agglutinating fungi with mannose specificity. **Fish & shellfish immunology**, [s.l.]. v. 30, n. 3, p. 836–844, 2011.

KRISSINEL, E. *et al.* CCP4 Cloud for structure determination and project management in macromolecular crystallography. **Biological Crystallography**, [s.l.]. v. 78, n. 9, p. 1079-1089, 2022.

KUO, T. H. *et al.* Ligand specificities and structural requirements of two *Tachypleus* plasma lectins for bacterial trapping. **Biochemical journal**, [s.l.]. v. 393, n. 3, p. 757–766, 2006.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. **Nature**, [s.l.]. v.227, p.680–683, 1970.

LANGER, G. *et al.* Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. **Nature protocols**, [s.l.]. v. 3, n. 7, p. 1171–1179, 2008.

LAURITANO, C.; IANORA, A. Grand challenges in marine biotechnology: Overview of recent EU-funded projects. **Grand challenges in marine biotechnology**, [s.l.]. p. 425-449, 2018.

LEFFLER, H. *et al.* Introduction to galectins. **Glycoconjugate journal**, [s.l.]. v. 19, n. 7, p. 433-440, 2002.

LEGRAND, P.; AISHIMA, J. **CV-GPhL. legrandp/xdsme**: March 2019 version working with the latest XDS version (0.6.6). Zenodo, 2019.

LEVENDOSKY, K. *et al.* Griffithsin and carrageenan combination to target HSV-2 and HPV. **Scientist**, [s.l.]. v. 1, p. 11, 2015.

LI, H. *et al.* A single-CRD C-type lectin from oyster *Crassostrea gigas* mediates immune recognition and pathogen elimination with a potential role in the activation of complement system. **Fish & shellfish immunology**, [s.l.]. v. 44, n. 2, p. 566-575, 2015.

LI, L. *et al.* High level expression, purification, and characterization of the shrimp antimicrobial peptide, Ch-penaeidin, in *Pichia pastoris*. **Protein expression and purification**, [s.l.]. v. 39, n. 2, p. 144-151, 2005.

LI, M.; SU, Z. G.; JANSON, J. C. *In vitro* protein refolding by chromatographic procedures. **Protein expression and purification**, [s.l.]. v. 33, n. 1, p. 1-10, 2004.

LINDER, S.; SCHLIWA, M.; KUBE-GRANDERATH, E. Direct PCR screening of *Pichia pastoris* clones. **Biotechniques**, [s.l.]. v. 20, n. 6, p. 980-982, 1996.

LOBSTEIN, J. *et al.* SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. **Microbial cell factories**, [s.l.]. v. 11, n. 1, p. 753, 2012.

LÖÖKE, M.; KRISTJUHAN, K.; KRISTJUHAN, A. Extraction of genomic DNA from yeasts for PCR-based applications. **Biotechniques**, [s.l.]. v. 50, n. 5, p. 325-328, 2011.

LORIAN, V. **Antibiotics in laboratory medicine**. Lippincott Williams & Wilkins, 2005.

MACAULEY-PATRICK, S. *et al.* Heterologous protein production using the *Pichia pastoris* expression system. **Yeast**, [s.l.]. v. 22, n. 4, p. 249-270, 2005.

MAGUIRE, C. M. *et al.* Characterisation of particles in solution: a perspective on light scattering and comparative technologies. **Science and technology of advanced materials**, [s.l.]. v. 19, n. 1, p. 732-745, 2018.

MAKINO, T.; SKRETAS, G.; GEORGIOU, G. Strain engineering for improved expression of recombinant proteins in bacteria. **Microbial cell factories**, [s.l.]. v. 10, n. 1, p. 32, 2011.

MARCO, A. Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. **Microbial cell factories**, [s.l.]. v. 8, n. 1, p. 26, 2009.

MARQUES, D. N. *et al.* Antibacterial activity of new lectin isolated from the marine sponge *Chondrilla caribensis*. **International Journal of Biological Macromolecules**, [s.l.]. v. 109, p. 1292-1301, 2018.

