



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE CIÊNCIAS FARMACÊUTICAS

PEDRO RAFAEL FIRMINO DIAS

Regulação da afinidade por interatores e localização celular da NEK6 pela serina 37 presente no N-terminal desordenado

CAMPINAS

2024

PEDRO RAFAEL FIRMINO DIAS

Regulação da afinidade por interatores e localização celular da NEK6 pela serina 37 presente no N-terminal desordenado

Tese apresentada à Faculdade de Ciências Farmacêuticas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Ciências, na área de Ciências Farmacêuticas - Insumos Farmacêuticos Naturais, Biotecnológicos e Sintéticos

Orientador: Prof. Dr. Jörg Kobarg

ESTE TRABALHO CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO PEDRO RAFAEL FIRMINO DIAS, ORIENTADA PELO PROF DR JÖRG KOBARG.

CAMPINAS

2024

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca da Faculdade de Ciências Médicas
Maristella Soares dos Santos - CRB 8402

D543r Dias, Pedro Rafael Firmino, 1991-
Regulação da afinidade por interatores e localização celular da NEK6 pela serina 37 presente no N-terminal desordenado / Pedro Rafael Firmino Dias. – Campinas, SP : [s.n.], 2024.

Orientador(es): Jörg Kobarg.
Tese (doutorado) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Ciências Farmacêuticas.

1. Mitose. 2. Ciclo celular. 3. Quinase. 4. Nek. I. Kobarg, Jörg, 1965-. II. Universidade Estadual de Campinas (UNICAMP). Faculdade de Ciências Farmacêuticas. III. Título.

Informações complementares

Título em outro idioma: Regulation of affinity for interactors and cellular localization of NEK6 by serine 37 present in the disordered N-terminus

Palavras-chave em inglês:

Mitosis

Cell cycle

Kinase

Nek

Área de concentração: Ciências Farmacêuticas: insumos farmacêuticos naturais, biotecnológicos e sintéticos

Titulação: Doutor em Ciências

Banca examinadora:

Jörg Kobarg [Orientador]

Déborah Schechtman

Nadja Cristhina de Souza Pinto

Wanda Pereira Almeida

Alessandra Sussulini

Data de defesa: 21-11-2024

Programa de Pós-Graduação: Ciências Farmacêuticas

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-1431-2617>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3419124985317566>



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE CIÊNCIAS FARMACÊUTICAS

Autor Pedro Rafael Firmino Dias

Orientador Prof. Dr. Jörg Kobarg

Tese aprovada em 21 de novembro de 2024

Comissão Examinadora

Prof. Dr. Jörg Kobarg

Prof^a Dr^a Déborah Schechtman

Prof^a Dr^a Nadja Cristhina de Souza Pinto

Prof^a Dr^a Alessandra Sussulini

Prof^a Dr^a Wanda Pereira Almeida

A ata de defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação da Faculdade de Ciências Farmacêuticas.

Campinas, 21 de novembro de 2024.

AGRADECIMENTOS

Este trabalho é um dos frutos dos meus esforços e dos esforços de meus pais, Célia Ap. de C. F. Dias e Nivaldo F. Dias. Com eles tive a oportunidade de ser atleta da seleção brasileira de Kung-fu/Wushu em diversas categorias (juvenil, adulto e principal), participar e ter bons resultados em campeonatos nacionais e internacionais, e tudo isso me possibilitou a compreender que nossa busca deve ser pelo autoconhecimento, autodesenvolvimento e fraternidade. O esporte é uma ótima ferramenta para isso, pois aprendemos que o nosso desenvolvimento não está acima do desenvolvimento do próximo e que devemos respeitar e ser fraternos em relação ao desenvolvimento e autoconhecimento do próximo. Como toda escolha, para iniciar na pesquisa foram necessárias diversas renúncias e sem o apoio familiar nada seria possível. O apoio é a melhor forma de dizer que amamos alguém sem querer ter o controle ou modificar essa pessoa. Essa é uma das lições mais preciosas que tive o privilégio de poder aprender com meus pais. Eu amo vocês.

Agradeço a minha nova família que tive o privilégio de construir com a minha esposa, Carolina C. Pereira. Ela é a pessoa que me conhece nos meus melhores e o piores momentos, e escolhe estar ao meu lado. Sua companhia e apoio foram importantes nesta etapa. Sou grato pelo que estamos construindo, e grato aos nossos filhos, Theodoro C. P. Dias e Matheus C. Pereira. Vocês são muito importantes para mim, eu amo vocês.

Durante o mestrado e doutorado tive a oportunidade de aprender e exercer um pouco mais a humanidade. Nós somos seres humanos, filhos, pais, avós, marido ou mulher e amigos antes de sermos alunos, professores, cientistas ou pesquisadores. Nosso desenvolvimento não é mais importante do que o desenvolvimento de nossos amigos, e ajudar o próximo sem querer algo em troca, apenas o bem estar e desenvolvimento do próximo, e vibrar com as conquistas de nossos amigos é o que nos torna humanos.

Agradeço aos meus professores de mestrado e doutorado, Profa. Dra. Denise Vaz de Macedo (LABEX-IB/UNICAMP) e Prof. Dr. Jörg Kobarg (LMS-FCF/UNICAMP). Vocês foram responsáveis por abrir as portas para um aluno que veio da rede municipal e estadual de ensino básico e médio, que se formou em educação física e que não tinha conhecimentos dos aspectos básicos e experimentais da bioquímica e biologia molecular. Além da oportunidade, vocês nunca me impediram de buscar informações e testá-las em seus laboratórios, permitindo minha autonomia dentro do laboratório e da ciência. Por isso eu sou muito grato a vocês e espero de alguma forma ter retornado a vocês um bom trabalho, muito obrigado.

Agradeço aos amigos que a ciência me proporcionou, Dra. Andressa P. de Oliveira e Dr. Ivan Rosa e Silva, dois dos melhores cientistas que tive a oportunidade de conhecer e que me proporcionaram boas risadas, discussões acadêmicas, aprendizados experimentais e trabalhos científicos, vocês foram importantes nessa etapa, muito obrigado. E aos colegas do laboratório, Dr. Fernando Riback, Dra. Camila Ferezin, Dr. Luidy Issayama, Dra. Fernanda Basei, Karoline, Kauany, Júlia, Lívia, Iago, Victor, Eloisa e Alessandro, agradeço pelas conversas e risadas, desejo tudo de melhor para vocês e que alcancem seus objetivos.

Agradeço aos membros da banca de qualificação (Profa. Dra. Catarina Raposo, Profa. Dra. Carmen V. Ferreira e Prof. Dr. Paulo G. Gandra) e aos membros da banca de defesa (Profa. Dra. Déborah Schechtman, Profa. Dra. Nadja C. de S. Pinto, Profa. Dra. Alessandra Sussulini e Profa. Dra. Wanda P. Almeida) pelas boas discussões, críticas e ideias sobre esse trabalho.

Agradeço aos professores, funcionários e terceirizados da FCF e UNICAMP, em especial ao Gustavo Teramatsu pela atenção em todas as minhas dúvidas relacionadas às formalidades dos processos do doutoramento. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001.

RESUMO

As NimA-related kinases (Neks) são uma família de serina/treonina e tirosina quinases altamente conservadas com funções relacionadas na integração do ciclo celular com sinalização de proliferação, resposta ao dano de DNA, metabolismo do cílio primário, progressão da mitose, além de funções relacionadas à inflamação, homeostase mitocondrial e redox. O eixo mitótico das NEKs está relacionado com as funções das NEK6, NEK7 e NEK9. A NEK6 é um dos menores membros das NEKs, sua ativação e regulação na progressão da mitose não estão completamente elucidados. A atividade quinase da NEK6 é regulada após estresse genotóxico através da fosforilação pela CHK1 e CHK2, no entanto o resíduo fosforilado não foi identificado, apenas foi observado que o resíduo fosforilado está localizado no N-terminal da NEK6 (aminoácido 1-80). Inicialmente hipotetizamos que o resíduo fosforilado da NEK6 pela CHK1 e CHK2 é a serina 37, por este ser o único resíduo dentro do N-terminal identificado como fosfopeptídeo experimentalmente, entretanto, não foram identificadas alterações na atividade quinase da NEK6 em mutantes fosfomiméticas e fosfodeficientes da serina 37. Dessa forma, hipotetizamos que a serina 37 pode regular a afinidade da NEK6 com alguns interatores sem alterar sua atividade quinase. Realizamos uma análise de sítios funcionais e regulatórios da NEK6 utilizando ferramentas *in silico*, e obtivemos um modelo da estrutura de alta resolução da NEK6 utilizando o software AlphaFold2. Foi possível observar através da análise dos modelos estruturais da NEK6 que a serina 37 não está na interface de homodimerização, corroborando a ideia de que a fosforilação da serina 37 não afeta a homodimerização e autoativação da NEK6, e consequentemente sua atividade quinase. Posteriormente, analisamos se a serina 37 possui relação com a regulação da homodimerização, autoativação, localização celular e estabilidade da NEK6. Utilizamos vetores de expressão em células de mamíferos clonados com a sequência selvagem da NEK6 de *Homo sapiens* (313 aminoácidos), com mutações fosfomiméticas e fosfodeficientes na serina 37 fusionadas com FLAG (pcDNA3.1) e GFP (pGFP). As construções fusionadas ao FLAG e GFP foram cotransfectadas em células HEK293T por 30 horas. Utilizamos a fluorescência do GFP para analisar a localização celular da NEK6 através do microscópio de fluorescência, e realizamos a imunoprecipitação das construções fusionadas ao FLAG utilizando beads de agarose conjugadas com IgG anti-FLAG. Analisamos a homodimerização pela coprecipitação da GFP-NEK6 com o imunoprecipitado de FLAG-NEK6 e a auto-ativação e estabilidade por western blot. Nossos dados indicam que a NEK6 possui capacidade de homodimerização e auto-ativação, nos quais não são afetadas pela serina 37. A estabilidade da NEK6 também não foi afetada pela serina 37. Identificamos que a serina 37 e a interação com a NEK9 regulam a localização celular da NEK6, e a serina 37 está relacionada com a regulação da afinidade da NEK6 com a Beta Tubulina e a Gama Tubulina. Em uma análise filogenética nós identificamos que a Nek6 surgiu em vertebrados após uma duplicação gênica, e a serina 37 surgiu nos mamíferos. Dessa forma, nossos dados indicam que a serina 37 da NEK6 regula sua localização celular e afinidade a interatores sem modificar sua atividade quinase, provavelmente sendo o resíduo fosforilado pela CHK1 e CHK2 após estresse genotóxico, no qual pôde ter conferido alguma vantagem evolutiva aos mamíferos. Identificamos também que a serina 37 regula um fenótipo que aparentemente são centrossomos amplificados, sendo necessário a confirmação através da análise da colocalização da NEK6 com proteínas centrossomais através de imunofluorescência em microscópio de fluorescência.

Palavras-chaves: Mitose; ciclo celular; quinase; Nek

ABSTRACT

NEKs are a family of serine/threonine and tyrosine kinases highly conserved with functions related in the the cell cycle integration with proliferation signaling, DNA damage response, primary cilium metabolism, mitotic progression, as well as functions related to inflammation, mitochondrial homeostasis and redox. The mitotic axis of NEKs is related to the NEK6, NEK7 and NEK9 functions. NEK6 is one of the smallest members of NEKs, its activation and regulation in mitotic progression are not completely elucidated. The kinase activity of NEK6 is regulated after genotoxic stress through phosphorylation by CHK1 and CHK2, however the phosphorylated residue was not identified, it was only observed that the phosphorylated residue is located at the N-terminus of NEK6 (amino acid 1-80). We initially hypothesized that serine 37 is the residue phosphorylated in NEK6 by CHK1 and CHK2, since this is the only residue identified experimentally as a phosphopeptide within the N-terminus. However, no changes in the NEK6 kinase activity were identified in serine 37 phosphomimetic and phosphodeficient mutants. Thus, serine 37 may regulate the affinity of NEK6 with some interactors without altering its kinase activity. We performed an analysis of functional and regulatory sites of NEK6 using *in silico* tools and obtained a high-resolution model of the NEK6 structure using AlphaFold2 software. It was possible to observe that serine 37 is not in the homodimerization interface of the modeled NEK6, corroborating the idea that serine 37 phosphorylation does not affect the homodimerization and autoactivation of NEK6, and consequently its kinase activity. Subsequently, we analyzed whether serine 37 is related to the regulation of homodimerization, autoactivation, cellular localization and stability of NEK6. We used mammalian cell expression vectors cloned with the wild type NEK6 sequence of the *Homo sapiens* (313 aminoacids) and with serine 37 phosphomimetic and phosphodeficient mutations fused to FLAG (pcDNA3.1) and GFP (pGFP). The constructions were cotransfected into HEK293T cells for 30 hours. We used GFP fluorescence to analyze the NEK6 cellular localization with fluorescence microscopy, and performed immunoprecipitation of the FLAG-fused constructs using agarose beads conjugated with anti-FLAG IgG. We analyzed homodimerization by coprecipitation of GFP-NEK6 with immunoprecipitation of FLAG-NEK6, and autoactivation and stability by western blot. Our data indicates that NEK6 can form homodimerization and autoactivation, which are not affected by serine 37. NEK6 stability was also not affected by serine 37. We identified that serine 37 and the interaction with NEK9 regulates NEK6 cellular localization, and serine 37 is related to the regulation of NEK6 affinity to beta tubulin and gamma tubulin. In a phylogenetic analysis, Nek6 arose in vertebrates after a gene duplication, and we identified that serine 37 arose in mammals. Thus, our data indicates that serine 37 of NEK6 regulates its affinity to interactors and cellular localization without modify its kinase activity, probably being the residue phosphorylated by CHK1 and CHK2 after genotoxic stress, in which it may have conferred some evolutionary advantage to mammals. We also identified that serine 37 regulates a phenotype that apparently are amplified centrosomes, which requires confirmation by analyzing the colocalization of NEK6 with centrosome proteins through immunofluorescence microscope.

Keywords: Mitosis; cell cycle; kinase; Nek

LISTA DE ILUSTRAÇÕES

| | |
|--|-----------|
| Figura 1. Características químicas e estruturais da ligação peptídica e da cadeia polipeptídica..... | página 17 |
| Figura 2. Forças relacionadas ao enovelamento de proteínas..... | página 18 |
| Figura 3. Estruturas secundárias e terciária das proteínas..... | página 19 |
| Figura 4. Estrutura de padrões, domínios e proteínas..... | página 20 |
| Figura 5. Modelo em funil rugoso de energia livre do enovelamento de proteínas..... | página 23 |
| Figura 6. Representação da radiação eletromagnética e espectro eletromagnético..... | página 26 |
| Figura 7. Radiação eletromagnética circular polarizada..... | página 28 |
| Figura 8. Espectro de dicroísmo circular..... | página 29 |
| Figura 9. Espalhamento de raio X a baixos ângulos..... | página 31 |
| Figura 10. Plot da intensidade do espalhamento de raio X a baixos ângulos..... | página 32 |
| Figura 11. Plot de Kratky dos dados de espalhamento e plot da dimensão máxima da proteína..... | página 33 |
| Figura 12. Difração de raio X..... | página 35 |
| Figura 13. Mapa de densidade eletrônica com moléculas alocadas com diferentes resoluções..... | página 36 |
| Figura 14. Plot de Ramachandran..... | página 37 |
| Figura 15. Exemplo de coevolução e comutações evolutivas de resíduos de aminoácidos dentro de cadeias polipeptídicas utilizadas pelos métodos de predição de estrutura de proteínas..... | página 41 |
| Figura 16. Contatograma das distâncias dos aminoácidos de uma proteína em sua conformação, nativa tridimensional..... | página 42 |
| Figura 17. Modelo das redes neurais e processos realizados pelo AlphaFold2..... | página 43 |
| Figura 18. Gráfico de pLDDT e exemplo de estrutura modelada pelo AlphaFold2..... | página 45 |
| Figura 19. Fosforilação de resíduos de serina, treonina e tirosina por proteínas quinase..... | página 46 |
| Figura 20. Representação estrutural de uma quinase típica..... | página 47 |
| Figura 21. Mapa filogenético do quinoma humano associado aos inibidores aprovados pela FDA no tratamento de tumores sólidos e hematológicos..... | página 50 |
| Figura 22. Alinhamento múltiplo dos domínios quinase das NEKs e Nima, e dendrograma de quinases eucarióticas..... | página 51 |

| | |
|---|-----------|
| Figura 23. Contribuição das NEKs para a progressão do ciclo celular, metabolismo do cílio primário e estabilidade do genoma nuclear..... | página 52 |
| Figura 24. Representação do envelope molecular, modelos alocados no envelope molecular e modelo estrutural de alta resolução da NEK6..... | página 55 |
| Figura 25. Funções e regulações do eixo mitótico das NEKs, formado pelas NEK6, NEK7 e NEK9..... | página 57 |
| Figura 26. Centrossomos e sua organização estrutural durante o ciclo celular..... | página 58 |
| Figura 27. Imagens ilustrativas e imunofluorescência de células normais e células cancerosas com centrossomos amplificados e sua organização ao longo das diferentes fases do ciclo celular..... | página 59 |
| Figura 28. Regulação da atividade quinase da NEK6..... | página 62 |
| Figura 29. Alinhamento da sequência de nucleotídeos do N-terminal da NEK6 (Q9HC98-1) com a sequência de nucleotídeos dos sequenciamentos dos vetores FLAG-NEK6 WT, S37A e S37D, e GFP-NEK6 S37A e S37D gerados por mutagênese sítio dirigida..... | página 67 |
| Figura 30. Imagens de campo claro e fluorescência de GFP de HEK293T transfectadas com pGFP-NEK6 WT por 24 horas nas concentrações de 1 µg, 2 µg e 4 µg de DNA por poço usando uma placa de 6 poços, com razão DNA:PEI 1 µg:3 µg..... | página 70 |
| Figura 31. Imagens de campo claro de HEK293T transfectadas com pcDNA3.1 FLAG-NEK6 WT e pcDNA3.1 FLAG-NEK9 WT por 24 horas nas concentrações de 1 µg, 2 µg e 4 µg de DNA por poço usando uma placa de 6 poços, com razão DNA:PEI 1 µg:3 µg..... | página 71 |
| Figura 32. Western blot das transfecções de FLAG-NEK6 WT, FLAG-NEK9 WT e GFP-NEK6 em HEK293T por 24 horas nas concentrações de 1 µg, 2 µg e 4 µg de DNA por poço usando uma placa de 6 poços, com razão DNA:PEI 1 µg:3 µg..... | página 72 |
| Figura 33. Alinhamento dos aminoácidos da NEK6 (Q9HC98-1) e NEK7 (Q8TDX7-1)..... | página 75 |
| Figura 34. Representações de alta resolução da estrutura e das celas unitárias dos cristais da NEK7 encontrados na literatura..... | página 76 |
| Figura 35. Gráfico de Ramachandran dos modelos estruturais de alta resolução da NEK7 determinados experimentalmente..... | página 78 |
| Figura 36. Superposição das estruturas da NEK7 determinadas experimentalmente identificadas na literatura..... | página 80 |
| Figura 37. Modelos estruturais de alta resolução da NEK6 e NEK7 preditos pelo AlphaFold2..... | página 81 |
| Figura 38. Gráfico de Ramachandran dos modelos estruturais de alta resolução da NEK6 e NEK7 gerados pelo AlphaFold2..... | página 82 |
| Figura 39. Superposição do modelo da NEK7 gerado pelo AlphaFold2 com as estruturas 5de2 e 2WQN da NEK7..... | página 84 |
| Figura 40. Superposição do modelo da NEK6 gerado pelo AlphaFold2 com as estruturas 5de2 e 2WQN da NEK7..... | página 85 |

| | |
|--|------------|
| Figura 41. Modelos das estruturas de alta resolução da NEK6 e NEK7 preditos pelo AlphaFold2..... | página 86 |
| Figura 42. Modelos estruturais da NEK6 e NEK7 gerados pelo AlphaFold2 com anotações de sítios funcionais..... | página 87 |
| Figura 43. Representações dos modelos estruturais da NEK6 e NEK7 gerados pelo AlphaFold2 com anotações de características estruturais e superfície eletrostática..... | página 89 |
| Figura 44. Representação da estrutura secundária, superfície e resíduos regulatórios do modelo estrutural de alta resolução da NEK6 gerado pelo AlphaFold2..... | página 90 |
| Figura 45. Superposição da estrutura secundária do homodímero da NEK7 determinada experimentalmente (PDB 5de2) com a estrutura secundária de dois modelos da estrutura da NEK6 gerado pelo AlphaFold2..... | página 91 |
| Figura 46. Representação da interface back-to-back da homodimerização da NEK7 (PDB 5de2) e NEK6 (AlphaFold2)..... | página 92 |
| Figura 47. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D..... | página 95 |
| Figura 48. Ponceau S dos western blots das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D..... | página 96 |
| Figura 49. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK9 WT com GFP-NEK6 WT, S37A, S37D e -NEK7 WT..... | página 97 |
| Figura 50. Imunofluorescências representativas do GFP e quantificação de células com focos de GFP semelhantes a centrosomos amplificados de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK9, FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D..... | página 100 |
| Figura 51. Imagens representativas da imunofluorescência de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT e S37A com GFP-NEK6 WT, S37A e S37D..... | página 101 |
| Figura 52. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D..... | página 102 |
| Figura 53. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-vazio..... | página 103 |
| Figura 54. Alinhamento múltiplo de sequências homólogas a NEK6 de Homo sapiens identificadas por BLASTP..... | página 106 |

LISTA DE TABELAS

| | |
|--|-----------|
| Tabela 1. Funções propostas para as quinases Neks de mamíferos..... | página 53 |
| Tabela 2. Porcentagem de estruturas secundárias preditas na estrutura da NEK6..... | página 56 |
| Tabela 3. Condições e tratamentos que afetam os níveis de mRNA e níveis proteicos da NEK6..... | página 63 |
| Tabela 4. Valores de R livre e resolução (angstroms) dos cristais da NEK7 encontrados na literatura..... | página 77 |
| Tabela 5. Porcentagem dos aminoácidos das estruturas da NEK7 determinadas experimentalmente nas regiões discriminadas pelo gráfico de Ramachandran..... | página 77 |
| Tabela 6. Porcentagem de aminoácidos encontrados em estruturas secundárias das estruturas da NEK7 determinadas experimentalmente..... | página 79 |
| Tabela 7. RMSD das superposições entre as cadeias principais dos monômeros das estruturas de alta resolução da NEK7 determinadas experimentalmente..... | página 80 |
| Tabela 8. Porcentagem dos aminoácidos das estruturas da NEK6 e NEK7 modeladas pelo AlphaFold2 nas regiões discriminadas pelo gráfico de Ramachandran..... | página 82 |
| Tabela 9. RMSD das superposições entre as cadeias principais da estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 5de2) e os 5 modelos da estrutura da NEK7 preditos pelo AlphaFold2 (AF)..... | página 84 |
| Tabela 10. RMSD das superposições entre as cadeias principais da estrutura de alta resolução determinada experimentalmente da NEK7 (PDB 5de2) e os 5 modelos da estrutura da NEK6 preditos pelo Alpha Fold 2 (AF)..... | página 85 |
| Tabela 11. Porcentagem de aminoácidos encontrados em estruturas secundárias das estruturas dos modelos da NEK6 e NEK7 gerados pelo AlphaFold2..... | página 86 |

LISTA DE ABREVIATURAS E SIGLAS

$[\theta]_{MR}$ = elipticidade molar residual média
6-4 PPs = 6-4 photoproduct
A = Absorção da radiação eletromagnética
AGC = cAMP-dependent, cGMP-dependent and protein kinase C
AMP = Adenosina monofosfato
ATM = ataxia-telangiectasia mutated
ATP = Adenosina trifosfato
ATP = Adenosina trifosfato
ATR = ataxia telangiectasia and Rad3-related
c = Concentração molar
CAMK = Ca^{2+} /calmodulin (CaM)-dependent protein kinase
CCP110 = Centriolar coiled-coil protein of 110 kDa
CDK1 = Cyclin-dependent kinase 1
CEP192 = Centrosome-associated protein 250
CEP250 = Centrosome-associated protein 250
CHK1 = Checkpoint kinase 1
CHK2 = Checkpoint kinase 2
CK = Casein kinase
CMGC = Cyclin-dependent kinases (CDKs), Mitogen-activated protein kinases (MAP kinases), Glycogen synthase kinases (GSK) and CDK-like kinases
c-Myc = cellular Myelocytomatosis
CPDs = cyclobutane pyrimidine dimers
CTCF = CCCTC-Binding Factor
 D_{max} = Dimensão máxima
DNA = Ácido desoxirribonucleico
DNMT1 = DNA (cytosine-5)-methyltransferase 1
Domínio RCC1 = Regulator of Chromosome Condensation 1
DSB = Double-strand break
DSKs = Dual-Specificity Kinases
EG5 = kinesin-5
EML4 = Echinoderm microtubule-associated protein-like 4
EVH1 = Enabled/Vasodilator-stimulated phosphoprotein Homology 1
FBS = Fetal bovine serum
FT = Flow-through
FUSED/ULK = Fused/Unc-51-like kinase
G = Energia livre de Gibbs
GFP = Green Fluorescent Protein
GM-CSF = granulocyte macrophage-colony stimulating factor
GST = Glutathione S-transferase
GTP = Guanosina trifosfato
GYF = Glycine-Tyrosine-Phenylalanine
HDAC = Histone deacetylase
HEK293 = Human embryonic kidney 293
HEK293T = Human embryonic kidney 293T
HeLa = Henrietta Lacks
HIF-1 α = Hypoxia-Inducible Factor 1alpha
HIS = Poli-histidina tag
HSP72 = Heat shock protein 72
IgG = Imunoglobulina G
IP = Imunoprecipitação
IR = Ionizing Radiation
 K_D = constante de dissociação
KIF11 = Kinesin Family Member 11

KIF14 = Kinesin Family Member 14
 l = comprimento do caminho óptico percorrido pela luz
LCP = Luz circular polarizada para a esquerda (left circular polarized)
LNX1 = ligand of numb-protein X 1
Lobo C = Lobo C-terminal
Lobo N = Lobo N-terminal
M = molar
M. tuberculosis = *Mycobacterium tuberculosis*
 Mg^{2+} = íon magnésio
MKLP2 = Mitotic kinesin-like protein 2
Motivo DFG = Motivo Aspartato (D), Fenilalanina (F) e Glicina (G)
Motivo DLG = Motivo Aspartato (D), Leucina (L) e Glicina (G)
Motivo HRD = Motivo Histidina (H), Arginina (R) e Aspartato (D)
mRNA = RNA mensageiro
N90R = mutação da Asparagina (N) 90 para Arginina (R)
NEDD1 = Neural precursor cell expressed, developmentally down-regulated 1
Nek = NimA-related kinases
NimA = Never in mitosis, gene A
NUP98 = Nucleoporin 98
OCT1 = Organic Cation Transporter 1
p70S6K = p70 Ribosomal protein S6 kinase
pm = picômetro
PCM = Pericentriolar material
PDB = Protein data bank
PKA = Protein Kinase A
pLDDT = Predicted local distance difference test
PLK1 = Polo-like kinase 1
R155E = mutação da Arginina (R) 155 para Glutamato (E)
RCP = Luz circular polarizada para a direita (right circular polarized)
 R_g = Raio de giro
RMN = Ressonância magnética nuclear
RMSD = Root Mean Square Deviation
RNA = Ácido ribonucleico
 R_s = Raio de Stokes
S206A = mutação da Serina (S) 206 para Alanina (A)
S37A = mutação da Serina (S) 37 para Alanina (A)
S37D = mutação da Serina (S) 37 para Aspartato (D)
SAXS = Small Angle X-ray Scattering
SDS-PAGE = Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis
Sequência Dead-box = sequência Aspartato (D), Glutamato (E), Alanina (A) e Aspartato (D)
Sequências PEST = sequências rica em Prolina (P), Glutamato (E), Serina (S) e Treonina(T)
SH3 = Src Homology 3
SMAD4 = Mothers Against Decapentaplegic homolog 4
SP1 = Specificity protein 1
STAT3 = Signal Transducers and Activators of Transcription 3
STE = Sterile Kinase
STK = Serine/Threonine Kinase
T = Temperatura em Kelvin
TGF-Beta = Transforming Growth Factor Beta
TK = Tyrosine Kinase
TKL = Tyrosine Kinase Like
TRB = Tribbles
tRNA = RNA transportador
TYR-down = Tyrosine-down
TYR-up = Tyrosine-up

UEV = Ubiquitin E2 variant
UV = Ultravioleta
WL = Whole Lysate
WT = Wild Type
 γ -TuRC = gamma-tubulin ring complex
 ΔG = Variação da energia livre de Gibbs
 ΔH = Variação da entalpia
 ΔS = Variação da entropia
 $\Delta \epsilon$ = Elipticidade molar
 $\Delta \epsilon_{MR}$ = Elipticidade molar residual média
 ϵ = Coeficiente de extinção molar
 θ = Ângulo
 λ = Comprimento de onda

SUMÁRIO

| | |
|--------------------------------|------------|
| Introdução..... | página 16 |
| Objetivos..... | página 65 |
| Materiais e métodos..... | página 66 |
| Resultados e discussão..... | página 74 |
| Conclusões e perspectivas..... | página 107 |
| Referências..... | página 110 |
| Anexos..... | página 120 |

1. Introdução

1.1. Estrutura e enovelamento de proteínas

Proteínas são moléculas biológicas produzidas por organismos vivos. Dentre os 4 principais tipos de moléculas existentes em organismos vivos, nos quais incluem os ácidos nucleicos (DNA e RNA), lipídeos e polissacarídeos, as proteínas são as moléculas mais abundantes em todos os organismos da terra. As proteínas são encontradas em grande variedade em relação à forma (estrutura), tamanho, dinâmica e função, sendo as moléculas responsáveis por diversas funções dentro dos organismos. Muitas dessas funções são conservadas, sendo possível encontrar proteínas com sequência, estrutura, dinâmica e funções semelhantes mas presentes em diferentes organismos de regiões distintas da árvore filogenética (Stolar & Smith, 2020).

As proteínas são polímeros formados por uma sequência de aminoácidos ligados por ligações peptídicas, ocorrendo a ligação entre o grupamento carboxila do primeiro aminoácido com o grupo amino do segundo aminoácido, posteriormente entre o grupo carboxila do segundo aminoácido com o grupo amino do terceiro, e assim sucessivamente. Dessa forma, o sentido de síntese da cadeia polipeptídica é do N-terminal ao C-terminal, sendo o N representativo do nitrogênio do grupo amina do primeiro aminoácido da cadeia polipeptídica e o C representativo do carbono do grupo carboxila do último aminoácido da cadeia polipeptídica (figura 1A). Durante a ligação peptídica o grupamento amina de um aminoácido age como nucleófilo e desloca a hidroxila do grupamento carboxila de outro aminoácido, causando a remoção de uma molécula de H₂O e formação de uma ligação covalente entre o carbono e o nitrogênio (figura 1A) (Pan, A., Pranavathiyani, G. & Chakraborty, S., 2021; Kabir, et al. 2023).

Dentro das células, a ligação peptídica e síntese de proteínas é um processo complexo e basicamente ocorre através de 5 estágios (Banik & Nandi, 2012): 1) a enzima aminoacil-tRNA-sintase forma um aminoacil-AMP através da reação de um ATP com o aminoácido, e posteriormente a mesma enzima realiza a ligação do grupamento carboxila do aminoacil-AMP com seu respectivo RNA transportador (tRNA), formando aminoacil-tRNA e liberando AMP; 2) ocorre a formação do complexo de iniciação, formado pelo RNA mensageiro (mRNA) que se liga à subunidade menor do ribossomo, pelo aminoacil-tRNA que se liga ao códon de iniciação do mRNA e pelos ribossomos, onde a subunidade maior se liga à subunidade menor (associada ao mRNA e tRNA) finalizando a formação do complexo; 3) ocorre o alongamento da cadeia polipeptídica através da formação de sucessivas ligações peptídicas, onde os aminoacil-tRNAs são orientados dentro dos sítios A (aminoacil), P (peptidil) e E (exit = saída) dos ribossomos por proteínas associadas aos ribossomos, ao passo que ocorrem as ligações peptídicas e síntese do polipeptídeo;

4) ocorre a terminação da síntese da cadeia polipeptídica pela interação do códon de parada do mRNA em tradução com fatores de terminação, nos quais contribuem com a hidrólise do peptidil-tRNA (tRNA ligado a cadeia polipeptídica sintetizada no sítio P dos ribossomos) e liberação da cadeia polipeptídica; 5) por fim, o polipeptídeo sofre seu enovelamento final (formação de pontes dissulfeto, adição de carboidratos, lipídeos) e é destinado ao seu local celular (citossol, mitocôndrias, núcleo, retículo endoplasmático, etc) ou são excretados, aonde realizarão suas funções.

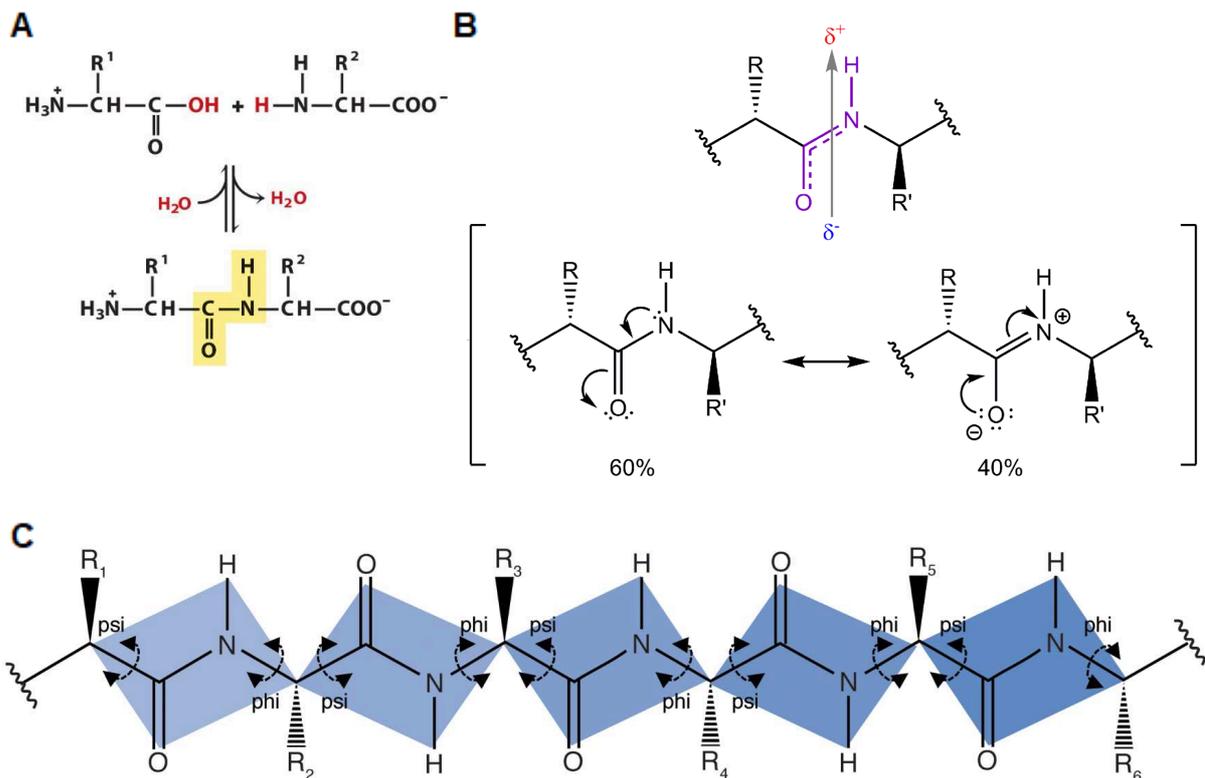


Figura 1. Características químicas e estruturais da ligação peptídica e da cadeia polipeptídica, removidos de Stolar & Smith (2020). A) Ligação peptídica entre dois aminoácidos. R representa o grupo radical (cadeia lateral) dos aminoácidos, a ligação peptídica está destacada em amarelo e em vermelho a hidroxila e o hidrogênio que são removidos na forma de H_2O durante a ligação peptídica. B) A ressonância da ligação peptídica induz a formação de um dipolo, representado pelos átomos em roxo. A ressonância eletrônica da ligação peptídica possui um equilíbrio, onde as espécies entre colchetes indicam um equilíbrio químico entre ligações peptídicas que possuem os elétrons na nuvem eletrônica do nitrogênio e do oxigênio. C) A cadeia polipeptídica linear, representada do N-terminal (esquerda) ao C-terminal (direita). Grupos radicais estão representados como R. Os retângulos indicam as ligações peptídicas e sua configuração planar (retângulo azul). Os ângulos de rotação ϕ e ψ estão indicados como setas tracejadas. Rotações ocorrem entre as ligações do carbono alfa (ligado ao grupo R) com o nitrogênio e carbono das ligações peptídicas.

A distribuição eletrônica e a ressonância dos elétrons dos átomos da ligação peptídica conferem características de ligação dupla para esta ligação, onde o oxigênio possui carga parcial negativa e o nitrogênio possui carga parcial positiva (figura 1B). Essas cargas parciais são oriundas da transferência do elétron do nitrogênio para o grupo carbonila, causando então a transferência deste elétron para o oxigênio e conferindo carga parcial negativa ao oxigênio (figura 1B). Essa distribuição eletrônica não permite que

ocorram torções da ligação peptídica, conferindo uma característica de ligação parcialmente dupla (Stolar & Smith, 2020; Kabir, et al. 2023).

Por outro lado, as ligações entre o carbono alfa (ligado a cadeia lateral ou grupo radical) e os grupos carboxila e amino da ligação peptídica são ligação simples, nos quais permitem rotações do carbono alfa com o grupo radical do aminoácido e da ligação peptídica (rotações na forma de planos rígidos) (figura 1C). Os ângulos rotacionais são denominados de ângulos ϕ (phi) e ψ (psi), sendo o ângulo ϕ formado entre o carbono alfa e o nitrogênio da ligação peptídica e o ângulo ψ formado entre carbono alfa e o carbono da ligação peptídica (figura 1C). Sendo assim, a ligação peptídica possui organização espacial planar, onde a cadeia principal dos polipeptídeos é formada por uma série de planos com eixos rotacionais no carbono alfa e nos grupos radicais dos diferentes aminoácidos. Entretanto, ocorrem impedimentos estéricos entre grupos radicais e ligações peptídicas, e tais impedimentos estéricos causam limitações no grau de liberdade entre os ângulos ϕ , ψ e dos grupos radicais (Stolar & Smith, 2020; Kabir, et al. 2023).

As forças predominantes no enovelamento das proteínas são decorrentes do efeito hidrofóbico, enquanto as interações fracas e covalentes definem a forma e estabilizam a estrutura das proteínas (figura 2A).

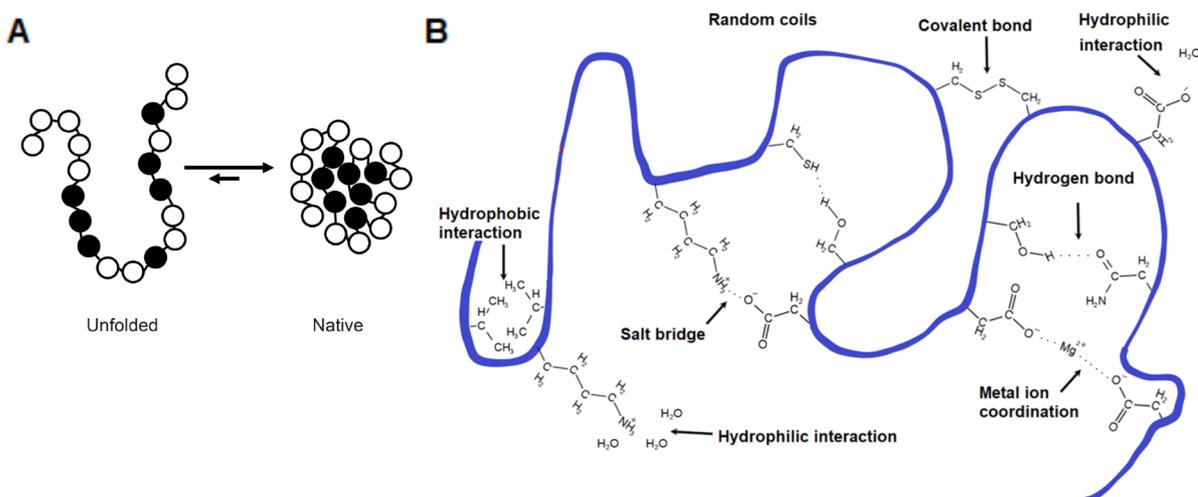


Figura 2. Forças relacionadas ao enovelamento das proteínas, removidos de Stolar & Smith (2020). A) Representação do colapso hidrofóbico no enovelamento de proteínas. Os aminoácidos hidrofóbicos (esferas escuras) sofrem colapso para o centro da proteína durante o enovelamento, sendo envolvidos pelos aminoácidos da camada de solvatação (esferas brancas). B) Interações intramoleculares entre as cadeias laterais dos aminoácidos de uma cadeia polipeptídica. Estão demonstrados exemplos de interações hidrofóbicas, hidrofílicas, iônica (ponte salina), covalente (ponte dissulfeto) e coordenação por íon metálico ilustradas dentro de uma cadeia polipeptídica (fita azul). O N-terminal e C-terminal das cadeias polipeptídicas não estão sendo discriminadas nas representações.

Os grupos radicais dos aminoácidos diferem em tamanho, formato, polaridade, carga e hidrofobicidade, onde os radicais e cadeia principal interagem entre si através de interações hidrofóbicas, interações fracas (van der Waals, ligação de hidrogênio e

eletrostática) e covalentes (pontes dissulfeto) (figura 2B). As pontes dissulfeto são formadas através da ligação covalente entre os grupos tióis de duas cisteínas da cadeia polipeptídica (Rico, F. et al. 2013; Stolar & Smith, 2020).

As proteínas sofrem de 3 a 4 níveis de organizações estruturais até alcançar seu estado conformacional nativo e realizar suas funções. Os níveis organizacionais são denominados de estrutura primária, secundária, terciária e quaternária (Rico, F. et al. 2013). A estrutura primária se refere a sequência dos aminoácidos de uma proteína de forma linear. A estrutura secundária se refere a formação de estruturas organizacionais com padrões regulares da cadeia principal do polipeptídeo, ocorrendo pela formação de ligações de hidrogênio entre o hidrogênio do grupo amina e o oxigênio do grupo carboxila das ligações peptídicas. Essas ligações de hidrogênio ocorrem devido a ressonância eletrônica da ligação peptídica. As estruturas secundárias mais comuns encontradas em proteínas são as alfa-hélices e folhas-beta, onde tais conformações evitam impedimentos estequiométricos e maximizam a ligação de hidrogênio entre a cadeia principal (Rico, F. et al. 2013; Stolar & Smith, 2020).

As alfa-hélices são caracterizadas por uma espiral com voltas para a direita (direção horária no sentido N-C terminal da cadeia polipeptídica) com ligações de hidrogênio entre a primeira e quarta ligação peptídica, segunda e quinta ligação peptídica, terceira e sexta ligação peptídica, quarta e sétima ligação peptídica, e assim sucessivamente. As alfa-hélices possuem aproximadamente 3,6 resíduos de aminoácidos por volta da espiral (aproximadamente 0,54 nm de distância entre a cadeia principal na espiral) como os grupos radicais voltados para fora da estrutura (figura 3A) (Rico, F. et al. 2013; Jiang, Q. et al. 2017).

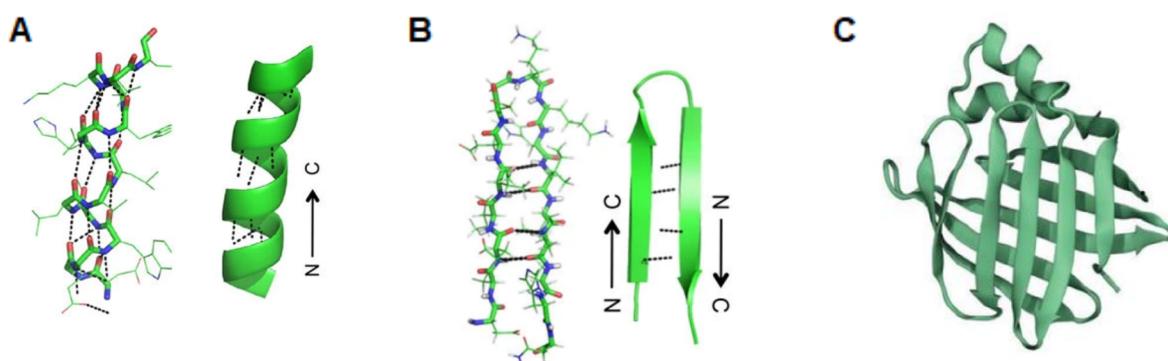


Figura 3. Estruturas secundárias e terciária das proteínas, removido de Rico, F. et al. (2013). Representações de uma alfa-hélice (A) e uma folha-beta antiparalela (B) na forma de “arestas” (com os grupos radicais voltados para fora da estrutura) e na forma de fitas (alfa-hélice como espirais e folha-beta como setas). As ligações de hidrogênio da cadeia principal estão representadas pelas linhas tracejadas. Os oxigênios, carbonos e nitrogênios da cadeia principal estão representados pelas cores vermelho, verde e azul, respectivamente. C) Uma representação típica da estrutura terciária ou tridimensional de uma proteína.

As folhas-beta são formadas por duas ou mais fitas beta em série, que se organizam em um padrão regular em zig-zag dentro de um mesmo polipeptídeo. A direção das fitas beta dentro de uma folha-beta podem ser organizadas de forma paralela (mesma direção) e antiparalela (direções opostas), ocorrendo a ligação de hidrogênio entre fitas beta adjacentes (figura 3B) (Cheng, P. et al. 2013; Rico, F. et al. 2013).

Neste sentido, a estrutura terciária de uma proteína está relacionada à sua estrutura tridimensional, onde é decorrente do efeito hidrofóbico, das interações realizadas pela cadeia principal e grupos radicais, e empacotamento das estruturas secundárias da cadeia polipeptídica (figura 3C) (Stolar & Smith, 2020). As cadeias laterais das estruturas secundárias e segmentos não ordenados do polipeptídeo interagem entre si, nos quais formam um arranjo de estruturas secundárias com conformações reconhecíveis, denominadas de padrões (figura 4A). Por exemplo, o padrão β - α - β consiste em duas folhas-beta conectadas por uma alfa-hélice de forma linear, nos quais durante o empacotamento destas estruturas secundárias as fitas beta formam uma folha-beta paralela com a alfa-hélice cruzando o centro da folhas-beta. Os padrões e estruturas secundárias interagem e formam domínios (figura 4B), nos quais sofrem enovelamento de forma independente da cadeia polipeptídica inteira. As proteínas são formadas por um ou diversos domínios, nos quais possuem função estrutural e/ou catalítica, funcionando como módulos funcionais que se relacionam ou não nas funções realizadas pela proteína (figura 4C) (Rico, F. et al. 2013; Stolar & Smith, 2020).

