

Escola Superior de Ensino do Instituto Butantan
Programa de Pós-graduação *Lato Sensu*
Biotecnologia para a Saúde - Vacinas e Biofármacos

Kauan Garcia de Miranda

**Expressão e purificação da proteína nucleocapsídica recombinante de SARS-
CoV-2**

São Paulo
2023

Kauan Garcia de Miranda

**Expressão e purificação da proteína nucleocapsídica recombinante de SARS-
CoV-2**

Trabalho de conclusão de curso apresentado ao Curso de Especialização em Biotecnologia para a Saúde - Vacinas e Biofármacos do Programa de Pós-graduação *Lato Sensu* da Escola Superior do Instituto Butantan como requisito básico para a obtenção do título de Especialista em Biotecnologia para a Saúde - Vacinas e Biofármacos.

Orientador (a): Dra. Giselle Pidde Marques Porto

São Paulo

2023

**Catálogo na Publicação
Instituto Butantan
Dados inseridos pelo(a) aluno(a)**

Miranda, Kauan Garcia de

Expressão e purificação da proteína nucleocapsídica recombinante de SARS-CoV-2 / Kauan Garcia de Miranda ; orientador(a) Giselle Pidde Marques Porto - São Paulo, 2023.

39 p. : il.

Monografia (Especialização) - Escola Superior do Instituto Butantan, Especialização na Área da Saúde - Biotecnologia Para a Saúde - Vacinas e Biofármacos.

1. Proteínas recombinantes 2. Sars-COV-2. 3. Proteínas do nucleocapsídeo de coronavírus. 4. Expressão em E. coli I. Porto, Giselle Pidde Marques. II. Escola Superior do Instituto Butantan. III. Especialização na Área da Saúde - Biotecnologia Para a Saúde - Vacinas e Biofármacos. IV. Título.

AUTORIZAÇÃO PARA ACESSO E REPRODUÇÃO DE TRABALHO

Eu, Kauan Garcia de Miranda, aluno(a) do Curso de Especialização em Biotecnologia para a Saúde - Vacinas e Biofármacos, autorizo a divulgação do meu trabalho de conclusão de curso por mídia impressa, eletrônica ou qualquer outra, assim como a reprodução total deste trabalho de conclusão de curso após publicação, para fins acadêmicos, desde que citada a fonte.

Prazo de liberação da divulgação do trabalho de conclusão de curso após a data da avaliação:

- Imediato
 06 meses
 12 meses
 Outro prazo _____ Justifique:

São Paulo, 20 de Dezembro de 2022

.....
aluno(a)

De acordo:.....
Orientador(a):

Giselle Pardo
Dra. Giselle Pardo
Especialista de Laboratório
Laboratório de Imunoquímica
Fundação Butantan

Para minha mãe e minha avó, por sempre acreditarem em mim, me darem todo o apoio e amor necessários e me incentivarem a buscar aperfeiçoamento naquilo que amo.

Para as minhas amigas Karoline, Nicole e Lavínia por sempre me ouvirem e me acolherem em todas as fases da minha vida em todos esses anos de amizade.
Obrigado!

AGRADECIMENTOS

A minha orientadora Dra. Giselle Pidde por todo o apoio, pela oportunidade e ensinamentos tanto na vida acadêmica como na vida pessoal.

Aos amigos e professores do Instituto Butantan e, principalmente, do laboratório de Imunoquímica pela incrível experiência proporcionada. Em especial para os amigos Nicole, Nathiê, Gabriella e Maurício por todas as horas de conversas e desabafos dentro e fora do Instituto, tornando essa experiência descontraída e divertida, os levarei para sempre comigo em meu coração.

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

Obrigado a todos!

"Dificuldades e obstáculos são fontes valiosas de saúde e força para qualquer sociedade".

Albert Einstein

RESUMO

GARCIA, Kauan. **Expressão e purificação da proteína nucleocapsídica recombinante de SARS-CoV-2**, 2022. 39 p. Trabalho de Conclusão de Curso (Especialização em Biotecnologia para a Saúde - Vacinas e Biofármacos) - Escola Superior do Instituto Butantan, São Paulo, 2023.

Denominada como Pandemia do Covid-19, causada pelo vírus SARS-Cov 2, ficou conhecida como um dos grandes desafios do século XXI e trouxe à indústria e à pesquisa a necessidade de criar parâmetros defensivos e de rápida efetividade para o estudo biológico de possíveis ameaças à saúde da população em geral. A proteína Nucleocapsídica do Sars-COV-2, responsável por estabilizar o nucleocapsídeo, além de estar envolvida na síntese e tradução do RNA, e atuar como antagonista do interferon e supressor de interferência do RNA, favorecendo a replicação viral, se tornou alvo deste estudo com o objetivo de criar parâmetros para a sua produção em laboratório em forma de proteína recombinante. Visando o alto rendimento e boa qualidade, através de sua expressão em bactéria competente *Escherichia coli* BL21(DE3) e a melhora nos passos de sua criação, contando com sonicação das bactérias expressas, sua purificação e concentração proteica. O presente estudo alcançou os objetivos para produção dessa proteína para o Laboratório de Imunoquímica do Instituto Butantan, também sendo capaz de guiar a produção dessa e de outras proteínas recombinantes que acabam se tornando de interesse para estudo e pesquisa em busca da saúde da população.

Palavras-chave: Proteínas recombinantes. Sars-COV-2. Proteínas do nucleocapsídeo de coronavírus. Expressão em *E. coli*.

ABSTRACT

GARCIA, Kauan. **Expression and purification of the recombinant SARS-CoV-2 nucleocapsid protein**, 2022. 39 p. Course Completion Work (Specialization in Biotechnology for Health - Vaccines and Biopharmaceuticals) - Escola Superior do Instituto Butantan, São Paulo, 2023.

Known as the Covid-19 Pandemic, caused by the SARS-Cov 2 virus, it became known as one of the great challenges of the 21st century and brought to industry and research the need to create defensive parameters and of rapid effectiveness for the biological study of possible threats to the health of the general population. The Sars-COV-2 Nucleocapsid protein, responsible for stabilizing the nucleocapsid, in addition to being involved in RNA synthesis and translation, and acting as an interferon antagonist and RNA interference suppressor, favoring viral replication, became the target of this study with the objective of creating parameters for its production in the laboratory in the form of recombinant protein. Aiming at high yield and good quality, through its expression in competent bacteria *Escherichia coli* BL21(DE3) and the improvement in the steps of its creation, counting on sonication of the expressed bacteria, its purification and protein concentration. The present study reached the objectives for the production of this protein for the Laboratório de Imunoquímica of the Instituto Butantan, also being able to guide the production of this and other recombinant proteins that end up becoming of interest for study and research in search of the health of the population.

Keywords: Recombinant proteins. SARS-CoV-2. Coronavirus nucleocapsid proteins. Expression in *E. coli*.