MARTIN, C. D. *et al.* A simple vector system to improve performance and utilisation of recombinant antibodies. **BMC biotechnology**, [s.l.]. v. 6, n. 1, p. 46, 2006.

MARTÍNEZ-ALARCÓN, D.; BLANCO-LABRA, A.; GARCÍA-GASCA, T. Expression of lectins in heterologous systems. **International journal of molecular sciences**, [s.l.]. v. 19, n. 2, p. 616, 2018.

MATSUZAWA, A. Thioredoxin and redox signaling: Roles of the thioredoxin system in control of cell fate. **Archives of biochemistry and biophysics**, [s.l.]. v. 617, p. 101-105, 2017.

MATTOX, D. E.; BAILEY-KELLOGG, C. Comprehensive analysis of lectin-glycan interactions reveals determinants of lectin specificity. **PLoS computational biology**, [s.l.]. v. 17, n. 10, p. e1009470, 2021.

MCCOY, A. J. Solving structures of protein complexes by molecular replacement with Phaser. **Biological Crystallography**, [s.l.]. v. 63, n. 1, p. 32-41, 2007.

MEIERS, J. *et al.* Lectin antagonists in infection, immunity, and inflammation. **Current Opinion in Chemical Biology**, [s.l.]. v. 53, p. 51-67, 2019.

MENEZES, Vinicius Paulino Pinto. **Estudo estrutural e avaliação do efeito antibacteriano de três lectinas de algas marinhas vermelhas da família OAAH em combinação com antibióticos**. 2024. Dissertação (Mestrado em Engenharia de Pesca) – Centro de Ciências agrárias, Universidade Federal do Ceará, Fortaleza, 2024. Disponível em: <https://repositorio.ufc.br/handle/riufc/21476>. Acesso em: 27 dez. 2025.

MENG, D. M. *et al.* Expression, purification and initial characterization of a novel recombinant antimicrobial peptide Mytichitin-A in *Pichia pastoris*. **Protein expression and purification**, [s.l.]. v. 127, p. 35–43, 2016.

MESQUITA, J. X. *et al.* Lectin from red algae *Amansia multifida* Lamouroux: extraction, characterization and anti-inflammatory activity. **International journal of biological macromolecules**, [s.l.]. v. 170, p. 532–539, 2021.

MEULEMAN, P. *et al.* Griffithsin has antiviral activity against hepatitis C virus. **Antimicrobial agents and chemotherapy**, [s.l.]. v. 55, n. 11, p. 5159-5167, 2011.

MICSONAI, A. *et al.* BeStSel: analysis site for protein CD spectra — 2025 update. **Nucleic acids research**, [s.l.]. v. 53, n. W1, p. W73–W83, 2025.

MIERENDORF, R. C. *et al.* Expression and purification of recombinant proteins using the pET system. In: **The nucleic acid protocols handbook**. Totowa, NJ: Humana Press, 2000. p. 947–977.

MILLET, J.K. *et al.* Middle East respiratory syndrome coronavirus infection is inhibited by griffithsin. **Antiviral research**, [s.l.]. v. 133, p. 1-8, 2016.

MISHRA, A. *et al.* Structure-function and application of plant lectins in disease biology and immunity. **Food and Chemical Toxicology**, [s.l.]. v. 134, p. 110827, 2019.

MITAL, S.; CHRISTIE, G.; DIKICIOGLU, D. Recombinant expression of insoluble enzymes in *Escherichia coli*: a systematic review of experimental design and its manufacturing implications. **Microbial cell factories**, [s.l.]. v. 20, n. 1, p. 208, 2021.