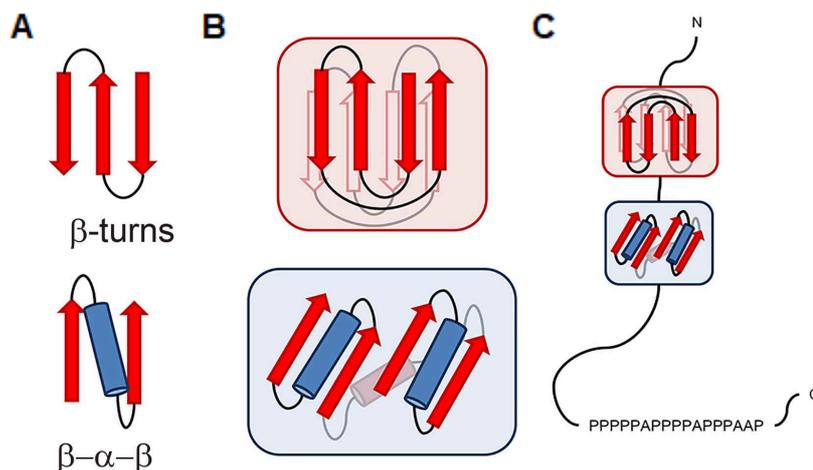


Figura 4. Estrutura de padrões, domínios e proteínas, removido de Stolar & Smith (2020). A) Exemplos de sequências padrões de proteínas, como voltas β e β - α - β . B) Domínios de proteínas são formados através do empacotamento das estruturas secundárias e padrões, dando origem a estrutura terciária no qual está relacionada a função da proteína. Os retângulos definem os limites de um domínio hipotético, formado pelos arranjos das estruturas secundárias e padrões alocados em seu interior. C) Exemplo de proteína formada por diferentes domínios. O segmento rico em prolina representa um segmento desordenado na estrutura da proteína. As setas vermelhas representam as fitas beta, os cilindros azuis representam as alfa-hélices.

A estrutura quaternária está relacionada com complexos de proteínas, nos quais possuem subunidades formadas por distintas proteínas que interagem através de ligações fracas, covalentes e em alguns casos ocorrem interações através de regiões hidrofóbicas (Rashid, M., Khatib, F. & Sattar, A., 2015).

Após o enovelamento final, as proteínas podem ser classificadas pelas suas características macromoleculares, como globulares, transmembranas, fibrosas e intrinsecamente desordenadas. Essas características macromoleculares estão correlacionadas com as características das sequências dos polipeptídeos e características estruturais dos domínios. No entanto, uma mesma proteína pode possuir segmentos globulares, transmembrana e intrinsecamente desordenados (Rico, F. et al. 2013; Stolar & Smith, 2020).

Aproximadamente um terço das proteínas humanas possuem algum segmento desordenado em sua estrutura. As proteínas intrinsecamente desordenadas possuem grande dinâmica, onde algumas proteínas ou regiões desordenadas sofrem modificações conformacionais após interagir com seu ligante, apresentando uma estrutura melhor definida após a interação. Dessa forma, os segmentos desordenados podem possuir função de scaffold na regulação da sinalização celular através de interações sequenciais, onde a interação do segmento desordenado de uma proteína com um interator induz modificações no segmento desordenado que são necessárias para a interação com um segundo interator (Wright & Dyson, 2015). De forma interessante, proteínas intrinsecamente desordenadas e segmentos desordenados de uma proteína estão sendo descritos com a função de separação de fase líquido-líquido, formando condensados de proteínas na fase líquida e agrupando conjuntos de reações, como ocorrem nos corpos de Cajal nos nucléolos, grânulos de estresse no citoplasma e centrossomos (Brocca, S. et al. 2020).

O enovelamento de proteínas é descrito pelas leis da termodinâmica (Stolar & Smith, 2020). O sistema é definido como a reação de interesse, onde as reações de enovelamento de proteínas incluem a proteína desenovelada e enovelada, e qualquer solvente ou soluto que interage com essa proteína durante as reações. O ambiente é todo o resto do universo que está fora do sistema. A primeira lei da termodinâmica atesta que a quantidade de energia do universo é constante, onde a energia não pode ser criada ou destruída mas pode ser transformada. Dessa forma, a quebra ou formação de novas ligações irão alterar a quantidade de energia do sistema na forma de calor, aumentando ou reduzindo a temperatura do sistema e induzindo resposta oposta no ambiente pela transferência do calor (Stolar & Smith, 2020).

A variação na quantidade de calor em uma reação é definida como entalpia (ΔH), o qual é uma mensuração da energia das ligações formadas e quebradas no sistema (Stolar & Smith, 2020). Se a reação libera calor para o ambiente o ΔH será negativo, exotérmico. Por

outro lado, se a reação absorve calor do ambiente o ΔH será positivo, endotérmico. Formação de ligações de hidrogênio, interações de van der Waals e eletrostáticas resultam na liberação de calor para o ambiente (exotérmica), enquanto a quebra dessas ligações resultam na absorção de calor do ambiente (endotérmica). Dessa forma, uma variação entre a formação e quebra dessas ligações irá determinar o valor positivo ou negativo do ΔH , onde a formação das ligações fracas causam a perda de calor e induzem a proteína ao seu estado de menor energia (Stolar & Smith, 2020).

Neste mesmo sentido, para o enovelamento protéico ocorrer de forma espontânea deve ocorrer redução na energia livre de Gibbs (G) e movimento conformacional da proteína para um estado de menor energia (Stolar & Smith, 2020). A variação na energia livre de Gibbs (ΔG) possui 2 componentes influenciados pela temperatura: a ΔH e na entropia (ΔS). O ΔS é uma mensuração da mudança na desordem do sistema, quanto maior a desordem menor é o ΔS , e quanto menor a desordem maior é o ΔS . A segunda lei da termodinâmica atesta que a entropia do universo está aumentando. Dessa forma, para que o enovelamento da proteína seja favorável a entropia do universo deve aumentar (Stolar & Smith, 2020).

A base do efeito hidrofóbico é o aumento no ΔS das moléculas de água associadas a proteína, sendo a força mais importante no enovelamento protéico (Stolar & Smith, 2020). Em um polipeptídeo desenovelado, as moléculas de água se organizam em volta dos aminoácidos hidrofóbicos em uma estrutura semelhante ao gelo, reduzindo a desordem das moléculas de água e aumentando a ordem do sistema (reduzindo o ΔS). O efeito hidrofóbico agrupa os aminoácidos hidrofóbicos no centro da estrutura da proteína, causando a formação da camada de solvatação pelos aminoácidos polares e alifáticos nos quais interagem com as moléculas de água. A interação das moléculas de água com os aminoácidos da camada de solvatação e a redução da organização em estrutura semelhante a gelo entre as moléculas de água induzidas pelos aminoácidos hidrofóbicos, causam o aumento de moléculas de água livres, e consecutivamente, aumentam a entropia do solvente/sistema (ΔS positivo). Adicionalmente, nas reações exotérmicas (ΔH negativo) a energia é liberada para o ambiente, aumentando a entropia do ambiente (ΔS positivo). Dessa forma, a entropia do ambiente pode ser aumentada pelo aumento da entropia do sistema ($\Delta S > 0$), ou pela transferência de calor do sistema para o ambiente ($\Delta H < 0$).

A energia livre de Gibbs é utilizada para monitorar a variação na entropia do universo, sendo $\Delta S_{\text{universo}} = \frac{-\Delta G}{T}$, onde T é temperatura em Kelvin (Stolar & Smith, 2020). Se o ΔG é negativo, o $\Delta S_{\text{universo}}$ é positivo e a reação será favorável de ocorrer, e o contrário é verdadeiro. O ΔG também é uma função da entalpia e entropia do sistema, onde $\Delta G = \Delta H - T\Delta S$. Uma proteína desenovelada terá mais energia livre (ΔH positivo e o ΔS negativo, causando valores positivos do ΔG) em relação a mesma proteína enovelada (ΔH

negativo e ΔS positivo, causando valores negativos do ΔG). Sendo assim, o aumento do ΔS do solvente pelo efeito hidrofóbico e interação com a camada de solvatação e a redução do ΔH pela liberação de calor pelas novas interações da cadeia polipeptídica causam redução do ΔG , tornando o enovelamento do polipeptídeo espontâneo. No entanto, o ΔG apenas indica que a reação de enovelamento é favorável, não indicando a velocidade do processo e são necessários experimentos calorimétricos para determinar os valores de ΔH e ΔS de um sistema.

Durante o enovelamento de uma proteína existem diversas vias de enovelamento que levam a proteína a alcançar seu estado nativo. Essas rotas microscópicas alternativas geram estados intermediários que são estáveis mas não enovelados completamente, que vão se assemelhando cada vez mais entre si até o enovelamento ser completo e alcançarem a conformação nativa da proteína. Esse fenômeno é representado pelo modelo em funil rugoso de energia livre (figura 5), o qual está relacionado com as interações intramoleculares e otimizações dessas interações durante o enovelamento da cadeia polipeptídica (Bartlett & Radford, 2009).

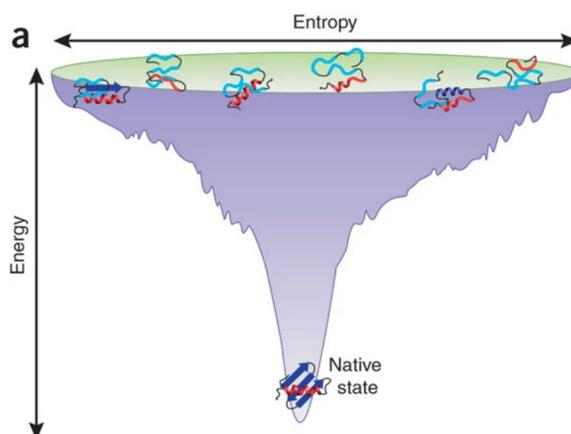


Figura 5. Modelo em funil rugoso de energia livre do enovelamento de proteínas, removido de Bartlett & Radford (2009). O topo do funil representa os mais variados estados conformacionais que uma cadeia polipeptídica pode apresentar quando suas interações intramoleculares não são formadas, sendo necessário alto aporte de energia para mantê-las nessas conformações. O efeito hidrofóbico causa grande redução da energia do sistema, reduzindo os possíveis estados conformacionais do polipeptídeo, e ao passo que as interações intramoleculares são feitas a energia livre do polipeptídeo é reduzida até seu estado de menor energia livre. As rugosidades do funil representam o nível energético de alguns estados conformacionais do polipeptídeo durante o enovelamento, nos quais podem necessitar de algum aporte energético para conseguir aumentar sua energia e “saltar” de uma “vala” e seguir até o seu estado nativo.

Durante o enovelamento dos polipeptídeos as interações intramoleculares aumentam a estabilidade das estruturas secundárias, nas quais induzem o enovelamento de estruturas terciárias até um estado conformacional de menor energia livre (termodinamicamente mais viável). No entanto, no modelo em funil rugoso existem diferentes possibilidades de níveis de enovelamento, nos quais necessitam de alguma energia adicional para prosseguir no enovelamento até o estado nativo. Essas regiões são

denominadas de valas e estão relacionadas à condição do meio em que o polipeptídeo está se enovelando (temperatura, variação de pH, parceiros moleculares, presença de determinados sais, etc.), onde em determinadas condições é necessário a adição de energia para desfazer e refazer, ou reordenar as interações intramoleculares para o polipeptídeo alcançar sua conformação nativa e realizar suas funções. Algumas mutações pontuais podem modificar os possíveis caminhos do enovelamento, podendo criar regiões onde a célula não é capaz de adicionar energia para o enovelamento prosseguir até o estado nativo, causando o aumento de proteínas mal enoveladas com função, dinâmica e estabilidade afetadas (Bartlett & Radford, 2009).

As proteínas não são estruturas estáticas, são moléculas dinâmicas que possuem movimentos contínuos devido às vibrações dos átomos, oscilações das ligações e mudanças entre estados conformacionais de domínios ao longo da catálise (Ellaway, J. 2024). Estes diferentes estados conformacionais se interconvertem, não sendo possível representar uma proteína como uma única estrutura. A população de uma mesma proteína com estrutura heterogênea é conhecida como conjunto estrutural. Tais estados conformacionais não podiam ser distinguidos experimentalmente, mas com o uso da microscopia crioeletrônica é possível analisar a distribuição dos diferentes estados conformacionais presentes em um conjunto estrutural. Algumas mutações pontuais podem induzir mais facilmente a proteína para um estado conformacional onde sua função, dinâmica e estabilidade estejam afetadas (Ellaway, J. 2024).

A atividade de uma proteína pode ser regulada de diversas formas (Lee & Yaffe, 2016). Algumas proteínas são reguladas por alterações nos níveis proteicos através do aumento ou diminuição da transcrição, tradução ou degradação. Podem ocorrer alterações nos níveis de isoformas através da regulação de processos de splicing. Por outro lado, existem regulações na atividade de proteínas que ocorrem após tradução e enovelamento da proteína, denominadas de modificações pós-traducionais, como fosforilação, SUMOilação e acetilação. Essas modificações pós-traducionais estão relacionadas com ativação ou inibição da atividade e alteração da localização celular de proteínas sem necessariamente modificar os níveis proteicos. De forma semelhante, existem regulações alostéricas da atividade de proteínas que ocorrem quando um ligante (molécula ou proteína) interage em uma região da estrutura da proteína e conseqüentemente ocorre uma modificação da conformação de um sítio distante da estrutura da proteína. Tal modificação causa alteração na afinidade por outros ligantes deste sítio distante, aumentando ou reduzindo a atividade da proteína.

1.2. Determinação da estrutura de proteínas através de dados experimentais

Os estudos sobre a estrutura de proteínas geralmente empregam procedimentos que utilizam métodos analíticos associados a modelos matemáticos e programações computacionais, possibilitando a conversão dos dados gerados pelos métodos analíticos em imagens tridimensionais de alta resolução da estrutura de proteínas (Dokholyan, N. 2020). Diversos são os métodos utilizados na obtenção dos dados biofísicos, abrangendo as áreas da bioinformática (procura por sequências de mRNAs através de bancos de dados gerados por sequenciamentos de genomas e transcritos de diversos organismos), biologia molecular (clonagem de uma sequência de mRNA do organismo de interesse, expressão em modelos celulares adequados para o tipo da proteína de escolha, extração e purificação e preparo da proteína para análise) e biofísica (obtenção e tratamento de dados através de métodos como difração de raio X, espalhamento de raio X em baixo ângulo (SAXs), microscopia crioelétrica, ressonância magnética nuclear (RMN), dicroísmo circular, etc) (Hartley, J. 2006; Mohd Salleh, et al. 2022).

A estrutura e dinâmica das proteínas estão relacionadas com as suas funções (Schwartz, S. 2023). Durante a catálise, ou interação entre diferentes proteínas, ocorrem movimentos dos diferentes domínios das proteínas com seus substratos ou interatores. Os grupos funcionais necessários para a catálise se alinham com os grupos funcionais do substrato, alterando a conformação do domínio catalítico e conseqüentemente do substrato. Tais mudanças conformacionais são necessárias para reduzir a energia de ativação da reação, alcançar o estado de transição do substrato e formar o produto (Schwartz, S. 2023). Dessa forma, para melhores conclusões é necessário a utilização de diversos métodos analíticos relacionados à estrutura, dinâmica e atividade das proteínas, onde os dados gerados pelos diferentes métodos são complementares.

Os estudos funcionais e estruturais de proteínas utilizam as diferentes propriedades das radiações eletromagnéticas (Stolar & Smith, 2020). A luz é um tipo de energia com dualidade onda-partícula, onde podemos utilizar as propriedades da luz como uma partícula, denominada de fóton, e podemos utilizar as propriedades da luz como ondas eletromagnéticas. Neste sentido, a espectroscopia é o estudo da interação das radiações eletromagnéticas com a matéria, e existem diferentes tipos de técnicas espectroscópicas utilizadas pelo campo da biologia estrutural. Entre as técnicas mais utilizadas para determinar estruturas de proteínas temos o dicroísmo circular, SAXS, RMN e difração de raio X.

A luz possui um campo elétrico e um campo magnético que oscilam perpendicularmente em picos e vales ao longo da direção que a luz percorre (figura 6A). A frequência (f) da radiação eletromagnética é mensurada através da distância entre os picos da mesma radiação, sendo definido como comprimento de onda (λ). A conversão entre frequência e comprimento de onda pode ser descrita pela equação $f = \frac{c}{\lambda}$, onde c é a velocidade da luz ($3 \times 10^8 \text{ ms}^{-1}$) (Stolar & Smith, 2020; Rouhi, et al. 2022).

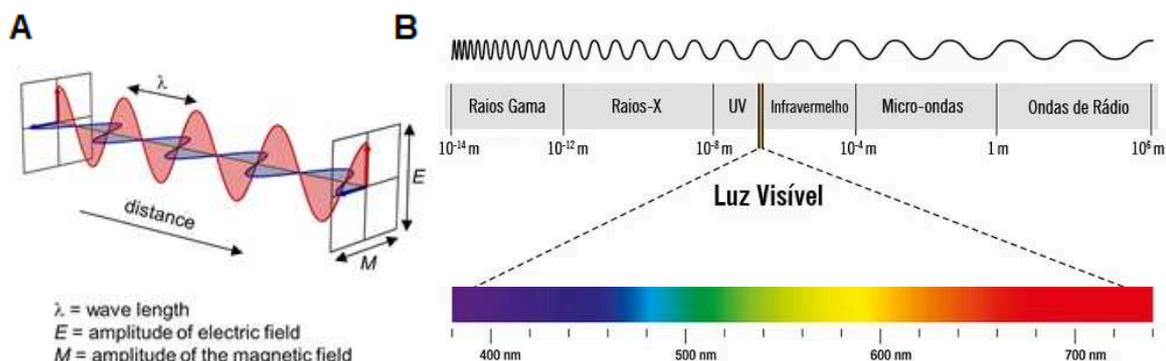


Figura 6. Representação da radiação eletromagnética e espectro eletromagnético, removido de Stolar & Smith (2020). A) Representação dos campos elétrico (E, em vermelho) e magnético (M, em azul) da luz. A distância entre os picos das ondas eletromagnéticas determina o comprimento de onda (λ) da radiação, e a amplitude é a altura dos picos das ondas. B) O espectro eletromagnético é a separação das radiações eletromagnéticas pelo seu comprimento de onda. O nível energético das radiações são inversamente proporcionais ao comprimento de onda, sendo as ondas do espectro das ondas de rádio as menos energéticas e as radiações do espectro do raio gama as mais energéticas. O espectro de luz visível são as radiações que o sistema visual dos humanos é capaz de identificar.

Os fótons são partículas de luz com discreta energia denominada de quanta (Stolar & Smith, 2020). A quantidade de energia carregada por um fóton depende de seu comprimento de onda. Quanto menor o comprimento de onda, maior será a energia carregada pelo fóton. O espectro eletromagnético é o intervalo de todas os comprimentos de ondas eletromagnéticas detectadas no universo até o momento (figura 6B) (Rouhi, et al. 2022). Os diferentes tipos de radiações eletromagnéticas (raio gama, raio X, etc) possuem diversos comprimentos de ondas em sua faixa do espectro eletromagnético. Podemos observar que os raios gama estão em um espectro mais energético do que os raios X, que estão em um espectro mais energético do que a ultravioleta (UV), e assim por diante.

Os níveis de energia da matéria são quantizados e a radiação eletromagnética pode causar transição de um nível energético para outro, onde a energia da luz é absorvida pela matéria. Os elétrons possuem diversos orbitais nos quais possuem um nível de energia associado. Quando uma molécula absorve um fóton com a energia correta (comprimento de onda), um elétron passa de seu estado fundamental para um estado excitado. Isto ocorre se a energia do fóton corresponder à diferença energética entre o estado fundamental e estado

excitado da matéria. O estado excitado é um nível energético superior ao do estado fundamental que o elétron pode alcançar, e é representado como um nível de energia vazio. Sendo assim, a absorção de uma onda eletromagnética por uma proteína converte a energia do fóton em energia interna, respeitando a primeira lei da termodinâmica. Após absorção da energia do fóton e excitação do elétron da molécula, o elétron retorna ao seu estado fundamental e a energia absorvida pelo elétron da matéria é transferida para o solvente na forma de calor (Stolar & Smith, 2020). A região de uma proteína que absorve radiação eletromagnética é denominada de cromóforo. Os aminoácidos aromáticos e as ligações peptídicas são os dois principais cromóforos encontrados nas proteínas (Stolar & Smith, 2020).

As ligações peptídicas são responsáveis pela absorção da radiação UV e infravermelho. A ressonância eletrônica da ligação peptídica possibilita absorver energia de fótons do espectro eletromagnético da faixa do UV, com máxima absorção a 214 nm. Os anéis aromáticos presentes nos radicais dos aminoácidos triptofano, tirosina e fenilalanina também podem absorver radiação eletromagnética na faixa do UV, com máxima absorção a 280 nm. O coeficiente de extinção molar é a capacidade de 1 mol de uma molécula absorver luz em um determinado comprimento de onda. Podemos utilizar a lei de Lambert-Beer, onde $A = \epsilon \cdot l \cdot c$, para determinar a concentração molar (c) de uma molécula através da absorção da radiação eletromagnética (A), seu coeficiente de extinção molar (ϵ) e o comprimento do caminho óptico percorrido pela luz na amostra (l) (Stolar & Smith, 2020).

O dicroísmo circular é uma técnica de espectroscopia de absorção que utiliza radiação eletromagnética circular polarizada para a direita e para a esquerda. Essa técnica infere informações básicas sobre a estrutura secundária da proteína (Rogers, et al. 2019). As radiações eletromagnéticas geradas por alguma fonte luminosa possuem diversos fótons com o eixo do campo elétrico posicionados em todos os planos (360°), nos quais nomeiam a radiação eletromagnética como linear não polarizada. Através de filtros é possível gerar radiação eletromagnética linear polarizada, onde o campo elétrico oscila em um único plano. A combinação de duas radiações eletromagnéticas com os campos elétricos polarizados, uma verticalmente e outra horizontalmente tendo um ângulo de 90° entre os dois planos, em fase (picos e vales sincronizados) e com a mesma amplitude induzem a geração de um campo elétrico que é resultante da média dos dois campos elétricos perpendiculares, o qual oscila a 45° (figura 7A). No entanto, ao combinar duas radiações eletromagnéticas de amplitudes iguais e campo elétrico polarizado, uma verticalmente e outra horizontalmente, mas diferindo na fase por um quarto (diferença entre a posição dos picos e vales em função do tempo), as forças eletromagnéticas resultantes tornam a luz linear polarizada em luz circular polarizada (figura 7B) (Rogers, et al. 2019; Stolar & Smith, 2020).

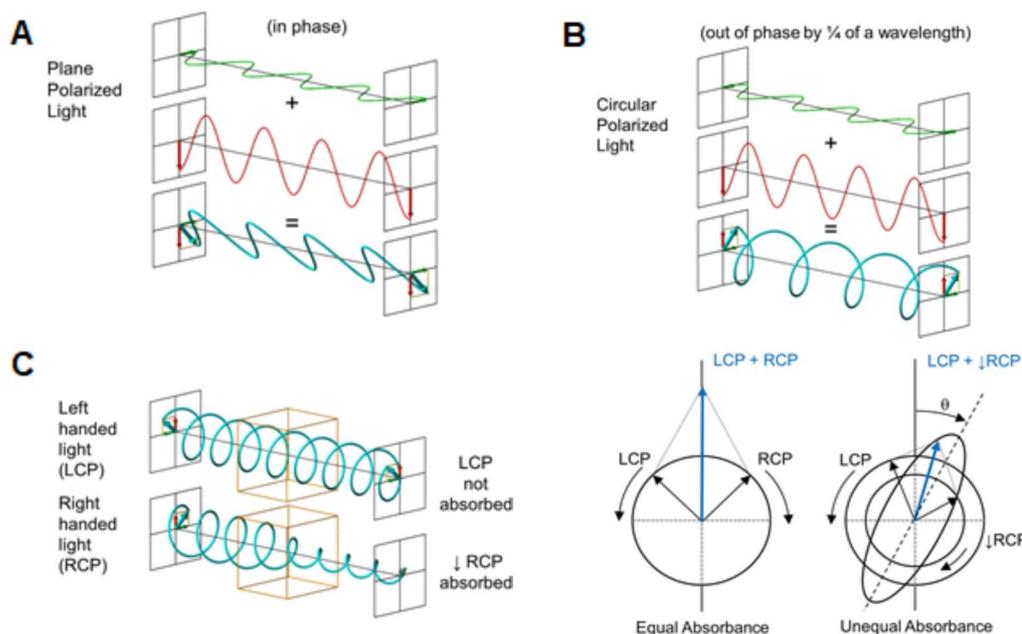


Figura 7. Radiação eletromagnética circular polarizada, removido de Stolar & Smith (2020). A) Exemplo de um campo elétrico de 45° resultante de dois campos elétricos de duas radiações eletromagnéticas polarizadas em fase posicionadas perpendicularmente a 90° . B) Exemplo de campo elétrico circular polarizado resultante entre dois campos elétricos polarizados posicionados perpendicularmente a 45° mas com uma das fases atrasadas por $1/4$ de distância entre os picos. C) Representação da absorção diferencial da luz polarizada para a esquerda e para a direita, e sua representação na forma de vetores pelo detector. LCP = luz circular polarizada para a esquerda, RCP = luz circular polarizada para a direita. θ é o ângulo resultante entre os vetores da LCP e RCP.

A transformação da radiação eletromagnética linear polarizada para circular polarizada é realizada através de um modulador fotoelástico, que rapidamente alterna a direção da rotação da luz circular polarizada para direita (anti horário) e para esquerda (horário), nas quais são projetadas na amostra. Moléculas simétricas absorvem igualmente a luz circular polarizada para a direita e para a esquerda. As moléculas assimétricas ou dicróicas, como as proteínas, absorvem diferencialmente a luz circular polarizada para a direita e para a esquerda. Proteínas que possuem grande quantidade de alfa-hélices, folhas-beta ou sequências intrinsecamente desordenadas em sua estrutura possuem diferenças no padrão da absorbância da luz circular polarizada para a direita e para a esquerda. Neste sentido, esses padrões na absorbância são utilizados para inferir aproximações da quantidade de alfa-hélices, folhas-beta e espirais aleatórias na estrutura das proteínas (Rogers, et al. 2019; Stolar & Smith, 2020).

A passagem da luz circular polarizada para a direita e para a esquerda na amostra são detectadas e representadas como combinação de vetores. Quando não ocorre absorção diferencial da luz circular polarizada para a direita e para a esquerda pela amostra ocorre a geração de um vetor linear, que oscila para cima e para baixo como resultante da interação entre os campos eletromagnéticos (figura 7C). Por outro lado, quando existe absorção

diferencial da luz circular polarizada para a direita e para a esquerda, o vetor resultante entre os campos eletromagnéticos apresenta uma forma elíptica (figura 7C) (Rogers, et al. 2019; Stolar & Smith, 2020).

O ângulo (θ) gerado entre o eixo da elipse em relação ao eixo do vetor linear (figura 7C) é mensurado em miligráus (1:1000 de 1°). Dessa forma, os sinais de dicroísmo circular são expressos em graus de elipticidade (θ), com unidades geralmente expressas em miligráus. A elipticidade molar ($\Delta\epsilon$) leva em consideração a concentração molar da molécula e pode ser calculada através da equação $\Delta\epsilon = \frac{\Delta A}{\text{concentração molar} \times \text{caminho óptico (cm)}}$ com unidades expressas em $\text{graus.M}^{-1}.\text{cm}^{-1}$. A absorção diferencial da luz circular polarizada para a esquerda e direita pode ser calculada por $\Delta A = A_{\text{luz circular polarizada para esquerda}} - A_{\text{luz circular polarizada para direita}}$. Dessa forma, os valores positivos do gráfico de absorbância do espectro de dicroísmo circular são referentes a absorção da luz circular polarizada para a esquerda, enquanto os valores negativos são referentes a absorção da luz circular polarizada para a direita (figura 8) (Rogers, et al. 2019; Stolar & Smith, 2020).

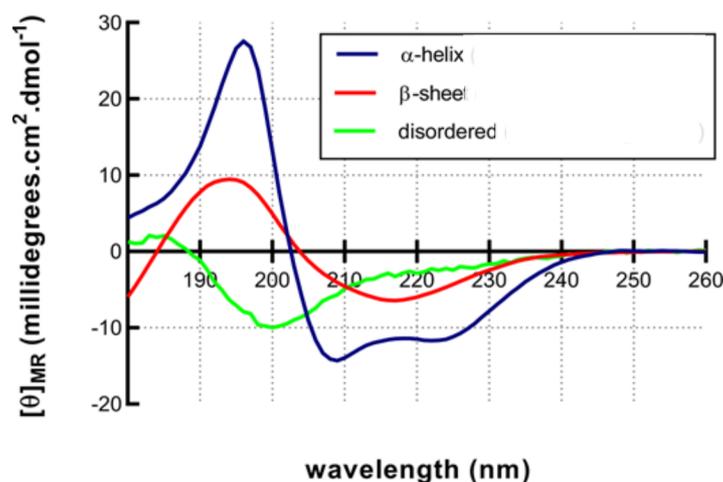


Figura 8. Espectro de dicroísmo circular, removido de Stolar & Smith (2020). Representação do espectro de absorção diferencial da luz circular polarizada para a esquerda e para a direita de 3 proteínas com estruturas compostas 100% de alfa-hélices (vermelho), folhas-beta (azul) e segmentos desordenados (verde).

Por outro lado, a elipticidade molar residual média ($\Delta\epsilon_{MR}$) é utilizada para reportar a elipticidade molar de proteínas individuais ao invés de proteínas no geral, onde é utilizado a concentração residual média que é definida pela equação $\Delta\epsilon_{mr} = \text{concentração molar} \times n^\circ \text{ de resíduos da proteína}$. A concentração residual média é utilizada na equação $\Delta\epsilon = \frac{\Delta A}{\Delta\epsilon_{mr} \times \text{caminho óptico (cm)}}$, onde substitui a concentração molar na equação que determina a elipticidade molar ($\Delta\epsilon$). Para conversão da ΔA para θ nós utilizamos a equação $\theta = 32,982 \times \Delta A$, e podemos obter a elipticidade molar residual

média ($[\theta]_{MR}$) através da equação $[\theta]_{mr} = \frac{100 \times \theta \text{ (em miligráus)}}{\Delta \epsilon_{mr} \times \text{caminho óptico (cm)}}$, onde as unidades são expressas em miligráus.cm².dmol⁻¹ (figura 8), equivalente a miligráus.m⁻¹.M⁻¹ (devido o fator 100 da equação converter centímetros para metros) (Rogers, et al. 2019; Stolar & Smith, 2020).

É necessário obter um espectro de referência (curva padrão) de dicroísmo circular de proteínas com 100% de folhas-beta, alfa-hélices e espirais aleatórias em suas estruturas, e com proteínas de diferentes proporções dessas estruturas secundárias. Sendo assim, é realizada uma varredura da absorvância da luz circular polarizada em diferentes comprimentos de onda do espectro de UV, de 200 a 280 nm. Proteínas com 100% de alfa-hélices em sua estrutura apresentam um pico positivo da $[\theta]_{MR}$ à ~190 nm e dois picos negativos da $[\theta]_{MR}$ à ~208 nm e ~222 nm (figura 8). Já proteínas com 100% de folhas-beta apresentam um pico positivo da $[\theta]_{MR}$ à ~195 nm (menor em relação ao pico observado em proteínas formadas por alfa-hélices) e um pico negativo a ~220 nm (figura 8). Por outro lado, proteínas intrinsecamente desordenadas que são 100% formadas por espirais aleatórias apresentam um baixo sinal negativo da $[\theta]_{MR}$ à ~195 nm, e apresentam sinais muito baixos em comprimentos de onda acima de 210 nm (figura 8). Posteriormente, é realizada a deconvolução do espectro de dicroísmo circular, um processo matemático que infere as frações relativas de cada estrutura secundária na estrutura da proteína, somando diferentes combinações dos espectros de referência (Rogers, et al. 2019; Stolar & Smith, 2020).

A absorvância da luz circular polarizada nos espectros de 250 a 350 nm pelas cadeias laterais dos resíduos de fenilalanina, tirosina e triptofano podem ser utilizadas para nos dar informações limitadas sobre a estrutura terciária, nos dando indicativos do quão empacotado estão as estruturas secundárias da proteína. Dessa forma, é possível utilizar o dicroísmo circular para inferir como alterações no ambiente (pH, solvente, calor, etc), presença de ligantes (inibidores, íons, DNA, RNA, polissacarídeos, etc) e mutações na sequência polipeptídica alteram as estruturas secundárias e terciárias das proteínas, sem entrar em detalhes no nível atômico devido ao tipo da radiação eletromagnética utilizada (espectro de UV) (Stolar & Smith, 2020).

O raio X possui comprimentos de onda de 1 nm até 5 pm, que são distâncias próximas ao comprimento de ligações covalentes. Dessa forma é possível discriminar a distância entre átomos de uma proteína utilizando o raio X. Existem técnicas que utilizam o raio X em amostras em solução e na forma de cristais (sólida) para determinar aspectos estruturais de proteínas. Os resultados dos experimentos que utilizam raio X em amostras de proteínas em solução não possibilitam a determinação da estrutura em alta resolução, devido às proteínas estarem em diversas posições dispersas no espaço. Já as amostras de

proteínas cristalizadas possibilitam a determinação da estrutura em alta resolução por estarem ordenadas de forma repetitiva no espaço (Stolar & Smith, 2020).

O espalhamento de raio X a baixos ângulos (*Small-Angle X-ray Scattering* - SAXS) utiliza feixes de raio X monocromático e colimado, onde é analisado o espalhamento da onda eletromagnética junto a uma varredura de baixo ângulo (centésimo de 1°) da amostra em solução (figura 9A) (Boldon, L., Laliberte, F. & Liu, L., 2015; Jeffries, et al. 2021).

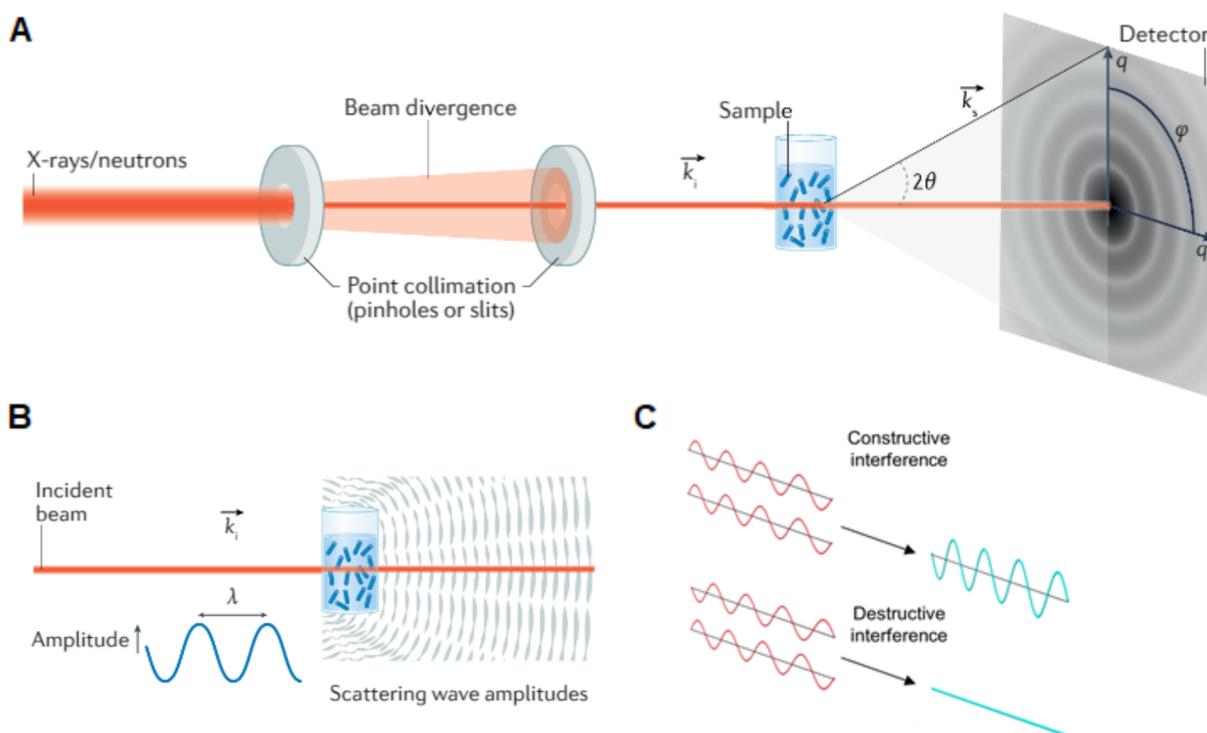


Figura 9. Espalhamento de raio X a baixos ângulos, removidos de Jeffries et al. (2021) e Stolar & Smith (2020). A) Um feixe de radiação eletromagnética monocromática no espectro do raio X é colimado e incidido em uma amostra. Os colimadores limitam o campo de incidência da radiação eletromagnética, direcionando o feixe através da absorção de parte da radiação. Como o SAXS é um método de contraste, é necessário haver diferença na densidade eletrônica entre a amostra e o solvente na aquisição dos dados. O espalhamento identificado pelo detector é analisado em função do ângulo de espalhamento do feixe de raio X. B) Espalhamento da radiação de forma coerente do feixe de radiação monocromática irradiado na amostra. C) Interferências construtivas e destrutivas ocorrem entre a interação de duas radiações eletromagnéticas espalhadas de forma coerente. Quando as radiações espalhadas de forma coerente estão em fase ocorre a interferência construtiva, causando o aumento da amplitude da onda eletromagnética. Quando as radiações espalhadas de forma coerente não estão em fase ocorre a interferência destrutiva, causando a redução da amplitude da onda eletromagnética.

A interação entre o raio X e os elétrons da matéria causam o espalhamento do raio X de forma elástica (figura 9B). Quando ocorre o espalhamento da radiação de forma elástica sem alterar o comprimento de onda, o espalhamento é denominado de coerente. O espalhamento coerente causa interferências construtivas e destrutivas na amplitude da onda eletromagnética, que estão relacionadas à fase, amplitude e comprimento de onda da radiação espalhada elasticamente. Quando duas radiações eletromagnéticas estão em fase (picos e vales no mesmo espaço ao longo do tempo) elas sofrem interferência construtiva, causando aumento na amplitude da radiação eletromagnética resultante (figura 9C). No

entanto, quando duas ondas eletromagnéticas não estão em fase (picos e vales em espaços diferentes ao longo do tempo) elas sofrem interferência destrutiva, causando queda na amplitude da radiação eletromagnética resultante (figura 9C). Dessa forma, é necessário passar a radiação eletromagnética em um monocromador para selecionar os comprimentos de onda e possibilitar o espalhamento coerente (Boldon, L., Laliberte, F. & Liu, L., 2015; Stolar & Smith, 2020; Jeffries, et al. 2021).

Sendo assim, um feixe de raio X monocromático colimado (K_i) é incidido sobre a amostra onde uma porção da radiação interage com a molécula e gera eventos de espalhamento (K_s), que são detectados em um detector 2D (figura 9A). A transmissão geométrica do espalhamento é analisada em termos do padrão de intensidade, $I(q)$, e o ângulo de rotação (φ) em torno das coordenadas do centro do feixe de raio X (Jeffries, et al. 2021). O padrão de intensidade I é expresso em função da magnitude do vetor de transferência do momentum q (figura 9A e 10), sendo $I(q)$, onde $q = \frac{4\pi\sin\theta}{\lambda}$, e θ é o ângulo de espalhamento e λ é o comprimento de onda da radiação (figura 9A). Como o espalhamento coerente é dependente da quantidade de elétrons na partícula, e as proteínas possuem tamanho e quantidade de átomos diferentes, e conseqüentemente, diferentes quantidades de elétrons, cada proteína apresenta um espalhamento coerente intrínseco que é refletido na geometria do espalhamento no detector (Jeffries, et al. 2021).

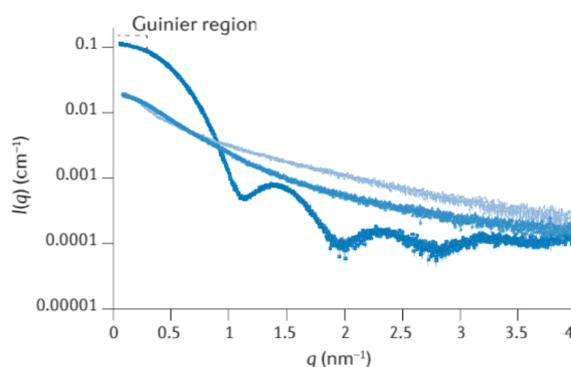


Figura 10. Plot da intensidade do espalhamento de raio X a baixos ângulos, removido de Jeffries et al. (2021). Está sendo representado o espalhamento de 3 proteínas distintas. Região utilizada na relação Guinier está demarcada pela linha tracejada.

Após realizar uma varredura do espalhamento do raio X em baixos ângulos, é obtido um perfil de decaimento do espalhamento de raio X (figura 10) o qual possibilita inferir informações sobre o volume (tamanho), área (formato) e rugosidade da partícula de 1 a 1000 nm (Boldon, L., Laliberte, F. & Liu, L., 2015; Jeffries, et al. 2021). O SAXS também extrai parâmetros estruturais de proteínas como raio de giro (R_g), dimensão máxima da partícula (D_{max}), e possibilita modelar o envelope molecular com baixa resolução e visualizar regiões desordenadas (Jeffries, et al. 2021).

Na relação Guinier (figura 10) são extraídas informações do tamanho e R_g da molécula. Quando $q = 0$, $I(0)$, é possível inferir o tamanho da partícula através do quadrado do seu volume, onde quanto maior for o sinal da intensidade maior é a partícula (Jeffries, et al. 2021). Por outro lado, a derivada do slope negativo do plot de espalhamento (inclinação da curva) pelo plot de Guinier pode determinar o R_g da partícula. É interessante notar que duas proteínas podem possuir R_g semelhantes mas com tamanhos diferentes. Dessa forma, uma proteína com R_g semelhante a uma outra proteína de tamanho maior, nos indicam que a primeira proteína possui uma estrutura mais estendida e não globular em relação à segunda (Jeffries, et al. 2021).

O plot de Kratky infere informações sobre a estrutura e D_{max} da partícula. O formato da curva do plot de Kratky infere conformações da partícula, onde um aumento da intensidade seguida de um platô indica que a proteína possui estrutura desordenada, por outro lado, um pico distinto (aumento e queda na intensidade) indica que a proteína possui estrutura compactada ou enovelada (figura 11A) (Jeffries, et al. 2021).

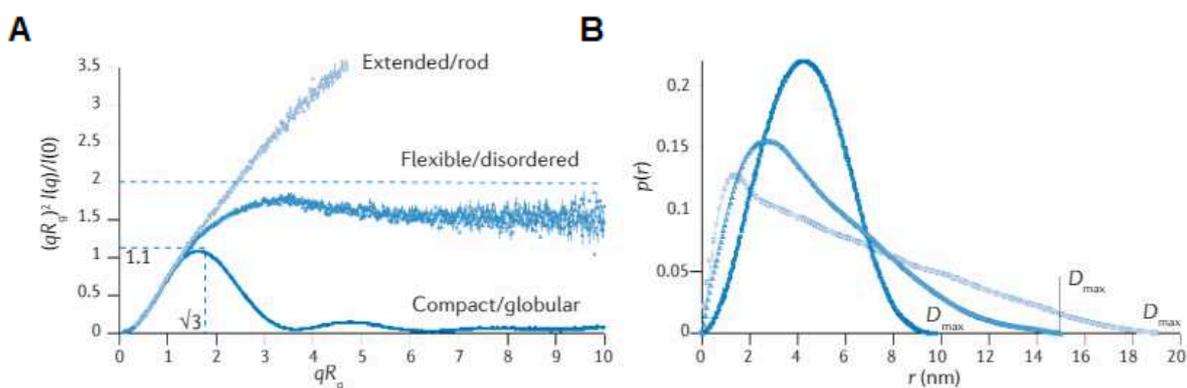


Figura 11. Plot de Kratky dos dados de espalhamento e plot da dimensão máxima da proteína, removido de Jeffries et al. (2021). A) Plot de Kratky utilizado para determinar o formato da proteína. B) Plot de Kratky da frequência provável de distâncias no espaço real para determinar a dimensão máxima da proteína.

A D_{max} é inferida pela frequência provável de distâncias no espaço real $p(r)$, sendo adquirida pela transformada inversa de Fourier da função $I(q)$, onde a I que está em função de q (figura 10) é transformada em frequência de distribuição p , no qual está em função da distância no espaço real (r). Uma vez determinada a função $p(r)$ é possível determinar a dimensão máxima da proteína em nanômetros (figura 11B) (Jeffries, et al. 2021).

O SAXS fornece informações sobre as distâncias de todos os pares de elétrons, onde essas informações não são acuradas o suficiente para gerar um modelo atômico. No entanto, é possível fornecer restrições espaciais para validar modelos estruturais, estados oligoméricos, flexibilidade e formato de proteínas (Jeffries, et al. 2021). É possível utilizar os dados de SAXS para obter o envelope molecular de proteínas através de métodos que calculam a área da proteína utilizando átomos fictícios *ab initio* (DAMMIN), e resíduos fictícios *ab initio* (GASBOR), por exemplo. O envelope molecular pode ser utilizado para

gerar limites na estrutura de proteínas modeladas computacionalmente, onde diversos modelos gerados são alocados dentro da área gerada pelos dados de SAXS. Para ajustar os modelos dentro do envelope molecular é realizado modelagem de corpo rígido, onde diversos modelos se encaixam dentro do envelope molecular gerado através dos dados de SAXS, demonstrando que uma proteína pode assumir diversas formas quando está em solução, nas quais possuem relação com a sua dinâmica (Jeffries, et al. 2021).

No entanto, o método que é padrão ouro na determinação da estrutura de alta resolução de proteínas é a cristalografia de raio X (Stolar & Smith, 2020). Diferentemente do difractoímetro circular, SAXS e RMN que utilizam amostras em solução no estado líquido, a difração de raio X utiliza amostras de proteínas purificadas e fixadas através de cristalização. O método mais comum para cristalização de proteínas é a difusão de vapor, denominada de *hanging drop*. Uma gota de solução tamponada de proteína purificada altamente concentrada é alocada no topo de um reservatório selado, que contém solução tamponada em seu interior. A osmolaridade do tampão presente na gota de proteína purificada é menor do que a osmolaridade da solução no reservatório, dessa forma, o equilíbrio osmótico irá ocorrer através da água no estado gasoso. Durante a osmose ocorre o crescente aumento da concentração de proteínas na gota, que inicia a organização entre as moléculas de proteína e por fim a formação de um cristal (Hou, et al. 2019).

