**Estudo da relação entre as características intrínsecas das proteínas e a sua propensão à cristalização**

Afonso Sá 1 and Filipa Castro 2

1 School of Engineering, Minho University, Campus de Azurém, 4800-019, Guimarães, Portugal

2 Centre of Biological Engineering, Minho University, Campus de Gualtar, 4710 - 057, Braga, Portugal, <https://www.ceb.uminho.pt/>

**Sumário.** A cristalização de anticorpos é crucial em diversas aplicações, desde a biologia estrutural à separação/purificação de biofármacos para o tratamento de várias doenças como o cancro. Este projeto tem como objetivo entender melhor como as propriedades (hidrofilicidade, carga, cadeia de aminoácidos...) dos anticorpos afetam a sua capacidade em cristalizar através do uso de ferramentas informáticas.

**Palavras-chave:** Crystallization · Antibodies · Intermolecular Forces · Solubility· Folding · Stability

**1. Introdução**

**1.1. Cristalização de proteínas**

A cristalização é um processo físico-químico que ocorre quando uma substância passa do estado líquido ou gasoso para o estado sólido, geralmente pela perda de calor ou solvente, no qual as moléculas de uma substância se organizam de forma ordenada e tridimensional para formar uma estrutura cristalina [1]. Essa organização está dependente de múltiplos fatores como a temperatura, pressão, agitação entre outros fatores [2]. Uma das características distintivas da cristalização é a formação de cristais com padrões de repetição definidos, resultando em estruturas sólidas com propriedades físicas e químicas específicas. Essas propriedades cristalinas são essenciais para uma ampla variedade de aplicações em diversos campos.

Na área da indústria a cristalização é fundamental para a síntese de materiais nano estruturados e semicondutores, que são utilizados em diversos dispositivos eletrônicos e na indústria energética [3]. A cristalização também desempenha um papel crucial na geociência, onde a formação de cristais naturais em minerais fornece informações sobre a evolução da Terra e de processos geológicos [4]. Na indústria farmacêutica, é utilizada para produzir medicamentos na forma de cristais sólidos, garantindo sua estabilidade e pureza [5]. A cristalização de proteínas é uma técnica

fundamental na biologia estrutural, permitindo a determinação da estrutura tridimensional das mesmas, e assim compreender melhor a sua função uma vez que a estrutura e função de uma dada proteína são intimamente relacionadas. Isso é fundamental para o desenvolvimento de medicamentos, bem como o desenvolvimento de terapias mais eficazes. A cristalização de proteínas pode também ser usada como técnica de separação e de purificação onde a formação de cristais permite a remoção de impurezas, resultando uma proteína altamente purificada [6].

A diagram of a dna structure

Description automatically generated

Fig.1. Estrutura geral das proteínas [7].

As proteínas são macromoléculas biológicas constituídas por aminoácidos que se unem formando 4 estruturas diferentes (Fig.1). A estrutura primária refere-se à sequência linear dos aminoácidos na cadeia polipeptídica, a secundária é a organização da cadeia polipeptídica em padrões como hélice alfa através da interação entre aminoácidos vizinhos, a terciária é a disposição tridimensional da proteína, resultante das interações entre os aminoácidos distantes na cadeia polipeptídica, e a quaternária é a junção de cadeias polipeptídicas formando uma estrutura funcional [8]. A capacidade de cristalização de uma proteína depende da sua estrutura (Fig.1) que influencia diretamente a capacidade de formar cristais ordenados e bem definidos. Proteínas com estruturas folded, estáveis e compactas tendem a cristalizar mais facilmente [9]. A capacidade em cristalizar também depende de fatores físicos, como a sua solubilidade que evita a precipitação das proteínas durante a formação dos cristais, a sua estabilidade que garante que permaneçam estruturalmente intactas durante o processo, e o folding adequado das proteínas que evita a agregação e promove a formação de cristais de alta qualidade [10].

No processo de cristalização de proteínas é necessário a formação de ligações cristalinas criadas através de interações intermoleculares que ocorrem principalmente entre os aminoácidos que se encontram à superfície da proteína [11] , estas interações podem ser feitas através de:

1. Pontes de hidrogênio, que podem ocorrer entre os grupos funcionais das cadeias polipeptídicas de proteínas como os grupos -OH de resíduos de aminoácidos como serina e treonina. Essas pontes de hidrogênio ajudam a organizar as moléculas numa estrutura cristalina ordenada.