- MOUTINHO, I. C. . *et al.* Bioprospecting and marine ‘omics’: Surfing the deep blue sea for novel bioactive proteins and peptides. **Frontiers in Marine Science**, [s.l.]. v. 11, p. 1362697, 2024.
- MUKHERJEE, S. *et al.* Antibacterial membrane attack by a pore-forming intestinal C-type lectin. **Nature**, [s.l.]. v. 505, n. 7481, p. 103–107, 2014.
- MUNSHI, C.; LEE, H. C. High-Level Expression of Recombinant Aplysia ADP-Ribosyl Cyclase in *Pichia pastoris* by Fermentation. **Protein expression and purification**, [s.l.]. v. 11, n. 1, p. 104-110, 1997.
- MURSHUDOV, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. **Biological crystallography**, [s.l.]. v. 67, n. 4, p. 355-367, 2011.
- NAGANO, C. S. *et al.* HCA and HML isolated from the red marine algae *Hypnea cervicornis* and *Hypnea musciformis* define a novel lectin family. **Protein science**, [s.l.]. v. 14, n. 8, p. 2167–2176, 2005.
- NAGARAJAN, D. Streptomycin. In: **Antibiotics and Their Mechanisms of Action**. Singapore: Springer Nature Singapore, p. 109-116, 2024.
- NAPOLÊÃO, T. H. *et al.* Lectins as natural antibiofilm agents in the fight against antibiotic resistance: a review. **Molecules**, [s.l.]. v. 30, n. 16, p. 3395, 2025.
- NASCIMENTO, Antônia Sâmia Fernandes do. **Lectinas recombinantes das algas marinhas *Hypnea musciformis* (Wulfen) J. V. Lamouroux e *Bryothamnion triquetrum* (S. G. Gmelin) M. Howe: produção heteróloga e caracterização bioquímica**. 2014. Tese (Doutorado em Bioquímica) – Centro de Ciências, Universidade Federal do Ceará, Fortaleza, 2014. Disponível em: <https://repositorio.ufc.br/handle/riufc/10459>. Acesso em: 15. nov. 2025.
- NASCIMENTO, A. S. F. *et al.* Algal lectin binding to core (α 1–6) fucosylated N-glycans: structural basis for specificity and production of recombinant protein. **Glycobiology**, [s.l.]. v. 25, n. 6, p. 607–616, 2015.
- NASCIMENTO-NETO, L. G. *et al.* Characterization of isoforms of the lectin isolated from the red algae *Bryothamnion seaforthii* and its pro-healing effect. **Marine drugs**, [s.l.]. v. 10, n. 9, p. 1936-1954, 2012.
- NASCIMENTO-NETO, L. G. *et al.* Halilectin-3, a lectin from the marine sponge *Haliclona caerulea*, induces apoptosis and autophagy in human breast cancer MCF7 cells through caspase-9 pathway and LC3-II protein expression. **Anti-Cancer Agents in Medicinal Chemistry**, [s.l.]. v. 18, n. 4, p. 521–528, 2018.
- NELSON, D. L.; COX, M. M. **Princípios de bioquímica de Lehninger**. Artmed Editora, 2022
- NEVES, S. A. *et al.* Antinociceptive properties in mice of a lectin isolated from the marine alga *Amansia multifida* Lamouroux. **Brazilian Journal of Medical and Biological Research**, [s.l.]. v. 40, p. 127-134, 2007.
- NEW ENGLAND BIOLABS. **Are SHuffle strains temperature sensitive?** Ipswich, MA: New England Biolabs, 2023. Disponível em: <https://www.neb.com/en/faqs/are-shuffle-strains-temperature-sensitive?>. Acesso em: 19 dez.2025.

NGUYEN, T. K. O. *et al.* Soluble prokaryotic overexpression and purification of human GM-CSF using the protein disulfide isomerase b'a' domain. **International journal of molecular sciences**, [s.l.]. v. 22, n. 10, p. 5267, 2021.

NIKAIDO, H. Molecular basis of bacterial outer membrane permeability revisited. **Microbiology and molecular biology reviews**, [s.l.]. v. 67, n. 4, p. 593–656, 2003.

NOVAGEN. **pET System Manual**. 11.ed. EUA: 2005. 80p.

O'CONNOR, B. F.; MONAGHAN, D.; CAWLEY, J. Lectin affinity chromatography (LAC). In: **Protein chromatography: Methods and protocols**. New York, NY: Springer New York, p. 411-420, 2016.