A cristalização de proteínas para difração de raio X possui limitações em relação a resolução de segmentos móveis presentes na estrutura, além de tornar difícil a cristalização de proteínas transmembranas, intrinsecamente desordenadas e complexos multiproteico. Proteínas transmembrana apresentam grandes dificuldades de expressão e purificação, devido aos segmentos transmembranares serem hidrofóbicos. Proteínas intrinsecamente desordenadas e complexos proteicos não formam cristais de difração facilmente, devido os segmentos móveis e a presença de diferentes polipeptídeos não permitirem a formação de celas unitárias por não formarem segmentos com periodicidade (Stolar & Smith, 2020).

Na difração de raio X, o componente elétrico da radiação eletromagnética do raio X interage com a nuvem eletrônica dos átomos das moléculas cristalizadas, gerando a difração da radiação. As reflexões (sinais que chegam ao detector na forma de “pontos”) são resultantes da difração do raio X em determinados ângulos entre o feixe de radiação e o cristal, e respeitam as leis geométricas da interferência construtiva descrita por Bragg (figura 12A) (Stolar & Smith, 2020). Apenas uma parte dos pontos de reflexão chegam ao detector, o qual possui forma planar e causa o padrão circular dos sinais no detector. As intensidades dos pontos de reflexão são determinadas pela quantidade de elétrons que existem em uma série de planos paralelos imaginários, denominados de planos de Bragg (figura 12B). Dessa forma, as reflexões das radiações eletromagnéticas irão ocorrer quando os elétrons

equivalentes repetidos em um cristal forem encontrados em um conjunto de planos com distância interplanar adequada.

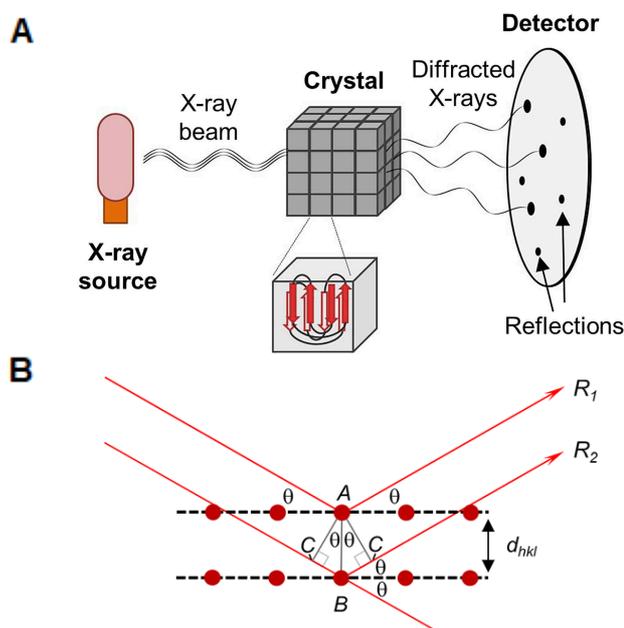


Figura 12. Difração de raio X, removido de Stolar & Smith (2020). A) Set up da cristalografia de raio X. B) Duas ondas em fase (R_1 e R_2) são espalhadas em um ângulo θ . Quando a distância percorrida por R_2 no segmento ABC for múltiplo inteiro do comprimento de onda, as ondas R_1 e R_2 ficarão em fase e gerarão interferência construtiva, aparecendo no sinal de reflexão. E quando a distância percorrida por R_2 não for múltiplo inteiro do comprimento de onda, as ondas R_1 e R_2 não ficarão em fase e gerarão interferência destrutiva, reduzindo o sinal de reflexão. Os pontos vermelhos são os átomos periódicos das células unitárias. d é a distância entre os planos de Bragg e hkl são os três planos dimensionais. θ é o ângulo entre a radiação eletromagnética e os elétrons presentes nos planos de Bragg durante a difração.

Na equação $n\lambda = 2d \cdot \text{sen}\theta$, d é a distância entre os planos paralelos e θ é o ângulo formado entre o feixe de raio X e os planos de Bragg. A Lei de Bragg determina que as interferências construtivas das radiações eletromagnéticas ocorrem apenas para certos valores de d e θ . Dessa forma, as interferências construtivas ocorrem quando o termo $2d \cdot \text{sen}\theta$ for múltiplo inteiro do comprimento de onda do raio X, ou seja, quando $n\lambda$ for 1λ , 2λ , 3λ , 4λ e assim por diante (figura 12B) (Stolar & Smith, 2020). É por isso que o cristal sofre rotações e o feixe de raio X muda de ângulo em relação ao cristal, para que possam ser encontrados os pontos que irão ocorrer interferências construtivas. Quanto menor for a distância entre os planos de Bragg, maior será o ângulo entre o raio X e o cristal para poder ocorrer a difração. Sendo assim, estruturas de alta resolução de proteínas que possuem grandes quantidades de elétrons próximos uns aos outros, necessitam de cristais que difratam a altos ângulos (Wlodawer, et al. 2008, 2013; Stolar & Smith, 2020).

A intensidade dos pontos de reflexão gerados pela difração são convertidos em densidades eletrônicas através da transformada de Fourier, utilizando a amplitude e fase de cada reflexão. Devido não ser possível inferir a fase das radiações difratadas que geram os pontos de reflexões, não é possível realizar a transformada de Fourier com os primeiros

dados da difração, isso é denominado de problema de fase (Wlodawer, et al. 2008, 2013; Stolar & Smith, 2020).

Para resolver o problema de fase é realizado “substituição molecular” (*molecular replacement*), onde é utilizado modelos estruturais de proteínas que assumem estrutura muito similar à estrutura da proteína cristalizada, que é desconhecida. É realizado o posicionamento da estrutura da proteína conhecida em todas as orientações possíveis dentro do cristal, gerando diversos modelos cristalográficos, onde é predito o padrão de difração de cada modelo (através de transformada de Fourier) e comparado com o padrão de difração gerado pelo cristal. Uma vez que os padrões de difração do modelo satisfazem as corretas orientações identificadas no cristal, o modelo é considerado cristalizado em um cristal desconhecido e os valores de fase gerados pelo modelo são utilizados. Dessa forma, através da transformada de Fourier utilizando os valores da amplitude (determinada experimentalmente) e os valores de fase (resolvidas por substituição molecular), é determinado o mapa de densidade eletrônica inicial (Stolar & Smith, 2020).

Posteriormente, a cadeia polipeptídica da proteína cristalizada é alocada dentro do mapa de densidade eletrônica inicial para construção de um modelo atômico (figura 13), onde são calculados as reflexões e fases teóricas utilizando a transformada de Fourier e é criado um novo mapa de densidade eletrônica (Wlodawer, et al. 2008, 2013; Stolar & Smith, 2020). Dessa forma, a fase está aprimorada por ser estimada da cadeia polipeptídica da proteína cristalizada, e não de uma proteína com estrutura semelhante. Em seguida, a cadeia polipeptídica da proteína cristalizada é alocada dentro do novo mapa de densidade eletrônica, sendo realizada uma nova transformada de Fourier e novamente criado um novo modelo atômico. Esse processo é denominado de refinamento, e os ciclos de refinamento são realizados até que não haja mais melhorias no modelo. Nesta etapa é possível definir os solventes, ligantes e corrigir erros como mal alocamento e sobreposição de átomos (Wlodawer, et al. 2008, 2013; Stolar & Smith, 2020).

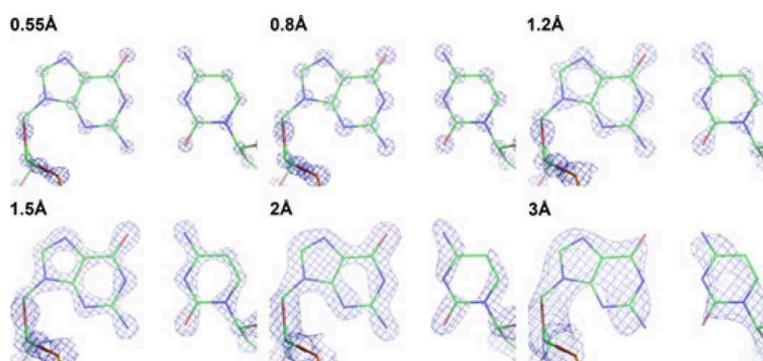


Figura 13. Mapa de densidade eletrônica com moléculas alocadas com diferentes resoluções, removido de Wlodawer, et al.(2013). A resolução do modelo cristalizado está relacionado ao quanto a molécula se encaixa dentro do mapa de densidade eletrônica (região azul quadriculada) e a menor distância necessária para distinguir dois átomos dentro do modelo. Å = Angstroms.

Após o refinamento, o modelo é validado através da análise dos ângulo diedros e posteriormente as coordenadas x,y e z dos átomos da proteína são depositados em um banco de dados de proteína (Protein Data Bank - PDB). Caso o modelo estrutural construído com dados de difração de raio X não sejam favoráveis pelas análises dos ângulos diedros, pode-se retornar na construção do modelo através da nuvem eletrônica ou realizar novas tentativas de cristalização (Wlodawer, et al. 2008, 2013). Estão sendo realizadas a formação de diversos cristais de uma mesma proteína com diversas mutações pontuais, para ser identificado as alterações nos cristais decorrentes das mutações (Wlodawer, et al. 2008, 2013; Stolar & Smith, 2020).

Devido ao processo de cristalização, a estrutura da proteína é semelhante a uma “foto”, onde é perdido informações sobre a sua dinâmica. Por outro lado, a associação dos dados de SAXS ou RMN com a cristalografia pode nos indicar níveis de movimentação. Como a cristalografia não possui boa resolução de segmentos desordenados, nos quais muitas vezes são removidos da construção para facilitar o processo de cristalização, a associação da estrutura de uma proteína determinada por cristalografia com o envelope molecular obtido por SAXS pode nos dar a dimensão da organização do segmento desordenado no espaço em relação ao segmento determinado por cristalografia (Stolar & Smith, 2020; Jeffries, et al. 2021).

Os modelos estruturais de proteínas gerados por SAXS, RMN, difração de raio X, microscopia crioelétrica e modelos computacionais (descritos posteriormente) precisam passar por validações (Wlodawer, et al. 2008, 2013; Stolar & Smith, 2020). A validação dos modelos pode ser através da análise dos ângulos diedros (ϕ e ψ) da cadeia polipeptídica pelo gráfico de Ramachandran (figura 14).

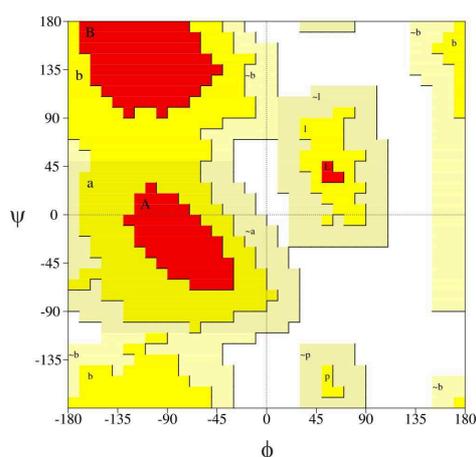


Figura 14. Plot de Ramachandran, removido do PDBsum (<https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>). As regiões delimitadas em vermelho são regiões favoráveis dos ângulos ϕ e ψ encontradas em estruturas secundárias, como folhas-beta (B), alfa-hélices orientadas para a mão direita (A) e mão esquerda (L). As regiões delimitadas em amarelo (b, a, e l) são para as mesmas estruturas secundárias descritas anteriormente, mas são regiões em que os ângulos ϕ e ψ são menos permissivas. Em ocre estão delimitadas as regiões em que os ângulos ϕ e ψ são menos favoráveis em relação às outras regiões. Já a região em branco são combinações de ângulos de torção que não são favoráveis de ocorrer na estrutura de uma proteína.

Os ângulos de torção foram definidos por Ramachandran, que considera as propriedades físico-químicas do grupo radical e sua interação com resíduos adjacentes. Os valores mais e menos favoráveis entre os ângulos ϕ e ψ encontrados em estruturas secundárias (alfa-hélice, folhas-beta e alças) de diferentes cadeias polipeptídicas determinadas experimentalmente foram identificados e representados dentro de um plano cartesiano. O eixo x representa a variação do ângulo ϕ e o eixo y representa a variação do ângulo ψ (Hollingsworth & Karplus, 2010; Wlodawer, et al. 2008, 2013).

O gráfico de Ramachandran gerado pelo PDBsum (<https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>) é dividido em 4 regiões (figura 14): região favorável (vermelho = A, B e L), região adicionais permissivas (amarelo = a,b e I), regiões permissivas (ocre = ~a, ~b, ~l e ~p) e regiões desfavoráveis/não permissivas (branco). O gráfico de Ramachandran dos aminoácidos isoleucina e valina são separados por possuírem ângulos ϕ e ψ mais restritos, isso ocorre devido serem alifáticos, ramificados e com cadeias laterais volumosas. O mesmo ocorre para o aminoácido prolina devido a sua estrutura cíclica, o qual confere ângulos ϕ e ψ mais restritos e que geralmente estão localizados em alças na estrutura de proteínas. A glicina, que é o aminoácido mais simples por possuir apenas um hidrogênio na cadeia lateral, apresenta baixo impedimento estérico e dessa forma possui alto grau de liberdade conformacional em relação aos demais aminoácidos. O ângulo ômega (Ω) é formado entre a ligação peptídica, onde a maioria dos aminoácidos possuem ângulo Ω de 180° , o qual determina sua forma trans. No entanto a prolina possui duas configurações que são relacionadas ao ângulo Ω , a cis-prolina e a trans-prolina. Na cis-prolina ocorre a formação de um ângulo Ω de 0° e na trans-prolina um ângulo Ω de 180° dentro da ligação peptídica. A pré-prolina se refere ao aminoácido que antecede um resíduo de prolina. Como o grau de liberdade da prolina é baixo, os ângulos ϕ e ψ do resíduo de aminoácido que antecede a prolina dentro de uma cadeia polipeptídica sofrem influência da prolina. Dessa forma, os ângulos ϕ e ψ dos resíduos pré-prolina são mais restritos (Hollingsworth & Karplus, 2010; Wlodawer, et al. 2008, 2013).

Quanto maior o número de ângulos ϕ e ψ dentro regiões favoráveis, maior é a qualidade do modelo experimental (ex. difração de raio X) ou hipotético (ex. modelagem por homologia) da estrutura de alta resolução da proteína. Para a validação do modelo pela análise do plot de Ramachandran o ideal é que $\geq 98\%$ dos resíduos estejam dentro de regiões favoráveis, ou que $\geq 99,8\%$ dos resíduos estejam em regiões permissivas (Hollingsworth & Karplus, 2010; Wlodawer, et al. 2008, 2013). O gráfico de Ramachandran pode ser gerado pelo PROCHECK (através da plataforma PDBsum - <https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>) e MOLPROBITY (<http://molprobity.biochem.duke.edu/>).

1.3. Determinação da estrutura de proteínas por modelagem

Determinar a estrutura de uma proteína é fundamental, pois a estrutura amplamente determina a função da proteína além de nos dar informações sobre regiões de interação proteína-proteína, proteína-DNA/RNA, que podem ser promissores alvos farmacêuticos (Senior, et al. 2020). Contudo, a estrutura de proteínas é muito laboriosa de ser determinada experimentalmente, e algumas proteínas podem ter estruturas complexas de serem determinadas.

Anfinsen observou que ao desnaturar a ribonuclease com agentes redutores e/ou calor, ela possui a capacidade de renaturar e voltar na conformação nativa. Isso demonstrou que a sequência polipeptídica ou a estrutura primária contém os aspectos necessários para o enovelamento e estrutura tridimensional da proteína. Dessa forma, as maiores questões na predição estrutural de proteínas estão em como uma sequência polipeptídica se enovela a partir da interação entre seus aminoácidos, qual é a cinética do enovelamento e como prever a estrutura tridimensional de uma proteína a partir de sua sequência primária (Seki, M. 2023).

Os métodos utilizados nas modelagens computacionais da estrutura de alta resolução de proteínas são escolhidos de acordo com a sequência polipeptídica de interesse e os dados relacionados a outras proteínas com estruturas resolvidas. Os métodos mais utilizados são o *ab initio* (ou *de novo*), threading e comparativa (homologia) (Senior, et al. 2020; Huang, et al. 2023).

Devido à estrutura de uma proteína ser mais conservada do que a sequência durante o processo evolutivo, proteínas homólogas compartilham semelhanças estruturais. Dessa forma, é possível modelar a estrutura de uma proteína utilizando estruturas determinadas de proteínas homólogas. Uma forma de realizar buscas por homólogos de uma proteína é através do alinhamento múltiplo de sequências polipeptídicas, onde uma identidade entre as sequências acima de 30% é o suficiente para tornar a proteína homóloga um bom template para modelagem. No entanto, quanto maior for a identidade entre as sequências polipeptídicas, melhores serão os modelos gerados (Huang, et al. 2023).

A modelagem comparativa ou por homologia é o método mais aceito para se determinar a estrutura de alta resolução de polipeptídeos computacionalmente. Ela é realizada através da busca por sequências polipeptídicas homólogas à sequência de interesse que possuem estrutura de alta resolução definidas experimentalmente. Dessa forma, é utilizada a estrutura determinada experimentalmente como molde, onde os aminoácidos do polipeptídeo de interesse são ajustados dentro das restrições determinadas pela estrutura da proteína homóloga. Quanto maior for a identidade e similaridade entre as cadeias polipeptídicas do template com a sequência de interesse, maiores são as chances

de se obter um modelo estrutural de alta resolução com boa confiabilidade, pois os aminoácidos são ajustados de acordo com suas propriedades físico-químicas (Muhammed & Aki-Yalcin, 2019; Huang, et al. 2023).

O Threading ou reconhecimento de enovelamento, é um método utilizado para modelagem por homologia de sequências que não possuem boa identidade com sequências de proteínas que possuem estrutura resolvida experimentalmente (entre 20 e 30% de homologia). Nesse método é realizada uma análise de possíveis estruturas secundárias formadas dentro da sequência polipeptídica, e posteriormente é procurado por padrões de estruturas secundárias semelhantes dentro dos modelos determinados experimentalmente. O modelo é gerado então pela orientação das estruturas secundárias das duas estruturas (Cao, et al. 2004).

Os métodos *ab initio* ou *de novo* são utilizados para modelagem de sequências polipeptídicas que não possuem homologia com nenhuma cadeia polipeptídica que possui estrutura determinada experimentalmente (abaixo de 20% de homologia). Os métodos *ab initio* não utilizam template no enovelamento, apenas é iniciado o dobramento do polipeptídeo em estruturas secundárias e terciárias considerando apenas as propriedades físico-químicas de cada aminoácido e o modelo em funil rugoso de energia livre (Bartlett & Radford, 2009; Huang, et al. 2023). Sendo assim, as predições de estruturas de proteínas são realizadas através da simulação do enovelamento e de funções que minimizam a energia, baseadas na tendência da proteína adotar sua conformação estrutural com a menor energia livre (Huang, et al. 2023).

Dessa forma, para serem realizadas predições com acurácia é necessário o entendimento sobre os aspectos biofísicos do enovelamento de proteínas, e a relação entre a sequência de aminoácidos e a estrutura nativa da proteína (baseada em busca por homologia). No entanto, uma análise da história evolutiva das proteínas através de alinhamentos múltiplos de sequências polipeptídicas disponibilizam informações nas quais podem inferir a estrutura nativa da proteína. É possível identificar resíduos de aminoácidos importantes para a estabilidade da proteína, onde resíduos que fazem contato para estabilizar a estrutura do domínio de uma proteína são co-evolutivos, ou seja, estão presentes nas sequências dos mesmos domínios de proteínas ortólogas e parálogas (figura 15) (Huang, et al. 2023).

Esses métodos que utilizam predições de contatos também exploram a comutação de pares de resíduos que ocorreram durante a evolução, onde dois pares de resíduos importantes para a estabilidade da estrutura da proteína tem uma forte tendência em comutar para outros dois pares de resíduos que também fazem interação forte o suficiente para manter a estrutura da proteína (figura 15). Dessa forma, a co-evolução de resíduos e as covariações evolutivas são encontradas através de alinhamento múltiplo de sequências

polipeptídicas, onde mudanças correlacionadas de dois aminoácidos podem inferir contato entre esses resíduos (Huang, et al. 2023).

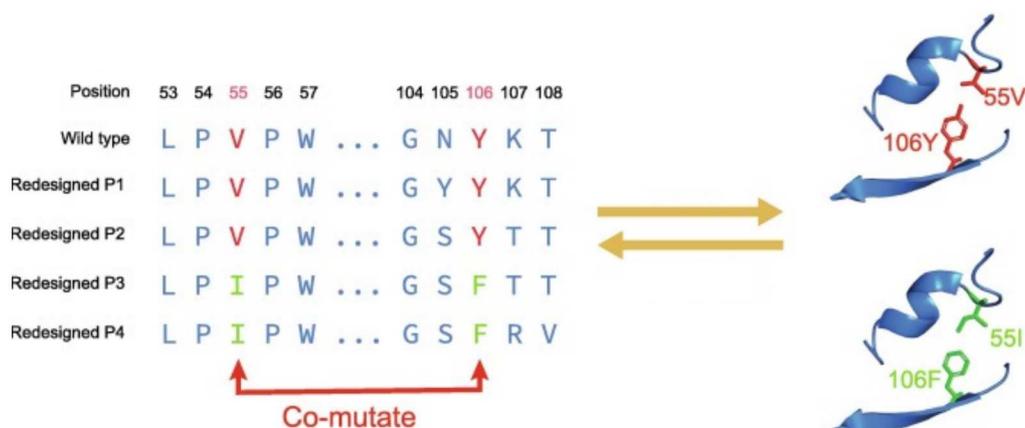


Figura 15. Exemplo de coevolução e comutações evolutivas de resíduos de aminoácidos dentro de cadeias polipeptídicas utilizadas pelos métodos de predição de estrutura de proteínas, removido de Huang, et al. (2023). Neste exemplo os resíduos de tirosina 106 e valina 55 fazem interação e estabilizam a estrutura da proteína de três organismos, e sofreram comutação para fenilalanina 106 e isoleucina 55 em outros dois organismos, nos quais podem também realizar interação e estabilizar a proteína.

Interessantemente, as predições de distância entre os resíduos fornecem informações mais específicas sobre a estrutura em relação às predições de contato (Senior, et al. 2020). Os modelos estruturais gerados computacionalmente nos permitem encontrar uma série de distâncias entre os aminoácidos, que possibilitam a observação de regiões da cadeia polipeptídica que estão próximas quando a estrutura está enovelada (figura 16). Resíduos são definidos como próximos quando dois átomos de carbono β (carbono do radical ligado ao carbono α) possuem a uma distância de 8 Å ou menos entre eles. Essas distâncias são plotadas em um gráfico da sequência polipeptídica pela sequência polipeptídica, sendo gerado um contatograma (figura 16). O contatograma gerado pelo modelo estrutural computacional é comparado aos contatogramas de outras proteínas com estrutura determinada experimentalmente, onde os contatos são utilizados na predição da estrutura para modificar o potencial estatístico de enovelamento e o processo de guia do enovelamento. Como muitas distâncias são analisadas, podem ser associadas às informações relacionadas às distâncias com informações de covariações, estruturas locais e identidade de resíduos adjacentes.

As estruturas modeladas que respeitam as distâncias preditas são geradas através da construção de um potencial total, que representa as melhores conformações em função dos ângulos de torção ϕ e ψ gerados pelas restrições impostas pelas distâncias. Um termo de torção e impedimentos estéricos são adicionados ao potencial total, nos quais estão em função dos ângulos ϕ e ψ e podem ser otimizados por gradiente descendente (algoritmo de otimização que minimiza algumas funções). Através do gradiente descendente realizado

pelo pareamento com estruturas resolvidas, são encontradas uma série de ângulos de torção que minimizam o potencial específico da proteína. As estruturas secundárias são previstas a partir da distribuição dos ângulos de torção preditos, e um conjunto de estruturas com baixo potencial são gerados a cada ciclo de otimização por gradiente descendente, que convergem até a estrutura de menor potencial (Senior, et al. 2020).

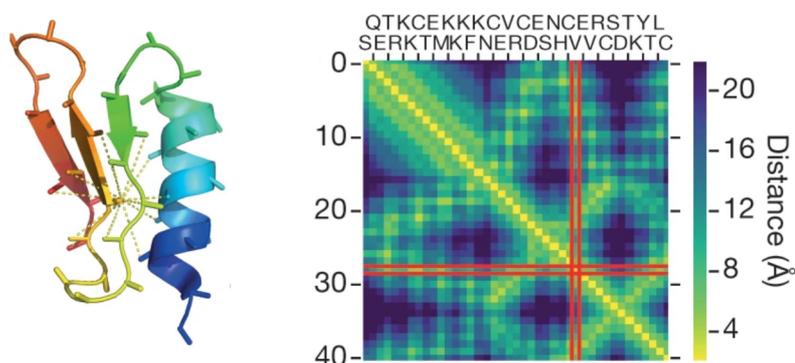


Figura 16. Contatograma das distâncias dos aminoácidos de uma proteína em sua conformação, nativa tridimensional, removido de Senior, et al. (2020). A predição das distâncias entre os aminoácidos ocorre através da determinação da distância entre os aminoácidos de uma estrutura tridimensional, nos quais são plotadas em uma matriz formada pela sequência de aminoácidos x sequência de aminoácidos, denominado de contatograma. O contatograma é utilizado na busca por proteínas homólogas para modelagem por homologia.

Sendo assim, as predições computacionais da estrutura de alta resolução de proteínas incluem os ângulos de torção do backbone e as distâncias entre os pares de resíduos. As probabilidades previstas baseadas em princípios são combinadas para formar um potencial específico da proteína, representando a melhor dentre as possíveis conformações estruturais que respeitam os princípios aplicados na modelagem (Senior, et al. 2020). Neste sentido, a procura pela conformação estrutural de proteínas é baseada em um problema de otimização, onde é procurado a conformação com menor energia livre e com a menor diferença entre a conformação predita e a estrutura nativa da proteína. Hoje em dia estão sendo utilizados “deep learning” e “machine learning” nas predições estruturais de proteínas, melhorando significativamente a acurácia das predições. O deep learning utiliza redes neurais para prever estruturas secundárias, inferir distâncias ou contato inter-resíduos, avaliar a qualidade dos modelos e construir estruturas utilizando a sequência de ponta a ponta (Huang, et al. 2023).

O AlphaFold2 é um software que utiliza uma combinação de duas redes neurais e modelagem por homologia para gerar modelos que podem ter acurácia experimental (Bertoline, et al. 2023). AlphaFold2 incorpora novas arquiteturas de redes neurais e procedimentos de machine learning baseadas nas restrições evolutivas, físicas e geométricas das estruturas de proteínas. É utilizada uma rede neural treinada nas estruturas do PDB para prever distâncias entre os átomos de carbono β dos resíduos de uma proteína. A rede neural do AlphaFold2 é capaz de aprender com dados do PDB e criar um

método indutivo físico e geométrico, assim as estruturas produzidas são treinadas para serem mais parecidas com as estruturas do PDB (Senior, et al. 2020; Jumper, et al. 2021).

A rede neural do AlphaFold2 processa a entrada de dados (input) através de repetidos blocos de novas redes neurais, denominado de Evoformer, permitindo o raciocínio entre as relações espaciais e evolutivas através da troca de informações entre o alinhamento múltiplo de sequências e representação de pares (figura 17A). Tais trocas de informações ocorrem através de componentes baseados em atenção e não atenção, criando hipóteses estruturais com contínuo refinamento através das repetições dos blocos Evoformer (Jumper, et al. 2021).

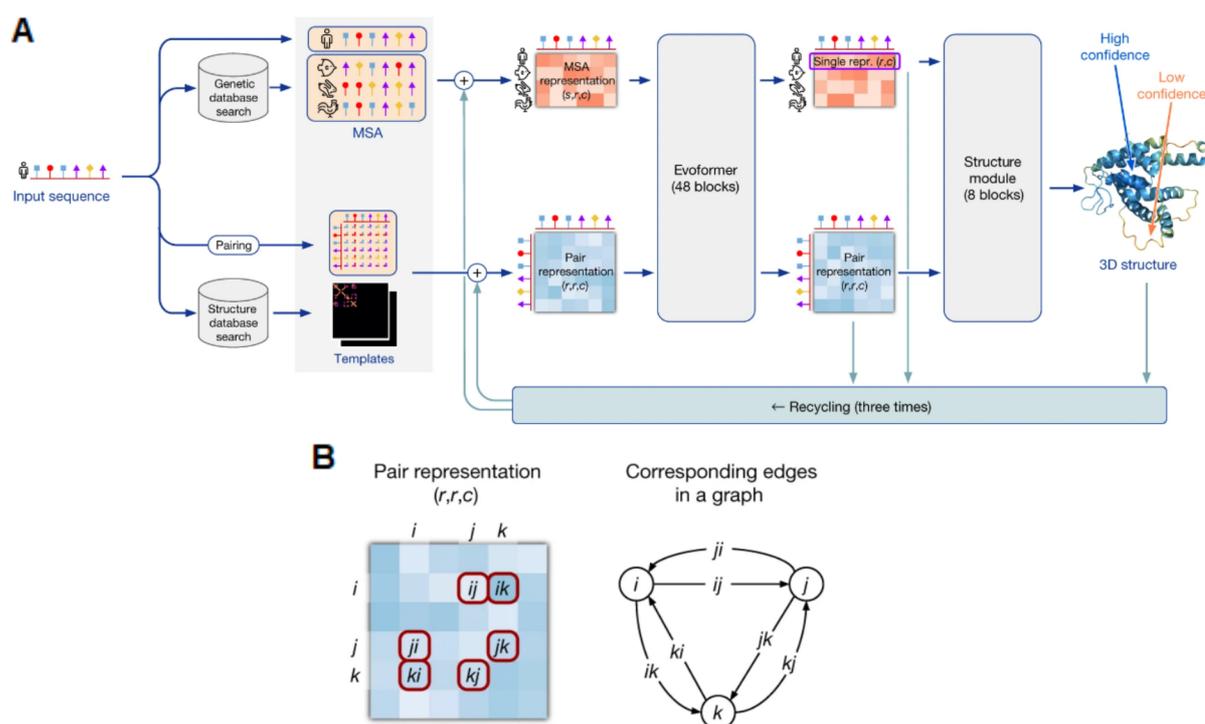


Figura 17. Modelo das redes neurais e processos realizados pelo AlphaFold2, removidos de Jumper, et al. (2021). A) Arquitetura do software AlphaFold2. As sequências de aminoácidos são inseridas no input e consecutivamente é realizado um alinhamento múltiplo de sequências para buscar por sequências homólogas e comutações. Ao mesmo tempo é realizada uma modelagem *ab initio* e gerado um contatograma da estrutura modelada, o qual é comparado com outros contatogramas gerados por machine learning. Os contatos e sequências homólogas passam por blocos Evoformer para refinamentos e por final é gerado um modelo estrutural de alta resolução. B) Representação dos pares por contatograma e triângulos multiplicativos de auto-atenção, onde os quadrados vermelhos (esquerda) e círculos (direita) representam os resíduos de aminoácidos.

Os elementos dos pares inferem informações sobre a relação entre os resíduos e realizam uma análise através de triângulos multiplicativos (figura 17B), onde analisam as distâncias entre os pares de aminoácidos e avaliam a distância utilizando os componentes de atenção e não atenção. Adicionalmente, a cada bloco Evoformer é realizado um novo alinhamento múltiplo de sequências, o qual é associado às informações dos componentes de atenção e não atenção gerados pela análise dos pares. Dessa forma, a repetição dos

blocos Evoformer garantem que os pares de contato sejam refinados através do refinamento do alinhamento múltiplo de sequências (Jumper, et al. 2021).

A partir da discriminação das distâncias entre os aminoácidos do polipeptídeo são realizados enovelamentos locais, os quais são utilizados para prever as consecutivas conformações possíveis até a conformação de menor energia livre. Dessa forma, um backbone de uma estrutura tridimensional é construída a partir dos pares de contato, alinhamentos múltiplos e enovelamentos locais, onde as representações da geometria do backbone (átomos do N-C α -C) são priorizadas em relação às cadeias laterais, nas quais as localizações são altamente restritas dentro da representação do backbone. A geometria da ligação peptídica não é restritiva pois a utilização da cadeia aberta permite os refinamentos locais, e durante o refinamento a geometria das ligações peptídicas são encorajadas (Jumper, et al. 2021).

Essas etapas são realizadas consecutivamente associadas aos mecanismos de machine learning, nos quais utilizam vias neurais que otimizam as informações processadas de acordo com os dados adquiridos ao longo da modelagem, fazendo ajustes ao longo do processo. Dessa forma, após os enovelamentos locais ocorrem buscas por proteínas que possuem padrões de enovelamento semelhantes, e posteriormente é realizada uma modelagem comparativa e novamente uma análise da distância entre os aminoácidos. Essas novas distâncias entre os aminoácidos são otimizadas e novamente são realizados enovelamentos locais até a conformação de menor energia livre, com posterior otimização do modelo tridimensional (figura 17A) (Jumper, et al. 2021).

A análise da qualidade do modelo gerado pelo AlphaFold2 é realizada através do teste de diferença de distância local predito (*predicted local distance difference test* - pLDDT) e o plot de Ramachandran (Jumper, et al. 2021; Huang, et al. 2023). O pLDDT avalia a distância local de todos os átomos do modelo com os átomos do template (Mariani, et al. 2013). Diferentemente da raiz quadrada do desvio quadrático médio (*Root Mean Square Deviation* - RMSD), o qual sobrepõe duas estruturas e analisa o RMSD dos átomos, o pLDDT não sobrepõe duas estruturas para avaliar os posicionamentos entre os átomos do template com os átomos do modelo. O pLDDT avalia as distâncias entre os átomos da mesma molécula e compara com as distâncias entre os átomos da molécula template. Utilizando a diferença entre os raios das distâncias entre os átomos da molécula template com os átomos da molécula modelo é obtido o raio de inclusão, onde caso as distâncias forem diferentes é um indicativo de que as estruturas são diferentes, por outro lado, se as distâncias estiverem dentro do raio de inclusão é um indicativo de que as estruturas são semelhantes. Sendo assim, os níveis de semelhança entre as estruturas é discriminado ao longo da cadeia polipeptídica, indicando as regiões que possuem desde ótima semelhança (azul escuro) a semelhança nenhuma (vermelho), discriminando as regiões com alto grau de

homologia das regiões com baixo grau de homologia (figura 18A) (Senior, et al. 2020; Jumper, et al. 2021; Huang, et al. 2023).

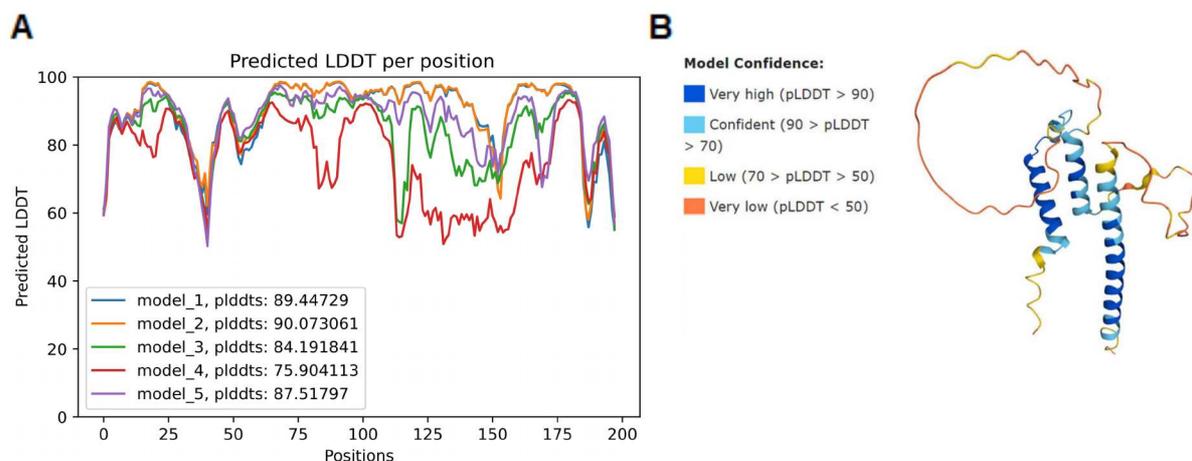


Figura 18. Gráfico de pLDDT e exemplo de estrutura modelada pelo AlphaFold2, removidos de Azzaz & Fantini (2022). A) O plot é gerado pela relação dos valores de pLDDT pela sequência de aminoácidos. É possível observar que no exemplo, o modelo 1 é o que possui maiores valores de pLDDT em relação aos outros modelos. No entanto, existem regiões que possuem baixos valores de pLDDT e em relação aos valores do restante da cadeia polipeptídica. B) Os scores do pLDDT são representados na estrutura tridimensional do modelo gerado pelo AlphaFold2 em escalas de cores. As estruturas secundárias geralmente apresentam bons scores de confiabilidade, por outro lado os segmentos desordenados não apresentam bons scores de confiabilidade por serem segmentos móveis com grande grau de movimentação.

Sendo assim, o gráfico do pLDDT mostra no eixo x as posições de cada aminoácido e no eixo y o pLDDT de cada aminoácido (figura 18A) (Senior, et al. 2020; Jumper, et al. 2021; Huang, et al. 2023). Quanto maior a semelhança na posição dos átomos entre o template e o modelo, maior é o valor de y (que é expresso em percentual, de 0 a 100 de semelhança). AlphaFold2 gera diversos modelos e realiza um rank entre os 5 melhores modelos levando em relação o pLDDT, quanto maior for a qualidade do modelo mais regiões com pLDDT alto terá (figura 18B). O pLDDT igual ou maior que 90 indica regiões de alto grau de precisão na predição e podem ser utilizados para identificar e caracterizar sítios de ligação, sítio ativo. O pLDDT entre 70 e 90 indica que as cadeias laterais não são bem modeladas e não possuem alto grau de predição, já o backbone pode ser fidedigno. Já o pLDDT entre 50 e 70 indicam baixa confiança na predição e as regiões devem ser tratadas com cautela (Senior, et al. 2020; Jumper, et al. 2021; Huang, et al. 2023). O AlphaFold2 já realiza a análise do gráfico de Ramachandran, não sendo necessário a validação dos ângulos diedros do modelo pelos programas PROCHECK e PROBABILITY.

O AlphaFold2 apresenta alta acurácia das predições das estruturas de proteínas, como a cadeia principal, domínios e das cadeias laterais (quando a cadeia principal possui alta confiabilidade). O aumento da confiabilidade dos modelos estruturais de proteínas pelo AlphaFold2 tem impactado o desenvolvimento de novos fármacos, uma vez que não possuímos estruturas de todos os alvos farmacológicos determinados de forma

experimental. O AlphaFold2 consegue modelar algumas interações proteína-proteína, apesar de não modelar bem complexos oligoméricos e proteínas intrinsecamente desordenadas (Senior, et al. 2020; Jumper, et al. 2021). O AlphaFold2 também pode ser utilizado para melhora da resolução do problema da fase na difração de raio X de proteínas (Huang, et al. 2023). Apesar do modelo gerado pelo AlphaFold2 ser uma foto da proteína, o pLDDT pode ser utilizado para discriminar informações sobre a dinâmica da proteína, onde regiões com baixos pLDDT podem ser atribuídos a alta flexibilidade da região e não a uma predição baixa da estrutura (Guo, et al. 2022).

1.4. Proteínas quinase

A fosforilação é um mecanismo reversível catalisado por quinases, sendo a primeira reação de fosforilação identificada no metabolismo do glicogênio. As quinases catalisam a transferência do fosfato gama do ATP ou GTP para o grupo hidroxila dos resíduos de serina e treonina, e grupo fenol de resíduos de tirosina do seu substrato, gerando monoéster de fosfato (figura 19) (Keck, et al. 2015). Já a remoção do fosfato ocorre através da hidrólise do monoéster de fosfato pelas fosfatases (Castelo-Soccio, et al. 2023).

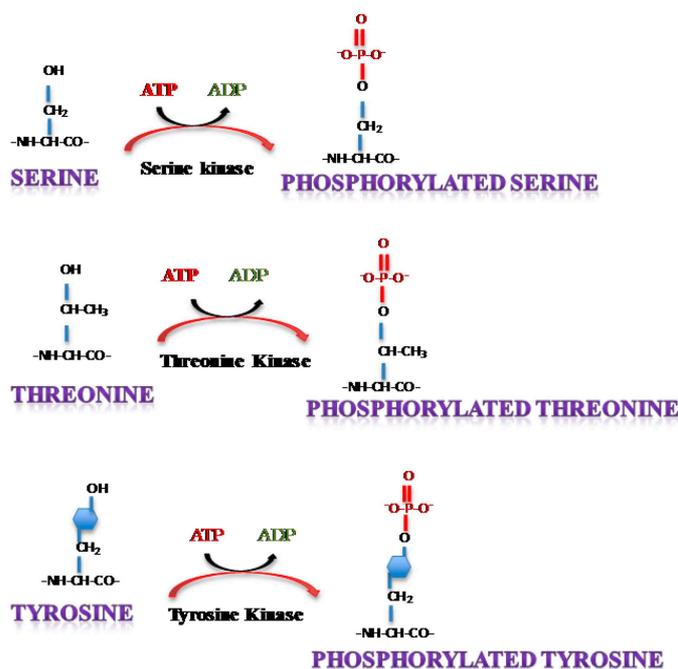


Figura 19. Fosforilação de resíduos de serina, treonina e tirosina por proteínas quinase, removido de Keck, et al. (2015). O fosfato gama da molécula de ATP é transferido para o grupo hidroxila dos aminoácidos serina, treonina e tirosina, com posterior liberação de ADP.

Aproximadamente um terço das transferências do fosfato gama ocorrem no grupo hidroxila dos resíduos de serina, treonina e tirosina, denominados de O-fosforilação. A fosforilação de tirosina é pouco frequente em relação à fosforilação de serina e treonina.

Alguns resíduos de histidina e aspartato também podem ser fosforilados, denominados de *N*-fosforilação. Neste sentido, as quinases são classificadas de acordo com os resíduos fosforilados por elas, como serina, treonina ou tirosina quinase. Geralmente as quinases que fosforilam resíduos de serina também fosforilam resíduos de treonina, sendo denominadas de serina/treonina quinase (Serine/Threonine Kinases - STKs). As quinases que fosforilam os resíduos de tirosina são denominadas de tirosina quinase (Tyrosine Kinases - TKs). E existem as quinases que fosforilam os três resíduos, denominadas de quinases de especificidade dual (Dual-Specificity Kinases - DSKs) (Keck, et al. 2015; Ardito, et al. 2017).

O domínio catalítico de proteínas quinases consistem de aproximadamente 250-300 aminoácidos, e o primeiro cristal de uma proteína quinase foi da subunidade catalítica da proteína quinase A (Protein Kinase A - PKA). Foi demonstrado que as quinases possuem uma estrutura globular dividida em dois lobos, lobo N e lobo C, com o sítio catalítico localizado entre os dois lobos o qual é regulado pelo looping de ativação e alfa-hélice C localizada no lobo N (figura 20) (Knighton, et al. 1991; Hanks & Hunter, 1995; Arter, et al. 2022).

O lobo N das quinases possui 5 folhas-beta e ao menos uma alfa-hélice, e o lobo C é composto por sua maioria de alfa-hélices e algumas folhas-beta (figura 20). Por convenção, as folhas-beta são numeradas e as alfa-hélices são nomeadas com letras alfabéticas de forma crescente, começando do N-terminal da proteína. O ATP fica localizado na interface entre os dois lobos. Os lobos são conectados por uma curta sequência denominada de Hinge, que também interage com a adenina do ATP através de duas ligações de hidrogênio. O sítio catalítico é formado pelo sítio de ligação do ATP e o sítio de ligação do polipeptídeo. O sítio de ligação do polipeptídeo é formado pela interface da alfa-hélice C do lobo N com o looping de ativação (Knighton, et al. 1991; Arter, et al. 2022).

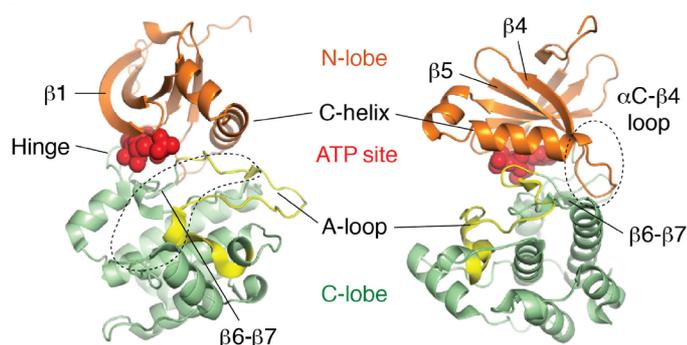


Figura 20. Representação estrutural de uma quinase típica, removido de Arter, et al. (2022). Imagens representativas do domínio quinase da Aurora A. O lobo N, lobo C e looping de ativação (A-loop) estão anotados pelas cores laranja, verde e amarela, respectivamente. As estruturas secundárias conservadas no domínio quinase como as folhas-beta 1, 4, 5, 6 e 7, alfa-hélice C, sítio de interação do polipeptídeo (área tracejada da esquerda) e o looping alfaC-Beta4 (área tracejada da direita) estão anotados na figura. As esferas vermelhas representam o sítio de ligação de ATP.

As proteínas quinases são ativadas através de diversos mecanismos, como a fosforilação do looping de ativação por outra quinase (*trans*-fosforilação) ou a autoativação (*cis*-fosforilação), ativação alostérica (interação entre duas quinases do mesmo tipo, a qual induz modificação no sítio catalítico para um estado ativo sem realizar a fosforilação do looping de ativação) e fosforilação de um segmento autoinibitório por uma segunda quinase (Ardito, et al. 2017).

A conformação estrutural do domínio quinase de diferentes proteínas quinases foram determinadas experimentalmente, possibilitando a identificação de subdomínios conservados dentro do domínio quinase de células eucarióticas. O domínio catalítico de quinases eucarióticas é subdividido em 12 subdomínios conservados que se enovelam em um core catalítico com estrutura semelhante (Hanks & Hunter, 1995).

