

A determinação de estruturas tridimensionais de enzimas pela cristalografia de raios-x

Mirian Cristina Feiten¹

Resumo

A cristalização de proteínas é um fenômeno em que as moléculas de uma certa proteína se associam de forma ordenada, em solução aquosa, para formar um arranjo periódico e em três dimensões, ou seja, um cristal. Nesta breve revisão, a aplicação da técnica de cristalografia de difração de raios-X é abordada com foco nos estudos estruturais de proteínas e, mais especificamente, de enzimas. São abordados detalhes das principais etapas envolvidas: purificação e obtenção dos cristais, coleta e processamento de dados, determinação e refinamento da estrutura. Em suma, esta é uma técnica de ponta que envolve conhecimentos multidisciplinares e com potencial para determinar de maneira bastante satisfatória a estrutura tridimensional de macromoléculas biológicas de interesse, não apenas nas áreas médica e farmacológica, bioquímica e biologia molecular, como também na biotecnologia, engenharia química e engenharia de alimentos.

Palavras-chave: Cristalografia; Enzima; Conformação.

Abstract

Protein crystallization is a phenomenon in which the molecules of a certain protein orderly associate, in aqueous solution, to form a periodic, three-dimensional arrangement, that is, a crystal. In this brief review, the application of the X-ray crystallography technique is approached focused on structural studies of proteins and, more specifically, enzymes. Details on the main steps involved are discussed: purification and obtaining of crystals, data collection and processing, model building and refinement. In short, this is a cutting-edge technique that involves multidisciplinary knowledge and presents a great potential to quite satisfactorily determine the three-dimensional structure of biological macromolecules of interest, not only in the medical and pharmacological, biochemistry and molecular biology areas, but also in biotechnology, chemical engineering and food engineering.

Keywords: Crystallography; Enzyme; Conformation.