OLIVEIRA, C.; TEIXEIRA, J. A.; DOMINGUES, L. Recombinant lectins: an array of tailor-made glycan-interaction biosynthetic tools. **Critical reviews in biotechnology**, [s.l.]. v. 33, n. 1, p. 66-80, 2013.

PACE, C. N.; SCHOLTZ, J. M.; GRIMSLEY, G. R. Forces stabilizing proteins. **FEBS letters**, [s.l.]. v. 588, n. 14, p. 2177-2184, 2014.

PEECHAKARA, B. V.; BASIT, H.; GUPTA, M. **Ampicillin**. 2018.

PINHEIRO, U. S.; HAJDU, E.; CUSTODIO, M. R. *Aplysina* Nardo (Porifera, Verongida, Aplysinidae) from the Brazilian coast with description of eight new species. **Zootaxa**, [s.l.]. v. 1609, n. 1, p. 1-51, 2007.

PINHO, S. S.; REIS, C. A. Glycosylation in cancer: mechanisms and clinical implications. **Nature Reviews Cancer**, [s.l.]. v. 15, p. 540–555, 2015.

POURESMAEIL, M.; AZIZI-DARGAHLOU, S. Factors involved in heterologous expression of proteins in *E. coli* host. **Archives of microbiology**, [s.l.]. v. 205, n. 5, p. 212, 2023.

PRIVALOV, P. L. Cold denaturation of protein. **Critical reviews in biochemistry and molecular biology**, [s.l.]. v. 25, n. 4, p. 281-306, 1990.

QI, N.; *et al.* Multiple truncated isoforms of MAVS prevent its spontaneous aggregation in antiviral innate immune signalling. **Nature communications**, [s.l.]. v. 8, n. 1, p. 15676, 2017.

QIU, W. *et al.* A new C-type lectin homolog SpCTL6 exerting immunoprotective effect and regulatory role in mud crab *Scylla paramamosain*. **Frontiers in immunology**, [s.l.]. v. 12, p. 661823, 2021.

RADOMAN, B. *et al.* The degree and length of O-glycosylation of recombinant proteins produced in *Pichia pastoris* depends on the nature of the protein and the process type. **Biotechnology journal**, [s.l.]. v. 16, n. 3, p. 2000266, 2021.

REILY, C. *et al.* Glycosylation in health and disease. **Nature Reviews Nephrology**, [s.l.]. v. 15, n. 6, p. 346-366, 2019.

RIBEIRO, J. P. *et al.* Characterization of a high-affinity sialic acid-specific CBM40 from *Clostridium perfringens* and engineering of a divalent form. **Biochemical Journal**, [s.l.]. v. 473, n. 14, p. 2109-2118, 2016.

RICCIO, G. *et al.* Ten-year research update review: Antiviral activities from marine organisms. **Biomolecules**, [s.l.]. v. 10, n. 7, p. 1007, 2020.

RIGOGLIUSO, S. *et al.* Recovery of bioactive compounds from marine organisms: focus on the future perspectives for pharmacological, biomedical and regenerative medicine applications of marine collagen. **Molecules**, [s.l.]. v. 28, n. 3, p. 1152, 2023.

RODRIGUEZ-BLANCO, J. D. *et al.* The role of pH and Mg on the stability and crystallization of amorphous calcium carbonate. **Journal of alloys and compounds**, [s.l.]. v. 536, p. S477–S479, 2012.

ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. **Frontiers in microbiology**, [s.l.]. v. 5, p. 172, 2014.

SAMBROOK, J. Molecular Cloning: A Laboratory Manual. **Cold Spring Harbor Laboratory Press**, p.2222, 2002.

SAMPAIO, A.H. *et al.* New affinity procedure for the isolation and further characterization of the blood group B specific lectin from the red marine alga *Ptilota plumose*. **Journal of Applied Phycology**, [s.l.]. v.14, p.489–495, 2002.