LISTA DE FIGURAS

Figura 1 - Expressão do cDNA codificante para a proteína N do vírus SARS-CoV-2.....	22
Figura 2 - Western Blotting da Sonicação de Padrão 3.....	24
Figura 3 - Cromatograma da purificação proteica utilizando protocolo 3.....	26
Figura 4 - SDS-PAGE das frações obtidas na cromatografia de protocolo 3.....	27
Figura 5 - Western Blotting das frações obtidas na cromatografia de protocolo 3....	28
Figura 6 - SDS-PAGE dos lotes finais da concentração proteica.....	29
Figura 7 - Western Blotting dos lotes finais da concentração proteica.....	30
Figura 8 - Espectrofotometria dos lotes finais da concentração proteica junto a uma amostra do filtro de corte.....	31
Figura 9 - Dados interpolados da leitura de espectrofotometria dos lotes da concentração proteica.....	31
Figura 10 - Dados finais da leitura de espectrofotometria dos lotes da concentração proteica e toda a quantidade de proteína nucleocapsídica obtida.....	32

LISTA DE ABREVIATURAS

BCIP - 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt

cDNA - DNA Complementar

DNA - Ácido Desoxirribonucleico

D.O. - Densidade Óptica

ICTV - International Committee on Taxonomy of Viruses

IPTG – Isopropil β -D-1-tiogalactopiranosídeo

NBT - Nitro-Blue tetrazolium chloride

PCR - *Polymerase Chain Reaction*

PM - Peso Molecular

PN - Proteína Nucleocapsídica

RNA - Ácido Ribonucléico

SARS-CoV-2 - Severe Acute Respiratory Syndrome-Associated Coronavirus 2

SUMÁRIO

1 INTRODUÇÃO	13
2 OBJETIVOS	16
3 MATERIAIS E MÉTODOS	17
3.1 Lista de Reagentes e Soluções	17
3.2 cDNA codificante para a Proteína Recombinante	17
3.3 Expressão Proteica	17
3.3.1 Transformação.....	18
3.3.2 Pré-inóculo.....	18
3.3.3 Inóculo.....	18
3.3.4 Expressão.....	18
3.4 Sonicação	19
3.5 Purificação	19
3.6 Concentração Proteica	20
3.7 Dosagem de Proteínas	20
3.8 Análise de Proteínas	20
3.8.1 Eletroforese em SDS-PAGE.....	21
3.8.2 Western Blotting.....	21
4 RESULTADOS	22
4.1 Expressão	22
4.2 Sonicação	23
4.3 Purificação	25
4.4 Concentração Proteica	29
4.4.1 Análise Qualitativa.....	29
4.4.2 Análise Quantitativa.....	31
5 DISCUSSÃO	33
6 CONCLUSÕES	36
7 DESAFIOS E APRENDIZADOS DESSE PROJETO	37
REFERÊNCIAS	38

1 INTRODUÇÃO

Denominada como Pandemia do Covid-19, causada pelo vírus SARS (*Severe Acute Respiratory Syndrome*), ficou conhecida como um dos grandes desafios do século XXI. Devido ao fato dessa pandemia perdurar até os dias em que este trabalho é escrito e sem previsão de término. Não é possível estimar o dano total causado por essa doença, porém se sabe que são extremos os impactos à economia e a saúde da população mundial, direta e/ou indiretamente (PIRES BRITO *et al.*, 2020).

A COVID-19 se trata de uma doença infectocontagiosa causada pelo vírus SARS-CoV-2 (*Severe Acute Respiratory Syndrome-Associated Coronavirus 2*), da família dos coronavírus, gênero Betacoronavirus, subfamília Coronavirinae, família Coronaviridae (GORBALENYA *et al.*, 2020), que tem como alvo principal o sistema respiratório (PIRES BRITO *et al.*, 2020). No geral, a morfologia dos coronavírus pode ser definida como vírus envelopados, que possuem genoma composto por uma molécula de RNA, fita simples, não segmentada e de polaridade positiva. As partículas virais são ditas como esféricas, com aproximadamente 80-220 nm (BORGES, 2020).

Nos coronavírus, o RNA genômico está associado a múltiplas cópias de nucleoproteínas, formando um nucleocapsídeo helicoidal. O envelope ao redor do nucleocapsídeo é formado por uma bicamada lipídica na qual as proteínas spike (S), envelope (E) e membrana (M) estão ancoradas (HELMY *et al.*, 2020; LI *et al.*, 2020). O trímero da proteína S do coronavírus forma picos na superfície do vírus, que são coroados ao microscópio eletrônico, por isso o nome "coronavírus" (HELMY *et al.*, 2020).

A proteína E é a menor proteína estrutural e, embora seja abundantemente expressa nas células hospedeiras durante a replicação, está presente nos vírions em níveis mais baixos do que outras proteínas estruturais. Embora sua função não tenha sido bem estabelecida, ela o faz desencadeando o processo de montagem e brotamento de partículas virais. A proteína M é a proteína mais abundante do envelope, participa da montagem da partícula viral, determina o local de brotamento do vírus e a forma do envelope viral; sua interação com a proteína N, foco de estudo deste presente trabalho, estabiliza o nucleocapsídeo (BORGES, 2020).

Além de estar associada ao RNA genômico que forma o nucleocapsídeo, a proteína N está envolvida na síntese e tradução do RNA e atua como antagonista do

interferon e supressor de interferência do RNA, o que favorece a replicação viral (BORGES, 2020).

Proteínas são estruturas polipeptídicas de extrema importância como componentes dos seres vivos, sendo utilizadas pelo organismo em praticamente todos os processos celulares, como por exemplo em processos enzimáticos, funções reguladoras, transporte, formação estrutural e sinalização (BILL, 2014). Algumas dessas proteínas são muito frágeis e difíceis de serem estudadas de forma isolada, sendo necessário sua expressão em um organismo vivo para a obtenção de um melhor desempenho de suas funções (BILL *et al.*, 2011). Alberts *et al.* (2002) explica que na biologia o termo “expressão” é universalmente dito como a “produção de um fenótipo observável por um gene – geralmente direcionando a síntese de uma proteína”.

A chegada da tecnologia de DNA recombinante nos anos 70, seguido da introdução da técnica de reação em cadeia polimerase (PCR), foi uma revolução na biologia molecular, principalmente no estudo de proteínas, permitindo uma visão mais aprofundada das funções proteicas que antes não eram possíveis (ROSANO; CECCARELLI, 2014). Foi demonstrado que o DNA eucariótico pode ser expresso em plasmídeos bacterianos, sendo utilizado inicialmente a bactéria *Escherichia coli* (MORROW *et al.*, 1974). Com o passar do tempo essas expressões proteicas puderam ser executadas em outros sistemas hospedeiros de expressão procariótica e eucariótica, pode se citar como exemplo: células de mamíferos, como células do ovário de Hamsters; bactérias, como a *Escherichia coli*; leveduras, como *Saccharomyces cerevisiae* e *Pichia pastoris*; células de insetos e plantas transgênicas (MORROW *et al.*, 1974; ROSANO e CECCARELLI, 2014). Logo foi percebido que as possibilidades de aplicação dessa técnica eram praticamente ilimitadas, com isso foi possível o nascimento da atual indústria biotecnológica (ROSANO e CECCARELLI, 2014).

Apesar da possibilidade do uso de diferentes tipos de hospedeiros para o processo de expressão proteica, a *E. coli* é a mais utilizada mundialmente devido ao seu baixo custo, além de oferecer uma taxa de crescimento acelerado e um alto rendimento (TRIPATHI e SHRIVASTAVA, 2019). Além disso, o vasto conhecimento de sua genética e sua biologia celular, permite uma maior segurança no seu uso. Esta bactéria se provou ser a melhor hospedeira no quesito de expressão de proteínas, devido ao fato de oferecer diversas vantagens sobre os outros tipos de hospedeiros,

principalmente as leveduras, além de seu fácil manuseio e um processo *upstream* simples que possibilita a produção final de uma grande quantidade de proteínas recombinantes, de forma econômica (GUPTA e SHUKLA, 2017). Porém, embora na teoria seja uma técnica com capacidade de grande efetividade, na prática a obtenção de um produto funcional não é tão simples (ROSANO e CECCARELLI, 2014). Algumas proteínas são mais difíceis de serem manipuladas e de serem produzidas em larga escala, principalmente as proteínas recombinantes de origem mamífera ou com múltiplas pontes de dissulfeto (PAPAGIANNI, 2012). Além disso o processo de expressão acaba por perturbar a homeostase celular, resultando no desequilíbrio do fluxo de carbono, drenagem energética e na regulação do metabolismo celular, gerando estresse bacteriano e dificultando o crescimento celular e a síntese de proteínas recombinantes, atrapalhando o rendimento na produção dessas proteínas (OW *et al.*, 2021).