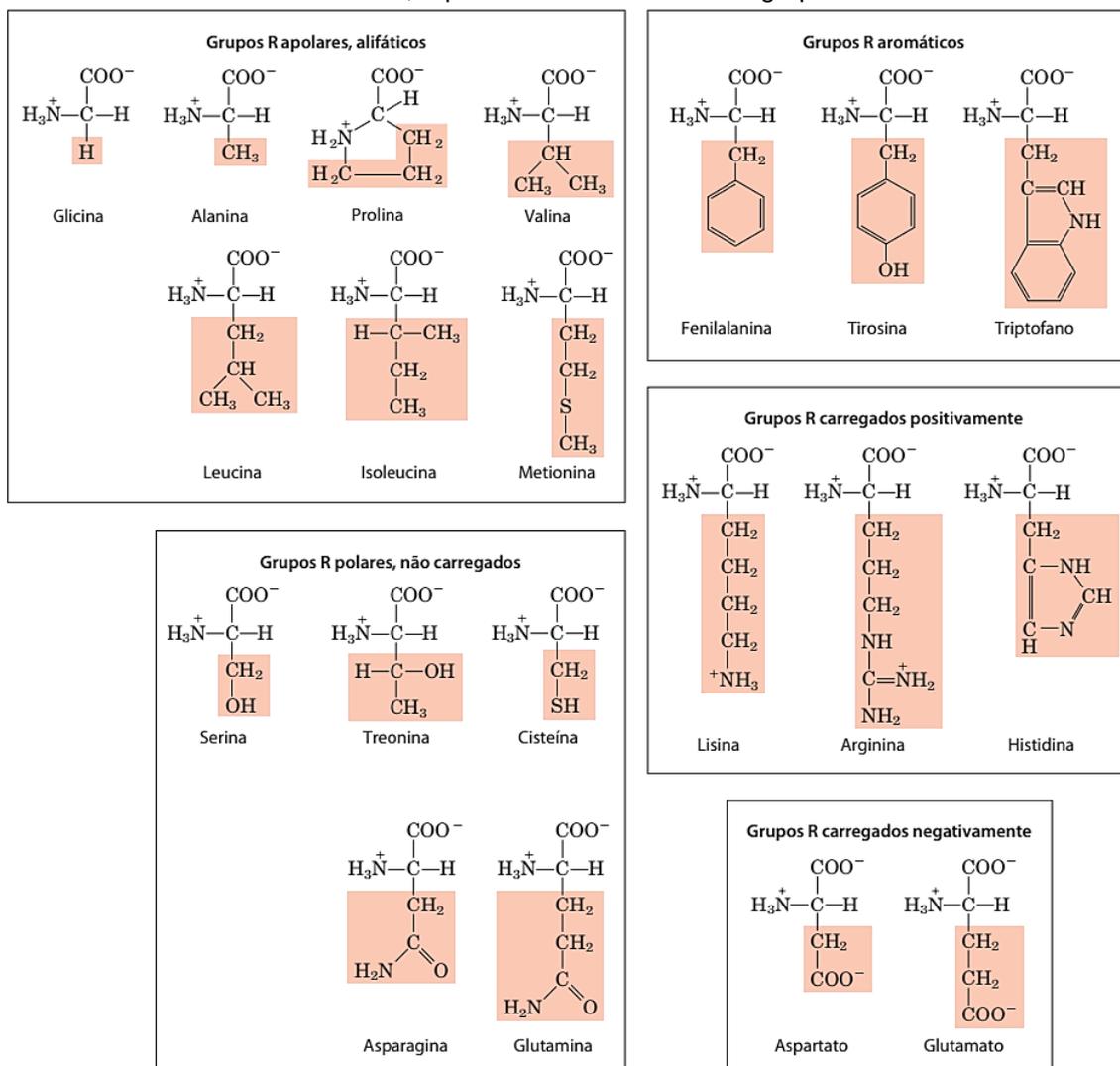
1. Introdução

Proteínas são as macromoléculas biológicas mais abundantes, ocorrendo em todas as células e em todas as partes das células. São biopolímeros

¹ Departamento de Tecnologia – UEM. E-mail: mirianfeiten.mf@gmail.com

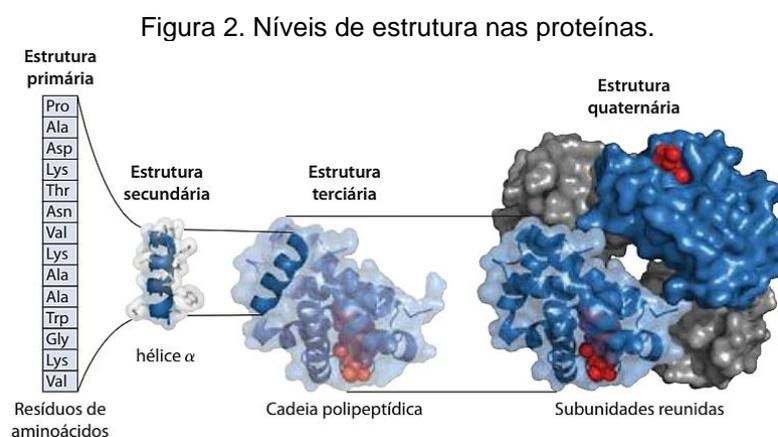
compostos por unidades monoméricas - resíduos de aminoácidos - ligados entre si por ligações peptídicas. Este tipo de ligação estabelece-se entre os resíduos de aminoácidos com consequente perda de uma molécula de água. Cada alfa-aminoácido é constituído por um grupo amino, um grupo carboxílico, um átomo de hidrogênio e um grupo R (cadeia lateral que difere de aminoácido para aminoácido), ligados a um carbono (carbono alfa). São apenas vinte os aminoácidos (Figura 1) que, combinados em número e sequência diferentes, dão origem a todas as proteínas existentes, conferindo-lhes a estrutura e capacidade funcional específica (BORZANI et al., 2001; COLNAGO; ALMEIDA; VALENTE, 2002).

Figura 1. Os 20 aminoácidos. As fórmulas estruturais mostram o estado de ionização que predomina em pH 7,0. As porções não sombreadas são aquelas comuns a todos os aminoácidos; aquelas sombreadas são os grupos R.



Fonte: NELSON; COX (2013).