SAMPAIO, A. H.; ROGERS, D. J.; BARWELL, C. J. A galactose-specific lectin from the red marine alga *Ptilota filicina*. **Phytochemistry**, [s.l.]. v. 48, n. 5, p. 765–769, 1998.

SANFELICE, D.; TEMUSSI, P. A. Cold denaturation as a tool to measure protein stability. **Biophysical chemistry**, [s.l.]. v. 208, p. 4-8, 2016.

SANTOS, J. V. O.; PORTO, A. L. F.; CAVALCANTI, I. M. F. Potential application of combined therapy with lectins as a therapeutic strategy for the treatment of bacterial infections. **Antibiotics**, [s.l.]. v. 10, n. 5, p. 520, 2021.

SHELLER, C. *et al.* A comparative study of CE-SDS, SDS-PAGE, and Simple Western: precision, repeatability, and apparent molecular mass shifts by glycosylation. **Electrophoresis**, [s.l.]. v. 42, n. 14–15, p. 1521–1531, 2021.

SCHLEGEL, S. *et al.* Optimizing heterologous protein production in the periplasm of *Escherichia coli* by regulating gene expression levels. **Microbial cell factories**, [s.l.]. v. 12, n. 1, p. 24, 2013.

SHARON, N; LIS, H. History of lectins: from hemagglutinins to biological recognition molecules. **Glycobiology**, [s.l.]. v. 14, n. 11, p. 53-62, 2004.

SHEVCHENKO, A. *et al.* In-gel digestion for mass spectrometric characterization of proteins and proteomes. **Nature protocols**, [s.l.]. v. 1, n. 6, p. 2859-2860, 2007.

SHIM, E. *et al.* Purification of a sex-specific lectin involved in gamete binding of *Aglaothamnion callophyllidicola* (rhodophyta). **Journal of phycology**, [s.l.]. v. 48, n. 4, p. 916-924, 2012.

SHIMIZU, Y. *et al.* Molecular cloning, functional expression, and characterization of isolectin genes of hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata*. **Bioscience, biotechnology, and biochemistry**, [s.l.]. v. 76, n. 2, p. 276–282, 2012.

SHIMIZU, Y.; YONEKURA, M.; KOUZUMA, Y. Functional expression and mutant analysis of thioredoxin-fused CEL-III, a hemolytic lectin from the marine invertebrate *Cucumaria echinata*. **Bioscience, Biotechnology, and Biochemistry**, [s.l.]. v. 86, n. 8, p. 1071-1074, 2022.

SILVA, R. R. S. *et al.* *Parkia platycephala* lectin enhances the antibiotic activity against multi-resistant bacterial strains and inhibits the development of *Haemonchus contortus*. **Microbial pathogenesis**, [s.l.]. v. 135, p. 103629, 2019.

SILVA, Suzete Roberta da. **Estrutura primária e análise do potencial biotecnológico de uma lectina da alga marinha vermelha *Amansia multifida* j.v. lamouroux**. 2016. Tese (Doutorado em Engenharia de Pesca) – Centro de Ciências agrárias, Universidade Federal do Ceará, Fortaleza, 2016. Disponível em: <https://repositorio.ufc.br/handle/riufc/21476>. Acesso em: 12 nov. 2025.

SIMAS, R. G.; PESSOA JUNIOR, A.; LONG, P. F. Mechanistic aspects of IPTG (isopropylthio- β -galactoside) transport across the cytoplasmic membrane of *Escherichia coli*—a rate limiting step in the induction of recombinant protein expression. **Journal of Industrial Microbiology and Biotechnology**, [s.l.]. v. 50, n. 1, p. 034, 2023.

SINDREWICZ, P. *et al.* Intrinsic tryptophan fluorescence spectroscopy reliably determines galectin-ligand interactions. **Scientific reports**, [s.l.]. v. 9, n. 1, p. 11851, 2019.

SINGH, A.; UPADHYAY, V.; PANDA, A. K. Solubilization and refolding of inclusion body proteins. **Insoluble proteins: methods and protocols**, [s.l.]. p. 283-291, 2015.