A indústria de produção de proteínas recombinantes expressas em bactérias tem avançado muito nos últimos anos, os itens produzidos por essa indústria incluem fatores de crescimento e hormônios, anticorpos, enzimas e vacinas (ROSANO e CECCARELLI, 2014). Esses avanços tecnológicos mudaram a indústria farmacêutica para melhor, tornando o rendimento de produto final muito maior e o custo muito menor, permitindo assim que fosse possível o tratamento de múltiplas doenças e distúrbios que assolam a humanidade (TRIPATHI e SHRIVASTAVA, 2019).

O estudo envolvendo tecnologia de DNA e proteínas recombinantes é de extrema importância há anos, porém o estado pandêmico atual permitiu que a comunidade científica tivesse a visão da importância do estudo da produção de insumos biológicos de forma autônoma dentro de laboratórios, permitindo que se haja liberdade e agilidade para a pesquisa de novos biofármacos e vacinas para o combate de doenças com potencial a se tornarem ameaças biológicas. A autonomia destes processos biológicos, como a expressão de proteínas recombinantes, em larga escala é um importante alvo para o desenvolvimento científico de qualidade e com rapidez.

2 OBJETIVOS

O objetivo geral do presente estudo visa a produção de insumo biológico para os diversos projetos em andamento no Laboratório de Imunoquímica envolvendo o vírus SARS-CoV-2.

Especificamente, melhorar o rendimento e a qualidade da proteína nucleocapsídica (PN) do vírus SARS-CoV-2, utilizando o sistema de expressão heteróloga em *E. coli*.

3 MATERIAIS E MÉTODOS

3.1 Lista de Reagentes e Soluções

Fosfatase Alcalina (IgG anti-cavalo em coelho, Sigma-Aldrich - EUA); LB Ágar (1% triptona; 1% NaCl; 0,5% extrato de levedura; e 2% Ágar); SOB (2YT 15,5g; KCL (1M) 0,095g; $MgCl_2$ (1M) 0,475g; $MgSO_4$ (1M) 0,6g; e Água MilliQ 500mL); Soro de Cavalo Anti-SARS-CoV-2 (1:2.000 em PBS-BSA 0,1% p/v); Tampão de Equilíbrio (Na_3PO_4 0,02M, NaCl 0,5M, pH 7,8); Tampão de Lavagem (Na_3PO_4 0,02M, NaCl 0,5M, pH 6,0); Tampão de Lise (Tris-HCl 100mM pH 8,0, NaCl 300mM); Tampão PBS (10mM Na_2HPO_4 , 150mM NaCl, pH 7.2).

5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt, (BCIP) (Molecular Probes - EUA); Ágar (Diagnolab - Espanha); Coomassie Brilliant Blue R-250 (Bio-Rad Lab., EUA); Glicose (Merck Millipore - EUA); IgG anti-cavalo em coelho conjugado com fosfatase alcalina (Sigma-Aldrich - EUA); Imidazol 1M (Sigma-Aldrich - EUA); isopropil- β -D-tio galactopiranosídeo (IPTG) (Invitrogen Life Technologies - EUA); Kanamicina (Sigma-Aldrich - EUA); KCL (Merck Millipore - EUA); Meio de cultura 2YT (Invitrogen Life Technologies - EUA); $MgCl_2$ (Sigma-Aldrich - EUA); $MgSO_4$ (Sigma-Aldrich - EUA); Na_2HPO_4 10mM (Sigma-Aldrich - EUA); NaCl (Merck Millipore - EUA); Na_3PO_4 0,02M (Sigma-Aldrich - EUA); Nitro-Blue tetrazolium chloride (NBT) (Molecular Probes - EUA); Triptona (Sigma-Aldrich - EUA); Tris-HCl (Thermo Fisher Scientific - EUA).

3.2 cDNA codificante para a Proteína Recombinante

O cDNA codificante para a proteína recombinante do nucleocapsídeo do SARS-CoV-2 foi sintetizado pela empresa *GenOne*. O desenho determinado pela Dra. Giselle Pidde Marques Porto. O cDNA está clonado no vetor de expressão pET-28a(+), com resistência a Kanamicina (inserto de 1269 pares de bases). Análises *in silico* da proteína N foram realizadas utilizando a plataforma online *ExPASy* (<http://www.expasy.ch>), tendo como características Físico-Químicas Teóricas o ponto isoelétrico de 10.11 e tamanho de 47.789,02 kDa.

3.3 Expressão Proteica

3.3.1 Transformação

A transformação foi feita na bactéria competente BL21(DE3) por choque térmico, utilizando 1µL de cDNA codificante para a proteína N (0,05µg). Incubação inicial em gelo por 30 minutos, seguida de 90 segundos a 42°C em Banho Maria e, novamente, em gelo por 5 minutos.

Adicionado então 1mL de meio SOC composto por 20µL de Glicose 1M a cada 1mL de Meio SOB autoclavado. Em seguida, as bactérias foram deixadas a 37°C por uma hora e, posteriormente, em meio sólido contendo LB Ágar com Kanamicina (30µg/mL), por 18 horas.

3.3.2 Pré-inóculo

As colônias isoladas foram selecionadas aleatoriamente e inoculadas em 3mL de meio 2YT com Kanamicina (20µg/mL). Foi realizada uma varredura para verificar quais as melhores colônias e condições ideais para a expressão da proteína recombinante. Para isso, foi utilizado diferentes concentrações de isopropil-β-D-tiogalactopiranosídeo (IPTG), bem como, diferentes tempos de induções, a 37°C em agitação. As culturas foram centrifugadas a 3500rpm por 20 minutos a 4°C. As amostras foram analisadas por eletroforese em gel de poliacrilamida (12,5%) na presença de SDS-PAGE, em condições redutoras (descrito no item 3.8).

3.3.3 Inóculo

Após a análise do pré-inóculo e definição das condições ideais de expressão, a(s) colônia(s) de interesse foi inoculada em 100mL de meio 2YT com Kanamicina (20µg/mL). Após atingir crescimento ideal $DO_{600nm} = 0,6/0,8$, a expressão das proteínas foi induzida pela adição de IPTG.

3.3.4 Expressão

Quando atingido a D.O. correta, foi adicionado concentrações crescentes de IPTG ao meio de cultura com duas diferentes horas de incubação, sendo 0,5mM e 1mM em 2 horas e em seguida em 4 horas, o melhor resultado se deu pelo uso de 1mM em 2 horas de incubação. Após incubação a 37°C em agitação constante, amostras de culturas de bactérias foram coletadas ao longo da expressão (tempo 0, 2 horas e 4 horas). A suspensão de bactérias foi submetida a centrifugação por 20 minutos a 3500rpm a 4°C, sendo concentrado 100mL de cultura bacteriana por tubo (50mL) e o *pellet* resultante de bactérias armazenado a -80°C.