As proteínas apresentam quatro aspectos distintos em suas estruturas: estrutura primária, secundária, terciária e quaternária (VOET; VOET, 2004; NELSON; COX, 2017). A Figura 2 ilustra estes aspectos estruturais. A estrutura primária é caracterizada pela sequência linear de aminoácidos e ligações peptídicas da molécula; é o nível estrutural mais simples e mais importante já que dele deriva todo o arranjo espacial da molécula. A estrutura secundária é caracterizada pelo estabelecimento de ligações de hidrogênio entre aminoácidos de uma cadeia proteica; os elementos deste tipo de estrutura mais comuns em proteínas são a α -hélice e a folha- β ou pregueada. A estrutura terciária é a organização “global” de uma única cadeia polipeptídica; é o arranjo tridimensional que a proteína assume para realizar suas funções. O principal fator que determina a estrutura terciária numa proteína globular é o efeito hidrofóbico, ou seja, as cadeias laterais dos aminoácidos não polares ficam como que “escondidas” dentro da estrutura e as cadeias laterais dos resíduos polares ficam expostos à superfície. As ligações de hidrogênio são fundamentais na estabilização da estrutura terciária. O termo estrutura terciária é frequentemente usado para se referir à proteína enovelada. Por fim, a estrutura quaternária é caracterizada pela associação de duas ou mais cadeias polipeptídicas numa estrutura com várias subunidades, resultando numa unidade ativa. Este tipo de estrutura é estabilizada por ligações de hidrogênio, forças de Van der Waals ou ligações iônicas (BORZANI et al., 2001; COLNAGO; ALMEIDA; VALENTE, 2002).



Fonte: NELSON; COX (2013).

As enzimas são proteínas biologicamente ativas e são responsáveis pela catálise de diversas reações bioquímicas. Em uma proteína enzimática, existe certo domínio chamado de “sítio ativo” que se liga ao substrato – a molécula reagente – e diminui a energia do estado de transição que leva ao produto desejado. A ligação entre o sítio ativo e o substrato é extremamente específica: a molécula precisa ter certas características eletrônicas e espaciais que permitam o seu “encaixe” com a proteína (BORZANI et al., 2001). A conformação e a estabilidade da estrutura molecular das enzimas são asseguradas por ligações de hidrogênio, interações hidrofóbicas, pontes de dissulfeto, ligações iônicas e forças de Van der Waals (NELSON; COX, 2017). O grande paradigma dos estudos estruturais de macromoléculas, principalmente em biologia molecular, é que o conhecimento da estrutura tridimensional é um pré-requisito para o entendimento de como essas moléculas funcionam (VOET; VOET, 2004; FEITEN et al., 2016). A atividade catalítica, bem como a estabilidade e a especificidade da enzima dependem da sua estrutura tridimensional. Condições ambientais, tais como pH, temperatura, força iônica do meio, pressão, cisalhamento, entre outros, afetam a estrutura da enzima e, em decorrência, suas propriedades (SCRIBAN, 1985).

O estudo da função de enzimas envolve uma aproximação multidisciplinar. Várias são as técnicas documentadas na literatura para determinação das estruturas secundária e terciária de enzimas, cada uma com suas particularidades. As técnicas espectroscópicas de Mossbauer, Raman, fluorescência, infravermelho, UV/visível e dicroísmo circular são bastante documentadas na literatura para a determinação da estrutura secundária de proteínas (FEITEN et al., 2016). Por outro lado, a determinação da estrutura tridimensional, usando técnicas espectroscópicas como a difração de raios-X em proteínas cristalizadas ou a ressonância magnética nuclear (RMN) em proteínas em solução, permite localizar a posição dos diferentes resíduos no espaço (NASCIMENTO; BLOCH Jr, 2001; COLNAGO; ALMEIDA; VALENTE, 2002; LAI et al., 2011). Este tipo de estudo revela muitas vezes detalhes sobre o mecanismo da reação, ou seja, a forma e sequência de ligações do substrato ao centro ativo para ser catalisado (MORRIS et al., 1992; MUNTE, 2001;

SILVESTRINI; CIANCI, 2020). Neste contexto, o objetivo deste trabalho é apresentar uma breve revisão bibliográfica sobre os princípios e o atual estado da arte do uso da técnica de cristalografia por difração de raios-X para a determinação da estrutura tridimensional de enzimas.

2. Revisão bibliográfica

2.1. Cristalografia: definição e conceitos

A cristalização é a etapa mais complexa quando queremos determinar a estrutura de uma macromolécula (GIACOVAZZO et al., 1992; RUPP, 2009). A formação de um monocristal de tamanho e qualidade adequados para os experimentos depende de vários parâmetros, tais como o grau de pureza da proteína; o tipo, a composição e o pH do tampão; concentração salina; temperatura do ambiente (RUPP, 2009; FEITEN et al., 2016), dentre outros.