2. Forças eletrostáticas, que podem surgir de interações entre resíduos carregados, contribuindo para a estabilidade da estrutura cristalina.

3. Forças de Van der Waals que ajudam a manter as moléculas unidas numa estrutura cristalina estável.

4. Resíduos hidrofóbicos que podem se agrupar no núcleo da estrutura cristalina, afastando-se do solvente aquoso circundante. Essas interações hidrofóbicas contribuem para a estabilidade da estrutura cristalina.

**1.2. Anticorpos**

Neste projeto irá ser abordado a cristalização de anticorpos. Os anticorpos são produzidos por células plasmáticas dentro do corpo humano que têm como função mediar uma resposta imunológica adaptativa contra patógenos invasores. Existem cinco classes de anticorpos predominantes, IgG, IgM, IgA, IgD, e IgE, cada um especializado em executar funções específicas (Fig.2 B). Os anticorpos adquirem a capacidade de identificar uma variedade diversa de antigénios por meio da recombinação genética de diferentes elementos de sua estrutura. Os anticorpos têm um elevado número de aplicações clínicas, sendo as mais importantes o seu uso no combate a condições autoimunes e cancros, conferindo imunidade passiva [12].

A diagram of different types of igg

Description automatically generatedA diagram of a blue and orange line

Description automatically generated

Fig.2. A. Classes dos diferentes tipos de anticorpos [13], B. Estrutura comum de um anticorpo [14]

Os diferentes tipos de anticorpos apresentam características comuns entre si. Todos os anticorpos possuem uma região constante e variável composta por quatro cadeias polipeptídicas - duas cadeias pesadas e duas cadeias leves, estes possuem regiões de fragmento de ligação ao antigénio que realizam diversas funções imunes, como ligação ao complemento, ligação a macrófagos e determinação dos isótopos do anticorpo (Fig.2 B).

Outras características em comum que os anticorpos podem apresentar incluem as cargas líquidas serem maioritariamente negativas devido à presença de grupos carboxilo nas suas cadeias laterais de aminoácidos [15]. Além disso, os anticorpos contêm regiões hidrofílicas geralmente nas superfícies externas da molécula, permitindo a interação eficiente com solventes aquosos [16]. No entanto, modificações pós-traducionais, como glicosilação, podem afetar negativamente a solubilidade e estabilidade dos anticorpos, devido ao aumento das regiões hidrofílicas da molécula, influenciando assim a capacidade de formar cristais [17].

**1.3. Objetivo**

Este projeto tem como objetivo relacionar as características (hidrofilicidade, carga, aminoácidos...) dos anticorpos com a sua propensão à cristalização. Para isso serão analisados artigos científicos provenientes do Pubmed e do Ncbi, e dados provenientes da base de dados “Protein data Bank [18]” que relatam a cristalização bem-sucedida de anticorpos recorrendo a diversas ferramentas de bioinformática.

**2. Materiais e métodos**

**2.1. Ferramentas computacionais**

Mendeley

O Mendeley é um software de criação de referências que tem como principal função criar e organizar automaticamente referências bibliográficas de diferentes bases de dados e artigos científicos. Este software também pode ser usado para criar bibliotecas personalizadas e categorizar as suas bibliografias em diferentes tópicos e ainda permite aos seus usuários aceder e fazer anotações nos seus documentos.

Python

O Python é uma linguagem de programação versátil e de alta capacidade utilizada em uma variedade de campos devido à sua facilidade de aprendizagem, legibilidade, vasta biblioteca padrão e tem um constante suporte da comunidade. Esta linguagem é normalmente utilizada no desenvolvimento de softwares, na análise de dados e na automação de diferentes tarefas.

Rstudio

O Rstudio é um ambiente de desenvolvimento integrado (IDE) para a linguagem de programação R utilizado para análise estatística, visualização de dados e desenvolvimento de modelos preditivos. O Rstudio fornece uma interface simplificada que facilita a análise exploratória de dados e a criação de relatórios interativos.

**2.2. Análise de dados**

Python

Recolha de informação

Mendeley

Criação de referências

Rstudio

Análise de dados

Fig.3. Metodologia de trabalho utilizada no projeto.

Durante o projeto vão ser utilizadas diferentes ferramentas computacionais a fim de otimizar o processo de obtenção e de análise de dados (Fig.3.).

Foi utilizado o software Mendeley para a organização e criação das referências bibliográficas. Através do Mendeley foram importados artigos científicos relevantes sobre o tema garantindo que todas as fontes utilizadas fossem devidamente documentadas.