3.4 Sonicação

Diferentes testes foram realizados, com configuração para alcance do melhor rendimento da proteína de interesse, sem perda de qualidade proteica, variando caracteres de amplitude, tempo ativo, tempo inativo e quantidade de ciclos. Durante todo o procedimento a amostra em processo foi mantida em gelo.

Os *pellets* resultantes (100mL de cultura bacteriana expressa por tubo) foram ressuspendidos em 10mL de Tampão de lise, sendo, posteriormente lisados por sonicação (*Vibra Cell - High Intensity Liquid Processors* da *Sonics & Materials, Inc*).

Protocolo 1: 5 ciclos de 40 segundos ativo e 1 minuto de descanso em gelo, a 70% de amplitude; Protocolo 2: 5 ciclos de 1 minuto ativo e 1 minuto de descanso em gelo, a 70% de amplitude; Protocolo 3: 10 ciclos de 1 minuto ativo e 1 minuto de descanso em gelo, a 100% de amplitude; Protocolo 4: 10 ciclos de 40 e 5 segundos ativo e 1 minuto de descanso em gelo, a 100% de amplitude.

As proteínas solúveis e insolúveis foram separadas por centrifugação a 4.000rpm por 20 minutos e a 4°C. A determinação da presença das proteínas recombinantes nas frações solúvel (sobrenadante) ou insolúvel (*pellet*) foi feita por eletroforese em SDS-PAGE e confirmada por *Western Blotting*, conforme descrito no item 3.8.

3.5 Purificação

Para obtenção da proteína recombinante de interesse de forma isolada, o sobrenadante do lisado bacteriano foi submetido à cromatografia de afinidade em

coluna de Ni⁺⁺ acoplado a sepharose, em equipamento *Äkta Start (GE)*, sendo a proteína recombinante eluída e submetida à análise por SDS-PAGE e *Western Blotting* (descrito no item 3.8).

Foram realizados diferentes protocolos de purificação para o melhor rendimento final da proteína sem contaminação bacteriana. Resumidamente, 9mL de sobrenadante do lisado bacteriano foi submetido a cromatografia em sistema FPLC, sob fluxo constante de 2,0mL/min, utilizando coluna HisTrapTMFF (Cytiva). Inicialmente, equilibrada com 5 volumes do tampão de equilíbrio, seguida de 10 volumes de tampão de lavagem. O material adsorvido foi eluído, em frações de 1mL, com o tampão de eluição contendo concentrações crescentes de imidazol 1M.

Protocolo 1: Após equilíbrio da coluna, a etapa de lavagem foi realizada com 3%B, seguida de eluição com as seguintes etapas: 10%B em 1,5 volumes; 30%B em 1,5 volumes; 50%B em 1,5 volumes; 75%B em 1,3 volumes; 100%B em 2 volumes.

Protocolo 2: Após equilíbrio da coluna, a etapa de lavagem foi realizada com 30%B, seguida de eluição com as seguintes etapas: 50%B em 2 volumes; 75%B em 1,3 volumes; 100%B em 2 volumes.

Protocolo 3: Após equilíbrio da coluna, a etapa de lavagem foi realizada com 10%B, seguida de eluição com as seguintes etapas: 30%B em 2 volumes; 50%B em 1,5 volumes; 100%B em 2 volumes.

3.6 Concentração Proteica

As amostras foram lavadas e concentradas utilizando o kit *Amicon Pro Affinity Concentrator* (Amicon, Alemanha) (com membrana de corte de 10kDa), utilizando as recomendações do fabricante. O conteúdo final foi de 4mL, acondicionados a -80°C.

3.7 Dosagem de Proteínas

A determinação da concentração das amostras foi feita utilizando o kit *BCA Protein Assay (PierceTM)* (Sigma-Aldrich, EUA), seguindo as recomendações do fabricante, a leitura feita com FLUOstar Omega (BMG LABTECH, Alemanha) e as análises no programa *GraphPad PRISM* versão 7.

3.8 Análise de Proteínas

3.8.1 Eletroforese em SDS-PAGE

A eletroforese em gel de poliacrilamida (12,5%) e dodecil-sulfato de sódio (SDS-PAGE) em condições redutoras foi realizada para identificar a proteína recombinante nucleocapsídica do SARS-CoV-2. As amostras foram aquecidas a 96°C por 10 minutos e aplicadas no gel sob corrente de 100 volts por aproximadamente 3 horas, em tampão Tris-Glicina, os géis foram corados em solução de *Coomassie Brilliant Blue R-250* (Bio-Rad Lab., EUA) ou pelo método de impregnação pela prata (MORRISSEY, 1981).

3.8.2 Western Blotting

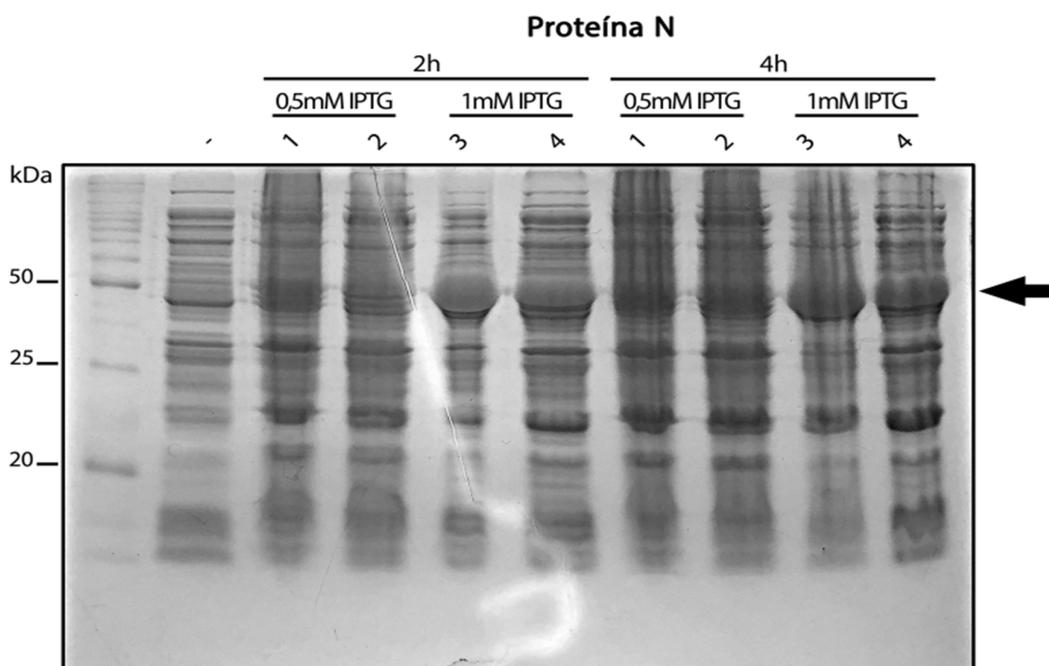
Para o Western blotting, foi realizada uma eletroforese das amostras em SDS-PAGE sem corar o gel e eletro transferidas para membrana de nitrocelulose. A transferência foi realizada durante 18 horas em câmara fria e corrente elétrica constante de 150 Miliamperes. Após a transferência, a membrana foi então bloqueada com tampão PBS-BSA 5% p/v, por 2 horas a 37°C. A membrana foi lavada com tampão PBS e em seguida incubada com o soro de cavalo anti-SARS-CoV-2 por 1 hora em temperatura ambiente. A membrana foi então lavada 3 vezes com tampão PBS-Tween 0.05% p/v e incubada com o anticorpo secundário anti-IgG de cavalo conjugado com fosfatase alcalina, diluído na proporção 1:10.000 em solução PBS-T/BSA 1% p/v. Após o período de incubação de 1 hora à temperatura ambiente. As reações foram reveladas com o substrato NBT/BCIP (Molecular Probes - EUA) de acordo com as instruções do fabricante. A reação foi interrompida com água destilada e, em seguida, as membranas foram secas à temperatura ambiente e, posteriormente, analisadas.