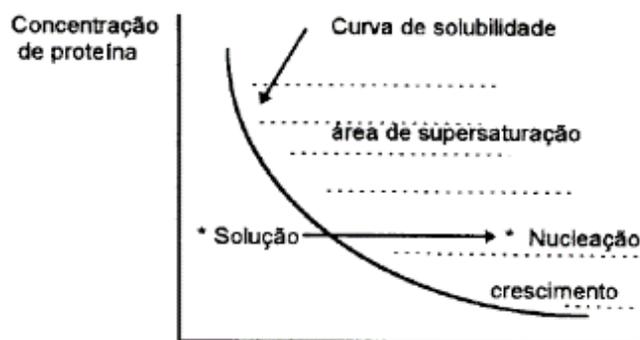
O crescimento de um cristal de proteína começa a partir de uma solução supersaturada da macromolécula, e evolui para um estado termodinamicamente estável, no qual a proteína é particionada entre uma fase sólida e a solução. O tempo necessário para que o equilíbrio seja atingido tem uma grande influência sobre o resultado final, que pode ir de um precipitado amorfo ou microcristalino até grandes monocristais. O estado de supersaturação é determinado pelas condições do meio, como a concentração do agente precipitante, concentração da proteína, temperatura, pH, força iônica, entre outros (GIACOVAZZO et al., 1992).

A cristalização começa com a fase de nucleação (formação dos primeiros agregados ordenados) que é seguida pela fase de crescimento. A nucleação requer um estado de supersaturação maior do que aquele da fase de crescimento. Consideremos uma proteína dissolvida em um tampão. Nessa situação, para que a proteína seja levada a formar cristais, é necessário que as moléculas da proteína dissolvidas na solução, sejam trazidas a uma situação onde as moléculas estejam próximas umas das outras. Para levar a proteína a essa situação podemos aumentar sua concentração, ou aumentar a concentração do sal presente na solução da proteína de modo a trazê-la na região de supersaturação, de forma lenta. Na região de supersaturação teremos

as moléculas da proteína próximas umas das outras, o que, em casos favoráveis, promoverá o aparecimento dos primeiros núcleos cristalinos. Esses microcristais servirão de base para o crescimento de cristais maiores, adequados para experimentos de difração de raios-X (RONDEAU; SCHREUDER, 2015).

O aumento da solubilidade de uma macromolécula biológica a baixa concentração salina ($< 0,5 \text{ M}$) é chamada *salting-in* (Figura 3). Este fenômeno é explicado pelas interações eletrostáticas não específicas entre a macromolécula carregada e as espécies iônicas do sal. Segundo a teoria de Debye-Hückel para soluções iônicas (SABADINI; CARVALHO, 2013), um aumento na força iônica reduz a atividade dos íons em solução e aumenta a solubilidade do composto iônico. Uma forma alternativa de se tratar este fenômeno é considerar o *salting-in* como o resultado da competição entre grupos carregados na superfície da macromolécula e os íons em solução. Na ausência de íons no solvente, a macromolécula precipita devido à atração eletrostática entre cargas opostas em diferentes partes da macromolécula. Se os íons são adicionados à solução eles blindam os grupos carregados na macromolécula e aumentam a sua solubilidade.

Figura 3. Típica curva de solubilidade de uma proteína, como uma função da concentração de sal ou qualquer outro parâmetro.



Fonte: NONATO (1997).

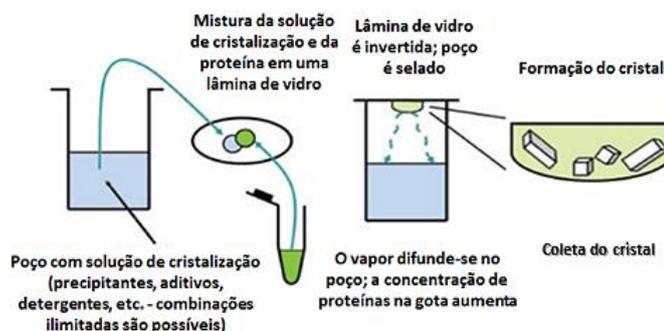
Outra forma de precipitar uma macromolécula é pela adição de sal (*salting-out*). O sal serve para imobilizar as moléculas de água, desta forma aumentando a concentração efetiva da macromolécula, ou seja, diminuindo a sua solubilidade. No fenômeno de *salting-out* a solubilidade da macromolécula é governada principalmente por efeitos hidrofóbicos. Os íons podem provir de um

agente precipitante. Um composto adicionado a um sistema macromolécula biológica+água (soluto + solvente) pode se ligar à macromolécula (ligação preferencial) ou ser excluído (exclusão preferencial). A interação líquida do *salting-out* é a exclusão preferencial (BORZANI et al., 2001).