Os subdomínios conservados do domínio quinase descritos por Hanks & Hunter (1995) estão descritos a seguir (figura 20): subdomínio I das quinases eucarióticas possuem a sequência consenso Gly-x-Gly-x-x-Gly-x-Val e está localizado em uma volta- β entre as folhas-beta 1 e 2 (formando uma estrutura que envolve e ancora os fosfatos não transferíveis do ATP); subdomínio II é caracterizado por uma Ala-x-Lys dentro da folha-beta 3, onde a alanina contribui em um bolsão hidrofóbico para o anel purínico do ATP, enquanto a lisina auxilia no ancoramento e orientação do ATP através da interação com os fosfatos α e β ; subdomínio III é representado pela alfa-hélice C, possuindo um glutamato localizado centralmente na hélice, auxiliando na estabilização da interação entre a Lisina do subdomínio II com os fosfatos α e β do ATP; subdomínio IV é caracterizado por uma folha-beta 4 com característica hidrofóbica; subdomínio V conecta o lobo N com o lobo C do domínio quinase, é composto por uma folha-beta 5 com característica hidrofóbica localizada no final do lobo N e conectada a alfa-hélice D localizada no início do lobo C. A região de conexão entre os lobos N e C auxilia na ancoragem do ATP através da formação de uma cavidade hidrofóbica para o anel da adenina, e realiza ligação de hidrogênio com a ribose do ATP além de estar envolvida na interação com o peptídeo-substrato; subdomínio VIA é caracterizado por uma grande e hidrofóbica alfa-hélice E, estendendo-se ao longo do lobo C e agindo como uma estrutura de suporte do domínio quinase; subdomínio VIB apresenta uma volta- β entre as folhas-beta 6 e 7, sendo presente o motivo consenso His-Arg-Asp-Leu-Lys-x-x-Asn (HRDLKxxN) e Asp-Iso-Lys-Pro-Ala-Asn (DIKPAN), nos quais são sequências atribuídas ao sítio ativo. Este subdomínio possui resíduos de aminoácidos relacionados com a interação com o peptídeo-substrato e com as reações de transferência do grupo fosforil (agindo como aceptores de prótons do substrato, como estabilizadores da ribose do ATP e como neutralizadores da carga negativa do fosfato γ e do segundo íon Mg^{2+} , o qual se liga ao fosfato α e γ do ATP); subdomínio VII apresenta uma volta- β entre as folhas-beta 8 e 9, que possui o motivo Asp-Phe/Leu-Gly (DFG ou DLG) e auxilia na

orientação do fosfato γ do ATP através da interação com íons Mg^{2+} que interage com os fosfatos β e γ do ATP; subdomínio VIII está relacionado com a estabilização do lobo C, reconhecimento dos substratos ou pseudo-substratos, localização do motivo conservado Ala-Pro-Glu (APE) e do looping de ativação (a reação de fosfotransferência para este segmento faz com que os oxigênios do grupamento fosfato formem ligações de hidrogênio com resíduos de Arg e Lys, estabilizando o looping em uma conformação que permite a orientação do peptídeo do substrato). Podem ocorrer modificações no subdomínio VIII após a interação da quinase com outro ligante (proteína/DNA/RNA), essa modificação movimentada o looping para uma conformação favorável para a fosforilação do looping, ou para uma posição que favoreça a atividade quinase sem modificação pós-traducional no looping; subdomínio IX corresponde a longa alfa-hélice F, realizando uma ligação de hidrogênio com uma sequência que precede o looping de ativação, estabilizando o looping; subdomínio X é pouco conservado com função variável dentro de cada tipo de quinase, possuindo diversas inserções entre o subdomínio X e XI em diferentes quinases; subdomínio XI se estende até o C-terminal do domínio quinase.

O genoma humano possui aproximadamente 518 genes de proteínas quinases, as quais são divididas em grupos, famílias e subfamílias. A maioria das quinases possuem outros domínios associados a sua estrutura, podendo regular a atividade quinase, localização ou oligomerização (Knighton, et al. 1991; Hanks & Hunter, 1995; Arter, et al. 2022). Interessantemente, mais de 30% de todas as proteínas humanas podem sofrer fosforilação. Dessa forma, proteínas quinases são componentes chaves na sinalização celular, induzindo respostas rápidas frente ao sinal intra e extracelular. A fosforilação é um mecanismo de regulação presente na maioria dos processos celulares, como regulação da síntese de proteínas, crescimento, divisão e morte celular, desenvolvimento e envelhecimento. Muitas enzimas e receptores são ativadas ou inativadas por fosforilação. O mecanismo de fosforilação não apenas regula vias de sinalização, as fosforilações controlam a amplitude dos sinais transduzidos pelas vias de sinalização devido ao fato de serem reações rápidas e reversíveis (Ardito, et al. 2017). Dessa forma, a desregulação da função de proteínas quinases possui grande impacto no desenvolvimento de patologias como doenças imunológicas, cardiovasculares, neurodegenerativas, metabólicas, inflamatórias, infecciosas e o câncer. Neste sentido, as proteínas quinases são alvos clínicos importantes onde existem diversos inibidores de quinases aprovados para uso clínico nos quais estão distribuídos dentro do quinoma humano (figura 21) (Kannaiyan & Mahadevan, 2018).

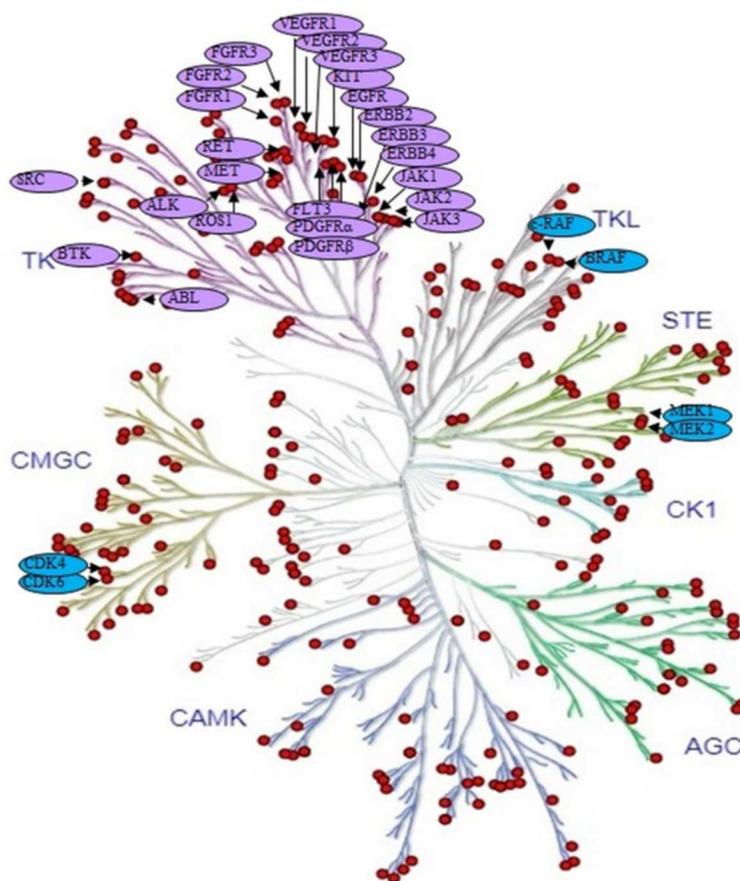


Figura 21. Mapa filogenético do quinoma humano associado aos inibidores aprovados pela FDA no tratamento de tumores sólidos e hematológicos, removido de Kannaiyan & Mahadevan (2018). Os círculos vermelhos indicam as quinases que possuem inibidores aprovados para uso clínico. As quinases dos círculos azuis representam as serina/treonina quinases e os círculos roxos representam as tirosinas quinases.

1.5. NEKs

As Neks (NimA-related kinases) são uma família de proteínas serina/treonina e tirosina quinases altamente conservadas. O protótipo da família Neks é a proteína NimA (Never in mitosis, gene A), uma serina-treonina quinase identificada e caracterizada funcionalmente nos fungos *Aspergillus nidulans*, sendo crítica para a progressão do ciclo celular e início da mitose (figura 22A) (Bergen, et al. 1984).

A homologia com o domínio quinase da NimA é o requisito para pertencer à família das Neks. No genoma de *Homo sapiens*, até o momento foram identificados e confirmados 11 membros desta família, NEK1 a NEK11, nos quais formam um ramo evolutivo independente no quinoma humano (figura 22B) (Kooij, et al. 2019). Os domínios quinase das NEKs possuem 40-50% de homologia entre si e em relação a NimA (Fry, et al. 2012), no entanto, as NEKs apresentam diferenças estruturais além do domínio quinase. Semelhante a NimA, as NEKs possuem anotações para segmentos coiled-coil e sequências PEST, além

do domínio RCC1, sequência Dead-box e repetições Armadillo (figura 22A) (Moniz, et al. 2011; Fry, et al. 2012).

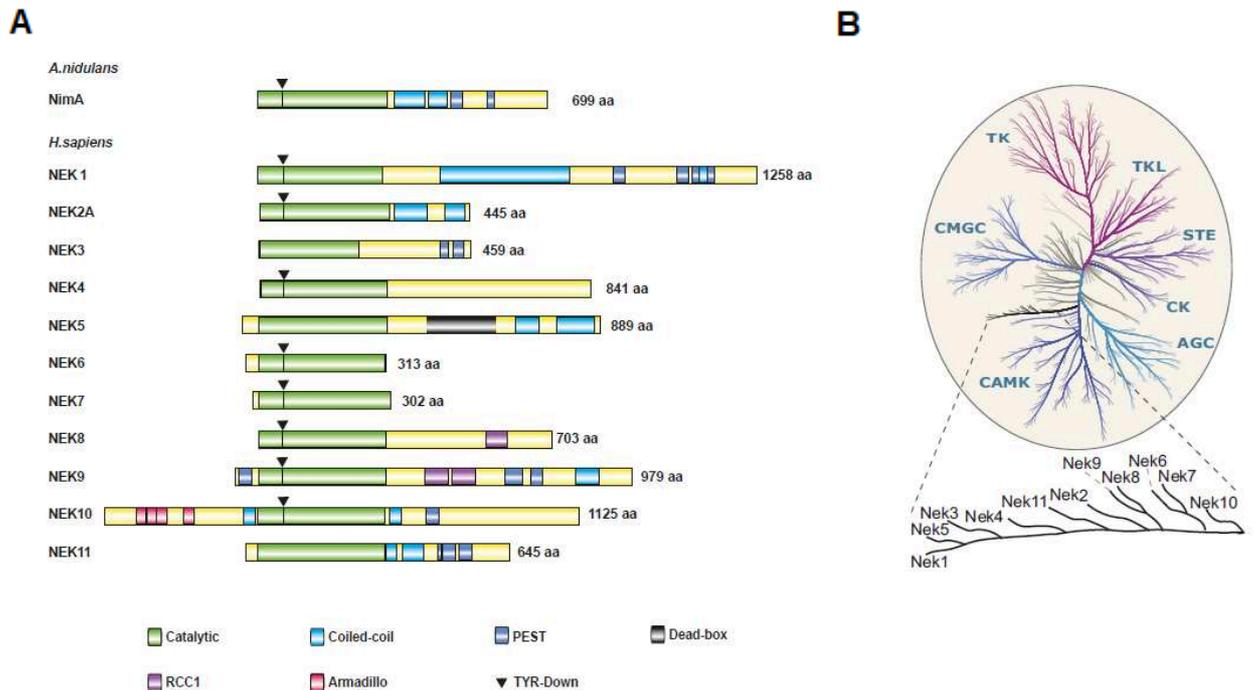


Figura 22. Alinhamento múltiplo dos domínios quinase das NEKs e Nima, e dendrograma de quinases eucarióticas, removidos de Moniz, et al. (2011) e Manning, et al. (2002). A) Alinhamento do domínio quinase das NEKs e NimA (retirado de Moniz, et al. 2011) com anotações dos domínios coiled-coil, RCC1, DEAD-BOX, sequências PEST, Armadillo e tirosina auto-inibitória (TYR-Down). B) Dendrograma do domínio catalítico de 491 proteínas quinases eucarióticas (retirado de van de Kooij, et al. 2019). Podemos observar grandes ramos filogenéticos nas famílias TK, TKL, CMGC, CAMK, AGC, CK e STE, onde o ramo das NEKs (ampliado) está localizado mais próximo ao ramo da família CAMK, entre as famílias de quinases FUSED/ULK e de pseudo-quinases TRB (Manning, et al. 2002).

Dessa forma as funções das NEKs podem ser quinase dependente e independente, e estão envolvidas na regulação do ciclo celular, metabolismo do cílio primário e estabilidade do genoma nuclear (figura 23) (Fry, et al. 2012; Pavan, et al. 2021). Entretanto, foram elucidadas novas funções das NEKs envolvendo a regulação de splicing de mRNAs, inflamação, metabolismo energético e homeostase mitocondrial (Basei, et al. 2024). A Tabela 1 sumariza as funções das NEKs descritas na literatura.

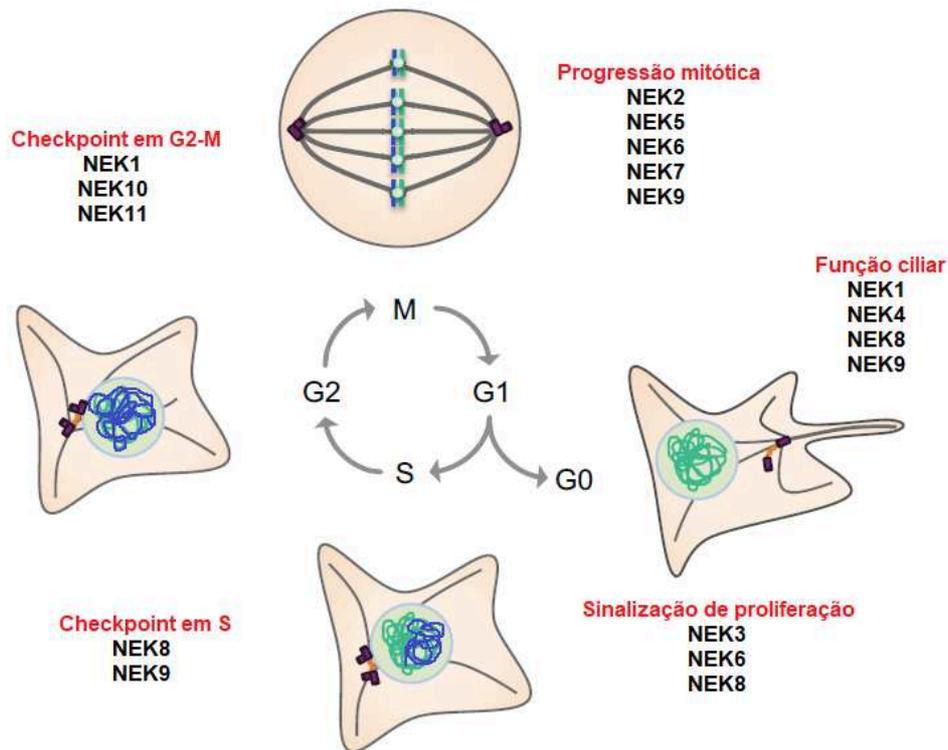


Figura 23. Contribuição das NEKs para a progressão do ciclo celular, metabolismo do cílio primário e estabilidade do genoma nuclear, adaptado de Fry, et al. (2012). Podemos observar que as funções relacionadas à progressão do ciclo celular das NEKs estão distribuídas entre seus membros.

Alterações na função ou no níveis proteicos das NEKs estão associados a diversas condições patológicas, como doenças neurodegenerativas, inflamatórias, ciliopatias e distrofias, além de estarem associados a formação e progressão dos cânceres. Neste sentido, as NEKs também são alvos farmacológicos e existem inibidores para alguns membros da família, nos quais ainda são alvos de estudos na busca de novos inibidores para uso clínico (Peres de Oliveira, et al. 2020).

Tabela 1. Funções propostas para as quinases NEKs de humanos, adaptada de Quarmby & Mahjoud (2005).

| Proteína | Funções descritas na literatura |
|-----------------|--|
| NEK1 | Alterações no gene <i>NEK1</i> em animais são modelos para a doença policística renal em humanos. Interage com proteínas envolvidas no ciclo celular e no reparo do DNA. Mecanismo de reparo por homologia. Papel na resposta a dano ao DNA induzido por radiação ionizante. |
| NEK2 | Localizado em centrossomos e cinetócoros. Fosforila a C-Napland Nlp no centrossomo. Regula a separação do centrossomo na transição G2/mitose. Possível papel na regulação do fuso mitótico. |
| NEK3 | Modula a sinalização dos receptores de prolactina. Aumento da expressão está relacionado com baixo prognóstico de pacientes com câncer gástrico. |
| NEK4 | Reparo de DNA, splicing RNA. |
| NEK5 | Respiração mitocondrial. Substrato da caspase 3, promovendo diferenciação das células do músculo esquelético. |
| NEK6 | Regulação dos telômeros através da fosforilação de TPP1. Estabilização das fibras do cinetócoro através da fosforilação de HSP72. Relacionado com câncer de próstata resistente à castração. Fosforila o fator de transcrição FOXJ2. Interator de RGS2. Centrossomo. |
| NEK7 | Ativação do inflamassoma NLRP3. Participa da mesma cascata de sinalização na mitose que Nek6. Interator de RGS2. Centrossomo. |
| NEK8 | Aumento da expressão em câncer de mama. Estabilidade na forquilha de replicação mediada por RAD51. |
| NEK9 | Associação ao Bicc2 <i>in vivo</i> , fosforilação de Bicc2. Regula o alinhamento e a segregação dos cromossomos na mitose. Mediador da organização dos cromossomos e do centrossomo. Ativa a Nek6 durante a mitose. Regula a progressão de G1 e S através da interação com o complexo facilitador da transcrição da cromatina (FACT) |
| NEK10 | Resposta a dano ao DNA induzido por UV |
| NEK11 | Quinase responsiva a estress replicativo e a danos ao DNA. É ativada pela Nek2a em células com defeito na progressão do ciclo celular em G1/S. |

1.6. NEK6

A NEK6 é paróloga da NEK7, e o locus Crb2-Dennd1A-Lhx2-Nek6 do cromossomo 9q33.3 e o locus Crb1-Dennd1B-Lhx9-Nek7 do cromossomo 1q31.3 são regiões parálogas do genoma humano (Katoh & Katoh, 2004). O gene da NEK6 possui aproximadamente 95.000 pb (NCBI), já o gene da NEK7 possui aproximadamente 165.000 pb (NCBI).

A NEK6 e a NEK7 são os menores membros da família NEK, onde a NEK6 possui 313 aminoácidos com massa molecular de 35 kDa e a NEK7 com 302 aminoácidos e 34 kDa. Elas consistem praticamente de um N-terminal regulatório curto e desordenado seguido do domínio quinase no C-terminal (figura 22A). As sequências da NEK6 e NEK7 possuem 77% de identidade, com 86% de identidade no domínio quinase e 20% de identidade no N-terminal curto e desordenado (Meirelles, et al. 2011).

O C-terminal e N-terminal da NEK6 possuem interatores diferentes, onde o mesmo ocorre com a NEK7. No entanto, apesar da alta homologia entre a sequência de aminoácidos e características estruturais entre a NEK6 e NEK7, os interatores do N-terminal desordenado e domínio quinase são diferentes entre elas, com alguns interatores no domínio quinase sendo compartilhados, nos quais estão relacionados com as funções compartilhadas entre NEK6 e NEK7 na mitose (Meirelles, et al. 2011; de Souza, et al. 2014).

Até o presente momento, não existem modelos estruturais de alta resolução gerados por difração de raio X de cristais da NEK6 na literatura, apenas para NEK7. No entanto, existem diversos trabalhos que realizaram modelagem por homologia da NEK6 utilizando os modelos estruturais de alta resolução da NEK7 como template (Meirelles et al. 2011; Moraes, et al. 2015; de Donato, et al. 2018). Dentre os modelos da estrutura de alta resolução da NEK6 gerados pelos diferentes trabalhos, o modelo gerado por Meirelles et al. (2011) apresenta as melhores previsões. Os autores identificaram que a NEK6 possui estrutura monomérica globular levemente alongada com o N-terminal flexível e desordenado, e realizaram modelagem de corpo rígido dos modelos estruturais gerados *in silico* com informações do envelope molecular da NEK6 obtidos por SAXS combinado com cálculos *ab initio* (Meirelles, et al. 2011).

Meirelles et al. (2011) utilizaram a construção 6xHis NEK6 S206A purificada de bactérias para a análise por SAXS, onde através da análise de Guinier e Kratky foi demonstrado um raio de giro de $32,0 \pm 1,0 \text{ \AA}$ e $32,4 \pm 0,8 \text{ \AA}$, respectivamente, e dimensão máxima de aproximadamente 110 \AA através da análise de Kratky. Os dados de SEC-MALS (Size-Exclusion Chromatography coupled to Multi-Angle Light Scattering) são associados aos dados de SAXS para caracterizar a amostra e demonstrar que não houve agregação ou formação de complexos da amostra. Neste sentido, a análise por SEC-MALS da 6xHis NEK6 S206A demonstrou a natureza monomérica desta construção. O raio de Stokes (R_s)

de polipeptídeos pode ser calculado por SEC-MALS a partir da construção de uma curva de calibração com diferentes polipeptídeos de massa molecular e R_s conhecidos. Para o cálculo do R_s é identificado o coeficiente de partição (K_{AV}) das proteínas padrão e da proteína de interesse através da relação dos volumes injetados, morto e total da coluna cromatográfica (Tiroli e Ramos, 2007), sendo é realizado uma regressão entre o R_s e K_{AV} .

Meirelles et al. (2011) purificaram NEK6 WT, S206A e com truncamento do N-terminal (Δ 1-44 NEK6 WT) expressas em bactérias com e sem a λ -fosfatase para a análise da relação do status de fosforilação e do N-terminal com o R_s da NEK6. A NEK6 WT desfosforilada apresenta um R_s de 2,1 nm, enquanto a NEK6 WT fosforilada apresenta um R_s de 2,6 nm e o truncamento do N-terminal (Δ 1-44 NEK6 WT) reduz o R_s para 1,8 nm. Isso indica que a fosforilação da NEK6 modifica sua conformação e seu N-terminal é responsável pelo seu alongamento (Meirelles, et al. 2011). A razão R_G/R_s é utilizada para indicar conformações estruturais, como esferas homogêneas (razão de 0,78) a espirais alongadas e elipsóides (razão acima de 2). O R_s por SEC-MALS (2,6 nm) e raio de giro por SAXS (3,2 nm) da NEK6 são valores próximos, dando uma razão R_G/R_s de 1,3, sendo mais um indicativo de que sua estrutura é globular levemente alongada (Meirelles, et al. 2011).

Meirelles et al. (2011) utilizaram diversos softwares para identificar qual é o melhor modelo da estrutura de alta resolução da NEK6. Os autores utilizaram os softwares I-TASSER, Geno3D, 3D-JIGSAW, SWISS-MODEL e o MODELLER 9v8 para modelagem por homologia da NEK6 sem o N-terminal desordenado e com a mutação S206A, utilizando o cristal 2WQM da NEK7 como template. Foi identificado através da análise da qualidade estereoquímica pelo plot de Ramachandran que o SWISS-MODEL gerou o modelo mais aceitável dentre os softwares utilizados (figura 24A).

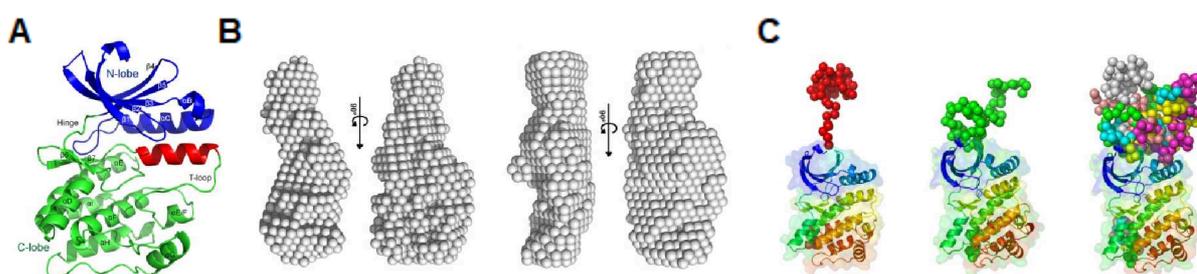


Figura 24. Representação do envelope molecular, modelos alocados no envelope molecular e modelo estrutural de alta resolução da NEK6, removidos de Meirelles, et al. (2011). A) Modelo molecular comparativo da NEK6 S206A baseado na estrutura da NEK7 do cristal 2WQM. Os lobos N e C estão representados de azul e verde, respectivamente. O looping de ativação (T-loop) foi modelado pois os cristais de NEK7 não possuem este segmento resolvido. Não foi realizada a modelagem do N-terminal desordenado. B) Visão ortogonal do envelope molecular médio do modelo da NEK6 de baixa resolução BUNCH. C) Dois modelos típicos e uma superposição de modelos de baixa resolução da NEK6 S206A obtidos a partir de dados de SAXS usando uma combinação de cálculos *ab initio* e modelagem de corpo rígido.

A reconstrução do envelope molecular através dos dados de SAXS é realizada por diversos softwares, onde utilizam reconstruções *ab initio*. Meirelles et al. (2011) realizaram a reconstrução do envelope molecular com os dados de SAXS utilizando os métodos de átomos fictícios *ab initio* (DAMMIN) e resíduos fictícios *ab initio* (GASBOR) (figura 24B). Os autores utilizaram os modelos BUNCH para modelar a estrutura da NEK6 modelada *in silico* com a reconstrução do envelope molecular construído com os dados de SAXS (figura 24C). O N-terminal desordenado não foi modelado, apenas foi representado alguns possíveis posicionamentos obtidos pelos métodos DAMMIN e GASBOR (figura 24C). O software BUNCH realiza modelagem de corpo rígido, sendo possível utilizar dados de SAXS para ajustar os modelos estruturais de alta resolução. Modelagem de corpo rígido utiliza estrutura de polipeptídeos modelados computacionalmente ou resolvidos experimentalmente (cristalografia), nos quais realizam movimentos rotacionais e translacionais em 3 eixos para determinar a melhor orientação da molécula dentro de um determinado molde.

O conteúdo de alfa-hélices, folhas-beta, voltas beta e espirais aleatórias presentes na 6xHis NEK6 S206A foram determinados por Meirelles et al. (2011) através da deconvolução do espectro de dicroísmo circular (Tabela 2). Meirelles et al. (2011) também realizaram uma análise da estrutura secundária através dos softwares PredictProtein/Prof, PSIPRED, SSpro, SOPMA e GOR4 e identificaram que a estrutura da NEK6 é composta por 34% de alfa-hélices, 12% folhas-beta e 54% de espirais, que são valores próximos aos observados pela deconvolução do espectro de dicroísmo circular (Tabela 2). Esses dados corroboram com o fato das quinases possuírem mais alfa-hélices em relação a folhas-beta na estrutura (figura 20), e a diferença nos valores de alfa-hélices observadas pela deconvolução podem ser explicadas pela dinâmica da estrutura e dispersão das moléculas em solução.

Tabela 2. Porcentagem de estruturas secundárias preditas na estrutura da NEK6, removido de Meirelles et al. (2011).

| | alfa-hélice | folhas-beta | Voltas beta | Espirais |
|--|-------------|-------------|-------------|----------|
| Predições <i>in silico</i>* | 34% | 12% | | 54% |
| Deconvolução do espectro de dicroísmo circular - software CDNN** | 41,7% | 13,2% | 15,7% | 25,8% |
| Deconvolução do espectro de dicroísmo circular - software K2d** | 41% | 17% | | 42% |
| Deconvolução do espectro de dicroísmo circular - Correia & Ramos (2009)** | 38,6% | | | |

*Softwares PredictProtein/Prof, PSIPRED, SSpro, SOPMA e GOR4 (Meirelles, et al. 2011)

**Deconvolução do espectro DC da NEK6 recombinante analisada a 4°C (Meirelles, et al. 2011)

A NEK6 é considerada uma NEK mitótica junto da NEK7 e NEK9, e foram descritas funções centrossomais durante o ciclo celular para a NEK2, NEK5, NEK6, NEK7 e NEK9 (Fry, et al. 2012). A NEK6 pode apresentar colocalização com Gama Tubulina nos centrossomos (Meirelles, et al. 2011). A localização da NEK6 nos centrossomos está relacionada à separação centrossomal durante a mitose, regulando a localização e atividade da EG5 nos centrossomos, o qual realiza atividade motora para movimentação dos centrossomos até os pólos da célula mitótica (Rapley, et al. 2008). Foi identificada a ativação da EG5 pela NEK6 e CDK1, nos quais induzem localização centrossomal da EG5 e geração de força mecânica para movimentação e separação dos centrossomos maduros para os pólos da célula (Rapley, et al. 2008).

Os centrossomos, diferentemente das mitocôndrias, retículo endoplasmático e lisossomos, são organelas que não são delimitadas por membranas, sendo formadas exclusivamente por proteínas (Conduit, P., Wainman, A. & Raff, J. 2015; Pihan, et al. 2013). Os centrossomos são formados por dois centríolos e material pericentriolar (*pericentriolar material* - PCM). Nas células em quiescência ou em G₀, os centríolos fazem parte do corpo basal do cílio primário, enquanto que durante o ciclo celular ocorre a biogênese (duplicação dos centríolos) e maturação (aumento do material pericentriolar) (figura 26A) onde o PCM possui diferentes camadas de organização proteica durante a mitose (figura 26B). Os centrossomos maduros são nucleados pelos microtúbulos para a separação dos centrossomos nos pólos da célula e separação das cromátides irmãs de forma bipolar (Conduit, P., Wainman, A. & Raff, J. 2015; Pihan, et al. 2013).

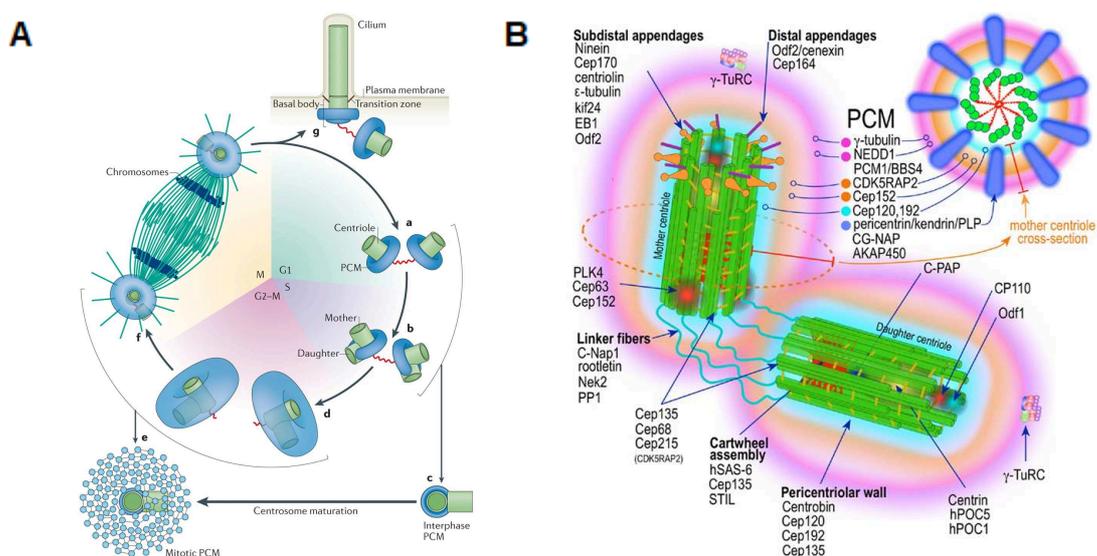


Figura 26. Centrossomos e sua organização estrutural durante o ciclo celular. A) Relação do ciclo celular e metabolismo do cílio primário e centrossomos, removidos de Conduit, P., Wainman, A. & Raff, J. (2015). B) Representação da organização centrossomal e do material pericentriolar (PCM) de vertebrados, removido de Pihan, et al. (2013).

Os processos que regulam a biogênese, maturação e separação dos centrosomos podem apresentar defeitos e induzir aberrações numéricas e morfológicas nos centrosomos, sendo a amplificação centrosomal uma das mais estudadas. Os principais mecanismos envolvidos na amplificação centrosomal estão relacionados ao excesso da biogênese centriolar (formação *de novo*), excesso de duplicação (over duplication), fragmentação por excesso de força mecânica sobre os centrosomos, fragmentação de centriolos super alongados e falhas na citocinese (Sabat-Pośpiech, et al. 2019; Mittal, K. et al. 2022).

Centrosomos amplificados são marcas do câncer nos quais induzem instabilidade genômica, sendo uma força formadora de novos subclones associados à agressividade das células de câncer (figura 27). Neste sentido, a amplificação de centrosomos é comumente observada em linhagens celulares de diverso tipos de câncer e em amostras clínicas de diversos tipos de tumores (biópsias), como mama, ovário, sistema nervoso central, cólon, pele, pulmão, próstata, renal e hematológicos (leucemia, linfoma e mieloma), sendo observadas também em lesões pré-cancerosas (Sabat-Pośpiech, et al. 2019; Mittal, K. et al. 2022).

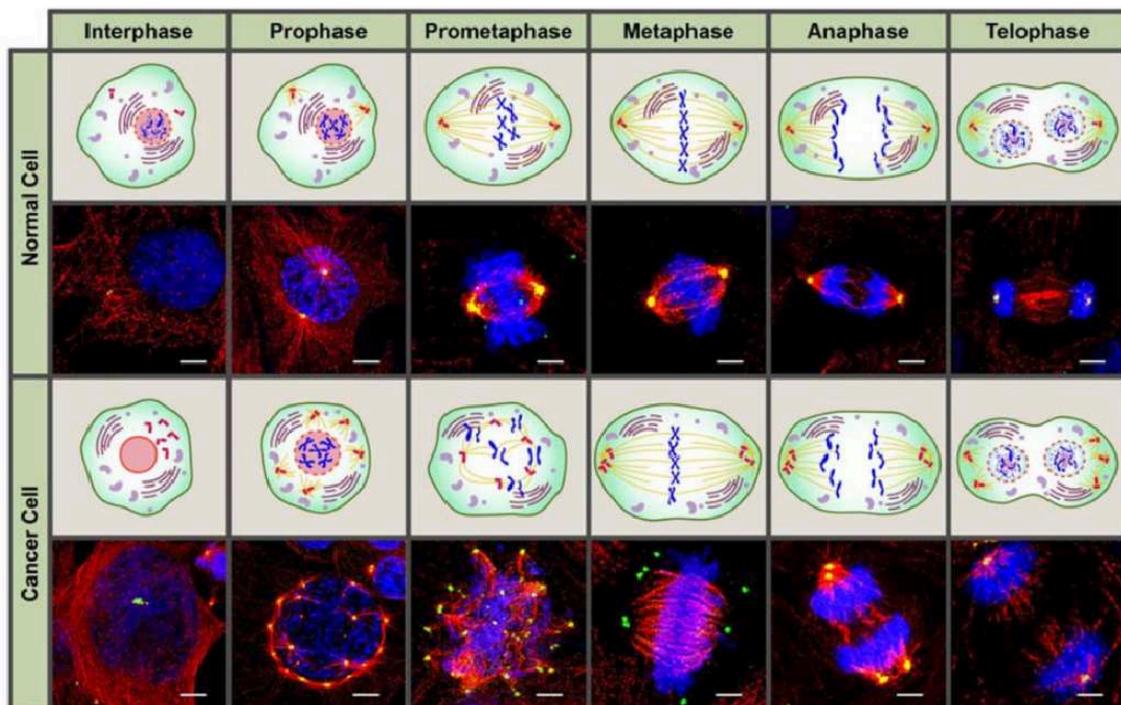


Figura 27. Imagens ilustrativas e imunofluorescência de células normais e células cancerosas com centrosomos amplificados e sua organização ao longo das diferentes fases do ciclo celular, removida de Mittal, K. et al. 2022. Microtúbulos estão marcados para Gama Tubulina (verde), os centrosomos estão marcados para α -tubulina (vermelho) e o DNA indicado com Hoechst (azul). Barra de escala = 5 μ m.

Quando as células que possuem amplificação centrossomal entram na mitose, elas tipicamente formam fuso mitótico multipolar (figura 27) o qual induz a morte celular caso não seja corrigido a multipolaridade do fuso, devido à instabilidade genômica gerada pelos diversos vetores de forças gerados sobre a cromatina (Sabat-Pośpiech, et al. 2019; Mittal, K. et al. 2022). Dessa forma, a morte celular pode ser induzida pela prolongada parada do ciclo celular e apoptose, uma vez que a amplificação centrossomal induz ativação da p53 e parada do ciclo celular induzida pela resposta ao dano no DNA. Por outro lado, a falta de genes essenciais nas células filhas devido a aneuploidia também pode induzir a morte celular (Sabat-Pośpiech, et al. 2019).

Uma das formas das células tumorais lidar com o excesso de centrossomos e reduzir a instabilidade genômica e morte celular, é através da inativação ou perda (degradação ou extrusão) de alguns centrossomos e do agrupamento desses centrossomos em dois pseudo-pólos (figura 27) (Sampson, et al. 2017; Sabat-Pośpiech, et al. 2019; Mittal, K. et al. 2022). Os dois pseudo-pólos do fuso mitótico são formados por mais de um centrossomo, nos quais o agrupamento gera dois vetores de força para a separação das cromátides irmãs, garantindo que a divisão mitótica ocorra de forma bipolar e possibilite a sobrevivência.

A NEK6 e a HSP72 são necessárias para o agrupamento (cluster) dos centrossomos amplificados em pseudo-pólos em células de tumores de mama (Sampson, et al. 2017). A HSP72 é necessária para a correta montagem do fuso bipolar e eficiente segregação dos cromossomos. Para isso, a HSP72 é substrato da NEK6 durante a mitose, onde a HSP72 é fosforilada na treonina 66 pela NEK6 e induz o aumento de HSP72 nos pólos do fuso mitótico e sítios do cinetócoro (O'Regan, 2015).

Entretanto, existem outras funções da NEK6 não relacionadas à mitose (Fry, et al. 2012). A NEK6 possui função na sinalização de proliferação celular, através da fosforilação e ativação de STAT3 e fosforilação e inibição da SMAD4 e OCT1 (Jeon, et al. 2010; Zuo, et al. 2015; Kan, et al. 2011). Recentemente, trabalhos do nosso grupo têm demonstrado a função da NEK6 na homeostase mitocondrial, sistema antioxidante e balanço redox (Basei, et al. 2024; Pavan, et al. 2023).

A função da NEK6 é regulada de forma semelhante à NimA. No início da mitose, ocorre o aumento dos níveis de NEK6 por síntese proteica junto à ativação induzida por ativação alostérica ou pela interação com a NEK9. Ao sair da mitose, ocorre redução dos níveis de NEK6 por degradação e conseqüentemente redução da atividade quinase (Fry, et al. 2012). Foram identificados interações de elementos regulatórios distais com a região promotora do gene da NEK6, onde sítios para CCCTC-Binding Factor (CTCF) localizados distantemente da região promotora formam contato com dois sítios de início de transcrição do gene da NEK6, formando um hub para “enhancers” e promotores de forma específica ao

tipo celular (Huang, et al. 2017). Entretanto, não foram identificados quais fatores de transcrição, coativadores, corepressores e receptores nucleares regulam a região promotora da NEK6. Por outro lado, os níveis de NEK6 são regulados pela E3 ligase LNX1 (ligand of numb-protein X 1) em macrófagos após infecção por *M. tuberculosis*, onde LNX1 promove poliubiquitinação da lisina 174 da NEK6 e consequentemente degradação proteassomal (Fu, et al. 2020). No entanto, não existem evidências da regulação da NEK6 pela LNX1 em outros contextos celulares sem ser pela infecção por *M. tuberculosis*.

De forma semelhante a NEK7, a ativação da atividade quinase da NEK6 ocorre tanto de forma alostérica quanto por fosforilação da serina 206 (S206) localizada no looping de ativação, podendo ocorrer ativação alostérica e fosforilação da S206 por outra NEK6 (autofosforilação) ou através da interação e fosforilação da S206 pela NEK9 (figura 25B) (Rapley, et al. 2008). A regulação alostérica da NEK6 foi identificada devido a alta homologia entre o domínio quinase da NEK6 e NEK7, e por uma série de estudos cristalográficos da NEK7 que identificaram regulação alostérica através da homodimerização da NEK7 (Haq, et al. 2015; Richards, et al. 2009). A homodimerização da NEK7 induz ativação pela liberação da tirosina 97 (Y97) da posição auto-inibitória (TYR-down para TYR-up), que está voltada para o sítio ativo do domínio quinase formando uma ponte de hidrogênio com a leucina 181 do motivo DLG, bloqueando uma conformação ativa da alfa-hélice C localizada no lobo N (Haq, et al. 2015). Na NEK6, o resíduo de tirosina 108 (Y108) foi identificado com função auto-inibitória (Richards, et al. 2009).

A homodimerização da NEK7 ocorre na orientação “back-to-back” e essa interação ocorre de forma passiva ou induzida pela NEK9 (figura 28A). A homodimerização passiva ocorre entre a interação de duas NEK7 na orientação back-to-back, causando sua ativação alostérica e posterior autofosforilação na serina 195 (S195) dentro do looping de ativação, mantendo sua ativação estável (Haq, et al. 2015). Na homodimerização da NEK7 induzida pela NEK9 ocorre interação de uma sequência curta do C-terminal de homodímeros de NEK9 (resíduos 810 ao 828) com o lobo C de duas NEK7, causando aproximação e dimerização da NEK7 (Haq, et al. 2015). Consequentemente, a dimerização da NEK7 induz a liberação da Y97 da posição auto-inibitória, ativação alostérica da NEK7 e posterior autofosforilação na S195. A NEK9 também pode fosforilar a S195 da NEK7 após a interação entre elas, ativando a atividade quinase da NEK7. Entretanto, Haq, et al. (2015) identificaram que a interface back-to-back ocorre exclusivamente no lobo N do homodímero de NEK7, e os resíduos envolvidos nessa interface são críticos para a ativação da NEK7 pela dimerização induzida pelo C-terminal da NEK9, onde mutações em resíduos de aminoácidos que estão localizadas na interface back-to-back, como N90R e R155E, causam inibição da atividade quinase da NEK7 mesmo na presença do C-terminal da NEK9. Por

outro lado, esses dados não indicam se os homodímeros de NEK7 são estáveis ou transientes.

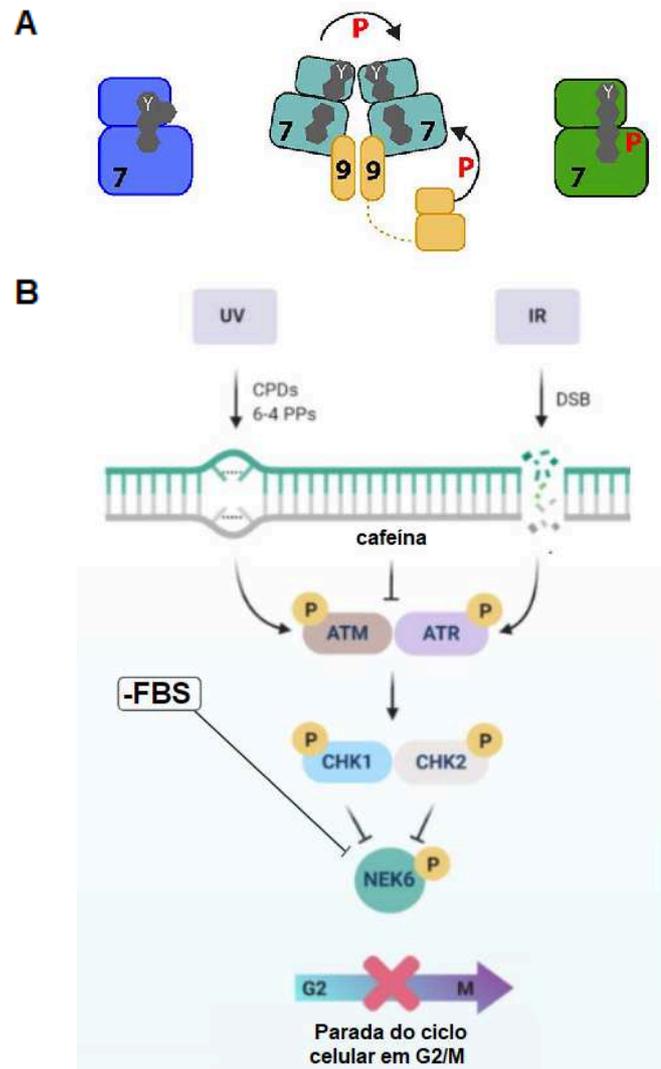


Figura 28. Regulação da atividade quinase da NEK6, adaptado de Pavan, et al. (2021). A) Mecanismos de ativação da NEK7. A NEK7 é ativada por homodimerização, a qual libera a Y7 da posição auto-inibitória causando sua ativação, e posteriormente a NEK7 pode ser autofosforilada ou fosforilada pela NEK9 na S195 o qual induz estabilidade de sua ativação. B) A atividade da NEK6 é inibida pela remoção de FBS e via ATM/ATR-CHK1/CHK2 sensível à cafeína após estresse genotóxico. A inibição da atividade da NEK6 após estresse genotóxico ocorre para haver parada do ciclo celular em G2/M.

A NEK6 possui a capacidade de autofosforilação *in vitro* (Belham, et al. 2003) e a Y108 como um resíduo auto-inibitório (Richards, et al. 2009; Haq, et al. 2015). Esses dados indicam que a NEK6 possui regulação alostérica da atividade quinase induzida pela formação de homodímeros na orientação back-to-back. Interessantemente, uma análise de cromatografia de exclusão acoplada à dispersão de luz multiângulo da 6xHis NEK6 S206A demonstrou uma natureza monomérica, enquanto a 6XHIS NEK6 WT foi identificada como monômeros com uma pequena fração de homodímeros. A expressão da NEK6 humana em células procarióticas causa a autofosforilação da NEK6, e quando a 6XHIS NEK6 WT é

desfosforilada pela λ -fosfatase ela deixa de apresentar a pequena fração de homodímeros (Meirelles, 2011). Dessa forma, esses dados indicam que existe uma pequena formação estável de homodímeros de NEK6 de forma dependente da fosforilação da S206, no entanto, não foi identificado se existem outros resíduos da NEK6 que são fosforilados após a fosforilação da S206 para ocorrer a formação do homodímero, e se a expressão em células de humanos causa alguma alteração na relação entre monômeros e dímeros uma vez que os experimentos relatados acima foram realizados em proteínas expressas e purificadas de bactérias.

O controle da atividade da NEK6 não está elucidado até o momento, mas existem evidências da regulação de sua atividade e nível protéico (figura 28B e Tabela 3), onde podem ser regulados positiva ou negativamente por alguns compostos.