SINGH, A. *et al.* Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. **Microbial cell factories**, [s.l.]. v. 14, n. 1, p. 41, 2015.

SINGH, A. *et al.* Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. **Microbial cell factories**, [s.l.]. v. 14, n. 1, p. 1-10, 2015.

SINGH, R. S.; WALIA, A. K. Lectins from red algae and their biomedical potential. **Journal of applied phycology**, [s.l.]. v. 30, n. 3, p. 1833-1858, 2018.

SIVAKAMAVALLI, J. *et al.* Purification and partial characterization of carbohydrate-recognition protein C-type lectin from *Hemifusus pugilinus*. **Carbohydrate Research**, [s.l.]. v. 499, p. 108224, 2021.

SKUBAK, P.; PANNU, N. S. Automatic protein structure solution from weak X-ray data. **Nature communications**, [s.l.]. v. 4, p. 2777, 2013.

SKUBÁK, P.; PANNU, N. S. Automatic protein structure solution from weak X-ray data. **Nature communications**, [s.l.]. v. 4, n. 1, p. 2777, 2013

SMITH, B. A. H; BERTOZZI, C. R. The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. **Nature Reviews Drug Discovery**, [s.l.]. v. 20, n. 3, p. 217-243, 2021.

SORENSEN, H. P.; MORTENSEN, K. K. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. **Microbial cell factories**, [s.l.]. v. 4, n. 1, p. 1, 2005.

SOUSA, A.R.O. *et al.* Structural characterization of a galectin isolated from the marine sponge *Chondrilla caribensis* with leishmanicidal potential. **Biochimica et Biophysica Acta (BBA)-General Subjects**, [s.l.]. v. 1865, n. 12, p. 129992, 2021.

SOUSA, Andressa Rocha de Oliveira. **Caracterização estrutural de uma lectina da esponja marinha *Chondrilla caribensis***. 2021. Dissertação (Mestrado em Engenharia de Pesca)-Centro de Ciências Agrárias, Universidade Federal do Ceará, Fortaleza, 2021. Disponível em: <https://repositorio.ufc.br/handle/riufc/69751>. Acesso em: 15. nov. 2025.

SOUSA, G. *et al.* Potential of marine biomolecules: Advances in extraction and applications of proteins, polysaccharides, and antioxidant compounds. **Foods**, [s.l.]. v. 14, n. 15, p. 2555, 2025.

SRIVASTAVA, A. *et al.* Role of computational methods in going beyond X-ray crystallography to explore protein structure and dynamics. **International journal of molecular sciences**, [s.l.]. v. 19, n. 11, p. 3401, 2018.

STEIN, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. **Applied Crystallography**, [s.l.]. v. 41, n. 3, p. 641-643, 2008.

STEWART, E. J.; ÅSLUND, F.; BECKWITH, J. Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. **The EMBO journal**, 1998.

STOCK, A. **Glicosilación Imágenes De Stock**. 2023. Disponível em: <https://www.alamy.es/imagenes/glicosilaci%C3%B3n.html?sortBy=relevant>. Acesso em: 24 mai. 2023.

STUDIER, F. W. Use of T7 RNA polymerase to direct expression of cloned genes. **Methods Enzymol**, [s.l.]. v. 185, p. 305-313, 1998.

STUDIER, F. W. Protein production by auto-induction in high-density shaking cultures. **Protein expression and purification**, [s.l.]. v. 41, n. 1, p. 207-234, 2005.

SULE, R.; RIVERA, G.; GOMES, A. V. Western blotting (immunoblotting): history, theory, uses, protocol and problems. **Biotechniques**, [s.l.]. v. 75, n. 3, p. 99–114, 2023.

TAKARA. **HAT Protein expression and purification system user manual**. USA: 2013. 23p.

TERADA, D. *et al.* Crystal structure of MytiLec, a galactose-binding lectin from the mussel *Mytilus galloprovincialis* with cytotoxicity against certain cancer cell types. **Scientific reports**, [s.l.]. v. 6, n. 1, p. 28344, 2016.