4 RESULTADOS

4.1 Expressão

O cDNA codificante para a proteína nucleocapsídica do vírus SARS-CoV-2, clonado no vetor de expressão pET28a(+), tem a construção final que codifica para a proteína em fusão com resíduos de histidina e um sítio de clivagem para enterokinase, resultando em um tamanho de proteína de, aproximadamente, 50kDa. A expressão da proteína recombinante foi induzida na linhagem *E. coli* BL21(DE3), utilizando diferentes concentrações de IPTG e diferentes tempos de incubação. Os ensaios levaram a utilização otimizada de 0,5mM ou 1mM de IPTG, com os tempos de indução de 2h ou 4h (Figura 1).

Figura 1 - Expressão do cDNA codificante para a proteína N do vírus SARS-CoV-2.



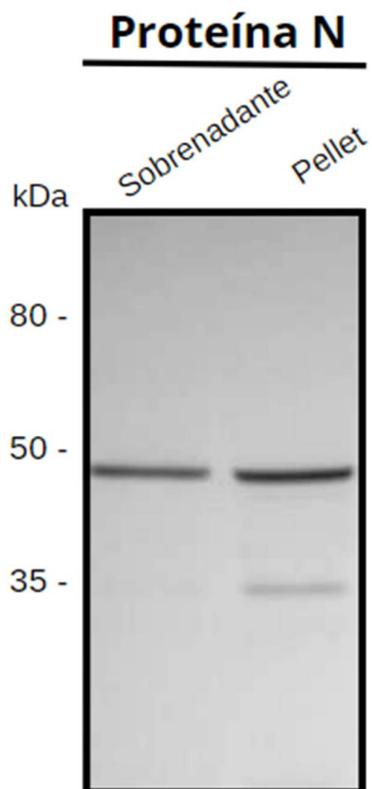
Fonte: Dra. Giselle Pidde, 2022.

Legenda: Expressão em linhagem bacteriana BL21(DE3). Amostras de 30µL dos extratos de bactérias foram submetidas à análise em gel de SDS-PAGE (12,5 %), em condições redutoras, corado por coomassie blue. (-) Sem indução por IPTG. 4 colônias isoladas aleatoriamente e a expressão induzida pela adição de IPTG (0,5mM ou 1mM) por um período de 2 horas ou 4 horas.

4.2 Sonicação

O sedimento bacteriano, obtido como descrito anteriormente (item 4.1), foi solubilizado em tampão de lise. As suspensões bacterianas foram lisadas por processo de sonicação e as amostras resultantes avaliadas quanto a solubilidade (sobrenadante e *pellet*). Para isso, diversos protocolos foram utilizados, sendo o mais eficiente apresentado na figura (Figura 2). O padrão 3 (1 minuto ativo, 1 minuto de descanso no gelo, por 10 ciclos a 100% de amplitude) foi o que gerou melhores resultados, permitindo a liberação da PN em maior parte no sobrenadante do que no pellet, mesmo que ainda tenha sobrado PN no pellet.

Figura 2 - Western Blotting da Sonicação de Padrão 3.



Fonte: O próprio autor, 2022.

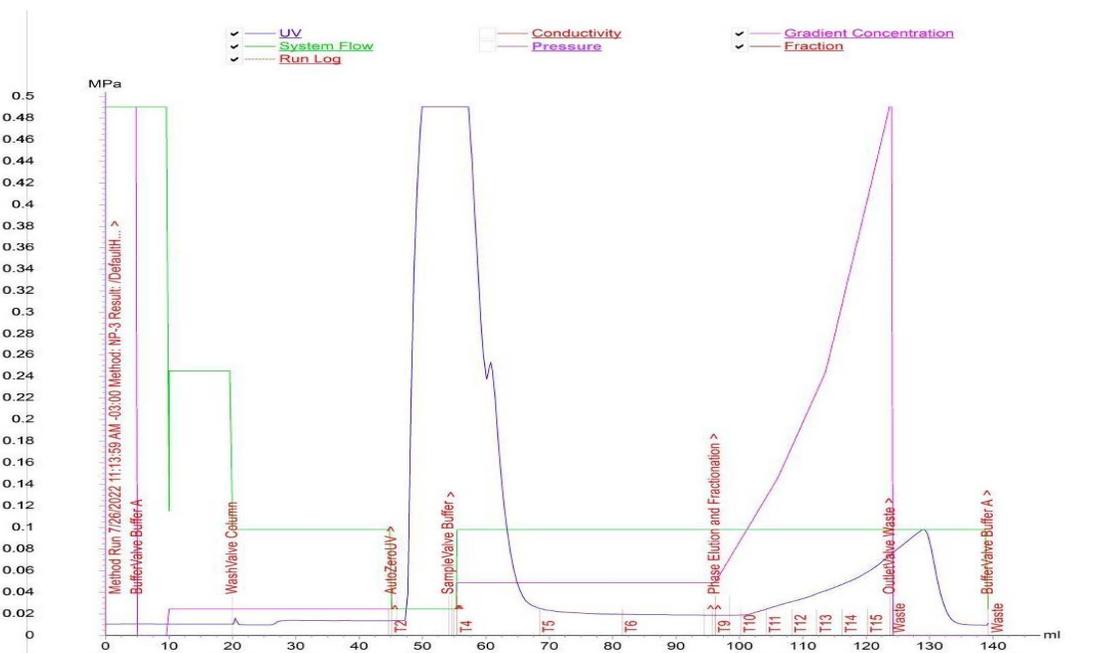
Legenda: Utilizando amostra de sobrenadante e Pellet da sonicação de padrão descrita no item 3.3, o sobrenadante é recolhido depois da centrifugação da amostra sonicada a 4000rpm por 3 minutos a 4°C. Utilizando processo de Western Blotting descrito no item 3.8.

4.3 Purificação

O sobrenadante do lisado bacteriano descrito no item anterior (item 4.2) foi submetido à cromatografia de afinidade em coluna de Ni⁺⁺ acoplado a sepharose, sendo a proteína recombinante eluída e submetida à análise por SDS-PAGE e *Western Blotting* (descrito no item 3.8) para obtenção da PN de forma isolada.

Nesse processo, 3 protocolos foram utilizados, sendo aquele que gerou melhores resultados (Protocolo 3: 9mL de sobrenadante do lisado bacteriano submetido a cromatografia em sistema FPLC, sob fluxo constante de 2,0mL/min, utilizando coluna HisTrapTMFF (Cytiva). Inicialmente, equilibrada com 5 volumes do tampão de equilíbrio, seguida de 10 volumes de tampão de lavagem. Após equilíbrio da coluna, a etapa de lavagem foi realizada com 10%B, O material adsorvido foi eluído, em frações de 1mL, com o tampão de eluição contendo concentrações crescentes de imidazol 1M, sendo 30%B em 2 volumes; 50%B em 1,5 volumes; 100%B em 2 volumes) apresentado na figura (Figura 3).

Figura 3 - Cromatograma da purificação proteica utilizando protocolo 3.