Tabela 3. Condições e tratamentos que afetam os níveis de mRNA e níveis proteicos da NEK6.

| Aumento dos níveis de mRNA e/ou níveis proteicos de NEK6 | Redução dos níveis de mRNA e/ou níveis proteicos de NEK6 |
|--|---|
| <ul style="list-style-type: none"> • Hipóxia moderada (1% O₂ por 24 horas) (Zuo, et al. 2015) • Hipóxia moderada (2% O₂ por 48-72 horas) (Donato, et al. 2015) • Desferroxamina (mimetiza indução da HIF1α pela hipóxia) (Zuo, et al. 2015) • Eritropoetina (Gillinder, et al. 2017) • Combinação de 5-Azacitidina (inibidor da DNMT1) com 4-fenilbutirato de sódio (pequena chaperona química e inibidor de HDAC) (Thayanithy, et al. 2012) • GM-CSF (granulocyte macrophage-colony stimulating factor) (Kimura, et al. 2009) | <ul style="list-style-type: none"> • Hipóxia severa (0,5% O₂ for 16 horas) (Brady, et al. 2017) • TGF-Beta 1 (Zuo, et al. 2015) • Berberina (Wang, et al. 2014) • Mitramicina (inibidor da Sp1/survivina) (Fajardo, et al. 2011) |

A regulação da atividade quinase da NEK6 sem alteração dos níveis protéicos ocorre após remoção do FBS do meio de cultivo e após estresse genotóxico (figura 28B) (Minoguchi, et al. 2003; Lee, et al. 2008). A redução da atividade da NEK6 ocorre de forma rápida após a remoção do FBS do meio de cultivo, e isso ocorre mesmo com adição de todos os aminoácidos no meio de cultivo, demonstrando que a redução da atividade da NEK6 ocorre pela remoção de fatores mitóticos no meio de cultivo, como hormônios, fatores de crescimento e alguns fosfolipídios (Minoguchi, et al. 2003).

Já após tratamento com UV e IR, a atividade da NEK6 é regulada pela via ATM/ATR-CHK1/CHK2 sensível a cafeína, onde ocorre fosforilação da NEK6 pela CHK1 e CHK2 com posterior parada do ciclo celular em G2/M. A inibição da atividade da NEK6 após estresse genotóxico está associada a resposta ao dano no DNA da célula (Pavan, et al. 2021), o qual garante que a célula não prossiga na mitose com danos no genoma. Por outro lado, a superexpressão da NEK6 impede a parada do ciclo celular em G2/M após estresse genotóxico, ocasionando aumento de células na fase G1 e sub-G1 (Lee, et al. 2008). O

mesmo ocorre em células superexpressando NEK6 após tratamento com etoposídeo (Minoguchi, et al. 2003). Lee et al. (2008) identificaram através de ensaios *in vitro* que a fosforilação da NEK6 pela CHK1 e CHK2 ocorre entre os aminoácidos 1 e 80 do N-terminal. De forma interessante, Belham et al. (2003) após superexpressão de NEK6 em células HEK293, imunoprecipitação, separação em SDS-PAGE, digestão das bandas da NEK6 e posterior análise de fosfopeptídeos por espectrometria de massas, identificaram que entre os aminoácidos 1 e 80, apenas a serina 37 (S37) da NEK6 se apresenta fosforilada. A S37 apareceu como fosfopeptídeo apenas na banda de migração lenta da NEK6 (NEK6 ativa, fosforilada na S206) no SDS-PAGE e em menor nível em relação ao fosfopeptídeo da S206. No entanto, os autores geraram mutações fosfomimética (S37D) e fosfodeficiente (S37A) da S37 da NEK6 e não identificaram alteração na atividade quinase através de ensaio *in vitro* (Belham, et al. 2003).

Sendo assim, a regulação da atividade quinase da NEK6 ocorre após estresse genotóxico através da fosforilação do N-terminal (entre os aminoácidos 1 e 80 da NEK6) pela CHK1 e CHK2, enquanto a S37 da NEK6 é o único resíduo fosforilado dentro do N-terminal encontrado experimentalmente, onde a mutação fosfomimética da S37 não altera sua atividade quinase. Se de fato a S37 for o resíduo fosforilado após o estresse genotóxico, tal contradição pode ser devido aos substratos utilizados para o ensaio quinase pelos dois grupos. Ambos grupos utilizaram para o ensaio quinase imunoprecipitados de NEK6 superexpressa em células HeLa e HEK293 em proliferação. No entanto Lee et al. (2008) utilizaram como substrato a proteína básica da mielina (*Myelin basic protein*, MBP), enquanto Belham et al. (2003) utilizaram GST-p70S6K Δ CT104 T252A, a qual consiste na proteína p70S6K sem os últimos 104 aminoácidos do C-terminal (que confere resistência à inibição da atividade quinase da p70S6K por rapamicina e remoção dos aminoácidos do meio de cultivo) associada a uma mutação da treonina 252 para alanina (que confere inibição de sua ativação por fosforilação). Como o N-terminal regulatório da NEK6 e NEK7 são funcionais e possuem interatores diferentes do C-terminal (Meirelles, et al. 2010; de Souza, et al. 2014), a localização da S37 da NEK6 no N-terminal regulatório pode estar associado a uma regulação dos interatores do N-terminal sem alteração da atividade quinase da NEK6. Dessa forma, é possível que a interpretação da inibição da atividade quinase da NEK6 que ocorre após o estresse genotóxico seja devido a alteração da afinidade da NEK6 pela MBP, e consequentemente sua fosforilação.

Sendo assim, para melhor compreensão e hipóteses sobre a S37 na regulação da afinidade por alguns interatores da NEK6, nós inicialmente realizamos uma caracterização estrutural da NEK6 utilizando ferramentas *in silico*, e posteriormente avaliamos se a capacidade de homodimerização, autoativação, localização celular e alteração em interatores da NEK6 podem ser reguladas pela S37.

2. Objetivos

2.1 Objetivos gerais

O presente trabalho teve como objetivo analisar se a S37 localizada no N-terminal desordenado da NEK6 possui função na regulação da homodimerização, autoativação, estabilidade, interação com outras proteínas e localização celular da NEK6.

2.2. Objetivos específicos

- Realizar análise da sequência, sítios funcionais e regulatórios da NEK6 utilizando ferramentas *in silico*, e visualizá-los de forma bidimensional e tridimensional através de um modelo de alta resolução da estrutura da NEK6 gerado computacionalmente pelo software AlphaFold2.
- Realizar uma análise com ferramentas *in silico* das características estruturais da homodimerização da NEK6, e identificar quais os possíveis papéis da S37 na regulação da função da NEK6.
- Identificar se a homodimerização, auto-ativação, localização celular e estabilidade da NEK6 podem ser reguladas pela S37 através da coexpressão, coprecipitação e imagens de fluorescência de construções da NEK6 WT, S37A e S37D com dois TAGs diferentes (FLAG e GFP) em células imortalizadas de mamíferos (HEK293T).

3. Materiais e métodos

3.1. Análise *in silico*

Analisamos através de ferramentas *in silico* o nível de conservação dos resíduos (servidor ConSurf e Jalview), predições de função e domínios (servidor SMART e BLASTP), sítios funcionais/regulatórios (servidor ELM), estruturas secundárias (PDBSum) da sequência polipeptídica da NEK6 e NEK7. Os modelos das estruturas de alta resolução da NEK6 e NEK7 foram realizadas pelo software AlphaFold2, através do Google Colab (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=kObIAo-xetgx>). As imagens e alinhamento das estruturas de alta resolução foram realizadas no software PyMol. O alinhamento múltiplo e conservação entre a sequência polipeptídica da NEK6 e NEK7 foi realizado pelo software Jalview.

3.2. Mutagênese sítio dirigida

Geramos as mutantes fosfodeficiente (serina para alanina) e fosfomimética (serina para aspartato) na serina 37 da NEK6 WT clonadas nos vetores pcDNA3.1 e pGFP. Os clones FLAG-NEK6 WT clonado no vetor pcDNA3.1 e GFP-NEK6 WT clonado no vetor pGFP foram gerados pela Dra. Gabriela Vaz Meirelles durante seu doutorado.

Modificamos o códon TCT da serina para GCT da alanina e GAT do aspartato. Dessa forma, os primers utilizados na reação de mutagênese sítio-dirigida para geração da NEK6 S37A foram: Foward 5'-GAGGCATCCCAACACGCTGGCTTTTCGCTGCTCGCTGGC-3' e Reverse 5'-GCCAGCGAGCAGCGAAAAGCCAGCGTGTGGGATGCCTC-3'; e para NEK6 S37D foram: Foward 5'-CACGCTGGATTTTCGCTGCTCGCTGGCGGACTTCCAGATCG-3' e Reverse 5'-GCAGCGAAAATCCAGCGTGTGGGATGCCTCTGTGGGTCAGG-3'. Os primers foram utilizados na reação para gerar FLAG-NEK6 S37A, FLAG-NEK6 S37D, GFP-NEK6 S37A e GFP-NEK6 S37D.

Foram realizadas reações de PCR com os seguintes componentes: reaction buffer, 50 ng do plasmídeo molde (FLAG-NEK6 WT ou GFP-NEK6 WT) e 125 ng do primer foward e 125 ng do primer reverse, 1 µL of dNTP mix, 1 µL do Pfu Ultra HF DNA polimerase (2.5 U/µL) e ddH₂O livre de nucleases até o volume final de 50 µL. As reações foram realizadas no termociclador na seguinte configuração: aquecimento inicial de 5 minutos a 95°C com repetição de 18 ciclos de 30 segundos a 95°C, 1 minuto a 55°C e 10 minutos a 72°C, e após os 18 ciclos foi realizado uma extensão final de 60 minutos a 72°C. As reações foram digeridas com DpnI, transformadas em *E. coli* termo-competentes e as colônias formadas

foram removidas para isolamento dos plasmídeos e posterior sequenciamento de Sanger para confirmação da mutação.

A qualidade e concentração dos vetores isolados foram analisadas por eletroforese em gel de agarose e espectrofotometria com posterior sequenciamento. Utilizamos o primer forward para o promotor do CMV 5'-GGTCTATATAAGCAGAGCTC-3' no sequenciamento das FLAG-NEK6 S37A e S37D pois o FLAG possui 1 kDa e está no N-terminal das construções, não atrapalhando o sequenciamento do início da sequência. Devido o GFP possuir 27 kDa (238 aminoácidos, traduzidos de um mRNA de aproximadamente 717 pb) e estar no N-terminal da construção de GFP-NEK6, ao utilizar o primer para o promotor do CMV para sequenciar apenas a parte GFP será sequenciada, pois o sequenciamento possui uma faixa de boa acurácia das leituras dos nucleotídeos após 50 pb do início do sequenciamento até aproximadamente 700 a 800 pb de leitura. Dessa forma, ao sequenciar o plasmídeo com GFP-NEK6 WT não teremos um bom sequenciamento da parte inicial da NEK6 após o GFP. Sendo assim, nós utilizamos um primer reverso para uma sequência interna da NEK6 (5'-CCACTAGGGCGTGGGCTGCGGTGGTC-3') disponível no laboratório, sequenciando do meio para o início até alcançar os códons da S37. O sequenciamento dos vetores com FLAG e GFP NEK6 WT, S37A e S37D utilizados nos experimentos estão demonstrados na figura 29 e estão mostrando apenas a região em torno da serina 37. Utilizamos o software BioEdit (<https://bioedit.software.informer.com/Download-gr%C3%A1tis/>) para realizar o alinhamento das sequências mutadas com a selvagem.

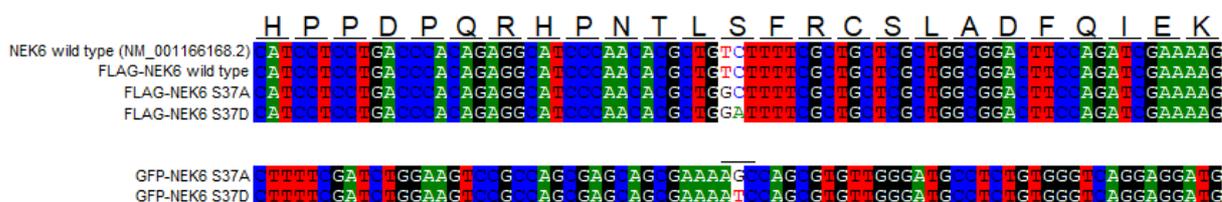


Figura 29. Alinhamento da sequência de nucleotídeos do N-terminal da NEK6 (Q9HC98-1) com a sequência de nucleotídeos dos sequenciamentos dos vetores FLAG-NEK6 WT, S37A e S3D, e GFP-NEK6 S37A e S37D gerados por mutagênese sítio dirigida. O sequenciamento das FLAG-NEK6 WT, S37A e S37D foram sequenciados utilizando um primer forward para o promotor. O sequenciamento das GFP-NEK6 S37A e S37D foram sequenciados utilizando um primer reverso para uma região no centro da NEK6 devido o GFP ser N-terminal, o qual impediria o sequenciamento passar pela sequência da GFP e chegar na região da S37 da NEK6 com uma leitura desta região se utilizado primer forward para o promotor. Os códons dos aminoácidos H25 ao K49 estão representados na figura. Serina = TCT, Alanina = GCT, Aspartato = GAT.

Posteriormente realizamos a transformação, inóculo e isolamento destes vetores e analisamos a qualidade e concentração para os próximos experimentos. A qualidade dos vetores foi analisada pela varredura da absorbância dos plasmídeos nos espectros de 240 a 300 nm, e através da eletroforese em gel de agarose para identificar se os plasmídeos estão íntegros ou possuem degradação.

3.3. Cultura de células

As células HEK293T foram cultivadas em meio DMEM suplementado com 10% de soro fetal bovino, mantidas a 37°C, sob atmosfera de 5% de CO₂.

3.4. Imunoprecipitação e Western Blot

Após as células HEK293T atingirem a confluência de 90%, foi realizada a troca do meio e a transfecção dos plasmídeos com agente carreador PEI e PBS (Sigma-Aldrich). Após a transfecção, as células foram cultivadas por 30 horas, retirado imagens no microscópio de fluorescência, tripsinizadas e centrifugadas a 300 g por 5 minutos à temperatura ambiente. A lise foi realizada em 200 µL de tampão TRIS (Tris-Cl 50 mM, pH 7,5, NaCl 137 mM, NP40 1%) por sonicação (10 ciclos sendo 1 segundo ligado e 5 segundos desligado em 20% da potência máxima da VibroCell) em banho de água/gelo. Após a sonicação, o lisado foi centrifugado por 10 minutos a 10.000xg a 4°C. O sobrenadante foi destinado para o lisado total (WL) e imunoprecipitação (IP).

A imunoprecipitação consistiu em remover 3 µg de proteína do lisado total e adicionar a 50 µL de ressuspensão de beads conjugadas com anticorpo anti-FLAG e completado para 300 µL de tampão TRIS, sendo incubado sob agitação por 4 horas na geladeira. Após a incubação, as amostras foram centrifugação a 10.000xg por 1 minuto a 4°C, o volume de 250 µL foi removido (flow through - FT) e as beads lavadas no tampão de TRIS com posterior incubação com 50 µL tampão Laemmli (eluído - IP). Adicionamos Laemmli nas amostras de LT e FT e juntamente a IP foram fervidos (95°C) por 5 minutos. As amostras foram submetidas a SDS-PAGE seguido da transferência para membrana de nitrocelulose, bloqueadas com 3% de leite em TBS e incubadas overnight com anticorpo primário para as proteínas de interesse (NEK6, NEK9, GAPDH, Beta Tubulina, Gama Tubulina, CEP250, NEK2, CCP110, CEP192 e CENTROBIN). Após as lavagens, as membranas foram incubadas por 1 hora à temperatura ambiente com anticorpo secundário conjugado a HRP e reveladas utilizando ECL (Thermo Fisher) no fotodocumentador (BioRad).

3.5. Fluorescência do GFP

As imagens de fluorescência do GFP fusionada a NEK6 foram adquiridas de células HEK293T transfectadas por 30 horas utilizando o microscópio de fluorescência (Leica) com o canal verde e ampliação de 10X e 20X.

3.6. Expressão e coexpressão FLAG NEK6/NEK9 e GFP NEK6 em células HEK293T

3.6.1. Padronização da concentração de DNA para transfecção em células HEK293T

Avaliamos a melhor concentração de DNA para transfecção dos vetores em células HEK293T. A razão da concentração DNA:PEI foi mantida em 1 µg de DNA para 3 µg para PEI. Aplicamos 1µg, 2µg e 4µg de DNA por poço na placa de 6 poços em duplicata por 24 horas, com as células a 70-80% de confluência no momento da transfecção. A toxicidade da transfecção foi analisada visualmente no microscópio de campo claro através da ocorrência ou não de desprendimento de células da placa e aparecimentos de células suspensas. Monitoramos a transfecção da GFP-NEK6 WT através da fluorescência do GFP e realizamos WB para análise da expressão dos vetores FLAG-NEK6 WT, FLAG-NEK9 WT e GFP-NEK6 WT.

A transfecção com 1 µg de DNA não apresentou boa eficiência, mostrando baixa fluorescência da GFP-NEK6 WT (figuras 30D e 30G) e baixa expressão de todos vetores quando analisado por western blot (figura 32). Houve maior toxicidade nas transfecções com 4 µg de DNA por poço. Observando as imagens de campo claro das transfecções com GFP-NEK6 WT (figura 30C), FLAG-NEK6 WT (figuras 31C) e FLAG-NEK9 WT (figura 31I) houve maior desprendimento das células transfectadas com 4 µg de DNA, apesar da fluorescência da GFP-NEK6 WT (figura 30F e 30I) apresentar ótimo sinal e alta quantidade de células transfectadas. Com 4 µg de DNA observamos a degradação da FLAG-NEK9 WT, apesar de mostrar boa expressão dos vetores FLAG-NEK6 WT e GFP-NEK6 WT quando analisado por western blot (figura 32).

As imagens nos microscópios em conjunto com o western blot mostraram que a melhor condição entre as testadas é a concentração de 2 µg por poço. Observamos alta quantidade de células fluorescentes (figura 30E e 30H), ínfimas regiões com pequenos desprendimento de células e poucas células suspensas (figura 30B, 31B, 31H). A expressão observada por western blot mostra ótimo sinal (figura 32), apresentando uma única banda para todos os vetores, indicando que não houve degradação das proteínas expressadas.

Assim, relativizando a concentração de 2 µg de DNA pela área do poço da placa de 6 poços (9,6 cm²), nós temos um valor de 0,2 µg DNA/cm² para transfecção. Essa relativização permite alterar a escala das transfecções (para placas de 60mm, 100mm, 24 poços) com a mesma eficiência e toxicidade.

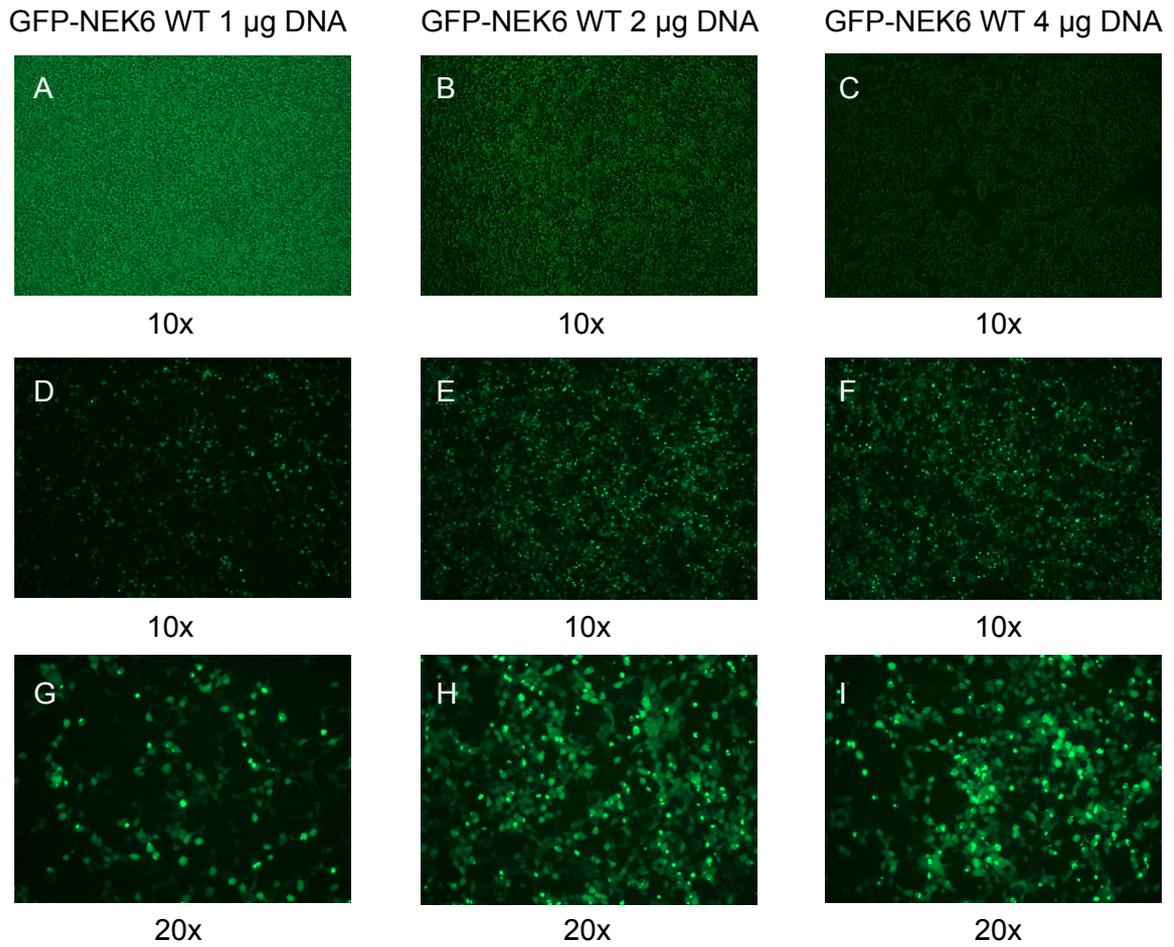


Figura 30. Imagens de campo claro e fluorescência de GFP de HEK293T transfectadas com pGFP-NEK6 WT por 24 horas nas concentrações de 1 μ g, 2 μ g e 4 μ g de DNA por poço usando uma placa de 6 poços, com razão DNA:PEI 1 μ g:3 μ g. Imagens representativas com ampliação de 10x e 20x. Parâmetros das imagens de fluorescência: exposição = 1 ms, ganho = 10x, gama = 0,6.

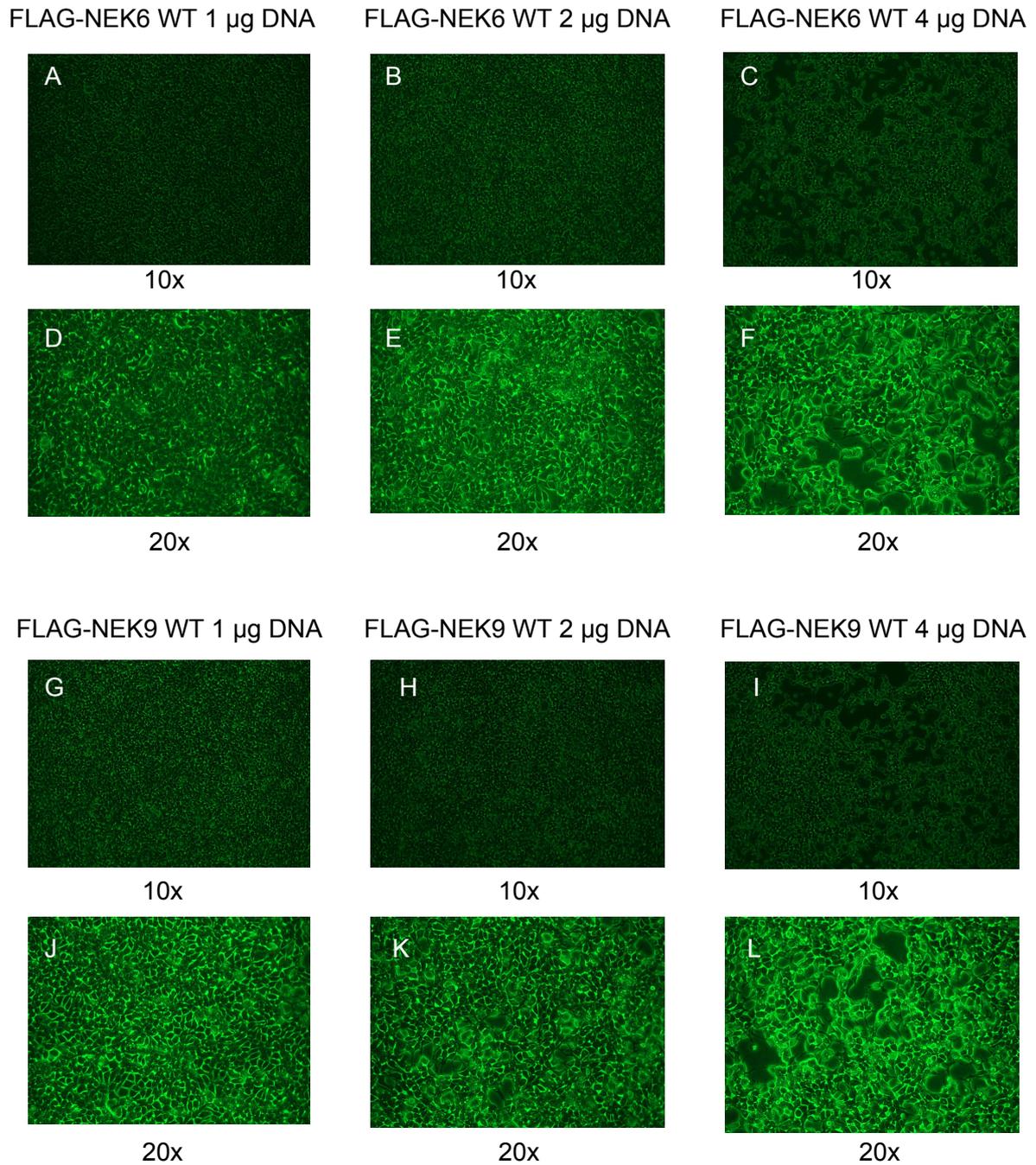


Figura 31. Imagens de campo claro de HEK293T transfectadas com pcDNA3.1 FLAG-NEK6 WT e pcDNA3.1 FLAG-NEK9 WT por 24 horas nas concentrações de 1 μg , 2 μg e 4 μg de DNA por poço usando uma placa de 6 poços, com razão DNA:PEI 1 μg :3 μg . Imagens representativas com ampliação de 10x e 20x.

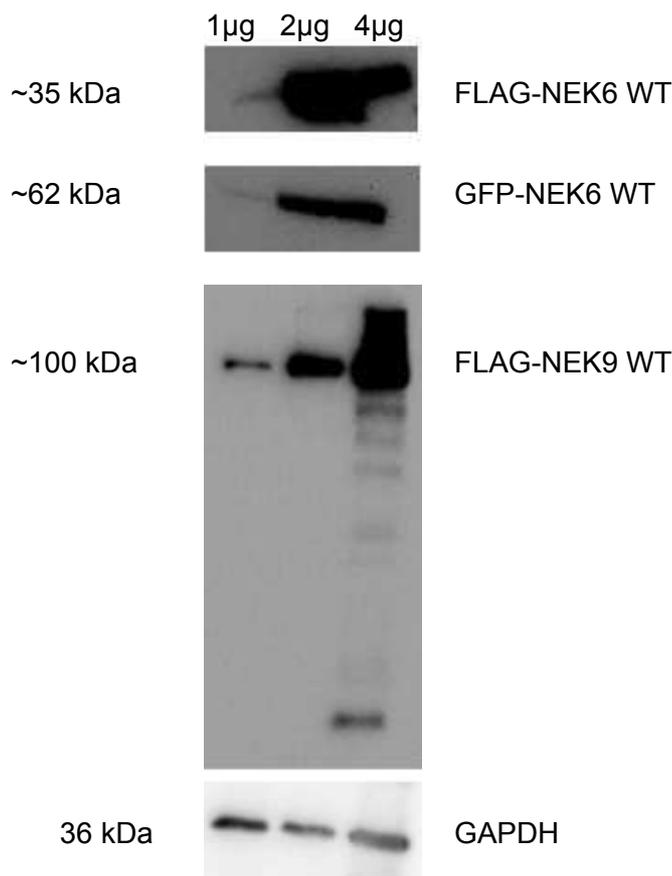


Figura 32. Western blot das transfecções de FLAG-NEK6 WT, FLAG-NEK9 WT e GFP-NEK6 em HEK293T por 24 horas nas concentrações de 1 µg, 2 µg e 4 µg de DNA por poço usando uma placa de 6 poços, com razão DNA:PEI 1 µg:3 µg. Foram carregadas 20 µg de proteína por poço para o SDS-PAGE. Anti-NEK6, anti-NEK9 e anti-GAPDH foram utilizados no experimento. A exposição das membranas foi otimizada automaticamente pelo software do fotodocumentador.

3.6.2. Padronização da relação DNA:PEI

Utilizamos uma placa de 6 poços e o GFP-NEK6 WT na concentração de 0,2 µg DNA/cm² para analisar a relação da concentração de DNA:PEI. Testamos a razão 1 µg:1 µg, 1 µg:2 µg, 1 µg:3 µg, 1 µg:4 µg, 1 µg:5 µg e 1 µg:6 µg de DNA:PEI. Monitoramos a eficiência da transfecção pela fluorescência do GFP-NEK6 WT através do microscópio de fluorescência. A toxicidade da transfecção foi analisada visualmente no microscópio de campo claro através da ocorrência ou não de desprendimento de células da placa e aparecimentos de células suspensas. As imagens (dados não apresentados) mostraram que a melhor razão DNA:PEI nas nossas condições foi de 1 µg:2 µg, com boa eficiência e baixa toxicidade. A razão 1:1 não apresentou boa fluorescência da GFP-NEK6, indicando baixa eficiência de transfecção. A razão 1:3 apresentou boa eficiência mas causou o desprendimento de mais células quando comparada com a razão 1:2. Já as razões de 1:4 a 1:6 causaram grande desprendimento de células da placa e aparecimentos de células

suspensas. Sendo assim, padronizamos nossas condições de transfecção em 0,2 μg DNA/ cm^2 na razão 1 μg DNA:2 μg PEI em células HEK293T por 24 horas.

3.6.3. Padronização da concentração de DNA para cotransfecção em HEK293T usando FLAG-NEK9 WT + GFP-NEK6 WT

Avaliamos a melhor relação da concentração de DNA para a cotransfecção dos vetores FLAG e GFP em HEK293T. Os experimentos foram realizados em placas de 60 mm em duplicata e as coexpressões foram por 24 horas com as células HEK293T a 70-80% de confluência no momento da transfecção. Utilizamos a FLAG-NEK9 WT e GFP-NEK6 WT para ver o efeito da expressão de 2 proteínas diferentes na toxicidade e eficiência da cotransfecção. Como identificamos que 0,4 $\mu\text{g}/\text{cm}^2$ de DNA é tóxico para a célula, nós utilizamos combinações que no total somam 0,25 $\mu\text{g}/\text{cm}^2$ a 0,35 $\mu\text{g}/\text{cm}^2$. Sendo assim, testamos as relações de 0,10 μg FLAG/ cm^2 + 0,15 μg GFP/ cm^2 ; 0,10 μg FLAG/ cm^2 + 0,20 μg GFP/ cm^2 ; 0,15 μg FLAG/ cm^2 + 0,15 μg GFP/ cm^2 e 0,15 μg FLAG/ cm^2 + 0,20 μg GFP/ cm^2 . A razão da concentração DNA:PEI foi 1 μg DNA:2 μg PEI, sendo a concentração de DNA a soma da massa dos 2 vetores. A toxicidade da transfecção foi analisada visualmente no microscópio de campo claro através da ocorrência ou não de desprendimento de células da placa e aparecimentos de células suspensas. Monitoramos a transfecção do GFP-NEK6 WT da fluorescência do GFP utilizando o microscópio de fluorescência. Todas as condições causaram o desprendimento de células de forma semelhante quando observadas no campo claro. As imagens fluorescentes (figuras não mostradas) mostraram que a melhor relação da concentração de DNA nas nossas condições foi de 0,15 $\mu\text{g}/\text{cm}^2$ + 0,15 $\mu\text{g}/\text{cm}^2$, com boa eficiência. Padronizamos então nossas cotransfecções com FLAG 0,15 μg DNA/ cm^2 + GFP 0,15 μg DNA/ cm^2 por 24 horas em células HEK293T.

4. Resultados e discussão

4.1. Análise de domínios, resíduos funcionais, conservação e alinhamento entre as sequências polipeptídicas da NEK6 e NEK7

Inicialmente realizamos o alinhamento entre os aminoácidos da cadeia polipeptídica da NEK6 (Q9HC98-1) e NEK7 (Q8TDX7-1) utilizando a ferramenta de alinhamento múltiplo Clustal Omega do software Jalview (<https://www.jalview.org/>), e anotamos o domínio quinase, os resíduos funcionais (sítio ativo, de ligação de ATP, de ligação de polipeptídeos e regulatórios) e a conservação entre os aminoácidos da NEK6 e NEK7 discriminados pelos softwares BLASTP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), SMART (<http://smart.embl-heidelberg.de/>), ELM (<http://elm.eu.org/>) e Jalview (figura 33). Podemos observar uma alta homologia dentro do domínio quinase entre NEK6 e NEK7, com os resíduos de aminoácidos anotados como domínio quinase apresentando cerca de 88% de homologia (figura 33). Os resíduos anotados como parte do sítio ativo (sítio de ligação de ATP e sítio de ligação de polipeptídeos) junto as sequências consenso de proteínas quinase (HRD, DLG e looping de ativação, resíduo de ativação e tirosina auto-inibitória) possuem aproximadamente 80% de conservação, diferenciando apenas em 5 aminoácidos na sequência da NEK6 em relação a sequência da NEK7, nos quais são anotados como sítio de ligação de ATP (V71, L73 e S131), um aminoácido anotado como sítio ativo (D190) e um aminoácido localizado no looping de ativação (E200) (figura 33). Podemos observar no lobo C da NEK6 uma sequência rica em prolina (PPLP) no qual é motivo para ligação do domínio WW da Fe65 (Lee, et al. 2007). Apesar do N-terminal desordenado da NEK6 e NEK7 apresentarem grandes diferenças, ambas possuem anotação de uma sequência rica em prolina (PPxPxRxP na NEK6 e PPxP na NEK7) no qual podem estar relacionadas ao docking do N-terminal com iteratores. Interessantemente, a S37 se encontra próxima a sequência rica em prolina anotada no N-terminal da NEK6, não possuindo conservação com a NEK7 (figura 33).

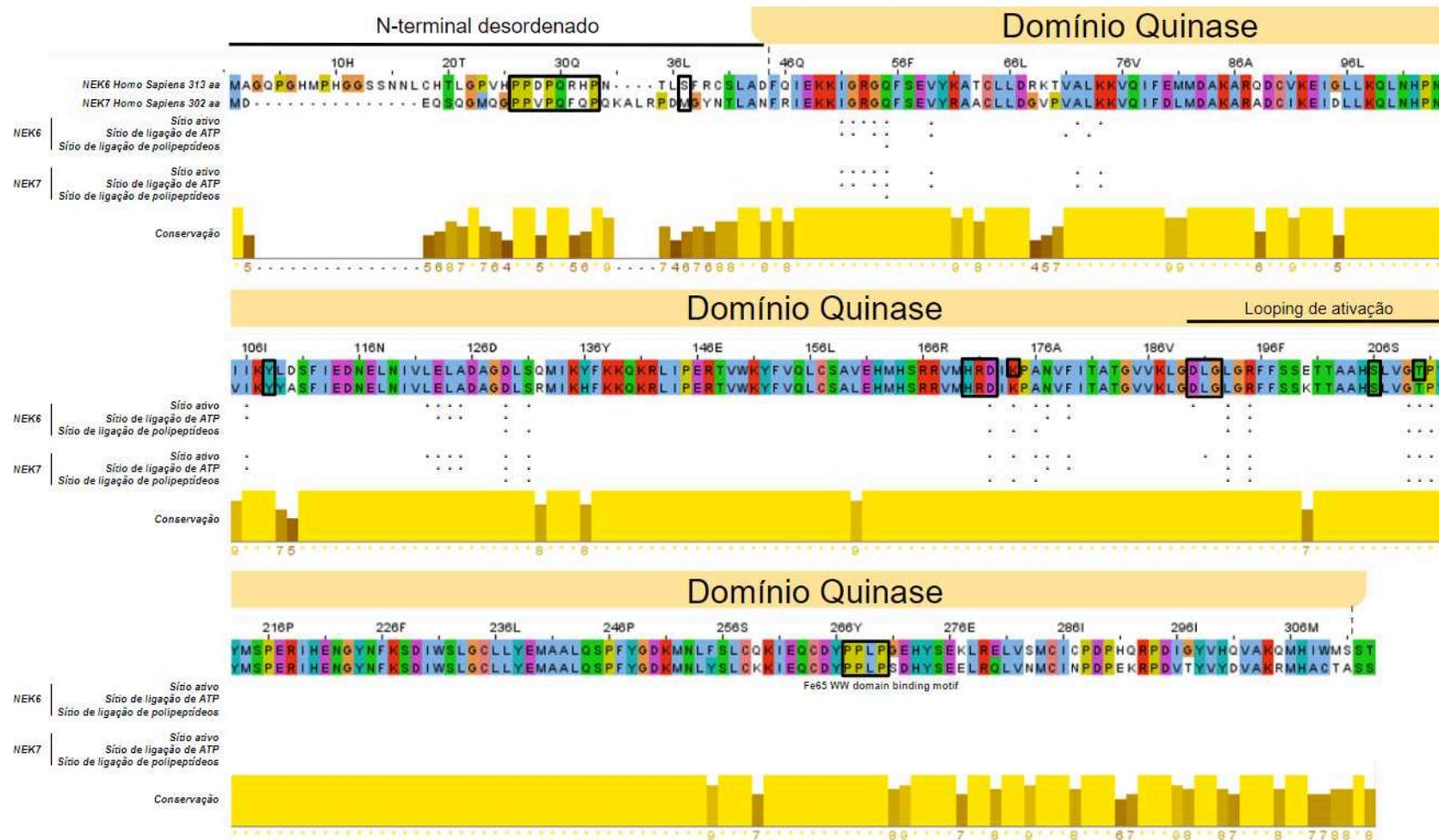


Figura 33. Alinhamento dos aminoácidos da NEK6 (Q9HC98-1) e NEK7 (Q8TDX7-1). Anotações do sítio ativo, sítio de ligação de ATP, sítio de ligação de polipeptídeos, conservação e resíduos regulatórios estão discriminados. Estão representados os aminoácidos regulatórios da NEK6 (Y108, T202, S206, T210, K174 e sequência motivo de ligação do domínio Fe65 WW). Sequências consenso de proteínas quinases estão discriminadas (HRD, DLG e Looping de ativação). Está representada a anotação de uma sequência rica em prolina e a S37 no N-terminal desordenado. A conservação foi analisada pelo Jalview através do alinhamento, representado com escala de 0 (-) a 10 (*).

4.2. Análise das estruturas de alta resolução da NEK7 determinadas experimentalmente identificadas na literatura

Na literatura existem diversas representações de alta resolução da estrutura da NEK7 determinadas experimentalmente por difração de raio X (PDB 5de2, 2WQM, 2WQN, 6S73, 6S75 e 6S76) e por microscopia crioeletrônica (PDB 6NPY) (figura 34). Interessantemente, a NEK7 pode ser cristalizada e formar cristais com celas unitárias distintas, podendo haver uma, duas e quatro unidades por cela unitária.

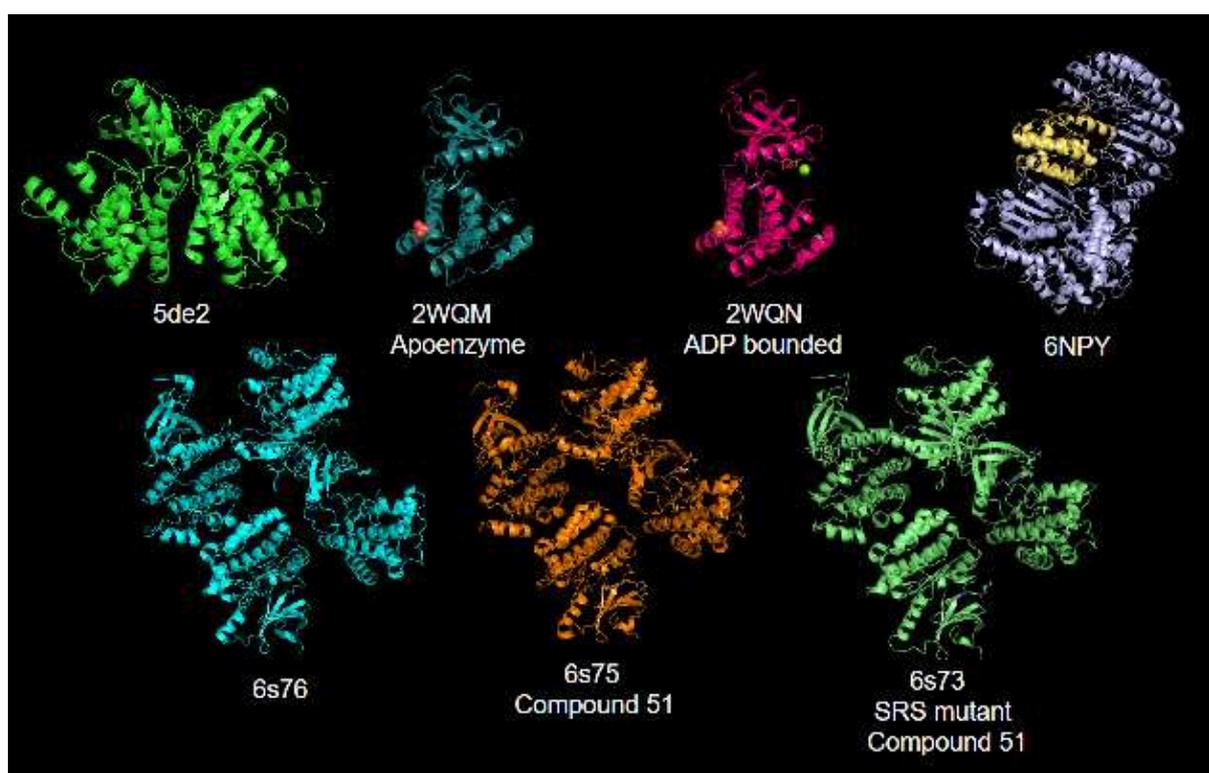


Figura 34. Representações de alta resolução da estrutura e das celas unitárias dos cristais da NEK7 encontrados na literatura. Imagens geradas no Pymol.

Os cristais da NEK7 não possuem resolução do N-terminal desordenado e do looping de ativação, pois são elementos móveis e não geram interferências construtivas devido à variação da posição dos aminoácidos nesses segmentos. Um modelo com boa qualidade possui resolução de $\leq 2 \text{ \AA}$, que se refere a distância mínima de resolução entre dois pontos, em outras palavras, dois objetos só serão observados como dois sinais a partir de uma distância mínima de $\leq 2 \text{ \AA}$ entre elas (lembrando que uma ponte de hidrogênio possui $\sim 3 \text{ \AA}$ de comprimento). O fator R, ou R livre, de um modelo cristalográfico está relacionado ao quanto o modelo estrutural da proteína se ajusta aos dados do mapa de densidade eletrônica. No R livre é considerado o nível de variação da sobreposição dos

átomos da molécula dentro do mapa da nuvem eletrônica e a permissividade dos ângulos diedros. Valores de R livre menores que 0,2 são observados em estruturas com boa resolução (Wlodawer, et al. 2008, 2013). Podemos observar que todos os cristais de NEK7 identificados na literatura possuem bons valores de R livre e de resolução (Tabela 4), indicando boa resolução e qualidade dos modelos gerados por esses cristais, onde o cristal 2WQM é o que possui a melhor qualidade em relação aos valores de R livre e resolução.

Tabela 4. Valores de R livre e resolução (angstroms) dos cristais da NEK7 encontrados na literatura.

| | 5de2 | 2WQM | 2WQN | 6S73 | 6S75 | 6S76 | 6NPY* |
|----------------------|-------|-------|-------|-------|-------|-------|-------|
| R livre | 0,254 | 0,215 | 0,221 | 0,271 | 0,317 | 0,286 | - |
| Resolução (Å) | 2,78 | 2,10 | 2,30 | 3,50 | 3,30 | 3,38 | 3,80 |

*Dados obtidos por microscopia crioeletrônica

Os gráficos de Ramachandran das estruturas de alta resolução da NEK7 identificadas através dos cristais 5de2, 2WQN e 2WQM possuem 100% dos aminoácidos com ângulos diedros em regiões permissivas, denotando alta acurácia do modelo (Tabela 5 e figura 35). Por outro lado, podemos observar que as estruturas identificadas através dos cristais 6s73, 6s75 e 6s76 apresentam alguns aminoácidos com ângulos diedros em regiões não permissivas (Tabela 5 e figura 35). A presença desses resíduos de aminoácidos em regiões não permissivas podem estar relacionadas com a estrutura dos cristais, onde nos cristais 6s73, 6s75 e 6s76 existem 4 monômeros em uma cela unitária (figura 34) nos quais podem forçar movimentação dos resíduos para regiões não permissivas durante a cristalização.