O método mais comum de cristalização de proteína é a difusão de vapor (RUPP, 2009). Possuindo duas variantes neste método, a “gota suspensa” (*hanging drop*) e a “gota sentada” (*sitting drop*), o processo consiste em uma pequena quantidade de solução (gota) contendo concentrações equivalentes de proteínas e reagentes precipitantes e uma solução reservatório com as mesmas substâncias em concentrações dobradas em um sistema isolado. Devido à diferença de concentração, a água da gota evapora lentamente para a solução reservatório. Assim, pode-se iniciar a cristalização das proteínas na gota.

Ainda segundo Rupp (2009), a configuração mais comum da difusão de vapor para crescer cristais de proteína é a técnica da gota suspensa (*hanging drop*) (Figura 4).

Figura 4. Cristalização de proteína por difusão de vapor – gota suspensa (*hanging drop*).



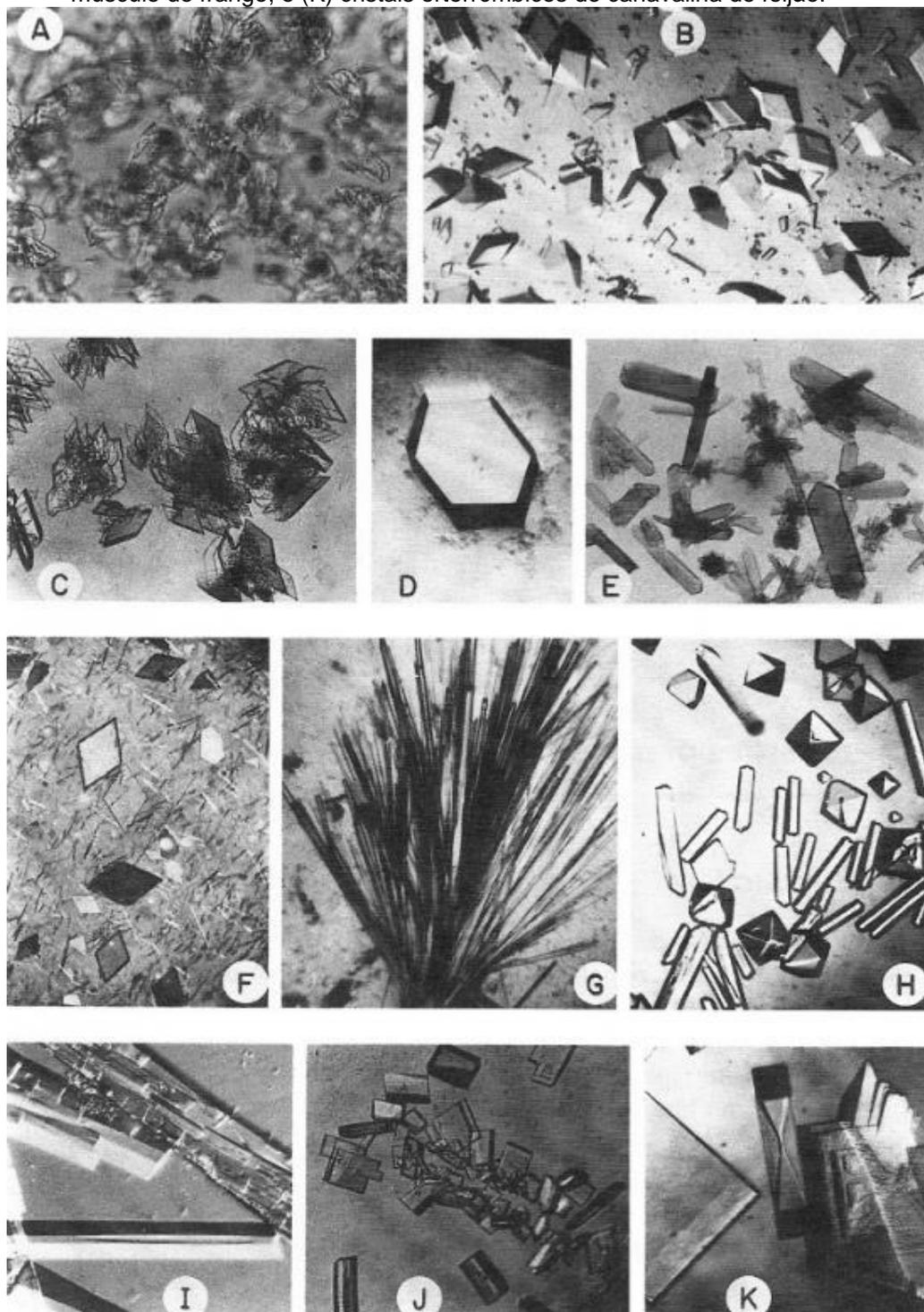
Fonte: Adaptado de RUPP (2009).