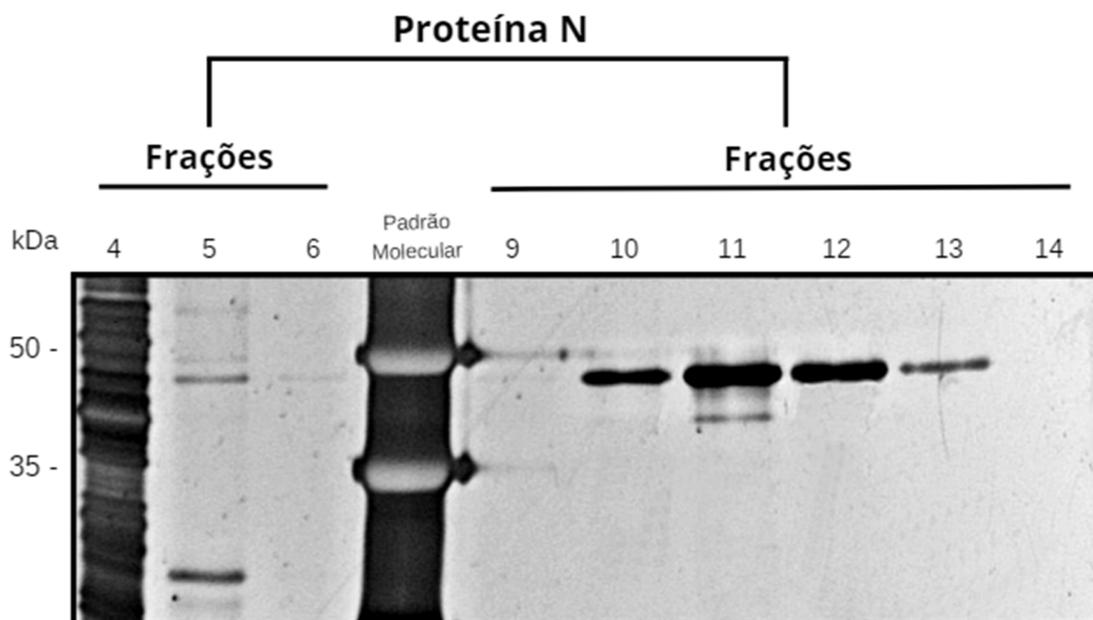


Fonte: O próprio autor, 2022.

Legenda: Cromatografia de afinidade em coluna de Ni⁺⁺ acoplado a sepharose em equipamento Äkta Start - GE. Protocolo utilizando 9 mL do sobrenadante do lisado bacteriano, com a obtenção de 15 frações.

A cromatografia gerou 15 frações que foram submetidas a análise em gel de SDS-PAGE (12,5%) (Figura 4) e *Western Blotting* (Figura 5) conforme descrito no item 3.8.

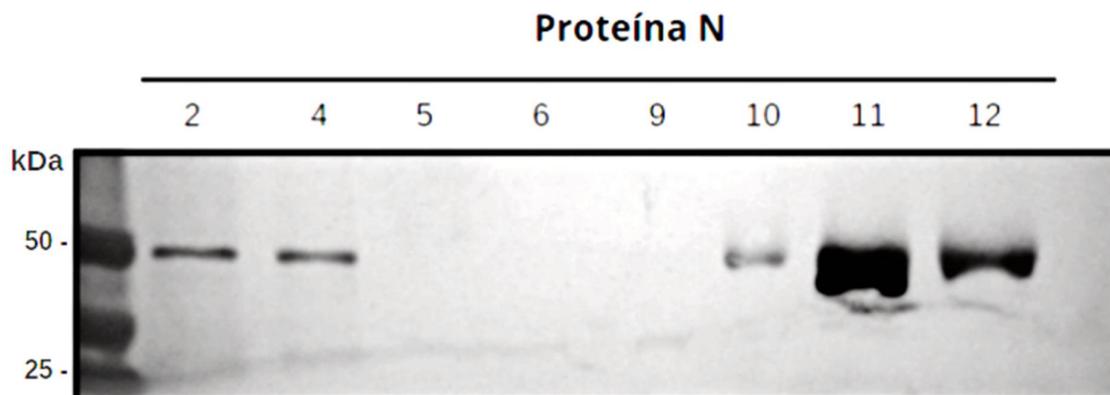
Figura 4 - SDS-PAGE das frações obtidas na cromatografia de protocolo 3.



Fonte: O próprio autor, 2022.

Legenda: Análise das frações recolhidas da cromatografia, as frações 1, 2, 3, 7, 8 e 15 foram consideradas frações de intervalo, contendo pouca amostra e apenas tampão em sua composição, por isso foram descartadas. Análise feita através de SDS-PAGE corado por Impregnação por Prata conforme descrito no item 3.8.

Figura 5 - Western Blotting das frações obtidas na cromatografia de protocolo 3.



Fonte: O próprio autor, 2022.

Legenda: Análise das frações recolhidas da cromatografia, as frações 1, 2, 3, 7, 8 e 15 foram consideradas frações de intervalo, contendo pouca amostra e apenas tampão em sua composição, por isso foram descartadas. Análise feita através de Western Blotting conforme descrito no item 3.8.

Após as análises, foi definido que as frações 10, 11, 12 e 13 eram as que continham PN de forma pura, sendo elas então as utilizadas nos lotes finais. Cada fração continha em média de 4mL, totalizando 16mL por cromatografia.

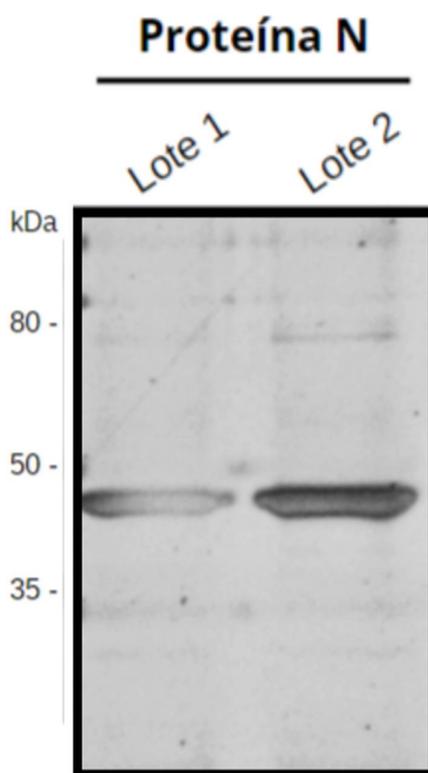
4.4 Concentração Proteica

O produto de 8 cromatografias, mais ou menos 128mL, foi lavado e concentrado utilizando membrana de corte de 10kDa, utilizando as recomendações do fabricante. O conteúdo final gerou dois lotes que foram aliquotados e acondicionados a -80°C.

4.4.1 Análise Qualitativa

Na análise qualitativa foi feito gel de SDS-PAGE (12,5%) (Figura 6) e *Western Blotting* (Figura 7), conforme descrito no item 3.8.