Tabela 5. Porcentagem dos aminoácidos das estruturas da NEK7 determinadas experimentalmente nas regiões discriminadas pelo gráfico de Ramachandran.

| | 5de2 | 6s76 | 6s75 | 6s73 | 2WQN | 2WQM |
|-------------------------------|-------|-------|-------|-------|-------|-------|
| Favorável | 88,6% | 81,8% | 85,0% | 89,8% | 90,0% | 90,8% |
| Adicionais permissivas | 11,0% | 17,1% | 14,1% | 9,9% | 9,6% | 8,8% |
| Permissivas | 0,4% | 0,4% | 0,7% | 0,1% | 0,4% | 0,4% |
| Desfavorável | 0,0% | 0,6% | 0,2% | 0,2% | 0,0% | 0,0% |

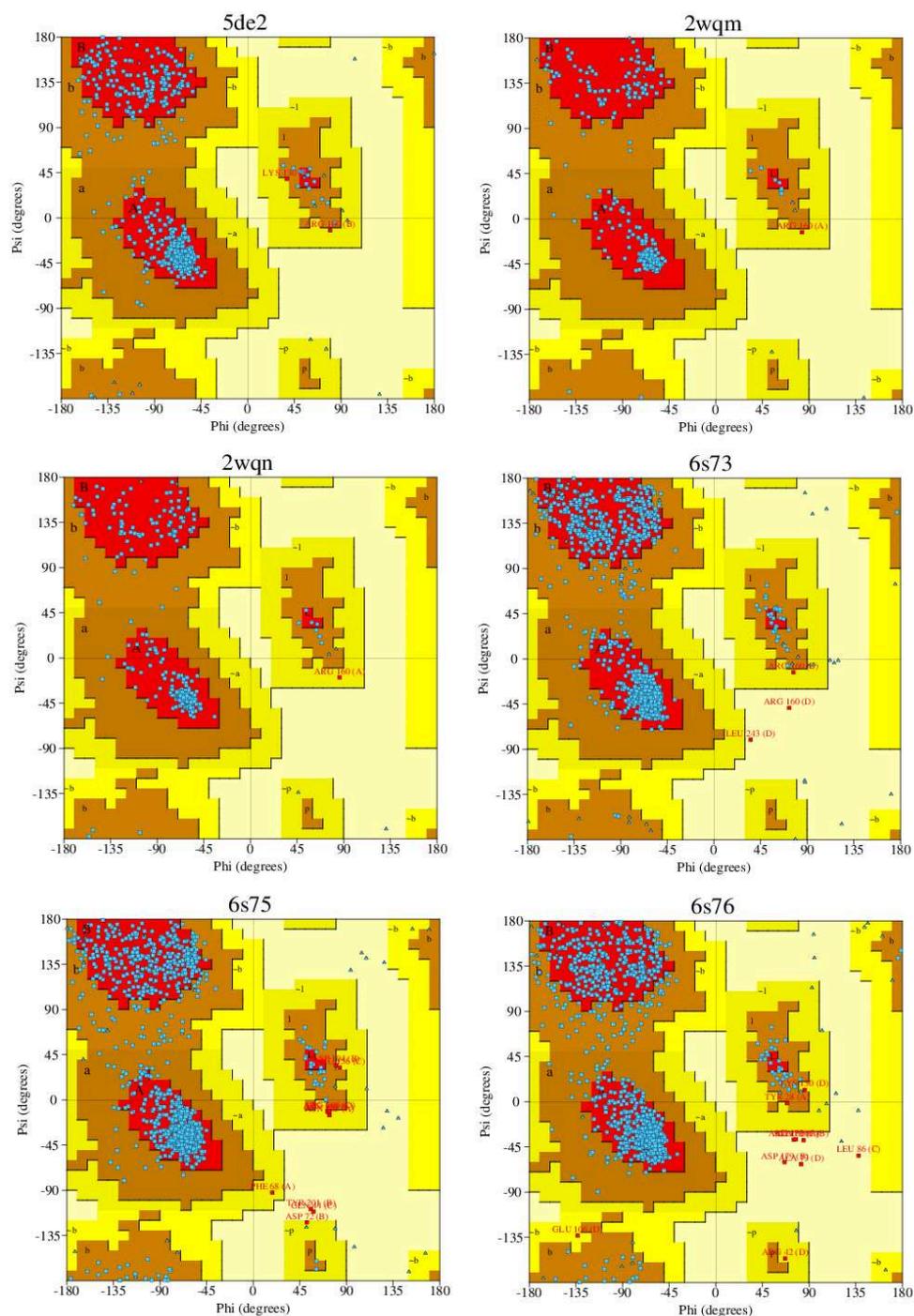


Figura 35. Gráfico de Ramachandran dos modelos estruturais de alta resolução da NEK7 determinados experimentalmente. Gráficos de Ramachandran gerados na plataforma PDBsum.

Todos os resíduos de aminoácidos localizados em estruturas secundárias de todas as estruturas da NEK7 determinadas experimentalmente possuem ângulos diedros em posições permissivas. A análise da estrutura secundária de todos os cristais da NEK7 através do PDBsum demonstrou uma média de $36,5 \pm 0,9$ % de alfa-hélice e $13,1 \pm 0,9$ % de folhas-beta na estrutura da NEK7 (Tabela 6). Os valores relativos de alfa-hélice e folhas-beta das estruturas da NEK7 são muito parecidos, demonstrando que apesar da

estrutura de alta resolução da NEK7 ter sido determinada por diferentes cristais e apresentarem resolução e fator R diferentes, a relação de estruturas secundárias se manteve próxima entre essas estruturas.

Tabela 6. Porcentagem de aminoácidos encontrados em estruturas secundárias das estruturas da NEK7 determinadas experimentalmente.

| | 5de2 | 6s76 | 6s75 | 6s73 | 2WQN | 2WQM |
|--------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| alfa-hélice | 36,1% | 35,4% | 38,1% | 36,1% | 36,4% | 37,1% |
| folhas-beta | 13,9% | 12,6% | 11,9% | 12,6% | 14,2% | 13,6% |

Separamos cada monômero da NEK7 presente nos arquivos PDB 5de2, 2WQM, 2WQN, 6S73, 6S75 e 6S76 e realizamos a superposição das estruturas secundárias e cadeias principais das estruturas de alta resolução da NEK7 determinadas experimentalmente através do software Pymol (figura 36A e 36B). A superposição dos modelos estruturais nos permite identificar a raiz quadrada do desvio quadrático médio (RMSD), no qual é uma medida de sobreposição de duas estruturas no mesmo plano que analisa a distância entre cada par de átomos correspondente entre as duas estruturas. Os valores de RMSD demonstrados na Tabela 6 estão relacionados à superposição dos monômeros do cristal 5de2 com todos os outros monômeros dos cristais da NEK7. Podemos observar uma boa superposição das estruturas secundárias e cadeia principal de cada monômero dos cristais da NEK7 (figura 36A e 36B), no qual refletem os valores baixos da RMSD das superposições das cadeias principais dos cristais, sendo entre 0,628 Å e 2,293 Å (Tabela 7). O mesmo ocorre para as cadeias laterais dos aminoácidos de cada monômero dos cristais da NEK7, onde podemos observar a superposição das cadeias laterais das estruturas da NEK7 determinadas pelos cristais 5de2 e 2WQN (figura 36C).

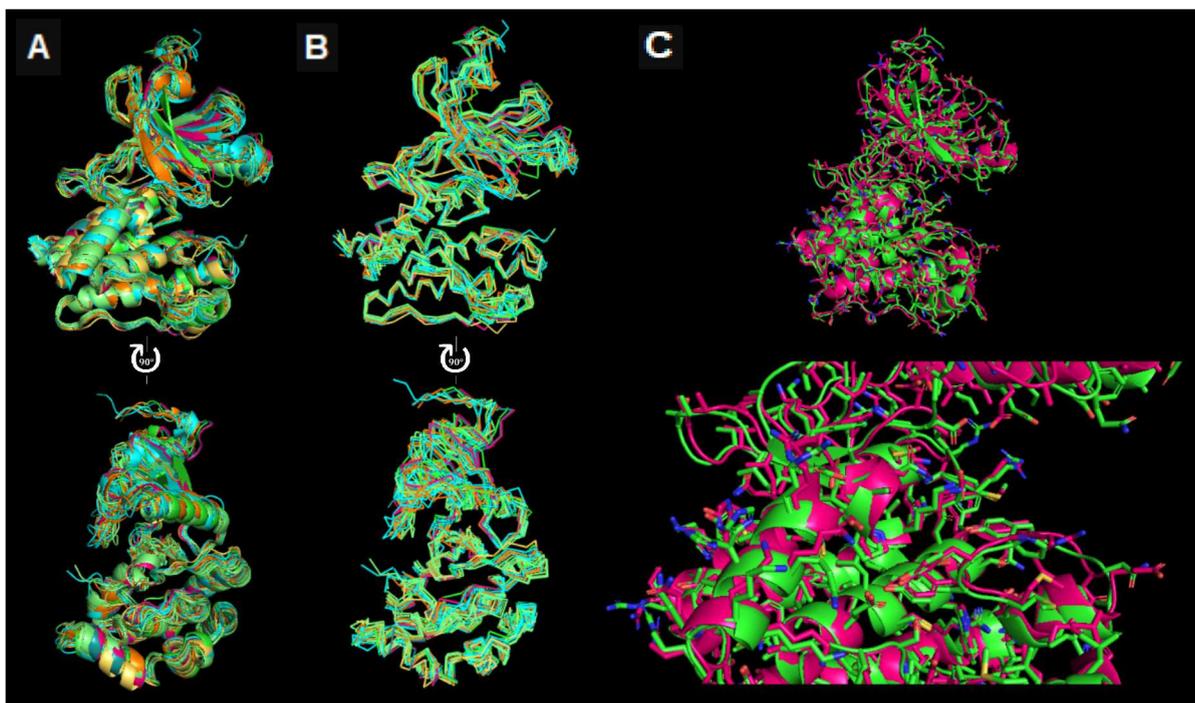


Figura 36. Superposição das estruturas da NEK7 determinadas experimentalmente identificadas na literatura. A) Superposição da estrutura secundária dos monômeros das estruturas de alta resolução da NEK7 determinadas experimentalmente. B) Superposição da cadeia principal dos monômeros das estruturas de alta resolução da NEK7 determinadas experimentalmente. C) Superposição da estrutura secundária e cadeia lateral dos aminoácidos das estruturas de alta resolução da NEK7 determinadas experimentalmente, PDB 5de2 (verde) e PDB 2WQN (rosa).

Tabela 7. RMSD das superposições entre as cadeias principais dos monômeros das estruturas de alta resolução da NEK7 determinadas experimentalmente.

| | 5de2 cristal 1 | 5de2 cristal 2 | 6s76 | 6s75 | 6s73 | 2WQM | 2WQN | 6NPY |
|----------------|----------------|----------------|--|--|--|---------|---------|---------|
| 5de2 cristal 1 | - | 0,707 Å | 1 - 1,164 Å 2 - 1,301 Å 3 - 1,529 Å 4 - 0,933 Å | 1 - 1,261 Å 2 - 1,342 Å 3 - 1,511 Å 4 - 0,956 Å | 1 - 1,969 Å 2 - 1,823 Å 3 - 1,764 Å 4 - 1,357 Å | 0,628 Å | 0,659 Å | 1,032 Å |
| 5de2 cristal 2 | 0,707 Å | - | 1 - 1,332 Å 2 - 1,177 Å 3 - 1,495 Å 4 - 1,015 Å | 1 - 1,406 Å 2 - 1,084 Å 3 - 1,482 Å 4 - 1,033 Å | 1 - 2,293 Å 2 - 2,070 Å 3 - 1,911 Å 4 - 1,633 Å | 0,756 Å | 0,913 Å | 0,941 Å |

As semelhanças entre as estruturas da NEK7 determinadas experimentalmente por difração de raio X de diferentes cristais aumentam a confiabilidade de sua estrutura. Devido a alta homologia entre NEK6 e NEK7, nós usamos os cristais da NEK7 como template para criar modelos de alta resolução da estrutura da NEK6 e NEK7 utilizando a ferramenta AlphaFold2.

4.3. Modelagem da estrutura de alta resolução da NEK6 e NEK7 pelo Alpha Fold 2

Realizamos a modelagem da estrutura de alta resolução da NEK6 e NEK7 através do software AlphaFold2. Os altos escores de confiabilidade (pLDDT) do domínio quinase dos modelos de alta resolução das estruturas da NEK6 e NEK7 preditos pelo AlphaFold2, indicam alta acurácia dos modelos (figura 37). A acurácia do modelo da NEK6 está do aminoácido 39 ao 313, já o modelo da NEK7 possui alta acurácia do aminoácido 22 ao 299.

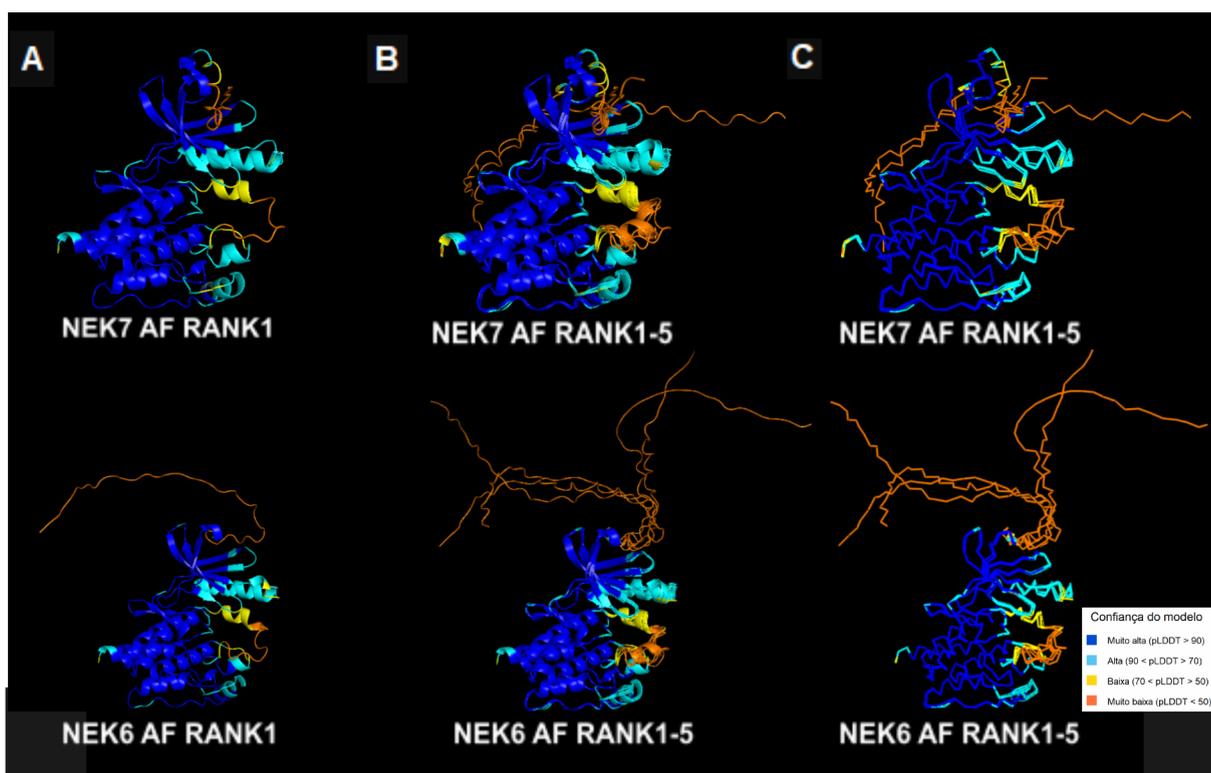


Figura 37. Modelos estruturais de alta resolução da NEK6 e NEK7 preditos pelo AlphaFold2. As predições estão coloridas pelos valores do pLDDT, mostrando o nível de confiança do modelo. A) Representação da estrutura secundária do modelo RANK1 da NEK6 e NEK7. B) Representação da superposição da estrutura secundária dos 5 modelos da NEK6 e NEK7. C) Representação da superposição da cadeia principal dos 5 modelos da NEK6 e NEK7.

Podemos observar que a alfa-hélice C e as voltas beta do lobo N da NEK6 e NEK7 apresentam pLDDT alto em relação aos valores de pLDDT muito alto no restante da estrutura do domínio quinase (figura 37). Esses menores valores de confiabilidade estão relacionados à dinâmica de ativação das quinases, onde ocorrem movimentações da alfa-hélice C e alças das voltas beta. O mesmo ocorre para os elementos móveis do N-terminal e looping de ativação, que apresentam baixos escores de confiabilidade (figura 37). A baixa confiabilidade do looping de ativação ocorre devido a dinâmica estrutural dessa região durante a ativação da proteína quinase. Após fosforilação no looping de ativação

ocorre grande movimentação da alça (looping) para expor o sítio ativo, possibilitando que a enzima quinase reconheça seus substratos e catalise reações de fosforilação.

O gráfico de Ramachandran do modelo da NEK6 gerado pelo AlphaFold2 apresenta 98,5% dos aminoácidos com ângulos diedros em regiões permissivas (Tabela 8 e figura 38). Já o modelo da NEK7 apresenta 100% dos aminoácidos com ângulos diedros em regiões permissivas (Tabela 8 e figura 38). Podemos observar que os aminoácidos do modelo da NEK6 que possuem ângulos diedros em regiões não permissivas no gráficos de Ramachandran estão localizados no N-terminal desordenado da NEK6, nos quais não apresentam boa confiabilidade na predição (figura 38). Já os ângulos diedros dos aminoácidos localizados no domínio quinase dos modelos da NEK6 e NEK7 se encontram em regiões permissivas no gráficos de Ramachandran (figura 38).

Tabela 8. Porcentagem dos aminoácidos das estruturas da NEK6 e NEK7 modeladas pelo AlphaFold2 nas regiões discriminadas pelo gráfico de Ramachandran.

| | AF-NEK6 | AF-NEK7 |
|-------------------------------|---------|---------|
| Favorável | 84,6% | 91,0% |
| Adicionais permissivas | 12,1% | 8,6% |
| Permissivas | 1,8% | 0,4% |
| Desfavorável | 1,5% | 0,0% |

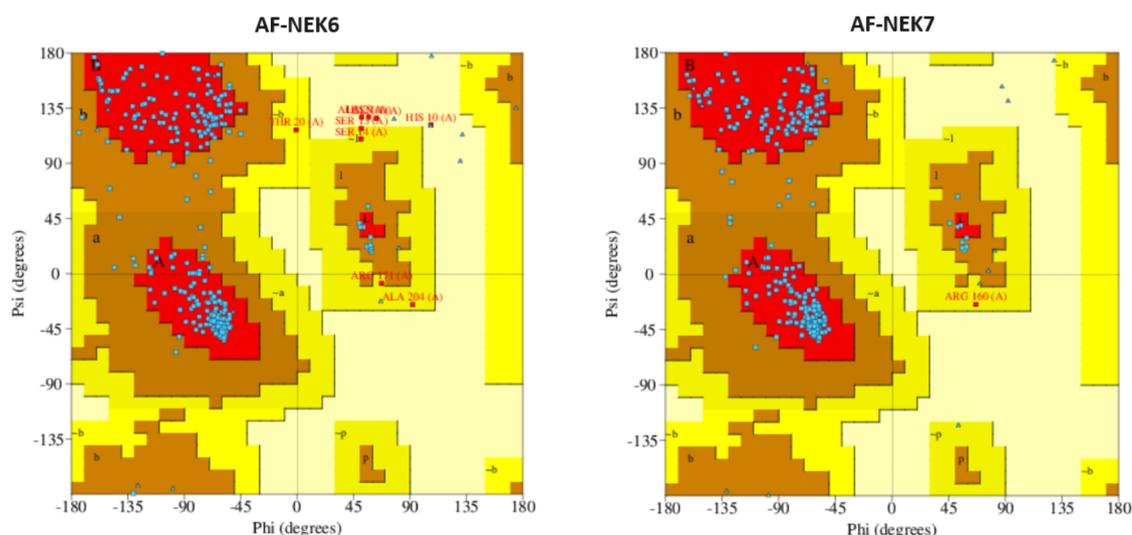


Figura 38. Gráfico de Ramachandran dos modelos estruturais de alta resolução da NEK6 e NEK7 gerados pelo AlphaFold2. Gráfico de Ramachandran gerado pela plataforma PDBsum.

O gráfico de Ramachandran do modelo da NEK6 S206A gerado pelo SWISS-MODEL por Meirelles et al. (2011) apresentou 86,6% dos resíduos em regiões favoráveis, 12,1% em regiões adicionais permissivas, 0,8% em regiões permissivas e 0,4% (1 aminoácido) em região não permissiva. O modelo gerado por Meirelles et al. (2011) não possui o N-terminal desordenado, reduzindo o número de aminoácidos com baixa confiabilidade da predição e consequentemente aumentando o valor relativo de aminoácidos em regiões permissivas no gráfico de Ramachandran. O SWISS-MODEL fez predição de uma alfa-hélice curta dentro do looping de ativação da NEK6, estando também presente no modelo gerado pelo AlphaFold2 (figura 37). Dessa forma, os modelos da NEK6 gerados pelo AlphaFold2 e pelo SWISS-MODEL são semelhantes em termos de validação.

Realizamos a superposição dos modelos da NEK6 e NEK7 com as estruturas de alta resolução da NEK7 determinadas experimentalmente através dos cristais 5de2 e 2WQN (figura 39 e 40, Tabela 9 e 10). Esses cristais foram escolhidos por serem compostos de homodímeros de NEK7 (5de2) e possuir um ADP no sítio ativo (2WQN), posteriormente possibilitando uma análise da interface de contato da homodimerização e ligação do ADP por superposição das estruturas.

Podemos observar ótimas superposições das estruturas secundárias, cadeia principal e cadeias laterais do domínio quinase dos modelos alta resolução das estruturas da NEK7 gerados pelo AlphaFold2 com as estruturas de alta resolução da NEK7 determinadas experimentalmente (PDB 2WQN e 5de2), no qual apresentam baixo valor de RMSD da superposição das cadeias principais, entre 0,665 Å e 0,728 Å (figura 39 e Tabela 9). Isso demonstra que a qualidade do domínio quinase dos modelos da NEK7 gerados pelo AlphaFold2 são comparáveis à estrutura dos modelos determinados experimentalmente.

Interessantemente, apesar do domínio quinase da NEK6 e NEK7 serem 12% não homólogos, podemos observar ótima superposição da estrutura secundária, cadeia principal e cadeia lateral dos aminoácidos homólogos entre os modelos de alta resolução das estruturas da NEK6 gerados pelo AlphaFold2 com as estruturas de alta resolução da NEK7 determinadas experimentalmente (PDB 2WQN e 5de2), no qual apresentam baixo valor de RMSD das superposições das cadeias principais, entre 0,507 Å e 0,671 Å (figura 40 e Tabela 10). Isto indica que os aminoácidos não homólogos do domínio quinase da NEK6 não são suficientes para causar alterações na estrutura, uma vez que o modelo possui altos escores de confiabilidade de predição da estrutura do domínio quinase da NEK6.

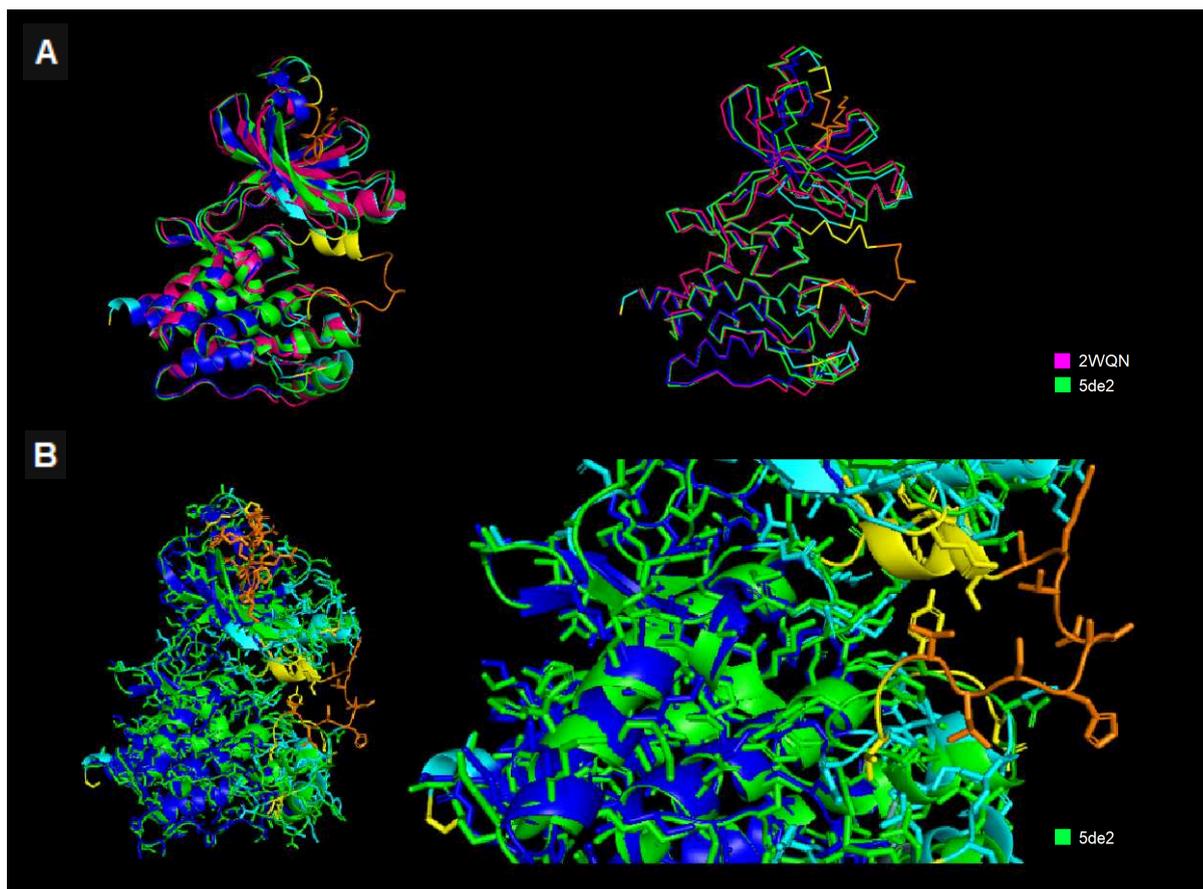


Figura 39. Superposição do modelo da NEK7 gerado pelo AlphaFold2 com as estruturas 5de2 e 2WQN da NEK7. A) Superposição da estrutura secundária e cadeia da principal do modelo 1 da NEK7 predito pelo AlphaFold2 (pLDDT) com as estruturas de alta resolução da NEK7 determinadas experimentalmente, PDB 5de2 (verde) e PDB 2WQN (rosa). B) Superposição da estrutura secundária e cadeia lateral dos aminoácidos do modelo 1 da NEK7 predito pelo AlphaFold2 (pLDDT) com a estrutura de alta resolução da NEK7 determinada experimentalmente, PDB 5de2 (verde).

Tabela 9. RMSD das superposições entre as cadeias principais da estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 5de2) e os 5 modelos da estrutura da NEK7 preditos pelo AlphaFold2 (AF).

| | NEK7 AF RANK1 | NEK7 AF RANK2 | NEK7 AF RANK3 | NEK7 AF RANK4 | NEK7 AF RANK5 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|
| 5de2 cristal 1 | 0,676 Å | 0,696 Å | 0,665 Å | 0,668 Å | 0,728 Å |

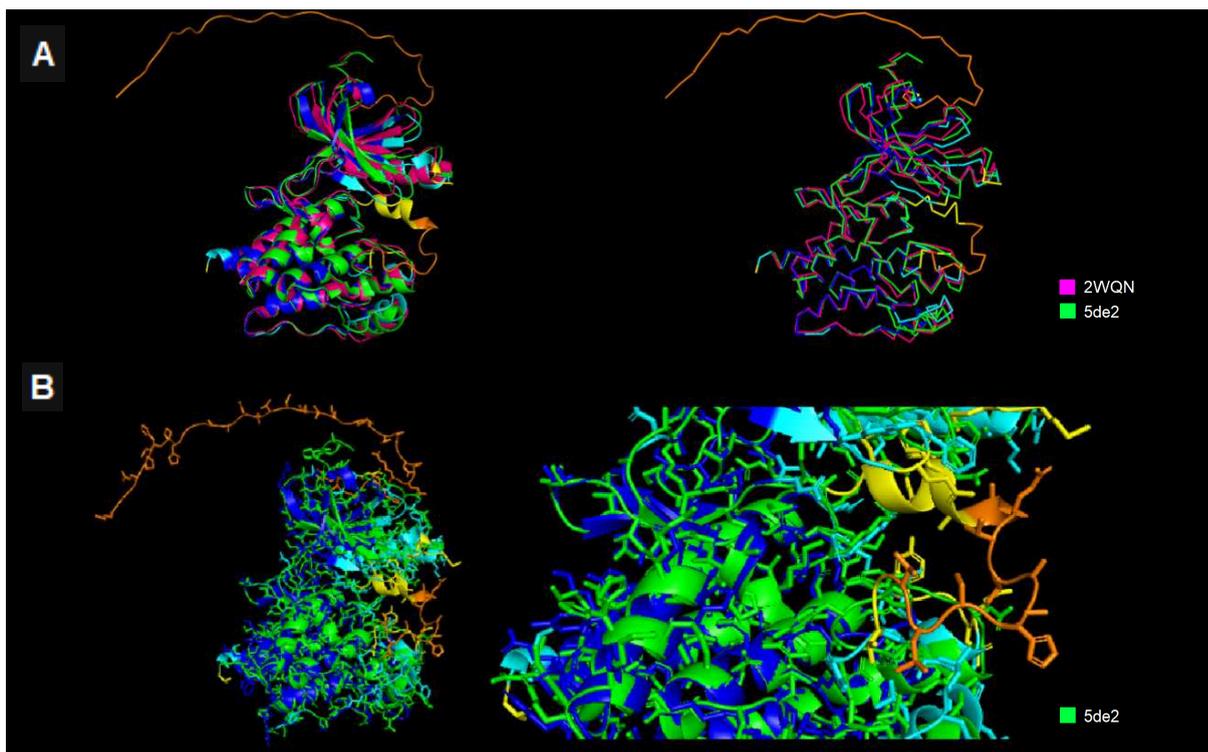


Figura 40. Superposição do modelo da NEK6 gerado pelo AlphaFold2 com as estruturas 5de2 e 2WQN da NEK7. A) Superposição da estrutura secundária e cadeia da principal do modelo 1 da NEK6 predito pelo AlphaFold2 (pLDDT) com as estruturas de alta resolução da NEK7 determinadas experimentalmente, PDB 5de2 (verde) e PDB 2WQN (rosa). B) Superposição da estrutura secundária e cadeia lateral dos aminoácidos do modelo 1 da NEK6 predito pelo AlphaFold2 (pLDDT) com a estrutura de alta resolução da NEK7 determinada experimentalmente, PDB 5de2 (verde).

Tabela 10. RMSD das superposições entre as cadeias principais da estrutura de alta resolução determinada experimentalmente da NEK7 (PDB 5de2) e os 5 modelos da estrutura da NEK6 preditos pelo Alpha Fold 2 (AF).

| | NEK6 AF RANK1 | NEK6 AF RANK2 | NEK6 AF RANK3 | NEK6 AF RANK4 | NEK6 AF RANK5 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|
| 5de2 cristal 1 | 0,671 Å | 0,596 Å | 0,507 Å | 0,578 Å | 0,590 Å |

A análise da estrutura secundária dos modelos da NEK6 e NEK7 gerados pelo AlphaFold2 através do PDBsum demonstraram 37,1% e 37,4% de alfa-hélices, e 13,7% e 14,2% de folhas-beta para NEK6 e NEK7, respectivamente (Tabela 11). Os valores relativos de alfa-hélice e folhas-beta das estruturas dos modelos da NEK6 são muito parecidos aos observados nas estruturas da NEK7 determinadas experimentalmente (Tabela 6), podendo ser consequência da alta homologia entre as sequências polipeptídicas da NEK6 e NEK7.

Tabela 11. Porcentagem de aminoácidos encontrados em estruturas secundárias das estruturas dos modelos da NEK6 e NEK7 gerados pelo AlphaFold2.

| | AF-NEK6 | AF-NEK7 |
|-------------|---------|---------|
| alfa-hélice | 37,1% | 37,4% |
| folhas-beta | 13,7% | 14,2% |

Podemos observar que os valores relativos de estruturas secundárias obtidas pelos modelos gerados pelo AlphaFold2 e através dos softwares utilizados por Meirelles et al. (2011) (Tabela 2) são semelhantes. No entanto, os valores relativos de alfa-hélice obtidos por deconvolução do espectro de dicroísmo circular são maiores em relação aos valores obtidos pelas previsões e modelos, exceto para a deconvolução do espectro de dicroísmo circular segundo Correia & Ramos (2009) realizado por Meirelles et al. (2011).

Dessa forma, obtivemos os modelos de alta resolução das estruturas da NEK6 e NEK7 com alto grau de confiabilidade dentro do domínio quinase (figura 41). O modelo da NEK6 gerado pelo AlphaFold2 possui semelhança ao modelo da NEK6 construído por Meirelles et al. (2011) utilizando o SWISS MODEL (figura 24C). Entretanto, o N-terminal desordenado do modelo da NEK6 gerado pelo AlphaFold2 possui maior grau de liberdade de movimento quando comparado com o grau de liberdade do N-terminal da NEK6 observado experimentalmente através dos dados de SAXS da NEK6 S206A, no qual utilizou combinação entre cálculos *ab initio* e modelagem de corpo rígido (Meirelles, et al. 2011). Isso indica um limite na movimentação do N-terminal da NEK6 no qual o algoritmo do AlphaFold2 não consegue realizar previsões acuradas.

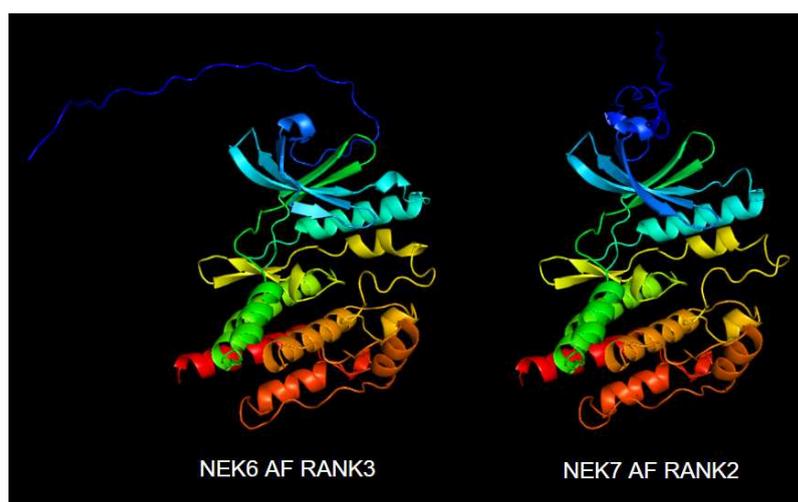


Figura 41. Modelos das estruturas de alta resolução da NEK6 e NEK7 preditos pelo AlphaFold2. As cadeias polipeptídicas estão coloridas pelo espectro visível, indo do azul no N-terminal para vermelho no C-terminal.

4.4. Análise das diferenças e semelhanças estruturais, dos sítios funcionais e regulatórios, conservação dos resíduos e homodimerização utilizando os modelos de alta resolução da NEK6 e NEK7 gerados pelo AlphaFold2

Como podemos observar, existem semelhanças topológicas do sítio ativo e do looping de ativação da NEK6 e NEK7, com essas regiões possuindo distribuição espacial semelhante (figura 42A e 42B). A superposição dos modelos da NEK6 e NEK7 com a estrutura de alta resolução da NEK7 ligada ao ADP no sítio ativo, mostra superposições das cadeias laterais dos resíduos do sítio ativo (sítio de ligação de ATP e sítio de ligação de polipeptídeos) e da leucina 122 da NEK6 (homóloga a L111 da NEK7, que estabiliza o ADP no sítio ativo) (figura 42C).

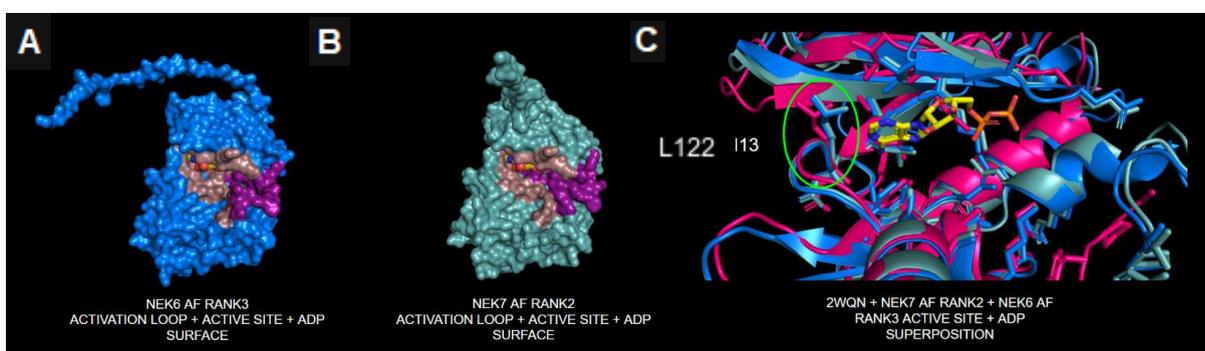


Figura 42. Modelos estruturais da NEK6 e NEK7 gerados pelo AlphaFold2 com anotações de sítios funcionais. A) Representação em superfície do modelo de alta resolução da NEK6 preditos pelo AlphaFold2 com marcação do sítio ativo (cinza), looping de ativação (roxo) e adição do ADP através de superposição com a estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 2WQN). B) Representação em superfície do modelo da estrutura de alta resolução da NEK7 preditos pelo AlphaFold2, com marcação do sítio ativo (cinza), looping de ativação (roxo) e adição do ADP através de superposição com a estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 2WQN). C) Superposição da estrutura secundária e cadeia lateral dos aminoácidos da região do sítio ativo dos modelos da estrutura de alta resolução da NEK6 e NEK7 preditos pelo AlphaFold2 com a estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 2WQN). Podemos observar a superposição da L122 da NEK6 com a L111 da NEK7, no qual estabiliza o ADP no sítio de ligação de ATP (Richards, et al. 2009).

O grau de superposição dessas regiões pode ser explicado pelo nível de conservação destes resíduos na família das Neks de diferentes espécies, onde podemos observar alto nível de conservação do sítio ativo da NEK6 (figura 43A). Podemos observar a presença de uma superfície eletrostática com potencial elétrico positivo no sítio ativo da NEK6 e NEK7 (figura 43B e 43C). Essa semelhança entre o sítio ativo da NEK6 e NEK7 pode ser devido a alta conservação e pode explicar a semelhança da preferência de resíduos ácidos (carregados negativamente) e alguns hidrofóbicos dentro das sequências reconhecidas pela NEK6 e NEK7 para fosforilação (Kooij, et al. 2019), podendo estar relacionado com a semelhança em alguns substratos entre elas, como a EML4 e EG5 (Adib, et al. 2019; Freixo, et al. 2018). No entanto, podemos observar que os aminoácidos não

homólogos entre o domínio quinase da NEK6 e NEK7 estão relacionados com uma alteração na distribuição do potencial eletrostático na camada de solvatação (figura 43B e 43C), podendo explicar as diferenças entre alguns interatores do domínio quinase da NEK6 e NEK7 (Meirelles, et al. 2010; de Souza, et al. 2014).

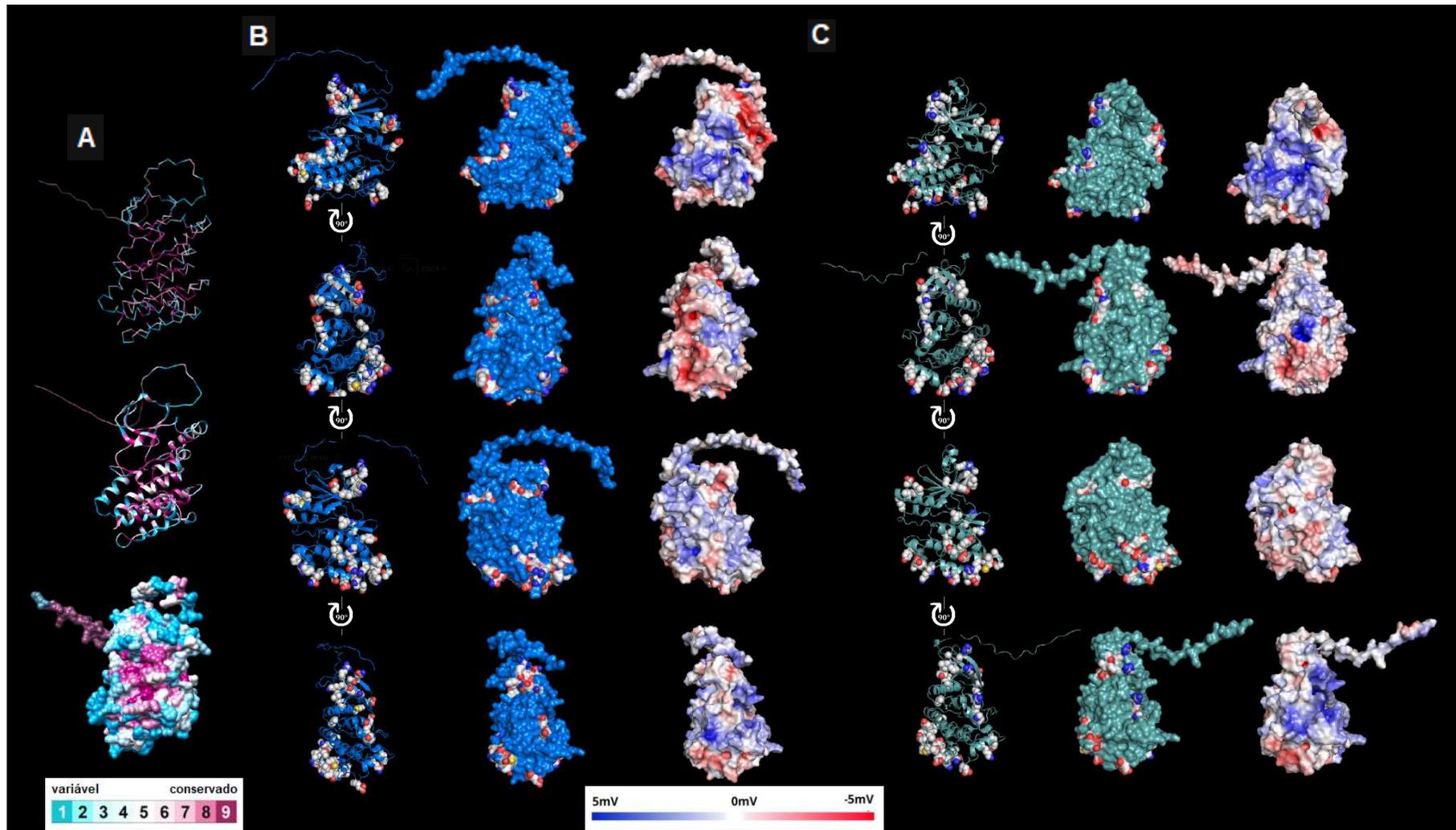


Figura 43. Representações dos modelos estruturais da NEK6 e NEK7 gerados pelo AlphaFold2. A) Representação da cadeia principal, estrutura secundária e superfície do modelo estrutural de alta resolução da NEK6 gerado pelo ConSurf. Coloração de acordo com os escores de conservação dos resíduos de aminoácidos gerados pelo ConSurf. B) Representação da estrutura secundária, superfície e superfície eletrostática do estrutural de alta resolução da NEK6 gerado pelo AlphaFold2. Os aminoácidos não homólogos com NEK7 (Figura 33) estão representados como esferas. C) Representação da estrutura secundária, superfície e superfície eletrostática do modelo estrutural de alta resolução da NEK7 gerado pelo AlphaFold2. Os aminoácidos não homólogos entre o domínio quinase da NEK6 e NEK7 (Figura 33) estão representados como esferas.

Posteriormente destacamos os resíduos regulatórios da NEK6 no modelo da estrutura de alta resolução gerado pelo AlphaFold2 (figura 44). Podemos observar que a K174 pertence a sequência anotada como sítio de ligação de polipeptídeo da NEK6 (figura 33), dessa forma, apesar de Fu et al. (2020) não investigarem se a poliubiquitinação da K174 altera a atividade quinase da NEK6, é possível que tal inibição ocorra. Interessantemente, a sequência rica em prolina (PPLP) está localizada na base do lobo C da NEK6 enquanto a S37 e outra sequência anotada como rica em prolina (PPxPxRxP) estão localizados no N-terminal desordenado da NEK6 (figura 44). Dessa forma, podemos observar que existe interação proteína-proteína em diferentes regiões da estrutura da NEK6, podendo ocorrer no lobo N, lobo C, sítio catalítico (envolve estruturas do lobo N e C) e N-terminal desordenado.

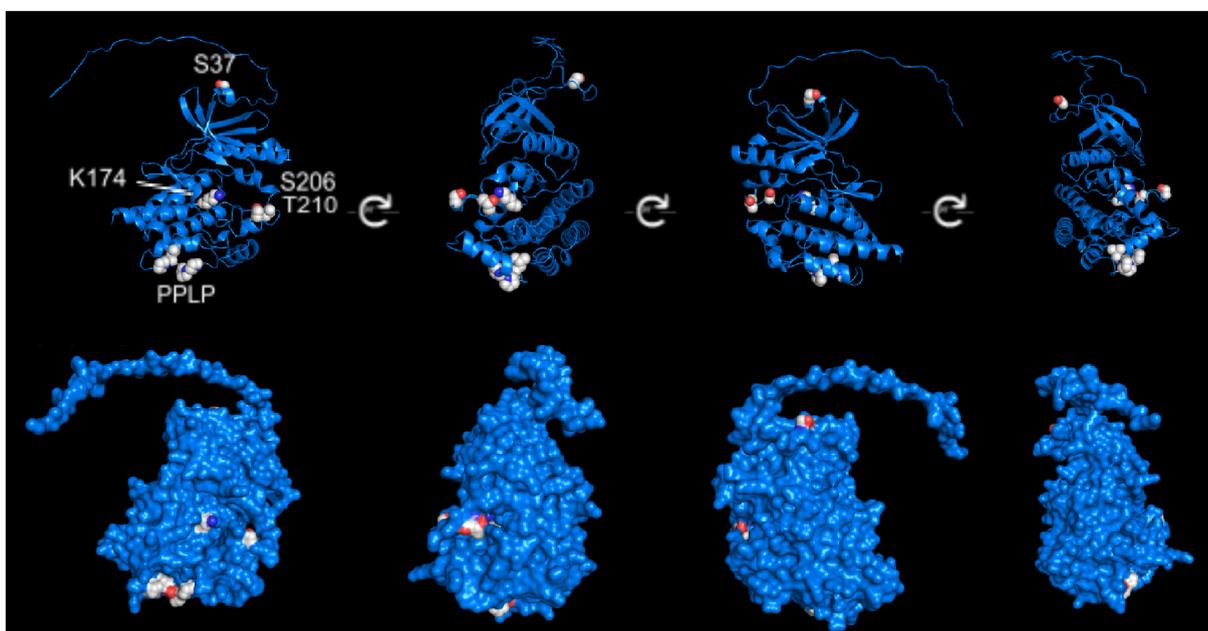


Figura 44. Representação da estrutura secundária, superfície e resíduos regulatórios do modelo estrutural de alta resolução da NEK6 gerado pelo AlphaFold2. Resíduos regulatórios estão demonstrados como esferas. S37 = serina 37; K174 = lisina 174; S206 = serina 206; T210 = treonina 210; PLPP = prolina 267, leucina 268, prolina 269, prolina 270.

Os aminoácidos localizados na interface back-to-back da NEK7 possuem cerca de 65% de homologia com os aminoácidos da NEK6 na mesma posição (Haq, et al. 2015). Dessa forma, realizamos a superposição de dois modelos da estrutura de alta resolução da NEK6 gerados pelo AlphaFold2 com o homodímero da estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 5de2) (figura 45). Os baixos valores de RMSD da superposição entre o dímero de NEK7 e os modelos da NEK6 (figura 40 e Tabela 10) associados a alta homologia entre a NEK6 e NEK7 nos permite especular que a homodimerização da NEK6 apresente a mesma característica estrutural.