Para retirar informações de artigos científicos relevantes sobre a cristalização de anticorpos vai ser utilizada a linguagem de programação Python. Utilizando bibliotecas específicas vão ser desenvolvidos scripts capazes de pesquisar automaticamente artigos no PubMed, utilizando palavras-chave como: “Crystallization, antibodies, intermolecular forces, solubility, folding, stability” fornecendo uma ampla base de informações para o desenvolvimento do projeto.

Por fim, para análise e visualização de dados obtidos, vai ser utilizada a ferramenta RStudio. Através do uso desta ferramenta vai ser possível criar gráficos que terão como função comparar as características dos anticorpos que influenciam à sua cristalização.

Dessa forma, ao integrar estas ferramentas computacionais no processo de pesquisa, poderei otimizar cada etapa do trabalho, desde a recolha de referências e informação até à análise e visualização de dados. Essas ferramentas vão desempenhar um papel fundamental na condução deste estudo e na obtenção de resultados precisos e significativos sobre arelação entre as características intrínsecas dos anticorpos e a sua propensão à cristalização.

**3. Referências bibliográficas**

1. McPherson A, Gavira JA. Introduction to protein crystallization. Acta Crystallogr F Struct Biol Commun. 2014 Jan;70(Pt 1):2-20. doi: 10.1107/S2053230X13033141. Epub 2013 Dec 24. PMID: 24419610; PMCID: PMC3943105.

2. C. N. Nanev, “Advancements (and challenges) in the study of protein crystal nucleation and growth; thermodynamic and kinetic explanations and comparison with small-molecule crystallization,” Prog. Cryst. Growth Charact. Mater., vol. 66, no. 2, 2020, doi: 10.1016/j.pcrysgrow.2020.100484.

3. Bučar DK, Lancaster RW, Bernstein J. Disappearing polymorphs revisited. Cryst Growth Des. 2015;15(7):3100-3106.

4. Hazen RM, Morrison SM, Downs RT. Mineral evolution: mineralogy in the fourth dimension. Elements. 2013;9(1):9-12

5. Aitipamula S, Chow PS, Tan RB. Polymorphs, salts, and cocrystals: What's in a name? Cryst Growth Des. 2012;12(5):2147-2152

6. Dos Santos, R., Carvalho, A. L., & Roque, A. C. A. (2017). Renaissance of protein crystallization and precipitation in biopharmaceuticals purification. In *Biotechnology Advances* (Vol. 35, Issue 1, pp. 41–50). Elsevier Inc. https://doi.org/10.1016/j.biotechadv.2016.11.005

7. CHAMPE, HARVEY & FERRIER. Bioquímica Ilustrada, 3° edição, Porto Alegre, Editora Artmed, 2006

8. Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. The Shape and Structure of Proteins. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26830/

9. Drenth, J. (2007). Principles of Protein X-Ray Crystallography. Springer.

10. Price II, W., Chen, Y., Handelman, S. et al. Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. Nat Biotechnol 27, 51–57 (2009). https://doi.org/10.1038/nbt.1514

11. J. Ferreira and F. Castro, “Advances in protein solubility and thermodynamics: quantification, instrumentation, and perspectives,” CrystEngComm, vol. 25, no. 46, pp. 6388–6404, 2023, doi: 10.1039/d3ce00757j.

12. Aziz M, Iheanacho F, Hashmi MF. Physiology, Antibody. [Updated 2023 May 1]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK546670/>

13. <https://bioxcell.com/educational-articles/antibody-structure/>

14. <https://lifesciences.danaher.com/us/en/library/antibodies.html>

15. Janeway CA Jr, Travers P, Walport M, et al. Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science; 2001. The interaction of the antibody molecule with specific antigen. Available from: https://www.ncbi.nlm.nih.gov/books/NBK27160/

16. Abdeldaim DT, Schindowski K. Fc-Engineered Therapeutic Antibodies: Recent Advances and Future Directions. Pharmaceutics. 2023 Sep 28;15(10):2402. doi: 10.3390/pharmaceutics15102402. PMID: 37896162; PMCID: PMC10610324.

17. Gupta SK, Shukla P. Glycosylation control technologies for recombinant therapeutic proteins. Appl Microbiol Biotechnol. 2018 Dec;102(24):10457-10468. doi: 10.1007/s00253-018-9430-6. Epub 2018 Oct 17. PMID: 30334089.

18. PROTEIN DATA BANK - <https://www.rcsb.org>