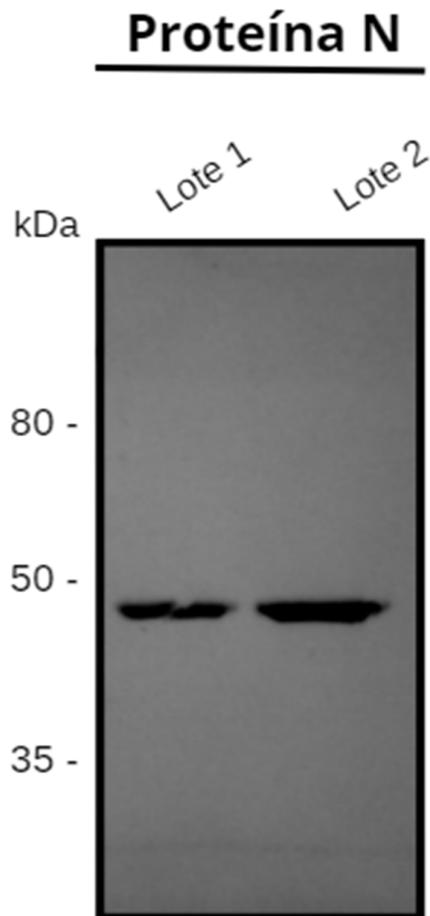
Figura 6 - SDS-PAGE dos lotes finais da concentração proteica.



Fonte: O próprio autor, 2022.

Legenda: Análise de SDS-Page dos lotes, com impregnação por Prata (item 3.8).

Figura 7 - *Western Blotting* dos lotes finais da concentração proteica.



Fonte: O próprio autor, 2022.

Legenda: Análise de Western Blotting feita conforme item 3.8.

4.4.2 Análise Quantitativa

Foi feito uma leitura de espectrofotometria, utilizando curva padrão do *kit Pierce™ BCA Protein Assay* (conforme item 3.7), 1 amostra de cada lote e 1 amostra do filtro de corte para verificar se não houve perda de proteína durante o processo de concentração, como se pode analisar na figura 8.

Figura 8 - Espectrofotometria dos lotes finais da concentração proteica junto a uma amostra do filtro de corte.

	1	2	3	4
A	0,67	0,66	0,116	0,1
B	0,414	0,38		
C	0,36	0,394	0,155	0,174
D	0,289	0,28	0,18	0,188
E	0,232	0,242		
F	0,156	0,173	0,125	0,116
G	0,138	0,138		
H	0,118	0,115		

Fonte: O próprio autor, 2022.

Legenda: Utilizado kit Pierce™ BCA Protein Assay (1 e 2 de A a H e A3 e 4), amostras dos lotes finais 1 (C3 e C4) e 2 (D3 e D4) e análise do filtro utilizado (F3 e F4) na concentração para avaliar perda proteica.

Os dados dessa leitura de espectrofotometria foram analisados no programa *GraphPad PRISM* versão 7, os dados dos lotes foram interpolados (Figura. 9) para que fosse possível a identificação da quantidade proteica presente em cada amostra e a quantidade final do lote (Figura 10) em parâmetros de $\mu\text{L}/\text{mL}$.

Figura 9 - Dados interpolados da leitura de espectrofotometria dos lotes da concentração proteica.

	X - $\mu\text{g}/\text{mL}$
Lote 1	240,80
Lote 2	317,92

Fonte: O próprio autor, 2022.

Legenda: Análise dos dados obtidos com a leitura de espectrofotometria dos lotes da concentração proteica, interpolados para avaliar quantidade em $\mu\text{g}/\text{mL}$ (X) por amostra, utilizando o programa *GraphPad PRISM* versão 7.

Figura 10 - Dados finais da leitura de espectrofotometria dos lotes da concentração proteica e toda a quantidade de proteína nucleocapsídica obtida.

Quantidade total por lote		
	Por tubo	Por lote
Lote 1	240,80	722,4
Lote 2	476,88	1430,64
Total	2153,04	

Fonte: O próprio autor, 2022.

Legenda: Cálculo feito considerando 3 alíquotas por lote sendo 1mL por tudo do lote 1 e 1,5mL por tudo do lote 2, então apresentado a quantidade de proteína nucleocapsídeo por tubo, por lote e o total obtido com toda a pesquisa feita neste trabalho.

5 DISCUSSÃO

O cenário pandêmico atual mostrou à comunidade científica a importância do estudo da produção de insumos biológicos em laboratórios, de forma rápida e eficaz, para que seja possível o estudo de agentes biológicos com potencial para se tornarem ameaças à saúde da população, assim se tornando possível a rápida produção de biofármacos e vacinas para o combate a doenças. Este trabalho teve como norteamento a estruturação de um protocolo adequado para a produção de proteína nucleocapsídica do SARS-CoV-2, expressa de forma heteróloga em bactéria *Escherichia coli*, visando uma boa qualidade e rendimento, para utilização em projetos em andamento no Laboratório de Imunoquímica do Instituto Butantan.

Para início de produção foi necessário a escolha do hospedeiro ideal para a expressão da PN. Foi decidido que esse processo ocorreria em um sistema heterólogo, mais especificamente em bactéria *Escherichia coli*. Sabe-se que a *E. coli* é considerada a forma mais econômica de expressão proteica atual (TRIPATHI e SHRIVASTAVA, 2019), além de ter um alto nível de proteína recombinante produzida e a rapidez nesse processo (PEREIRA, 2009). Para o desenho da proteína foi utilizado o vetor pET-28a(+), LI *et al.* (2022) cita que os vetores pET são a primeira escolha em muitos trabalhos para que seja feita clonagem e expressão proteica, isso se dá por sua alta capacidade de expressão, além de possuírem um promotor T7 potente. A expressão da PN foi feita na cepa BL21(DE3), caracterizada por ter aspectos que a fazem ser uma ótima escolha para expressão de proteínas recombinantes, como por exemplo por possuir defeitos na capacidade de degradação de proteínas estranhas no citoplasma e extracelulares (ROSANO e CECCARELLI, 2014; LI *et al.* 2022). A combinação desse vetor com essa cepa é o sistema de expressão mais utilizado pela indústria e na produção de biofármacos (HAYAT *et al.*, 2018). Para o controle da indução por IPTG foi feito testes, assim pudemos ter noção da quantidade de IPTG e de tempo que seria necessário para produção de uma quantidade satisfatória de proteína recombinante, novamente tendo em vista a necessidade de um processo que fosse de baixo custo e tempo investido.

Para que fosse possível ser realizada a purificação desta proteína, antes foi necessário submeter as culturas bacterianas contendo a expressão da PN a um processo de lise celular, o método utilizado foi o de lise física através de um processador ultrassônico (sonicador) que utiliza ondas de pressão sônica para criar

microbolhas que quebram a parede celular. Durante os testes de protocolos, o que mostrou melhor resultados foi o protocolo 3 que utiliza 1 minuto ativo, 1 minuto de descanso no gelo, por 10 ciclos a 100% de amplitude, na figura (Figura 2) é possível perceber a presença de PN ainda no pellet após o processo, o aumento de ciclos ou tempo ativo poderia gerar desnaturação proteica fazendo com que se perdesse PN, por isso uma possível solução seria uma segunda ressuspensão do pellet seguido por uma nova sonicação seguindo os mesmos padrões, isso aumentaria a quantidade de PN obtida, mas aumentaria o tempo de sua produção e o custo envolvido, além da possibilidade de não se conseguir uma quantidade considerável com esse processo, por esse motivo o processo de sonicação foi mantido com apenas uma sonicação no protocolo apresentado anteriormente.

Na etapa seguinte era necessário a obtenção da PN de forma pura, ou seja, sem a presença da degradação das estruturas bacterianas devido a sonicação, para isso foi utilizado o processo de cromatografia por afinidade a coluna de níquel (Ni⁺⁺), esse íon metálico é geralmente utilizado para purificação de proteínas que contém histidina em sua composição, nesse caso a PN foi desenhada contendo uma cauda de histidina facilitando a purificação da proteína alvo, essa técnica é utilizada frequentemente já que a cauda é pouco imunogênica, não carregada em pH fisiológico e raramente interfere na estrutura ou função da proteína alvo (BRESOLIN, MIRANDA e BUENO, 2009). Além disso, esse íon interage com o nitrogênio aromático do grupamento imidazol (BRESOLIN, MIRANDA e BUENO, 2009), presente no tampão B utilizado nesse processo.