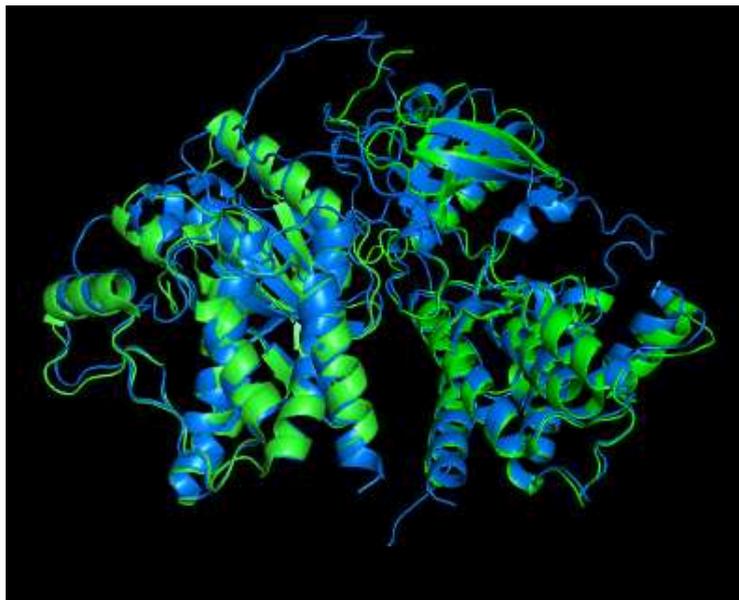


Figura 45. Superposição da estrutura secundária do homodímero da NEK7 determinada experimentalmente (PDB 5de2, em verde) com a estrutura secundária de dois modelos da estrutura da NEK6 gerado pelo AlphaFold2 (em azul).

Foi identificado que a Y28 do N-terminal desordenado da NEK7 junto da Y98 do mesmo monômero formam um cage hidrofóbico para a R155 presente no segundo monômero da interface back-to-back do homodímero de NEK7 (figura 46) (Haq, et al. 2015). Mutações como R155E e truncamento do N-terminal da NEK7 por deleção dos primeiros 30 aminoácidos, M1 ao T30, causam redução da atividade quinase da NEK7 (Haq, et al. 2015; Richards, et al. 2009). A remoção dos primeiros 30 aminoácidos da NEK7, no qual são removidos os aminoácidos P24 ao N29 que estão presentes na interface back-to-back, provavelmente causam alterações no cage hidrofóbico para a R155 formado pela Y28 e Y98, no qual impede a homodimerização e ativação da NEK7, e conseqüentemente redução da atividade quinase. O mesmo ocorre com a mutação N90R da NEK7, no qual faz interação com a cadeia lateral da K96 e cadeia principal da Y97 do segundo monômero do homodímero, auxiliando na modificação da posição auto-inibitória da Y97 (TYR-down para TYR-up) (figura 46) (Haq, et al. 2015).

Observamos que os resíduos N90, K96, Y97 e R155 da NEK7 são homólogos aos resíduos N101, K107, Y108 e R166 da NEK6 (figura 46). No entanto, não existe homologia entre os resíduos na NEK6 que estão na mesma posição da Y28 e Y98 da NEK7. Na mesma posição estão a R39 e L109 da NEK6. A leucina é um aminoácido hidrofóbico, dessa forma a L109 da NEK6 pode ter função na formação do cage hidrofóbico para a R166 da interface back-to-back do homodímero. Por outro lado, a arginina é um aminoácido polar com carga positiva no grupo radical, e dessa forma a R39 da NEK6 poderia impedir a formação do cage hidrofóbico para a R166.

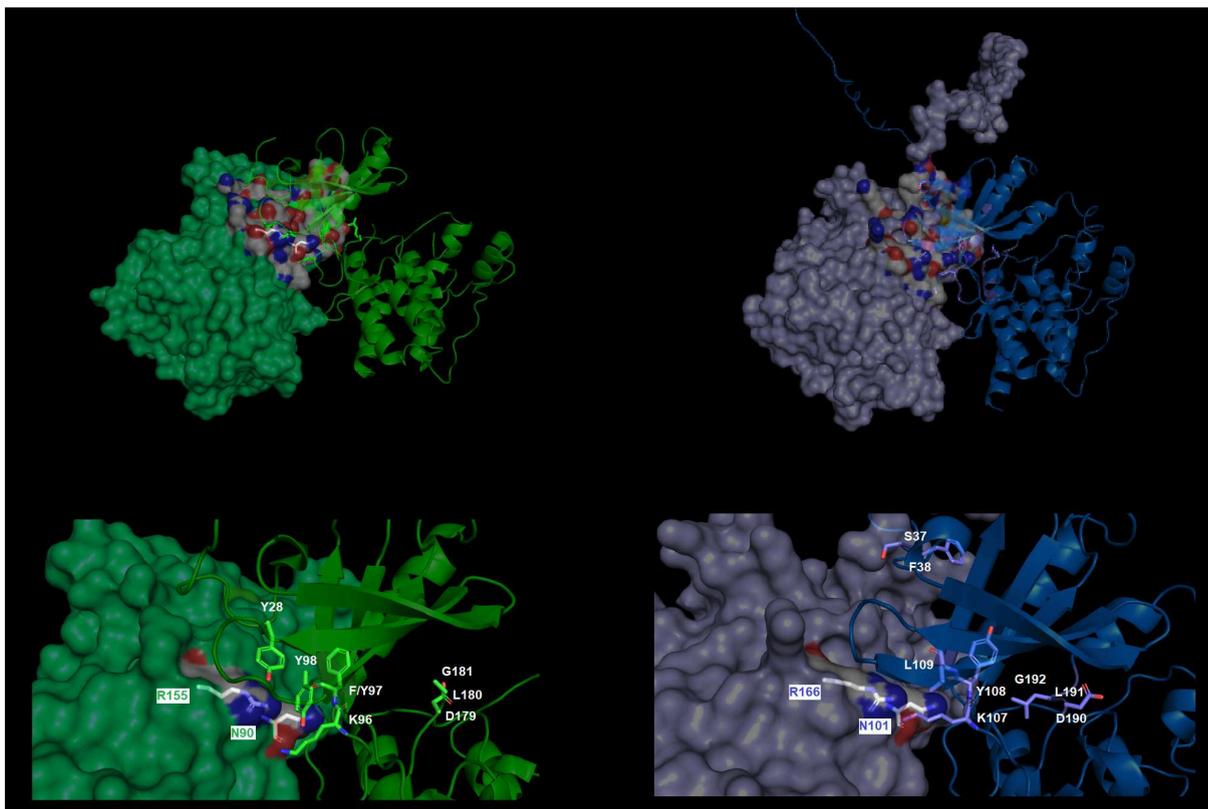


Figura 46. Representação da interface back-to-back da homodimerização da NEK7 (PDB 5de2, verde) e NEK6 (AlphaFold2, azul). Um dos monômeros está representado como estrutura secundária e o outro como superfície. A cadeia lateral dos resíduos Y28, N90, K96, Y97, Y98 e R155 da NEK7 e dos resíduos S37, F38, N101, K107, Y108, L109 e R166 da NEK6 estão destacadas como arestas nos monômeros representados como estrutura secundária. Todos os resíduos da interface back-to-back estão coloridos de acordo com os átomos das cadeias laterais nos monômeros representados com superfície.

Interessantemente, a S37 está dentro da região do N-terminal desordenado da NEK6 (T35, L36, S37, F38, R39 e C40) no qual está alinhada com os resíduos de aminoácidos do N-terminal desordenado da NEK7 (P24, D25, M26, G27, Y28 e N29) que estão voltados para a interface back-to-back do homodímero (figura 46 e 33). E como foi observado por Belham et al. (2003), mutações da S37 da NEK6 não altera sua atividade *in vitro*, existe a possibilidade da sequência do N-terminal desordenado da NEK6 (T35, L36, S37, F38, R39 e C40) não possuir função na interface back-to-back da NEK6, como ocorre com a NEK7. Como a posição dos aminoácidos localizados na interface back-to-back modelada com a NEK6 possuem alta acurácia (alto pLDDT do aminoácido 39 ao 313), é possível que a S37 da NEK6 regule interatores apenas do N-terminal regulatório, sem modificar a homodimerização, ativação e atividade. E como a NEK6 possui uma sequência rica em prolina no N-terminal desordenado, próximo a S37 (figura 33), é possível que os interatores regulados pela S37 realizem o docking na sequência rica em prolina. Essas características do modelo da NEK6 corroboram com a hipótese de que a afinidade da NEK6 pela MBP foi regulada após o estresse genotóxico, e não sua atividade.

4.5. S37 da NEK6 não possui relação com a sua autoativação, coprecipitação e estabilidade

As proteínas quinases que possuem auto-ativação podem ser analisadas através da superexpressão e modificação da mobilidade eletroforética por SDS-PAGE. A homodimerização da NEK6 possui uma constante de dissociação (K_D), a qual prevê um equilíbrio entre as formas de monômeros e homodímeros de NEK6. Dessa forma, a superexpressão da NEK6 aumenta a formação de homodímeros e auto-ativação por ação das massas, onde o aumento da expressão de monômeros inativos causa o aumento da homodimerização e auto-ativação. Com o aumento da homodimerização e auto-ativação ocorre o aumento da banda que possui enzimas ativadas por fosforilação, que migram mais lentamente no SDS-PAGE em comparação às enzimas inativas, que se encontram desfosforiladas e migram mais rapidamente no SDS-PAGE. Esse fenômeno foi demonstrado na NEK6 por Belham et al. (2003), onde a atividade quinase é observada apenas na banda de migração lenta (superior) no qual possui à presença de fosfopeptídeos na S206, T202 e S37. Por outro lado, a banda de migração lenta não apresenta atividade quinase e não possui presença de fosfopeptídeos na S206, T202 e S37. Meirelles et al. (2011) também demonstraram que a NEK6 S206A (mutação no qual impede a ativação por fosforilação do looping de ativação) migra rapidamente como uma única banda.

Nós geramos mutantes fosfodeficiente (S37A) e fosfomimético (S37D) da S37 da NEK6 em construções fusionadas com FLAG e GFP através da mutagênese sítio dirigida (figura 29), coexpressamos as combinações entre FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D em células HEK293T por 30 horas e imunoprecipitamos a FLAG-NEK6 utilizando beads de agarose conjugadas com IgG anti-FLAG. Analisamos a homodimerização da NEK6 pela coprecipitação da GFP-NEK6 com o imunoprecipitado da FLAG-NEK6, e a auto-ativação pela presença da banda de migração lenta no western blot.

A NEK6 possui massa molecular de 35 kDa e o GFP possui massa molecular de 27 kDa, dessa forma a fusão GFP-NEK6 gera uma construção com aproximadamente 62 kDa. Por outro lado, o peptídeo FLAG possui apenas 1 kDa, alterando a massa molecular da fusão FLAG-NEK6 para 36 kDa. A revelação dos sinais das proteínas NEK6 e Beta Tubulina durante o western blot das imunoprecipitações (figura 47) foram realizadas juntas (revelação das 4 membranas marcadas para as duas proteínas ao mesmo tempo). Dessa forma podemos observar que o nível endógeno da NEK6 presente no lisado total (WL) da cotransfecção dos controles (FLAG-vazio com GFP-vazio), possui o sinal muito baixo em relação ao sinal da Beta Tubulina (figura 47). Podemos ver a banda de migração rápida da NEK6 endógena, indicando que a maior parte se apresenta na forma inativa, no qual corrobora com os baixos níveis e baixa atividade da NEK6 observados durante a interfase

(Yin, et al. 2003). Por outro lado, todas as coexpressões de NEK6 causaram aumento dos níveis de NEK6 de forma semelhante, onde podemos observar duas bandas na altura da FLAG-NEK6 (36 kDa) e GFP-NEK6 (62 kDa) no lisado total de todas as construções cotransfectadas, com sinais muito maiores em comparação a Beta Tubulina (figura 47). Todas as coexpressões da NEK6 causaram o surgimento da banda de migração lenta (superior) no lisado total, no qual possui sinal maior em relação a banda de migração rápida (inferior), indicando que a maior parte da NEK6 superexpressa sofreu auto-ativação e que a S37 não altera sua capacidade de auto-ativação (figura 47).

Devido os maiores níveis de FLAG-NEK6 ativa (banda superior) em relação a inativa (banda inferior) no lisado total das cotransfecções, todas as imunoprecipitações de FLAG-NEK6 (WT, S37A e S37D) ocorreram praticamente com a forma ativa da FLAG-NEK6, onde podemos observar apenas a banda de migração lenta da FLAG-NEK6 nas imunoprecipitações (figura 47). Todas as imunoprecipitações capturaram de forma semelhante a maior parte da FLAG-NEK6 superexpressa, indicado pela redução do sinal da banda de migração lenta da FLAG-NEK6 no flow-through (FT) em relação ao lisado total (WL) (figura 47). Podemos observar a coprecipitação da GFP-NEK6 WT com a FLAG-NEK6 WT, no qual a GFP-NEK6 WT está na forma ativa devido a presença apenas da banda de migração lenta (figura 47), indicando que a NEK6 possui a capacidade de homodimerização. Como a ativação alostérica da NEK6 através da homodimerização ocorre anteriormente a fosforilação da S206, a ausência de bandas de migração lenta (desfosforilada, inativa) de GFP-NEK6 coprecipitada com FLAG-NEK6 é um indicativo da natureza transiente da homodimerização entre duas NEK6 desfosforiladas/inativas, e que a ativação através da fosforilação estabiliza a dimerização. Esses dados corroboram com os dados de Meirelles et al. (2011), nos quais indicam que a fosforilação é importante para a homodimerização. E como não houve diferença entre os sinais de GFP-NEK6 WT, S37A e S37D coprecipitados com FLAG-NEK6 WT, S37A e S37D (figura 47), nossos dados indicam que a S37 não possui interferência na homodimerização da NEK6.

Dessa forma, nossos dados indicam que a NEK6 pode formar homodímero no qual é transiente e a ativação por fosforilação aumenta a estabilidade do homodímero, e a S37 não está relacionada ao controle da homodimerização e ativação da NEK6, corroborando com Belham et al. (2003) ao não encontrarem alterações na atividade da NEK6 S37A e NEK6 S37D *in vitro*.

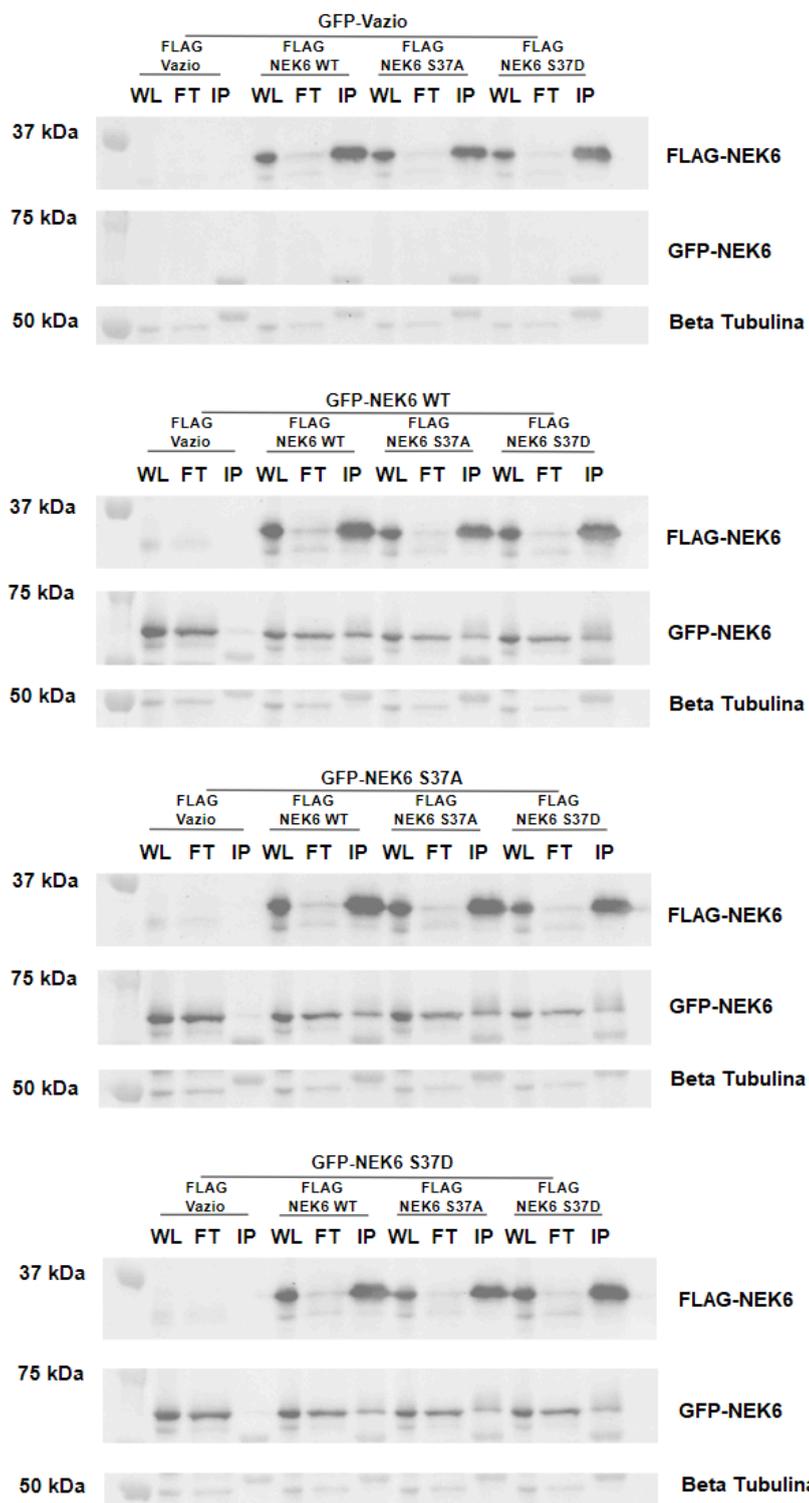


Figura 47. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D. Lisado total (WL), Flow-Through (FT) e imunoprecipitado (IP). Foi utilizado anticorpos anti-NEK6 e anti-Beta Tubulina como controle endógeno neste WB. A SDS-PAGE foi realizada em gel com 10% de acrilamida até a saída da frente de azul de bromofenol. O ponceau S das membranas está demonstrado na figura 48.

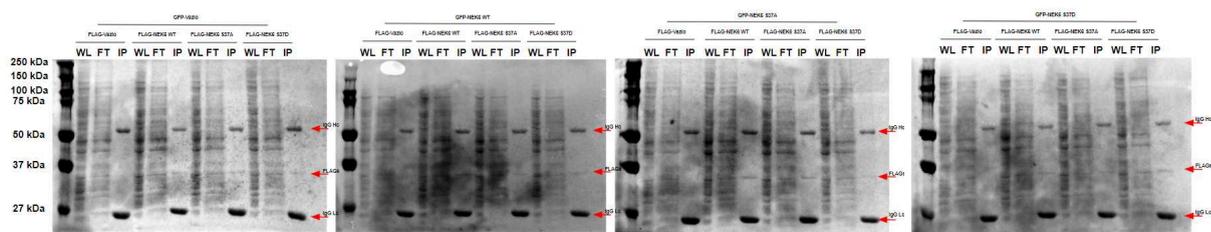


Figura 48. Ponceau S dos western blots das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D da figura 47.

Interessantemente, os níveis de GFP-NEK6 coprecipitadas com FLAG-NEK6 representam uma pequena fração dos níveis de FLAG-NEK6 imunoprecipitado, no qual podemos observar uma leve redução no sinal de GFP-NEK6 no flow through (FT) em comparação ao lisado total (WL) das imunoprecipitações das coexpressões de NEK6 (figura 47). Dessa forma, uma pequena parte da GFP-NEK6 coexpressa se encontra como homodímero ativo (coprecipitado), enquanto a maior parte está na forma de monômeros no estado ativo com uma pequena fração no estado inativo (flow-through) (figura 47). No entanto, a interpretação do equilíbrio entre monômeros e homodímeros de NEK6 possui viés na forma que analisamos, podendo haver maiores níveis de homodímeros do que o observado através da coprecipitação de GFP-NEK6 com FLAG-NEK6 (figura 47). Isso é devido ocorrer a homodimerização de forma transiente entre duas NEK6 inativa/desfosforiladas e provavelmente deve ocorrer homodimerização entre duas FLAG-NEK6 e entre duas GFP-NEK6 durante as cotransfecções, no qual não são possíveis de identificar através do nosso experimento. Sendo assim, a homodimerização da NEK6 ativa/fosforilada pode representar ao menos metade do sinal da NEK6 ativa após superexpressão, com a segunda metade na forma de monômeros ativos. Interessante notar que Belham et al. (2003) encontraram um equilíbrio entre os peptídeos fosforilados em ambos S206/T202 e fosforilados apenas na S206 na banda de migração lenta da NEK6. Por outro lado, fosfopeptídeos apenas na T202 não foram encontrados. Dessa forma, como possivelmente metade da NEK6 superexpressa se encontre na forma de homodímero ativo, é provável que a T202 tenha função na regulação homodimerização da NEK6, podendo aumentar ou reduzir a afinidade entre os homodímeros. Diferentemente do observado por Meirelles et al. (2011) onde foi identificado que a NEK6 é encontrada em sua maioria como monômeros, no entanto esses dados são de proteínas expressas e purificadas de bactérias e essa diferença pode ser devido ao processo de purificação que pode causar a separação dos homodímeros durante as lavagens, ou devido a alguma fosforilação que ocorre apenas em células de eucariotos.

No entanto, a coprecipitação da GFP-NEK6 poderia estar ocorrendo através da interação com a NEK9 endógena, no qual pode estar em um oligômero de NEK9 junto de uma FLAG-NEK6. Sendo assim, nós coexpressamos a FLAG-NEK9 WT com as GFP-NEK6 WT, S37A e S37D em células HEK293T por 30 horas e realizamos a imunoprecipitação da FLAG-NEK9 utilizando beads de agarose conjugadas a IgG anti-FLAG. Como podemos observar no lisado total (WL) na figura 49, a coexpressão de FLAG-NEK9 com GFP-NEK6 causou ativação da NEK6 WT, S37A e S37D devido a presença das duas bandas de GFP-NEK6. Boa parte das GFP-NEK6 coprecipitaram com FLAG-NEK9, indicado pelo sinal no flow through (FT) em relação ao lisado total (WL) da GFP-NEK6 (figura 49). E podemos observar duas bandas de GFP-NEK6 de intensidades semelhantes que coprecipitam com FLAG-NEK9, indicando que interação entre NEK6 e NEK9 possui um equilíbrio entre hetero oligômeros com NEK9 e NEK6 inativas e ativas (figura 49). Como a coprecipitação das GFP-NEK6 com as FLAG-NEK6 ocorreu com a GFP-NEK6 ativa (banda superior) (figura 47), a falta da banda de GFP-NEK6 inativa na coprecipitação é um forte indicativo de que a coprecipitação das GFP-NEK6 com as FLAG-NEK6 não ocorreram de forma mediada pela NEK9 endógena.

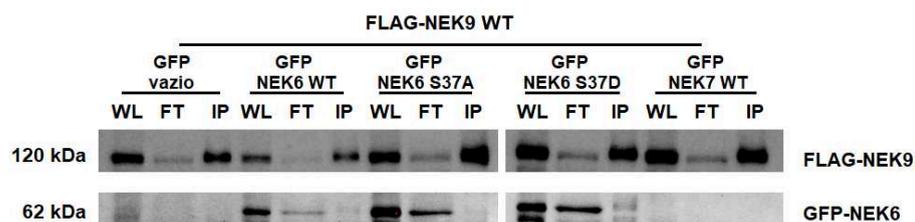


Figura 49. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK9 WT com GFP-NEK6 WT, S37A, S37D e -NEK7 WT. Lisado total (WL), Flow-Through (FT) e imunoprecipitado (IP). Foi utilizado anticorpos anti-NEK9 e anti-NEK6 neste WB. A SDS-PAGE foi realizada em gel com 10% de acrilamida até a saída da frente de azul de bromofenol.

Como a S37 não afetou a homodimerização e auto-ativação da NEK6, é provável que a região do N-terminal desordenado que está voltado para a interface back-to-back do modelo da homodimerização da NEK6 não possua função na dimerização. Sendo assim, é possível que a R166 não possua função na homodimerização da NEK6 (figura 46). Por outro lado, os resíduos N101, K107 e Y108 podem ter função na homodimerização e outros resíduos na interface back-to-back podem ser responsáveis pela dimerização back-to-back e ativação da NEK6. A D67, Q99 e R167 do modelo da NEK6 são resíduos de aminoácidos voltados para a interface back-to-back, no qual possuem superposição com a D56, Q88 e R156 da estrutura de alta resolução da NEK7 determinada experimentalmente (PDB 5de2). Estes resíduos estão próximos de outros resíduos que permitem interação eletrostática e ponte de hidrogênio na interface back-to-back entre os homodímeros. A R167 da NEK6

pode interagir com a cadeia principal da D67 do segundo monômero, e a Q99 pode interagir com a K69 do segundo monômero. No entanto, os resíduos D56, Q88 e R156 da NEK7 não foram modificados e a capacidade de ativação não foi testada experimentalmente.

Identificamos também que as mutantes S37A e S37D não afetam a estabilidade da NEK6 (figura 47). Podemos observar a presença apenas de 2 bandas de NEK6 na altura da FLAG-NEK6 e na altura da GFP-NEK6 no lisado total de todas as cotransfecções, sem a presença do padrão em escada da banda que é característico de degradação da proteína. Também não foi possível observar mudanças na mobilidade eletroforética da NEK6 no qual seria resultante da incorporação de ubiquitinas em sua estrutura, sendo assim, a S37 não possui relação com a poliubiquitinação da K174 e degradação da NEK6.

4.6. S37 da NEK6 regula sua localização celular e afinidade a interatores

Interessantemente, nós identificamos que a localização celular da NEK6 é regulada pela S37 e pelos níveis de NEK9. Durante as cotransfecções de NEK6 nós identificamos o surgimento de células com focos de GFP-NEK6 (figura 50). Podemos observar que a NEK6 WT e S37A induzem grande número de células a formarem focos de GFP. As cotransfecções de GFP-NEK6 S37A com as FLAGs-NEK6 WT, S37A e S37D causaram aumento na quantidade de células com focos de GFP. O mesmo ocorre na cotransfecção de GFP-NEK6 WT com FLAG-NEK6 WT. Entretanto, ocorreu um leve aumento do número de células com focos de GFP na cotransfecção de GFP-NEK6 WT com FLAG-NEK6 S37A, contrário ao observado quando cotransfectamos GFP-NEK6 S37A com FLAG-NEK6 WT. Como o sinal observado é apenas da NEK6 conjugada ao GFP, a menor formação de focos pelo GFP-NEK6 WT quando cotransfectado com FLAG-NEK6 S37A pode ser devido à FLAG-NEK6 S37A estar presente nos focos, causando a redução da quantidade de células com GFP-NEK6 WT nos focos. Isso pode indicar que a NEK6 S37A possui maior afinidade pela estrutura relacionada aos focos em relação a NEK6 WT. Apenas a expressão do GFP-NEK6 WT, S37A e S37D com o FLAG-vazio não foi suficiente para ter grandes aumentos na quantidade de células com focos de GFP, demonstrando que esse fenômeno é dependente dos níveis de NEK6.

Por outro lado, a NEK6 S37D possui inibição no aumento do número de células com focos de GFP (figura 50). As cotransfecções de GFP-NEK6 WT e S37D com FLAG-NEK6 S37D não causam o aumento na quantidade de células com focos de GFP induzidos por elevados níveis de NEK6. Como os níveis dos FLAGs são maiores em relação aos GFPs no lisado total das coexpressões (figura 47), é possível que os maiores níveis de FLAG-NEK6 S37D apresentem efeito antagônico na formação dos focos maior do que o efeito agonista da GFP-NEK6 WT. Nós observamos grande número de células com focos de

GFP nas cotransfecções da GFP-NEK6 S37D com as FLAG-NEK6 WT e S37A (figura 50 e imagens não mostradas), no qual podem também serem explicados pelos maiores níveis dos FLAGs em relação aos GFPs no lisado total das coexpressões (figura 47). Dessa forma, os maiores níveis de FLAG-NEK6 WT e FLAG-NEK6 S37A possuem maior efeito positivo na formação dos focos em relação ao efeito negativo da GFP-NEK6 S37D, e devido a S37 não influenciar na coprecipitação da GFP-NEK6 S37D com as FLAG-NEK6 WT e S37A (figura 47), a GFP-NEK6 S37D pode estar interagindo com a FLAG-NEK6 WT e S37A nos focos formados pela maior expressão delas e assim gerando focos de GFP-NEK6 S37D nessas cotransfecções (figura 50).

Como não observamos a formação de focos de GFP durante a cotransfecção de GFP-NEK6 S37D com FLAG-NEK6 S37D, e observamos focos de GFP durante as cotransfecções de GFP-NEK6 S37D com FLAG-NEK6 WT e S37A (figura 50 e imagens não mostradas), podemos especular que a NEK6 pode estar presente como homodímero nos focos gerados por elevados níveis de NEK6. Dessa forma, é possível que dímeros entre FLAG-NEK6 WT e GFP-NEK6 S37D e entre FLAG-NEK6 S37A e GFP-NEK6 S37D estejam presentes nessas estruturas. Por outro lado, o impedimento da localização da GFP-NEK6 S37D nesses focos durante a cotransfecção com FLAG-NEK6 S37D nos indica de que a regulação da formação dessas estruturas ocorre através de homodímeros de NEK6 ambos fosforilados na S37.

Interessantemente, quando cotransfectamos as GFP-NEK6 WT, S37A e S37D com a FLAG-NEK9 WT (figura 50) ocorreu aumento na localização citosólica da NEK6, onde não é possível observar o surgimento de focos de GFP pela NEK6. Como a maior parte da NEK9 se encontra inativa no citosol e uma pequena fração se encontra ativa nos centrossomos (Roig, et al. 2002; Bertrand, et al. 2011), é provável que a interação entre NEK6 e NEK9 ocorra em maior parte no citosol, causando assim o aumento da localização citosólica da NEK6 na cotransfecção com NEK9. Interessantemente, como não houve aumento nos focos de GFP-NEK6 na cotransfecção com FLAG-NEK9 WT, é provável que a via mitótica da NEKs que ocorre através da hetero oligomerização entre NEK9-NEK6 (figura 25) não ocorra nas estruturas relacionadas com os focos de GFP-NEK6. Dessa forma, a ausência de focos de GFP nas cotransfecções de NEK6 WT, S37A e S37D com NEK9 (figura 50) e coprecipitação de aproximadamente metade da GFP-NEK6 corexpressa com FLAG-NEK9 (figura 49), nos indicam que a localização da NEK6 nessas estruturas ocorrem na forma de homodímeros independentemente de NEK9, e são regulados pela fosforilação da S37 e aumento dos níveis de NEK9, no qual aparentemente sequestra a NEK6 no citosol.

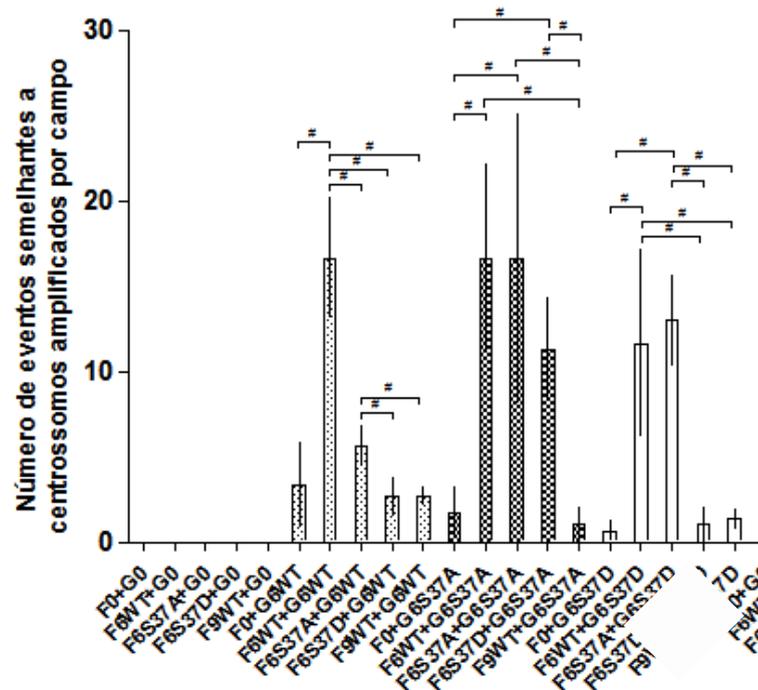
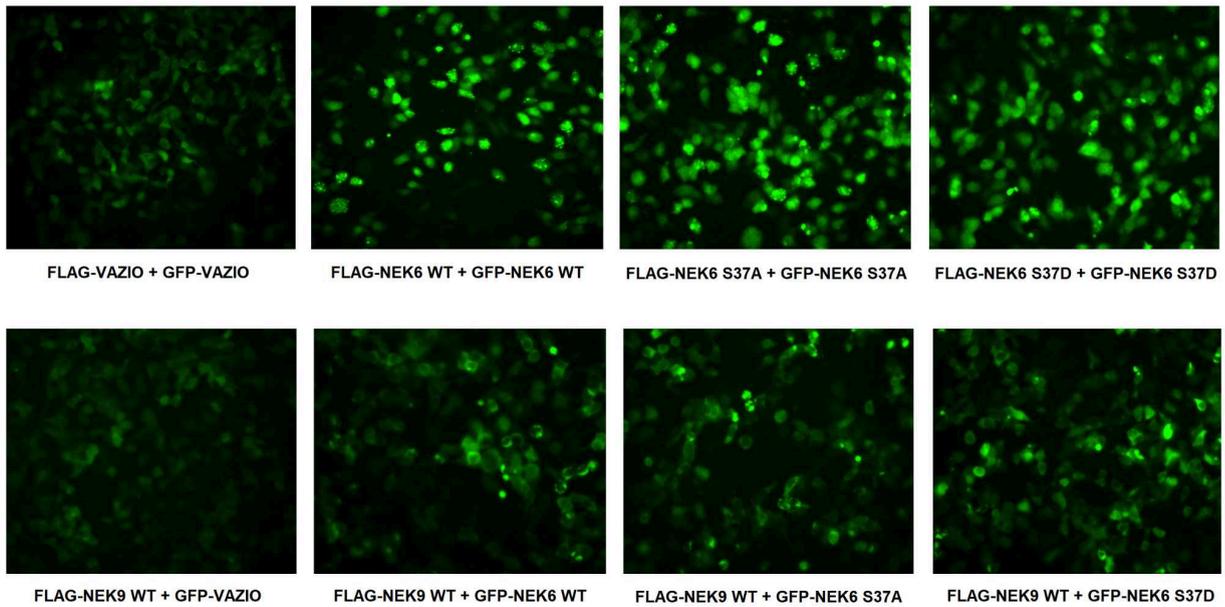


Figura 50. Imunofluorescências representativas do GFP e quantificação de células com focos de GFP semelhantes a centrosomos amplificados de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK9, FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D. O número de células com eventos semelhantes a centrosomos amplificados foram contadas manualmente de imagens com ampliação de 20X. Gráfico em barra representa as médias e desvio padrão dos focos de GFP semelhantes a centrosomos amplificados. Os dados estão apresentados como média \pm desvio padrão, foi realizado o teste ANOVA com pós-teste de Tukey. # = $p < 0.05$. As imagens foram adquiridas de 2 experimentos independentes, onde analisamos 3 campos aleatórios de cada experimento. Parâmetros das imagens de fluorescência: exposição = 1 ms, ganho = 10x, gama = 0,6.

Meirelles et al. (2011) identificaram que a NEK6 possui localização centrossomal através de imunofluorescência, e acreditamos que esses focos (figura 50) podem ser centrossomos amplificados. Nós identificamos células com sinais de GFP-NEK6 semelhantes a presença de 3, 4 e 5 centrossomos (figura 51) nas células observadas na figura 50. Ao mesmo tempo foi possível observar células com distribuição dos focos de GFP-NEK6 semelhantes aos observados nas diferentes fases da mitose (figura 51).

Foi demonstrado que a downregulação dos níveis de NEK6 em células MDA-MB-231 utilizando siRNA causou o aumento de células com centrossomos amplificados com fusos multipolares (Sampson, et al. 2017). Esses dados estão em contraponto às nossas observações, uma vez que nossas condições representam excesso de NEK6 no sistema celular e seria esperado que esses focos estivessem agrupados em pseudo-pólos. Por outro lado, a EG5 pode induzir a amplificação de centrossomos através da geração de excessiva força mecânica e ruptura da estrutura dos centrossomos. Esses fragmentos de centrossomos são formados por material pericentriolar (PCM) e podem conter ou não centríolos. Dessa forma, é possível que em nossas condições, o excesso de NEK6 no sistema celular induziu o aumento de forças atuantes nas estruturas mitóticas, como o fuso mitótico e centrossomos.

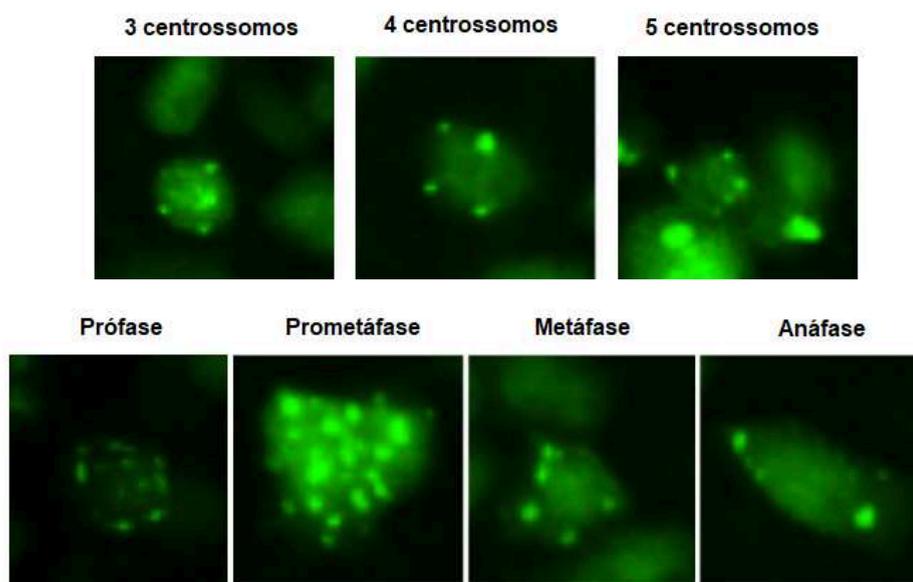


Figura 51. Imagens representativas da imunofluorescência de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT e S37A com GFP-NEK6 WT, S37A e S37D. Imagens de células apresentando três, quatro e cinco focos de GFP semelhantes a centrossomos amplificados. E Imagens representativas da imunofluorescência de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D. Imagens de células apresentando focos de GFP semelhantes a centrossomos amplificados com padrões semelhantes ao observado ao longo das diferentes fases do ciclo celular de células cancerosas (figura 27). Parâmetros das imagens de fluorescência: exposição = 1 ms, ganho = 10x, gama = 0,6.

Como foi identificado que a NEK6 possui localização centrossomal, em colocalização com a Gama Tubulina, e possui interação com proteínas do fuso mitótico e função na regulação de proteínas relacionadas a montagem e movimento dos fusos na separação dos centrossomos e separação das cromátides irmãs, nós analisamos se as FLAG-NEK6 WT, S37A, S37D e FLAG-NEK9 das imunoprecipitações mostradas nas figuras 47 e 49, coprecipitam com proteínas dos centríolos, dos diferentes níveis de organização do PCM e do fuso mitótico (microtúbulos). Nós analisamos a coprecipitação das proteínas do linker centrossomal (C-Nap1 ou CEP250, NEK2), do terminal distal do centríolo-mãe (CCP110), da parede pericentriolar (CEP192 e CENTROBIN), do complexo γ TuRC (Gama Tubulina) e do fuso mitótico (Beta Tubulina). Não foi possível identificar sinal das CEP250, NEK2, CCP110, CEP192 e CENTROBIN nas imunoprecipitações (dados não apresentados), apenas foi possível observar uma variação no sinal de Gama Tubulina e Beta Tubulina (figura 52) indicando que a NEK6 e NEK9 estão localizados na camada mais externa do PCM, provavelmente junto ao complexo γ TuRC.

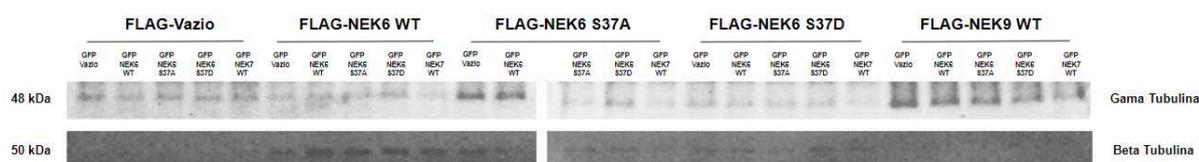


Figura 52. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-NEK6 WT, S37A e S37D. A imagem apresenta apenas o imunoprecipitado (IP) da preparação analisada na figura 47. Foi utilizado anticorpos anti-Gama Tubulina e anti-Beta Tubulina neste WB. A SDS-PAGE foi realizada em gel com 10% de acrilamida até a saída da frente de azul de bromofenol.

Podemos observar que todos os controles das imunoprecipitações das cotransfecções realizados com FLAG-vazio apresentaram sinal moderado de Gama Tubulina, e sinal muito fraco de Beta Tubulina (figura 52). Como a nucleação de microtúbulos ocorre através da interação do terminal negativo dos microtúbulos (no qual está associado a presença da α -Tubulina) com o terminal positivo do γ -TuRC (associado a presença de anéis de Gama Tubulina) (Sulimenko, et al. 2022), é possível que o peptídeo FLAG (no qual é carregado negativamente pela presença de 5 aspartatos na sequência FLAG, DYKDDDDK) interaja de forma inespecífica com essas estruturas. Dessa forma, devido o sinal de Gama Tubulina não variar entre as cotransfecções de FLAG-vazio com os GFP-vazio, NEK6 WT, S37A, S37S e NEK7 WT, é provável que esteja ocorrendo interação inespecífica do peptídeo FLAG com a Gama Tubulina. O mesmo deve ocorrer para Beta Tubulina, no qual é o monômero associado ao terminal positivo dos microtúbulos (Howard, et al. 2003). As imunoprecipitações da FLAG-NEK9 WT das cotransfecções com GFP-vazio, NEK6 WT, S37A e S37D apresentaram aumento do sinal de Gama Tubulina em relação ao

controle, mas não houve sinal de Beta Tubulina (figura 52). O mesmo ocorre para o sinal de Gama Tubulina nas imunoprecipitações da FLAG-NEK6 S37A cotransfectada com GFP-vazio e GFP-NEK6 WT em relação ao controle (figura 52). Por outro lado, as imunoprecipitações da FLAG-NEK6 WT com os GFP-vazio, NEK6 WT, S37A, S37 e NEK7 WT apresentaram aumento do sinal de Beta Tubulina em relação as imunoprecipitações de FLAG-NEK6 S37A e S37D (figura 52). Sendo assim, a presença de sinal de Gama Tubulina e Beta Tubulina nos controles das imunoprecipitações indicam interação inespecífica com o peptídeo FLAG, e aparentemente não apresentam interferência nas imunoprecipitações com as construções de NEK6 e NEK9 devido existir variações nos sinais de Gama Tubulina e Beta Tubulina entre as imunoprecipitações das cotransfecções (figura 52). Para confirmar isto será necessário realizar as cotransfecções e imunoprecipitações com beads conjugadas a anti-GFP, ou utilizar outras tags como His ou c-Myc.

Como discutido, as imunoprecipitações das cotransfecções entre duas NEK6 e entre NEK6 e NEK9 possuem variação nos sinais de Gama Tubulina e Beta Tubulina, dessa forma apresentaremos apenas as imunoprecipitações dos FLAG-NEK6 WT, S37A, S37D e NEK9 WT contranfectados com GFP-vazio, no qual são as condições que não apresentam variação induzida pela cotransfecção por não haver competição entre as FLAG-NEK6 ou FLAG-NEK9 coexpressa com GFP-NEK6 pela Gama Tubulina e Beta Tubulina, no qual interferem nas interpretações (figura 53).



Figura 53. Western blot das imunoprecipitações utilizando beads de agarose conjugada com anti-FLAG de células HEK293T cotransfectadas por 30 horas com combinações de FLAG-NEK6 WT, S37A e S37D com GFP-vazio, removidos da figura 52. A imagem apresenta apenas o imunoprecipitado (IP) da preparação analisada na figura 47 e 49. Foi utilizado anticorpos anti-Gama Tubulina e anti-Beta Tubulina neste WB. A SDS-PAGE foi realizada em gel com 10% de acrilamida até a saída da frente de azul de bromofenol.

Podemos observar na figura 53 que a NEK6 S37A e NEK9 WT possuem maior coprecipitação com Gama Tubulina em relação a imunoprecipitação de NEK6 WT e S37D. Já foi demonstrado que uma pequena porção da NEK9 se encontra ativa durante a prófase, no qual está localizada nos centrossomos junto da CDK1 e PLK1 ativas (Bertran, et al. 2011). Dessa forma, a presença de forte sinal de Gama Tubulina na imunoprecipitação de FLAG-NEK9 com GFP-vazio (figura 53), corrobora com sua localização centrossomal. O

mesmo foi demonstrado para NEK6, no qual foi observado através de imunofluorescência a colocalização com Gama Tubulina nos centrossomos (Meirelles, et al. 2010).

Podemos observar que a cotransfecção das GFP-NEK6 WT, S37A, S37D e GFP-NEK7 WT com FLAG-NEK9 WT causam redução do sinal de Gama Tubulina no imunoprecipitado de FLAG-NEK9 WT (figura 52), indicando que a NEK6 e NEK7 no citosol compete com a Gama Tubulina pela NEK9, reduzindo a interação entre NEK9 e Gama Tubulina. Dessa forma, a redução do sinal de Gama Tubulina na imunoprecipitação de FLAG-NEK9 coexpressa com GFP-NEK6 (figura 52), associado a indução da localização citosólica de NEK6 com a inibição da formação de focos de NEK6 através da coexpressão com NEK9 (figura 50), nos indicam que a localização centrossomal da NEK6 é independente da hetero oligomerização com NEK9.

A mutante NEK6 S37A apresenta maior afinidade pela Gama Tubulina possivelmente devido a impossibilidade de fosforilação da S37 e redução da afinidade pela Gama Tubulina no PCM. Por outro lado, como a NEK6 WT pode ser fosforilada na S37 no PCM, os baixos níveis de Gama Tubulina na imunoprecipitação de NEK6 WT e S37D em oposição aos altos níveis na NEK6 S37A e NEK9 WT nos indicam que a localização da NEK6 no PCM ocorre através da NEK6 desfosforilada na S37 e a fosforilação da S37 regula os níveis de NEK6 no PCM, impedindo o acúmulo de NEK6 nessa estrutura. Dessa forma, uma vez que os focos de GFP são formados por elevados níveis de NEK6 (figura 47 e 50) e a NEK6 S37D impede o aumento de focos semelhantes a centrossomos amplificados, é possível que o acúmulo de NEK6 S37A e WT no PCM ocorra devido à incapacidade das proteínas centrossomais regularem o excesso de NEK6 nos centrossomos através da fosforilação da S37, causando o acúmulo de NEK6 nos centrossomos e aumento de focos semelhantes a centrossomos amplificados em condições de elevados níveis de NEK6.