THEURETZBACHER, U.; PIDDOCK, L. J. V. Non-traditional antibacterial therapeutic options and challenges. **Cell host & microbe**, [s.l.]. v. 26, n. 1, p. 61-72, 2019.

TOMMASONE, S. *et al.* The challenges of glycan recognition with natural and artificial receptors. **Chemical Society Reviews**, [s.l.]. v. 48, n. 22, p. 5488-5505, 2019.

TORRES, R. C. F.; *et al.* Structural insights and antimicrobial synergy of a proto-galectin from the marine sponge *Aiolochoia crassa*. **Comparative biochemistry and physiology part B: biochemistry and molecular biology**, [s.l.]. v. 275, p. 111034, 2025.

TRIPATHI, N. K. Production and purification of recombinant proteins from *Escherichia coli*. **ChemBioEng reviews**, [s.l.]. v. 3, n. 3, p. 116–133, 2016.

TUTTLE, A. R.; TRAHAN, N. D.; SON, M. S. Growth and maintenance of *Escherichia coli* laboratory strains. **Current protocols**, [s.l.]. v. 1, n. 1, p. e20, 2021.

VAN DAMME, E. J. M. 35 years in plant lectin research: a journey from basic science to applications in agriculture and medicine. **Glycoconjugate journal**, [s.l.]. v. 39, n. 1, p. 83-97, 2022.

VARKI, A. Biological roles of glycans. **Glycobiology**, [s.l.]. v. 27, n. 1, p. 3-49, 2017.

VASILENKO, A. *et al.* Obtaining and refolding a recombinant mannan-binding lectin from the holothurian *Apostichopus japonicus*. **Russian journal of marine biology**, [s.l.]. v. 38, n. 1, p. 72–78, 2012.

VASTA, G. R. *et al.* Structural and functional diversity of the lectin repertoire in fish: Relevance to innate and adaptive immunity. **Developmental Comparative Immunology**, [s.l.]. v. 3, p. 1388–1399, 2011.

VASTA, G. R.; WANG, J. X. Galectin-mediated immune recognition: Opsonic roles with contrasting outcomes in selected shrimp and bivalve mollusk species. **Developmental & Comparative Immunology**, [s.l.]. p. 103721, 2020.

VISWAMBARI DEVI, R.; BASILROSE, M. R.; MERCY, P. D. Prospect for lectins in arthropods. **Italian Journal of Zoology**, [s.l.]. v. 77, n. 3, p. 254-260, 2010.

WALVEKAR, P.; GANNIMANI, R.; GOVENDER, T. Combination drug therapy via nanocarriers against infectious diseases. **European Journal of Pharmaceutical Sciences**, [s.l.]. v. 127, p. 121-141, 2019.

WANG, H. *et al.* Cloning and characterization of a novel C-type lectin from Zhikong scallop *Chlamys farreri*. **Molecular immunology**, [s.l.]. v. 44, n. 5, p. 722–731, 2007.

WANG, H. *et al.* A novel C-type lectin from the sea cucumber *Apostichopus japonicus* (AjCTL-2) with preferential binding of D-galactose. **Fish & shellfish immunology**, [s.l.]. v. 79, p. 218–227, 2018.

WANG, L. *et al.* The immune role of C-type lectins in molluscs. **Invertebrate Survival Journal**, [s.l.]. v. 8, n. 2, p. 241-246, 2011.

WANG, W. H.; *et al.* The role of galectins in virus infection-A systemic literature review. **Journal of Microbiology, Immunology and Infection**, [s.l.]. v. 53, n. 6, p. 925-935, 2020.

WANG, X. W.; WANG, J. X. Pattern recognition receptors acting in innate immune system of shrimp against pathogen infections, **Fish Shellfish Immunology**, [s.l.]. v. 34, n. 4, p. 981–989, 2013.

WARFEL, K. F. *et al.* Cell-free expression and characterization of multivalent rhamnose-binding lectins using bio-layer interferometry. **Glycobiology**, [s.l.]. v. 33, n. 5, p. 358–363, 2023.