Durante os testes com os protocolos de purificação, o que demonstrou ter o melhor desempenho gerando uma quantidade boa de proteínas puras finais foi o Protocolo 3 (9mL de sobrenadante do lisado bacteriano submetido a cromatografia em sistema FPLC, sob fluxo constante de 2,0mL/min, utilizando coluna HisTrapTMFF (Cytiva). Inicialmente, equilibrada com 5 volumes do tampão de equilíbrio, seguida de 10 volumes de tampão de lavagem. Após equilíbrio da coluna, a etapa de lavagem foi realizada com 10%B, O material adsorvido foi eluído, em frações de 1mL, com o tampão de eluição contendo concentrações crescentes de imidazol 1M, sendo 30%B em 2 volumes; 50%B em 1,5 volumes; 100%B em 2 volumes). É possível observar (Figura 4 e Figura 5) que há presença da PN nas amostras que foram descartadas, consideradas frações de lavagem, isso poderia ser resolvido alterando os níveis de imidazol no processo de lavagem, porém não foi executado neste trabalho devido ao

fato de que a quantidade de PN perdida não era consideravelmente grande para novas alterações no protocolo.

Analisando as quantidades finais obtidas pelas análises quantitativas e qualitativas, foi possível definir como quantidades aceitáveis para as demandas do Laboratório de Imunoquímica do Instituto Butantan, sendo então definido o processo executado com esse trabalho por satisfazer a necessidade imposta. Vale ressaltar que ainda seria possível fazer alterações nos procedimentos aqui presentes para melhorar ainda mais a quantidade de PN produzida, além de que os protocolos aqui presentes podem servir de base para expressão de outras proteínas recombinantes de interesse que se tornem necessárias.

6 CONCLUSÕES

Bons resultados foram alcançados, podendo manter os padrões aqui estabelecidos para a produção da proteína recombinante do Nucleocapsídeo (N) do Sars-COV-2. Conseguindo reproduzir os resultados aqui presentes é possível obter boas quantidades de lotes dessa proteína em sua forma purificada, de maneira rápida, efetiva e barata. Ainda, pode-se concluir que as etapas aqui contidas podem ser replicadas para a produção de outras proteínas recombinantes que venham a ter necessidade de uso.

7 DESAFIOS E APRENDIZADOS DESSE PROJETO

Com a produção desse projeto e os ensinamentos obtidos durante sua realização, foi possível acessar conhecimento sobre a produção de proteínas recombinantes e os processos biotecnológicos envolvidos, obtendo aprendizado desde os mais básicos como a produção de gel SDS-Page, Western Blotting, colorações, produção de químicos para utilização em laboratório, entre outros testes, até processos mais complexos como a cultura bacteriana, a purificação, a identificação de aspectos presentes na amostra para resolução de sua qualidade, entre outros. Se faz necessário o conhecimento da importância de projetos como esse para o avanço tecnológico da comunidade científica e da saúde da população.

REFERÊNCIAS¹

ALBERTS, B. *et al.* **Molecular biology of the cell**. 4. ed. [s.l.] Garland Science, 2002. ISBN: 9780815340720

BILL, R. M. *Playing catch-up with Escherichia coli: using yeast to increase success rates in recombinant protein production experiments*. **Frontiers in Microbiology**, v. 5, 5 mar. 2014. <https://doi.org/10.3389/fmicb.2014.00085>

BILL, R. M. *et al.* *Overcoming barriers to membrane protein structure determination*. **Nature Biotechnology**, v. 29, n. 4, p. 335–340, abr. 2011. <https://doi.org/10.1038/nbt.1833>

BORGES, A. A. **SARS-CoV-2: origem, estrutura, morfogênese e transmissão. Construção de conhecimento no curso da pandemia de COVID-19: aspectos biomédicos, clínico-assistenciais, epidemiológicos e sociais**, 2020.

BRESOLIN, I. T. L.; MIRANDA, E. A.; BUENO, S. M. A. Cromatografia de afinidade por íons metálicos imobilizados (IMAC) de biomoléculas: aspectos fundamentais e aplicações tecnológicas. **Química Nova**, v. 32, n. 5, p. 1288–1296, 2009.

GORBALENYA, A. E. *et al.* **Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group**. 11 fev. 2020.

GUPTA, S. K.; SHUKLA, P. Sophisticated Cloning, Fermentation, and Purification Technologies for an Enhanced Therapeutic Protein Production: A Review. **Frontiers in Pharmacology**, v. 8, 4 jul. 2017.

HAYAT, S. M. G. *et al.* Recombinant Protein Expression in Escherichia coli (E.coli): What We Need to Know. **Current Pharmaceutical Design**, v. 24, n. 6, p. 718–725, 10 maio 2018.

HELMY, Y. A. *et al.* The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. **Journal of Clinical Medicine**, v. 9, n. 4, p. 1225, 24 abr. 2020.

LI, L. *et al.* Expression and purification of soluble recombinant β -lactamases using Escherichia coli as expression host and pET-28a as cloning vector. **Microbial Cell Factories**, v. 21, n. 1, 23 nov. 2022.

LI, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. **The New England journal of medicine**, v. 382, n. 13, p. 1199–1207, 2020.

¹ ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: Informação e documentação - referências - elaboração. Rio de Janeiro: ABNT, 2018.

MORRISSEY, J. H. *Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity*. **Analytical Biochemistry**, v. 117, n. 2, p. 307–310, nov. 1981.

MORROW, J. F. *et al.* *Replication and Transcription of Eukaryotic DNA in Escherichia coli*. **Proceedings of the National Academy of Sciences**, v. 71, n. 5, p. 1743–1747, maio, 1974.

OW, D. S.-W. *et al.* Editorial: *Technological Advances Improving Recombinant Protein Production in Bacteria*. **Frontiers in Microbiology**, v. 12, 26 jul. 2021.

PAPAGIANNI, M. METABOLIC ENGINEERING OF LACTIC ACID BACTERIA FOR THE PRODUCTION OF INDUSTRIALLY IMPORTANT COMPOUNDS. **Computational and Structural Biotechnology Journal**, v. 3, n. 4, p. e201210003, out. 2012.

PEREIRA, L. INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES
Autarquia associada à Universidade de São Paulo **CLONAGEM, EXPRESSÃO,
PURIFICAÇÃO E CARACTERIZAÇÃO ESTRUTURAL DA PROTEÍNA
RIBOSSOMAL L10 HUMANA RECOMBINANTE**. [s.l.: s.n.]. Disponível em:
<<https://www.teses.usp.br/teses/disponiveis/85/85131/tde-22092011-101810/publico/2009PereiraClonagem.pdf>>.

PIRES BRITO, S. B. *et al.* Pandemia da COVID-19: o maior desafio do século XXI. **Vigilância Sanitária em Debate**, v. 8, n. 2, p. 54–63, 29 maio 2020.

ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. **Frontiers in Microbiology**, v. 5, n. 172, 17 abr. 2014.

TRIPATHI, N. K.; SHRIVASTAVA, A. Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. **Frontiers in Bioengineering and Biotechnology**, v. 7, n. 420, 20 dez. 2019. [https://doi: 10.3389/fbioe.2019.004](https://doi.org/10.3389/fbioe.2019.004)