Por outro lado, NEK6 WT e S37A possuem maior coprecipitação com Beta Tubulina em relação a imunoprecipitação de NEK6 S37D e NEK9 WT (figura 53). Foi demonstrado que a NEK6 possui capacidade de interagir de forma fraca com a Beta Tubulina, imunoprecipitar com feixes de microtúbulos e fosforilar Beta Tubulina *in vitro* (Adib, et al. 2019; Takatani, et al. 2017). Nos microtúbulos, a função da NEK6 está relacionada com a regulação de proteínas regulatórias da dinâmica de microtúbulos (ex. EML4), e na ativação da EG5 para realizar força mecânica para a separação das cromátides irmãs na transição da metáfase para anáfase (Adib, et al. 2019; Bertran, et al. 2011). Sendo assim, os maiores níveis de Beta Tubulina nas imunoprecipitações de NEK6 WT e S37A em relação a NEK6 S37D (figura 53), nos indica que a fosforilação da S37 ao menos regula a interação da NEK6 com Beta Tubulina e provavelmente sua localização nos microtúbulos. Como não existem informações da localização da NEK9 nos microtúbulos, a ausência de Beta Tubulina na imunoprecipitação de NEK9 é um indicativo de que a regulação da função mitótica da NEK6

pela NEK9 aparentemente ocorre no citosol, com posterior translocação da NEK6 para estruturas mitóticas, como centrossomo e fuso mitótico através da interação com Gama Tubulina e Beta Tubulina, respectivamente.

Nós testamos se a EG5 coprecipitou nas imunoprecipitações, no entanto não foi possível identificar o sinal da EG5 no western blot devido a um sinal inespecífico no western blot na altura da EG5, e possivelmente pela interação entre NEK6 e EG5 ser transiente, uma vez que a interação entre elas foi observada através da coexpressão e imunoprecipitação da EG5 e não da NEK6.

Nós realizamos a transfecção da NEK6 WT, S37A e S37D em células HEK293T e desafiamos com a remoção do FBS até que todas as células estivessem mortas visualmente no microscópio de luz. Surpreendentemente, as células transfectadas com NEK6 S37D sobreviveram por 4 meses sem qualquer alteração do meio, enquanto as outras células morreram dentro de alguns dias (dados não mostrados). Meirelles et al. (2010) demonstraram os interatores da NEK6 identificados por duplo híbrido e na plataforma da BioGrid possuem funções no metabolismo, reparo de DNA, resposta ao estresse, síntese proteica, entre outros processos celulares, além de alguns interatores não terem funções descritas. Dessa forma, a S37 da NEK6 aparentemente possui relação com a regulação de vias de sobrevivência e senescência celular, além de proteínas de estruturas mitóticas.

Centrossomos são consideradas organelas sem membranas, onde a formação da organela é governada pela separação líquido-líquido. Segmentos desordenados possuem função na separação líquido-líquido. Dessa forma, é possível que a fosforilação da S37 no N-terminal desordenado da NEK6 esteja relacionada com a separação líquido-líquido e sua localização nas estruturas semelhantes a centrossomos. No entanto, é necessário a confirmação de que os focos de GFP são centrossomos amplificados através de imagens de imunofluorescência, analisando a colocalização das GFP-NEK6 WT e S37A com a Gama Tubulina outras proteínas centrossomais.

4.7. Análise filogenética da S37 da NEK6 de humanos

Nós realizamos uma busca por sequências polipeptídicas homólogas a NEK6 de humanos através do BLASTP e posteriormente o alinhamento múltiplo entre as sequências polipeptídicas com maior nível de homologia. Nós selecionamos as sequências homólogas encontradas em dois mamíferos, uma ave, um réptil, um peixe, um tunicado sésil e um verme (figura 54).

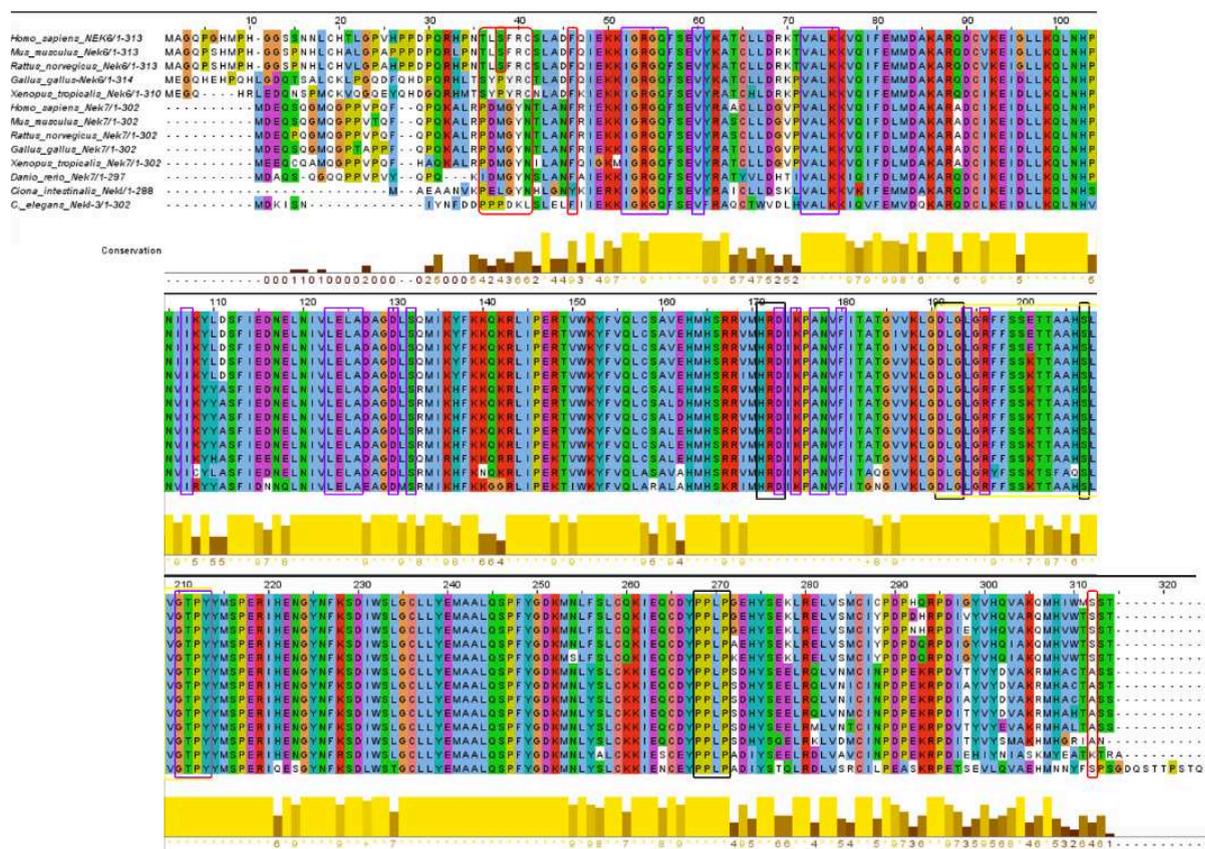


Figura 54. Alinhamento múltiplo de sequências homólogas a NEK6 de *Homo sapiens* identificadas por BLASTP. Estão demonstradas as sequências de proteínas homólogas encontradas em répteis, aves, peixes, mamíferos, tunicado e verme. Estão destacados nos quadrados roxo os resíduos relacionados ao sítio catalítico, nos quadrados pretos os sítios funcionais e regulatórios, e em amarelo o looping de ativação (figura 33). Os quadrados vermelhos estão anotando o N-terminal que faz interação na interface back-to-back da NEK7 (sendo onde está localizado a S37 da NEK6 de *Homo sapiens*), e os aminoácidos que iniciam (fenilalanina e tirosina) e terminam (serina, alanina e lisina) o domínio quinase.

Nós identificamos sequências polipeptídicas da NEK6 e NEK7 apenas nos vertebrados. Não foi identificada a sequência da NEK6 no banco de dados de cDNA de *Danio rerio* (zebrafish), no entanto este organismo possui os locus da NEK6 e NEK7 em seu genoma. A *Ciona intestinalis* é um cordado utilizado em estudos evolutivos dos vertebrados, como apenas a sequência da NEK7 foi identificado nos bancos de dados de cDNA de *Ciona intestinalis*, é possível que o evento de duplicação gênica no qual originou as NEK6 e NEK7 ocorreu após o surgimento do vertebrados. Interessantemente, *C. elegans* não apresenta a duplicação gênica que gerou os locus de NEK6 e NEK7 no genoma por possuir um ortólogo, nekl-3, que é semelhante a NEK6 e NEK7, corroborando a idéia de que a duplicação gênica ocorreu após o surgimento dos vertebrados.

Todas as quinases analisadas no alinhamento múltiplo possuem alto nível de homologia no domínio quinase (figura 54). Podemos observar quase 100% de conservação do sítio do sítio ativo, looping de ativação, sítios funcionais e regulatórios (serina fosforilada

no looping de ativação, sequência HDR, DLG e PLPP localizado no lobo C) entre as quinases analisadas. No entanto, o N-terminal desordenado possui baixa homologia, e interessante apenas nos mamíferos é possível observar a conservação da S37 da NEK6, não estando presente na NEK6 de aves e répteis. Dessa forma, como a S37 da NEK6 possui relação da regulação dos interatores relacionados a estrutura mitótica (gama e Beta Tubulina), regulação de vias de sobrevivência e senescência celular, aparentemente por estar presente apenas nas sequências da NEK6 de mamíferos é possível que a fosforilação da S37 trouxe alguma vantagem para essas espécies em relação a sinalização induzida por estresse. No entanto, são necessárias análises para identificar se algum outro resíduo de serina, tirosina ou treonina presente no N-terminal das ortólogas da NEK6 e da NEK7 que podem ser fosforilados e realizarem a função de regulação dos interatores, localização celular e sinalização de sobrevivência.

5. Conclusões e perspectivas

Nós conseguimos demonstrar que o AlphaFold2 modela o domínio quinase da NEK6 e NEK7 com alta acurácia. A alta homologia entre a NEK6 e NEK7 possibilitou a análise da homodimerização através da superposição do modelo da NEK6 gerado pelo AlphaFold2 com a estrutura da NEK7 (5de2), onde observamos que o N-terminal desordenado da NEK6 não está posicionado na interface do homodímero, sendo um indicativo de que o N-terminal desordenado não possui função na homodimerização e autoativação da NEK6, diferentemente da NEK7. Nossos dados indicam que a NEK6 possui capacidade de homodimerização e auto-ativação, nos quais não foram afetadas por mutações fosfomimética e fosfodeficiente da S37 da NEK6, corroborando com a observação de que a S37 não está próxima a interface de dimerização. A estabilidade da NEK6 também não foi afetada por mutações fosfomimética e fosfodeficiente da S37, não possuindo relação com a poliubiquitinação e degradação da NEK6.

Por outro lado, identificamos que a S37 da NEK6 e a interação com NEK9 regulam a localização celular e um novo fenótipo induzido pela superexpressão de NEK6. A superexpressão de NEK6 induziu o aumento de focos semelhantes a centrossomos amplificados, e a mutação fosfomimética da S37 impediu a formação desses focos. A maior afinidade da NEK6 S37A com Gama Tubulina é um indicativo de sua localização na camada mais externa do PCM, e a localização da NEK6 no PCM aparentemente ocorre com homodímeros de NEK6 de forma independente de NEK9 e é regulado pela fosforilação da S37 de homodímeros de NEK6. No entanto, é necessário confirmar a colocalização da NEK6 com as diferentes proteínas dos centríolos e PCM (CEP250, NEK2, CCP110, CEP192, CENTROBIN e Gama Tubulina) através de imagens de imunofluorescência, para

demonstrar que NEK6 possui colocalização com proteínas centrossomais mas apenas interage com Gama Tubulina, confirmando o novo fenótipo e que a localização centrossomal da NEK6 ocorre na camada mais externa do PCM.

O mesmo ocorre com Beta Tubulina, onde identificamos que a NEK6 S37A possui maior afinidade pela Beta Tubulina, enquanto a NEK6 S37D possui menor afinidade com a Beta Tubulina. A maior afinidade pela Beta Tubulina e Gama Tubulina pela NEK6 S37A em relação a NEK6 S37D é um indicativo de que 37 regula os interatores da NEK6 associados a processos mitóticos.

Devido à função na sobrevivência celular frente à ausência de FBS e alteração na afinidade por interatores pela NEK6 S37D associados à presença da S37 apenas na sequência da Nek6 de mamíferos, são indicativos de que essa regulação trouxe vantagens para os mamíferos frente a situações de estresse. Neste sentido, acreditamos que a S37 é o resíduo da NEK6 fosforilado após o estresse genotóxico. A regulação da localização e interatores sem modificação na atividade quinase da NEK6 pode ser devido a outras funções reguladas pela NEK6, como a homeostase mitocondrial e balanço redox. Dessa forma, a inibição da atividade da NEK6 após o estresse genotóxico poderia induzir disfunção mitocondrial e estresse oxidativo.

Esse conjunto de dados indicam que a inibição da atividade quinase da NEK6 após estresse genotóxico observada por Lee et al. (2008), pode estar relacionada à alteração da afinidade da NEK6 pelo substrato (proteína básica da mielina) após a fosforilação via ATM/ATR-CHK1/CHK2 sensível à cafeína. Dessa forma, é possível que após estresse genotóxico ocorra a fosforilação da S37 da NEK6 para reduzir sua afinidade e localização com estruturas mitóticas sem modificar sua atividade quinase. No entanto, são necessários experimentos utilizando diversos quimioterápicos para analisar o efeito das mutações da S37 na resposta celular após o estresse genotóxico, além de experimento que identifiquem quais são os interatores regulados pela S37 e quais são as proteínas que podem fosforilar a S37 da NEK6.

A NEK6 é considerada uma oncoproteína por estar superexpressa em diversos tipos de tumores sólidos e hematológicos, e a superexpressão induz transformação independente de ancoragem em células JB6 C141 (Jeon, et al. 2010). O câncer de mama apresenta grande incidência de amplificação centrossomal, e pode ser um alvo farmacológico promissor para o tratamento de câncer de mama triplo negativo. Interessantemente, foi identificado que os níveis de NEK6 aumentam ao longo da progressão do câncer de mama, enquanto os níveis de NEK9 sofrem redução (He, et al. 2018; Xu, et al. 2020). Em células MDA-MB-231, na qual é uma célula de câncer de mama triplo negativa, Anuraga et al. (2021) observaram que entre diversos tipos de células de linhagem de câncer de mama, os níveis de transcrito da NEK6 é superior nas células

MDA-MB-231 em relação às demais linhagens analisadas, enquanto os níveis de transcrito da NEK9 são baixos nessa linhagem. Dessa forma, como identificamos que a NEK9 induz a localização citosólica da NEK6 e impede a formação de focos semelhantes a centrosomos amplificados pela superexpressão de NEK6, é possível que a redução dos níveis de NEK9 associado ao aumento dos níveis de NEK6 durante a progressão do câncer de mama possa ter ação no desenvolvimento deste fenótipo, tornando o câncer de mama agressivo.

Sendo assim, concluímos que a S37 localizada no N-terminal desordenado da NEK6 regula a afinidade de alguns interatores e sua localização, e esse mecanismo aparentemente surgiu nos mamíferos.

6. Referências

- Adib, R., Montgomery, J. M., Atherton, J., O'Regan, L., Richards, M. W., Straatman, K. R., Roth, D., Straube, A., Bayliss, R., Moores, C. A., & Fry, A. M. (2019). Mitotic phosphorylation by NEK6 and NEK7 reduces the microtubule affinity of EML4 to promote chromosome congression. *Science Signaling*, 12(594). <https://doi.org/10.1126/scisignal.aaw2939>
- Anuraga, G., Wang, W.-J., Phan, N. N., An Ton, N. T., Ta, H. D. K., Berenice Prayugo, F., Minh Xuan, D. T., Ku, S.-C., Wu, Y.-F., Andriani, V., Athoillah, M., Lee, K.-H., & Wang, C.-Y. (2021). Potential Prognostic Biomarkers of NIMA (Never in Mitosis, Gene A)-Related Kinase (NEK) Family Members in Breast Cancer. *Journal of Personalized Medicine*, 11(11), 1089. <https://doi.org/10.3390/jpm11111089>
- Ardito, F., Giuliani, M., Perrone, D., Troiano, G., & Muzio, L. Io. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *International Journal of Molecular Medicine*, 40(2), 271–280. <https://doi.org/10.3892/ijmm.2017.3036>
- Arter, C., Trask, L., Ward, S., Yeoh, S., & Bayliss, R. (2022). Structural features of the protein kinase domain and targeted binding by small-molecule inhibitors. *Journal of Biological Chemistry*, 298(8), 102247. <https://doi.org/10.1016/j.jbc.2022.102247>
- Azzaz, F., & Fantini, J. (2022). The epigenetic dimension of protein structure. *Biomolecular Concepts*, 13(1), 55–60. <https://doi.org/10.1515/bmc-2022-0006>
- Banik, S. D., & Nandi, N. (2012). Chirality and Protein Biosynthesis (pp. 255–305). https://doi.org/10.1007/128_2012_369
- Bartlett, A. I., & Radford, S. E. (2009). An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nature Structural & Molecular Biology*, 16(6), 582–588. <https://doi.org/10.1038/nsmb.1592>
- Basei, F. L., e Silva, I. R., Dias, P. R. F., Ferezin, C. C., Peres de Oliveira, A., Issayama, L. K., Moura, L. A. R., da Silva, F. R., & Kobarg, J. (2024). The Mitochondrial Connection: The Nek Kinases' New Functional Axis in Mitochondrial Homeostasis. *Cells*, 13(6), 473. <https://doi.org/10.3390/cells13060473>
- Belham, C., Roig, J., Caldwell, J. A., Aoyama, Y., Kemp, B. E., Comb, M., & Avruch, J. (2003). A Mitotic Cascade of NIMA Family Kinases. *Journal of Biological Chemistry*, 278(37), 34897–34909. <https://doi.org/10.1074/jbc.M303663200>
- Bergen, L. G., Upshall, A., & Morris, N. R. (1984). S-phase, G2, and nuclear division mutants of *Aspergillus nidulans*. *Journal of Bacteriology*, 159(1), 114–119. <https://doi.org/10.1128/jb.159.1.114-119.1984>
- Bertoline, L. M. F., Lima, A. N., Krieger, J. E., & Teixeira, S. K. (2023). Before and after AlphaFold2: An overview of protein structure prediction. *Frontiers in Bioinformatics*, 3. <https://doi.org/10.3389/fbinf.2023.1120370>

Bertran, M. T., Sdelci, S., Regué, L., Avruch, J., Caelles, C., & Roig, J. (2011). Nek9 is a Plk1-activated kinase that controls early centrosome separation through Nek6/7 and Eg5. *The EMBO Journal*, 30(13), 2634–2647. <https://doi.org/10.1038/emboj.2011.179>

Boldon, L., Laliberte, F., & Liu, L. (2015). Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Reviews*, 6(1), 25661. <https://doi.org/10.3402/nano.v6.25661>

Brady, L. K., Wang, H., Radens, C. M., Bi, Y., Radovich, M., Maity, A., Ivan, C., Ivan, M., Barash, Y., & Koumenis, C. (2017). Transcriptome analysis of hypoxic cancer cells uncovers intron retention in EIF2B5 as a mechanism to inhibit translation. *PLOS Biology*, 15(9), e2002623. <https://doi.org/10.1371/journal.pbio.2002623>

Brocca, S., Grandori, R., Longhi, S., & Uversky, V. (2020). Liquid–Liquid Phase Separation by Intrinsically Disordered Protein Regions of Viruses: Roles in Viral Life Cycle and Control of Virus–Host Interactions. *International Journal of Molecular Sciences*, 21(23), 9045. <https://doi.org/10.3390/ijms21239045>

Cao, H., Ihm, Y., Wang, C.-Z., Morris, J. R., Su, M., Dobbs, D., & Ho, K.-M. (2004). Three-dimensional threading approach to protein structure recognition. *Polymer*, 45(2), 687–697. <https://doi.org/10.1016/j.polymer.2003.10.091>

Castelo-Soccio, L., Kim, H., Gadina, M., Schwartzberg, P. L., Laurence, A., & O’Shea, J. J. (2023). Protein kinases: drug targets for immunological disorders. *Nature Reviews Immunology*, 23(12), 787–806. <https://doi.org/10.1038/s41577-023-00877-7>

Cheng, P.-N., Pham, J. D., & Nowick, J. S. (2013). The Supramolecular Chemistry of β -Sheets. *Journal of the American Chemical Society*, 135(15), 5477–5492. <https://doi.org/10.1021/ja3088407>

Conduit, P. T., Wainman, A., & Raff, J. W. (2015). Centrosome function and assembly in animal cells. *Nature Reviews Molecular Cell Biology*, 16(10), 611–624. <https://doi.org/10.1038/nrm4062>

de Donato, M., Righino, B., Filippetti, F., Battaglia, A., Petrillo, M., Pirolli, D., Scambia, G., de Rosa, M. C., & Gallo, D. (2018). Identification and antitumor activity of a novel inhibitor of the NIMA-related kinase NEK6. *Scientific Reports*, 8(1), 16047. <https://doi.org/10.1038/s41598-018-34471-y>

de Souza, E. E., Meirelles, G. V., Godoy, B. B., Perez, A. M., Smetana, J. H. C., Doxsey, S. J., McComb, M. E., Costello, C. E., Whelan, S. A., & Kobarg, J. (2014). Characterization of the Human NEK7 Interactome Suggests Catalytic and Regulatory Properties Distinct from Those of NEK6. *Journal of Proteome Research*, 13(9), 4074–4090. <https://doi.org/10.1021/pr500437x>

de Souza, E. E., Meirelles, G. V., Godoy, B. B., Perez, A. M., Smetana, J. H. C., Doxsey, S. J., McComb, M. E., Costello, C. E., Whelan, S. A., & Kobarg, J. (2014). Characterization of the Human NEK7 Interactome Suggests Catalytic and Regulatory Properties Distinct from Those of NEK6. *Journal of Proteome Research*, 13(9), 4074–4090. <https://doi.org/10.1021/pr500437x>

Dokholyan, N. v. (2020). Experimentally-driven protein structure modeling. *Journal of Proteomics*, 220, 103777. <https://doi.org/10.1016/j.jprot.2020.103777>

Donato, M. de, Fanelli, M., Mariani, M., Raspaglio, G., Pandya, D., He, S., Fiedler, P., Petrillo, M., Scambia, G., & Ferlini, C. (2015). Nek6 and Hif-1 α cooperate with the cytoskeletal gateway of drug resistance to drive outcome in serous ovarian cancer. *American Journal of Cancer Research*, 5(6), 1862–1877.

Ellaway, J. I. J., Anyango, S., Nair, S., Zaki, H. A., Nadzirin, N., Powell, H. R., Gutmanas, A., Varadi, M., & Velankar, S. (2024). Identifying protein conformational states in the Protein Data Bank: Toward unlocking the potential of integrative dynamics studies. *Structural Dynamics*, 11(3). <https://doi.org/10.1063/4.0000251>

Fajardo, O. A., Thompson, K., Parapuram, S. K., Liu, S., & Leask, A. (2011). Mithramycin reduces expression of fibro-proliferative mRNAs in human gingival fibroblasts. *Cell Proliferation*, 44(2), 166–173. <https://doi.org/10.1111/j.1365-2184.2011.00738.x>

Fry, A. M., Bayliss, R., & Roig, J. (2017). Mitotic Regulation by NEK Kinase Networks. *Frontiers in Cell and Developmental Biology*, 5. <https://doi.org/10.3389/fcell.2017.00102>

Fry, A. M., O'Regan, L., Sabir, S. R., & Bayliss, R. (2012). Cell cycle regulation by the NEK family of protein kinases. *Journal of Cell Science*. <https://doi.org/10.1242/jcs.111195>

Fu, B., Xue, W., Zhang, H., Zhang, R., Feldman, K., Zhao, Q., Zhang, S., Shi, L., Pavani, K. C., Nian, W., Lin, X., & Wu, H. (2020). MicroRNA-325-3p Facilitates Immune Escape of Mycobacterium tuberculosis through Targeting LNX1 via NEK6 Accumulation to Promote Anti-Apoptotic STAT3 Signaling. *MBio*, 11(3). <https://doi.org/10.1128/mBio.00557-20>

Gillinder, K. R., Tuckey, H., Bell, C. C., Magor, G. W., Huang, S., Ilesley, M. D., & Perkins, A. C. (2017). Direct targets of pSTAT5 signalling in erythropoiesis. *PLOS ONE*, 12(7), e0180922. <https://doi.org/10.1371/journal.pone.0180922>

Guo, H.-B., Perminov, A., Bekele, S., Kedziora, G., Farajollahi, S., Varaljay, V., Hinkle, K., Molinero, V., Meister, K., Hung, C., Dennis, P., Kelley-Loughnane, N., & Berry, R. (2022). AlphaFold2 models indicate that protein sequence determines both structure and dynamics. *Scientific Reports*, 12(1), 10696. <https://doi.org/10.1038/s41598-022-14382-9>

Haq, T., Richards, M. W., Burgess, S. G., Gallego, P., Yeoh, S., O'Regan, L., Reverter, D., Roig, J., Fry, A. M., & Bayliss, R. (2015). Mechanistic basis of Nek7 activation through Nek9 binding and induced dimerization. *Nature Communications*, 6(1), 8771. <https://doi.org/10.1038/ncomms9771>

- Hartley, J. L. (2006). Cloning technologies for protein expression and purification. *Current Opinion in Biotechnology*, 17(4), 359–366. <https://doi.org/10.1016/j.copbio.2006.06.011>
- He, Z., Ni, X., Xia, L., & Shao, Z. (2018). Overexpression of NIMA-related kinase 6 (NEK6) contributes to malignant growth and dismal prognosis in Human Breast Cancer. *Pathology - Research and Practice*, 214(10), 1648–1654. <https://doi.org/10.1016/j.prp.2018.07.030>
- Hollingsworth, S. A., & Karplus, P. A. (2010). A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *BioMolecular Concepts*, 1(3–4), 271–283. <https://doi.org/10.1515/bmc.2010.022>
- Hou, H., Shi, M., Hu, S.-Y., Ahmad, F., Zhang, B., Chen, Z.-H., & Yin, D.-C. (2019). A systematic comparison of sitting and hanging-drop crystallization using traditional and cross-diffusion microbatch crystallization plates. *Journal of Crystal Growth*, 521, 1–8. <https://doi.org/10.1016/j.jcrysgro.2019.05.011>
- Howard, J., & Hyman, A. A. (2003). Dynamics and mechanics of the microtubule plus end. *Nature*, 422(6933), 753–758. <https://doi.org/10.1038/nature01600>
- Huang, B., Kong, L., Wang, C., Ju, F., Zhang, Q., Zhu, J., Gong, T., Zhang, H., Yu, C., Zheng, W.-M., & Bu, D. (2023). Protein Structure Prediction: Challenges, Advances, and the Shift of Research Paradigms. *Genomics, Proteomics & Bioinformatics*, 21(5), 913–925. <https://doi.org/10.1016/j.gpb.2022.11.014>
- Huang, B., Kong, L., Wang, C., Ju, F., Zhang, Q., Zhu, J., Gong, T., Zhang, H., Yu, C., Zheng, W.-M., & Bu, D. (2023). Protein Structure Prediction: Challenges, Advances, and the Shift of Research Paradigms. *Genomics, Proteomics & Bioinformatics*, 21(5), 913–925. <https://doi.org/10.1016/j.gpb.2022.11.014>
- Huang, Y., Koues, O. I., Zhao, J., Liu, R., Pyfrom, S. C., Payton, J. E., & Oltz, E. M. (2017). cis -Regulatory Circuits Regulating NEK6 Kinase Overexpression in Transformed B Cells Are Super-Enhancer Independent. *Cell Reports*, 18(12), 2918–2931. <https://doi.org/10.1016/j.celrep.2017.02.067>
- Jeffries, C. M., Ilavsky, J., Martel, A., Hinrichs, S., Meyer, A., Pedersen, J. S., Sokolova, A. v., & Svergun, D. I. (2021). Small-angle X-ray and neutron scattering. *Nature Reviews Methods Primers*, 1(1), 70. <https://doi.org/10.1038/s43586-021-00064-9>
- Jeon, Y. J., Lee, K. Y., Cho, Y.-Y., Pugliese, A., Kim, H. G., Jeong, C.-H., Bode, A. M., & Dong, Z. (2010). Role of NEK6 in Tumor Promoter-induced Transformation in JB6 C141 Mouse Skin Epidermal Cells. *Journal of Biological Chemistry*, 285(36), 28126–28133. <https://doi.org/10.1074/jbc.M110.137190>
- Jiang, Q., Jin, X., Lee, S.-J., & Yao, S. (2017). Protein secondary structure prediction: A survey of the state of the art. *Journal of Molecular Graphics and Modelling*, 76, 379–402. <https://doi.org/10.1016/j.jmgm.2017.07.015>

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>

Kabir, M. W. U., Alawad, D. M., Mishra, A., & Hoque, M. T. (2023). TAFPred: Torsion Angle Fluctuations Prediction from Protein Sequences. *Biology*, 12(7), 1020. <https://doi.org/10.3390/biology12071020>

Kang, J., Goodman, B., Zheng, Y., & Tantin, D. (2011). Dynamic Regulation of Oct1 during Mitosis by Phosphorylation and Ubiquitination. *PLoS ONE*, 6(8), e23872. <https://doi.org/10.1371/journal.pone.0023872>

Kannaiyan, R., & Mahadevan, D. (2018). A comprehensive review of protein kinase inhibitors for cancer therapy. *Expert Review of Anticancer Therapy*, 18(12), 1249–1270. <https://doi.org/10.1080/14737140.2018.1527688>

Kannaiyan, R., & Mahadevan, D. (2018). A comprehensive review of protein kinase inhibitors for cancer therapy. *Expert Review of Anticancer Therapy*, 18(12), 1249–1270. <https://doi.org/10.1080/14737140.2018.1527688>

Katoh, M., & Katoh, M. (2004). Identification and characterization of Crumbs homolog 2 gene at human chromosome 9q33.3. *International Journal of Oncology*. <https://doi.org/10.3892/ijo.24.3.743>

Keck, F., Ataey, P., Amaya, M., Bailey, C., & Narayanan, A. (2015). Phosphorylation of Single Stranded RNA Virus Proteins and Potential for Novel Therapeutic Strategies. *Viruses*, 7(10), 5257–5273. <https://doi.org/10.3390/v7102872>

Kimura, A., Rieger, M. A., Simone, J. M., Chen, W., Wickre, M. C., Zhu, B.-M., Hoppe, P. S., O'Shea, J. J., Schroeder, T., & Hennighausen, L. (2009). The transcription factors STAT5A/B regulate GM-CSF-mediated granulopoiesis. *Blood*, 114(21), 4721–4728. <https://doi.org/10.1182/blood-2009-04-216390>

Knighton, D. R., Zheng, J., ten Eyck, L. F., Ashford, V. A., Xuong, N.-H., Taylor, S. S., & Sowadski, J. M. (1991). Crystal Structure of the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase. *Science*, 253(5018), 407–414. <https://doi.org/10.1126/science.1862342>

Lee, E. J., Hyun, S. H., Chun, J., & Kang, S. S. (2007). Human NIMA-related kinase 6 is one of the Fe65 WW domain binding proteins. *Biochemical and Biophysical Research Communications*, 358(3), 783–788. <https://doi.org/10.1016/j.bbrc.2007.04.203>

Lee, M. J., & Yaffe, M. B. (2016). Protein Regulation in Signal Transduction. *Cold Spring Harbor Perspectives in Biology*, 8(6), a005918. <https://doi.org/10.1101/cshperspect.a005918>

- Lee, M.-Y., Kim, H.-J., Kim, M.-A., Jee, H. J., Kim, A. J., Bae, Y.-S., Park, J.-I., Chung, J. H., & Yun, J. (2008). Nek6 is involved in G2/M phase cell cycle arrest through DNA damage-induced phosphorylation. *Cell Cycle*, 7(17), 2705–2709. <https://doi.org/10.4161/cc.7.17.6551>
- Maiato, H., & Logarinho, E. (2014). Mitotic spindle multipolarity without centrosome amplification. *Nature Cell Biology*, 16(5), 386–394. <https://doi.org/10.1038/ncb2958>
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The Protein Kinase Complement of the Human Genome. *Science*, 298(5600), 1912–1934. <https://doi.org/10.1126/science.1075762>
- Mariani, V., Biasini, M., Barbato, A., & Schwede, T. (2013). IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, 29(21), 2722–2728. <https://doi.org/10.1093/bioinformatics/btt473>
- Meirelles, G. v, Silva, J. C., Mendonça, Y. de A., Ramos, C. H., Torriani, I. L., & Kobarg, J. (2011). Human Nek6 is a monomeric mostly globular kinase with an unfolded short N-terminal domain. *BMC Structural Biology*, 11(1), 12. <https://doi.org/10.1186/1472-6807-11-12>
- Minoguchi, S., Minoguchi, M., & Yoshimura, A. (2003). Differential control of the NIMA-related kinases, Nek6 and Nek7, by serum stimulation. *Biochemical and Biophysical Research Communications*, 301(4), 899–906. [https://doi.org/10.1016/S0006-291X\(03\)00049-4](https://doi.org/10.1016/S0006-291X(03)00049-4)
- Mittal, K., Kaur, J., Jaczko, M., Wei, G., Toss, M. S., Rakha, E. A., Janssen, E. A. M., Søiland, H., Kucuk, O., Reid, M. D., Gupta, M. v., & Aneja, R. (2021). Centrosome amplification: a quantifiable cancer cell trait with prognostic value in solid malignancies. *Cancer and Metastasis Reviews*, 40(1), 319–339. <https://doi.org/10.1007/s10555-020-09937-z>
- Mohd Salleh, M. H., Fitri Peli, A., Ngalimat, M. S., & Sim, K. J. (2022). A Mini Literature Review on Current Advancements in Protein Purification Techniques. *The 2nd International Electronic Conference on Biomolecules: Biomacromolecules and the Modern World Challenges*, 12. <https://doi.org/10.3390/IECBM2022-13507>
- Moniz, L., Dutt, P., Haider, N., & Stambolic, V. (2011). Nek family of kinases in cell cycle, checkpoint control and cancer. *Cell Division*, 6(1), 18. <https://doi.org/10.1186/1747-1028-6-18>
- Moraes, E., Meirelles, G., Honorato, R., de Souza, T., de Souza, E., Murakami, M., de Oliveira, P., & Kobarg, J. (2015). Kinase Inhibitor Profile for Human Nek1, Nek6, and Nek7 and Analysis of the Structural Basis for Inhibitor Specificity. *Molecules*, 20(1), 1176–1191. <https://doi.org/10.3390/molecules20011176>
- Muhammed, M. T., & Aki-Yalcin, E. (2019). Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chemical Biology & Drug Design*, 93(1), 12–20. <https://doi.org/10.1111/cbdd.13388>

O'Regan, L., & Fry, A. M. (2009). The Nek6 and Nek7 Protein Kinases Are Required for Robust Mitotic Spindle Formation and Cytokinesis. *Molecular and Cellular Biology*, 29(14), 3975–3990. <https://doi.org/10.1128/MCB.01867-08>

O'Regan, L., Sampson, J., Richards, M. W., Knebel, A., Roth, D., Hood, F. E., Straube, A., Royle, S. J., Bayliss, R., & Fry, A. M. (2015). Hsp72 is targeted to the mitotic spindle by Nek6 to promote K-fiber assembly and mitotic progression. *Journal of Cell Biology*, 209(3), 349–358. <https://doi.org/10.1083/jcb.201409151>

Pan, A., Pranavathiyani, G., & Chakraborty, S. sen. (2021). Computational Modeling of Protein Three-Dimensional Structure: Methods and Resources. In *Molecular Docking for Computer-Aided Drug Design* (pp. 155–178). Elsevier. <https://doi.org/10.1016/B978-0-12-822312-3.00023-0>

Pavan, I. C. B., Basei, F. L., Severino, M. B., Rosa e Silva, I., Issayama, L. K., Mancini, M. C. S., Góis, M. M., da Silva, L. G. S., Bezerra, R. M. N., Simabuco, F. M., & Kobarg, J. (2023). NEK6 Regulates Redox Balance and DNA Damage Response in DU-145 Prostate Cancer Cells. *Cells*, 12(2), 256. <https://doi.org/10.3390/cells12020256>

Pavan, I. C. B., Peres de Oliveira, A., Dias, P. R. F., Basei, F. L., Issayama, L. K., Ferezin, C. de C., Silva, F. R., Rodrigues de Oliveira, A. L., Alves dos Reis Moura, L., Martins, M. B., Simabuco, F. M., & Kobarg, J. (2021). On Broken Ne(c)ks and Broken DNA: The Role of Human NEKs in the DNA Damage Response. *Cells*, 10(3), 507. <https://doi.org/10.3390/cells10030507>

Peres de Oliveira, A., Kazuo Issayama, L., Betim Pavan, I. C., Riback Silva, F., Diniz Melo-Hanchuk, T., Moreira Simabuco, F., & Kobarg, J. (2020). Checking NEKs: Overcoming a Bottleneck in Human Diseases. *Molecules*, 25(8), 1778. <https://doi.org/10.3390/molecules25081778>

Pihan, G. A. (2013). Centrosome Dysfunction Contributes to Chromosome Instability, Chromoanagenesis, and Genome Reprograming in Cancer. *Frontiers in Oncology*, 3. <https://doi.org/10.3389/fonc.2013.00277>

Quarmby, L. M., & Mahjoub, M. R. (2005). Caught Nek-ing: cilia and centrioles. *Journal of Cell Science*, 118(22), 5161–5169. <https://doi.org/10.1242/jcs.02681>

Rapley, J., Nicolàs, M., Groen, A., Regué, L., Bertran, M. T., Caelles, C., Avruch, J., & Roig, J. (2008). The NIMA-family kinase Nek6 phosphorylates the kinesin Eg5 at a novel site necessary for mitotic spindle formation. *Journal of Cell Science*, 121(23), 3912–3921. <https://doi.org/10.1242/jcs.035360>

Rashid, M. A., Khatib, F., & Sattar, A. (2015). Protein preliminaries and structure prediction fundamentals for computer scientists. <https://doi.org/10.48550/arXiv.1510.02775>

- Richards, M. W., O'Regan, L., Mas-Droux, C., Blot, J. M. Y., Cheung, J., Hoelder, S., Fry, A. M., & Bayliss, R. (2009). An Autoinhibitory Tyrosine Motif in the Cell-Cycle-Regulated Nek7 Kinase Is Released through Binding of Nek9. *Molecular Cell*, 36(4), 560–570. <https://doi.org/10.1016/j.molcel.2009.09.038>
- Rico, F., Rigato, A., Picas, L., & Scheuring, S. (2013). Mechanics of proteins with a focus on atomic force microscopy. *Journal of Nanobiotechnology*, 11(Suppl 1), S3. <https://doi.org/10.1186/1477-3155-11-S1-S3>
- Rogers, D. M., Jasim, S. B., Dyer, N. T., Auvray, F., Réfrégiers, M., & Hirst, J. D. (2019). Electronic Circular Dichroism Spectroscopy of Proteins. *Chem*, 5(11), 2751–2774. <https://doi.org/10.1016/j.chempr.2019.07.008>
- Roig, J., Mikhailov, A., Belham, C., & Avruch, J. (2002). Nercc1, a mammalian NIMA-family kinase, binds the Ran GTPase and regulates mitotic progression. *Genes & Development*, 16(13), 1640–1658. <https://doi.org/10.1101/gad.972202>
- Rouhi, M., Hajizadeh, Z., Taheri-Ledari, R., Maleki, A., & Babamoradi, M. (2022). A review of mechanistic principles of microwave absorption by pure and composite nanomaterials. *Materials Science and Engineering: B*, 286, 116021. <https://doi.org/10.1016/j.mseb.2022.116021>
- Sabat-Pośpiech, D., Fabian-Kolpanowicz, K., Prior, I. A., Coulson, J. M., & Fielding, A. B. (2019). Targeting centrosome amplification, an Achilles' heel of cancer. *Biochemical Society Transactions*, 47(5), 1209–1222. <https://doi.org/10.1042/BST20190034>
- Sampson, J., O'Regan, L., Dyer, M. J. S., Bayliss, R., & Fry, A. M. (2017). Hsp72 and Nek6 Cooperate to Cluster Amplified Centrosomes in Cancer Cells. *Cancer Research*, 77(18), 4785–4796. <https://doi.org/10.1158/0008-5472.CAN-16-3233>
- Schwartz, S. D. (2023). Protein Dynamics and Enzymatic Catalysis. *The Journal of Physical Chemistry B*, 127(12), 2649–2660. <https://doi.org/10.1021/acs.jpcc.3c00477>
- Seki, M. (2023). On the origin of the genetic code. *Genes & Genetic Systems*, 98(1), 22–00085. <https://doi.org/10.1266/ggs.22-00085>
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792), 706–710. <https://doi.org/10.1038/s41586-019-1923-7>
- Stollar, E. J., & Smith, D. P. (2020). Uncovering protein structure. *Essays in Biochemistry*, 64(4), 649–680. <https://doi.org/10.1042/EBC20190042>
- Sulimenko, V., Dráberová, E., & Dráber, P. (2022). γ -Tubulin in microtubule nucleation and beyond. *Frontiers in Cell and Developmental Biology*, 10. <https://doi.org/10.3389/fcell.2022.880761>

Takatani, S., Ozawa, S., Yagi, N., Hotta, T., Hashimoto, T., Takahashi, Y., Takahashi, T., & Motose, H. (2017). Directional cell expansion requires NIMA-related kinase 6 (NEK6)-mediated cortical microtubule destabilization. *Scientific Reports*, 7(1), 7826. <https://doi.org/10.1038/s41598-017-08453-5>

Thayanithy, V., Park, C., Sarver, A. L., Kartha, R. v., Korpela, D. M., Graef, A. J., Steer, C. J., Modiano, J. F., & Subramanian, S. (2012). Combinatorial Treatment of DNA and Chromatin-Modifying Drugs Cause Cell Death in Human and Canine Osteosarcoma Cell Lines. *PLoS ONE*, 7(9), e43720. <https://doi.org/10.1371/journal.pone.0043720>

van de Kooij, B., Creixell, P., van Vlimmeren, A., Joughin, B. A., Miller, C. J., Haider, N., Simpson, C. D., Linding, R., Stambolic, V., Turk, B. E., & Yaffe, M. B. (2019). Comprehensive substrate specificity profiling of the human Nek kinome reveals unexpected signaling outputs. *ELife*, 8. <https://doi.org/10.7554/eLife.44635>

Vaz Meirelles, G., Ferreira Lanza, D. C., da Silva, J. C., Santana Bernachi, J., Paes Leme, A. F., & Kobarg, J. (2010). Characterization of hNek6 Interactome Reveals an Important Role for Its Short N-Terminal Domain and Colocalization with Proteins at the Centrosome. *Journal of Proteome Research*, 9(12), 6298–6316. <https://doi.org/10.1021/pr100562w>

Wang, N., Zhu, M., Wang, X., Tan, H.-Y., Tsao, S., & Feng, Y. (2014). Berberine-induced tumor suppressor p53 up-regulation gets involved in the regulatory network of MIR-23a in hepatocellular carcinoma. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1839(9), 849–857. <https://doi.org/10.1016/j.bbagr.2014.05.027>

Wlodawer, A., Minor, W., Dauter, Z., & Jaskolski, M. (2008). Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS Journal*, 275(1), 1–21. <https://doi.org/10.1111/j.1742-4658.2007.06178.x>

Wlodawer, A., Minor, W., Dauter, Z., & Jaskolski, M. (2013). Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *The FEBS Journal*, 280(22), 5705–5736. <https://doi.org/10.1111/febs.12495>

Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1), 18–29. <https://doi.org/10.1038/nrm3920>

Xu, Z., Shen, W., Pan, A., Sun, F., Zhang, J., Gao, P., & Li, L. (2020). Decreased Nek9 expression correlates with aggressive behaviour and predicts unfavourable prognosis in breast cancer. *Pathology*, 52(3), 329–335. <https://doi.org/10.1016/j.pathol.2019.11.008>

Yin, M.-J., Shao, L., Voehringer, D., Smeal, T., & Jallal, B. (2003). The Serine/Threonine Kinase Nek6 Is Required for Cell Cycle Progression through Mitosis. *Journal of Biological Chemistry*, 278(52), 52454–52460. <https://doi.org/10.1074/jbc.M308080200>

Zuo, J., Ma, H., Cai, H., Wu, Y., Jiang, W., & Yu, L. (2015). An inhibitory role of NEK6 in TGF β /Smad signaling pathway. *BMB Reports*, 48(8), 473–478. <https://doi.org/10.5483/BMBRep.2015.48.8.225>

Zuo, J., Ma, H., Cai, H., Wu, Y., Jiang, W., & Yu, L. (2015). An inhibitory role of NEK6 in TGF β /Smad signaling pathway. *BMB Reports*, 48(8), 473–478. <https://doi.org/10.5483/BMBRep.2015.48.8.225>

Anexo



INFORMAÇÃO

INFORMAMOS que o Projeto CIBIO/FCF UNICAMP (CQB 7436/2020) nº 01/2020: **ESTUDO FUNCIONAL E MOLECULAR DE NOVOS CANDIDATOS DE ONCO-PROTEÍNAS E TUMOR SUPRESSORES VISANDO O DESENVOLVIMENTO DE NOVOS INIBIDORES ANTI-CÂNCER**, cujo pesquisador responsável é o Prof. Dr. Jorg Kobarg, sub-projeto: “Caracterização funcional e molecular da Nek6 humana selvagem e contendo mutações encontradas em câncer no contexto da resposta à danos de DNA”, do Pós Graduando Pedro Rafael Firmino Dias, RA 180634, encontra-se devidamente aprovado e regularizado junto a CIBio FCF UNICAMP e CTNBio, conforme legislação vigente.

O professor responsável se compromete em incluir o nome do aluno no próximo relatório da CIBIO FCF UNICAMP.

Cidade Universitária “Zeferino Vaz”, 03 de julho de 2023.

Prof. Dr. Marcelo Lancellotti
Presidente da CIBio

Faculdade de Ciências Farmacêuticas - Secretaria de Pesquisa
Rua Cândido Portinari, nº 200 – Cidade Universitária “Zeferino Vaz”
Telefone: (19) 3521-8126

Documento assinado. Verificar autenticidade em sigad.unicamp.br/verifica
e-mail: pesquisa@fcf.unicamp.br
Informar código 7D952350 E2984D66 BD8C0048 436DCDD

Documento assinado eletronicamente por **Marcelo Lancellotti**, PRESIDENTE CIBIO FCF, em 04/07/2023, às 13:55 horas, conforme Art. 10 § 2º da MP 2.200/2001 e Art. 1º da Resolução GR 54/2017.



A autenticidade do documento pode ser conferida no site:
sigad.unicamp.br/verifica, informando o código verificador:
7D952350 E2984D66 BD8C0048 436DCDD