WATARI, H. *et al.* A marine sponge-derived lectin reveals hidden pathway for thrombopoietin receptor activation. **Nature Communications**, [s.l.]. v. 13, n. 1, p. 7262, 2022.

WAUGH, D. S. An overview of enzymatic reagents for the removal of affinity tags. **Protein expression and purification**, [s.l.]. v. 80, n. 2, p. 283–293, 2011.

WEI, X. *et al.* Critical roles of sea cucumber C-type lectin in non-self recognition and bacterial clearance. **Fish & shellfish immunology**, [s.l.]. v. 45, n. 2, p. 791–799, 2015.

WEIDENMAIER, C.; PESCHEL, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. **Nature reviews microbiology**, [s.l.]. v. 6, n. 4, p. 276–287, 2008.

WILSON, N.; SIMPSON, R.; COOPER-LIDDELL, C. Introductory glycosylation analysis using SDS-PAGE and peptide mass fingerprinting. **Glycomics: methods and protocols**, [s.l.]. p. 205-212, 2008.

WURM, D. J. *et al.* How to trigger periplasmic release in recombinant *Escherichia coli*: a comparative analysis. **Engineering in life sciences**, [s.l.]. v. 17, n. 2, p. 215–222, 2017.

XUE, Z. *et al.* A review of the immune molecules in the sea cucumber. **Fish & shellfish immunology**, [s.l.]. v. 44, n. 1, p. 1-11, 2015.

YAMASHITA, K. *et al.* GEMMI and Servalcat restrain REFMAC5. **Biological Crystallography**, [s.l.]. v. 79, n. 5, p. 368-373, 2023.

YAN, C.; ZHANG, J.; LIANG, T. Diorganotin (IV) complexes with 4-nitro-N-phthaloylglycine: Synthesis, characterization, antitumor activity and DNA-binding studies. **Biomedicine & Pharmacotherapy**, [s.l.]. v. 71, p. 119-127, 2015.

YANG, M. *et al.* Role of N-linked glycosylation in the secretion and enzymatic properties of *Rhizopus chinensis* lipase expressed in *Pichia pastoris*. **Microbial cell factories**, [s.l.]. v. 14, n. 1, p. 40, 2015.

YANG, W. *et al.* A fibrinogen-related protein mediates the recognition of various bacteria and haemocyte phagocytosis in oyster *Crassostrea gigas*. **Fish & shellfish immunology**, [s.l.]. v. 114, p. 161–170, 2021.

YOUNG, C. L.; BRITTON, Z. T.; ROBINSON, A. S. Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. **Biotechnology journal**, [s.l.]. v. 7, n. 5, p. 620–634, 2012.

YOUSSEF, D. T. A. Marine-Derived Biomolecules. **Biomolecules**, [s.l.]. v.11, n.1, 2021.

YU, W. W.; WANG, Y. A.; PENG, X. Formation and stability of size-, shape-, and structure-controlled CdTe nanocrystals: ligand effects on monomers and nanocrystals. **Chemistry of Materials**, [s.l.]. v. 15, n. 22, p. 4300-4308, 2003.

ZHANG, C. *et al.* A tandem-repeat galectin-1 from *Apostichopus japonicus* with broad PAMP recognition pattern and antibacterial activity. **Fish & shellfish immunology**, [s.l.]. v. 99, p. 167-175, 2020.

ZHANG, Q. *et al.* Expression and antimicrobial characterization of rhamnose-binding lectin in *Pinctada fucata martensii*. **Fish & shellfish immunology**, [s.l.]. v. 157, p. 110079, 2025.

ZHANG, Y.; YANG, B. In vivo optimizing of intracellular production of heterologous protein in *Pichia pastoris* by fluorescent scanning. **Analytical biochemistry**, [s.l]. v. 357, n. 2, p. 232-239, 2006.

ZHAO, Q. *et al.* A C-type lectin (AjSJL-1) containing DPN/WVD motifs in *Apostichopus japonicus* recognizes multiple microbes. **Invertebrate survival journal**, [s.l]. p. 86–97, 2021.