Alguns microlitros de solução de proteína são misturados com uma quantidade aproximadamente igual de uma solução reservatório contendo os agentes precipitantes. Uma gota desta mistura é colocada numa lâmina de vidro que cobre o reservatório. Como a mistura proteína/precipitante na gota é menos concentrada do que a solução reservatório (mistura da solução de proteína com a solução reservatório de 1:1 na gota), a água evapora da gota para dentro do reservatório. O equilíbrio prossegue por meio da difusão das espécies voláteis (água e/ou solventes orgânicos) até que a pressão de vapor na gota se iguale

àquela do reservatório. Como resultado, a concentração de ambos, proteína e agente precipitante, aumenta lentamente na gota, e os cristais podem se formar. Para espécies químicas com pressão de vapor maior que a pressão de vapor da água, a troca ocorre do reservatório para a gota.

Os cristais de macromoléculas biológicas são diferentes dos cristais de pequenas moléculas em vários aspectos. Eles diferem em tamanho: cristais de macromoléculas são normalmente bem menores, dificilmente excedem 1 mm^3 em volume. Outra diferença é que cristais de macromoléculas biológicas são bem mais frágeis e possuem um alto conteúdo de solvente. A fragilidade dos cristais de macromoléculas biológicas é uma consequência das fracas interações entre as macromoléculas biológicas dentro da rede cristalina e do alto conteúdo de solvente nestes cristais (de 20% a mais de 75%) (MATTHEWS, 1968). Por esta razão, os cristais de macromoléculas biológicas devem ser mantidos em um ambiente saturado de solvente, caso contrário, a desidratação conduzirá à quebra destas ligações fracas e, conseqüentemente, à destruição dos cristais. O alto conteúdo de solvente, contudo, tem conseqüências úteis, pois os canais de solvente nos cristais de macromoléculas biológicas permitem a difusão de pequenas moléculas, uma propriedade usada na preparação de derivados isomorfos e de complexos binários de proteínas e ligantes. Uma outra característica dos cristais de macromoléculas biológicas são as grandes dimensões de suas celas unitárias, se comparadas com as dimensões das celas unitárias das pequenas moléculas. Cristais de macromoléculas biológicas podem ter parâmetros de cela unitária de até algumas dezenas de milhares de angstroms (DUCRUIX; GIEGÉ, 2000; RHODES, 2006). A Figura 5 apresenta alguns exemplos de cristais de proteínas obtidos por diferentes técnicas e empregando diferentes agentes de precipitação.

Figura 5. Alguns cristais de proteínas crescidos por diferentes técnicas e usando diferentes agentes de precipitação: (A) catalase animal; (B) forma trigonal da frutose-1,6 difosfatase de fígado de galinha; (C) proteína ligante do cortisol de soro de porquinho-da-índia; (D) concanavalina B de feijão; (E) catalase de fígado bovino; (F) proteína desconhecida de abacaxis; (G) forma ortorrômbica do fator de alongamento Tu de *Escherichia coli*; (H) cristais hexagonais e cúbicos de tRNA-fenilalanina de levedura; (I) bastões monoclônicos da proteína desenrolada do gene 5 do DNA de bacteriófago fd; (J) glicerol-3-fofato desidrogenase de músculo de frango; e (K) cristais ortorrômbicos de canavalina de feijão.



Fonte: RHODES (1993).

Por mais de meio século, a determinação de estrutura tridimensional (3D) de proteínas vem sendo feita por métodos de cristalografia de difração de raios-X (BARCELOS et al., 2008). É inquestionável a importância e a contribuição que aplicações dessa técnica têm proporcionado à Bioquímica (NASCIMENTO; BLOCH Jr, 2001; NELSON; COX, 2017). A técnica de cristalografia de raios-X pode ser definida de maneira simples como a submissão de uma partícula cristalizada à uma radiação de ondas de tamanho conhecido. A cristalografia de proteínas é um ramo especializado da cristalografia que investiga, através de técnicas de difração em cristais individuais, a estrutura tridimensional destas macromoléculas biológicas. Conhecendo o comprimento de onda e o ângulo de difração do raio é possível determinar a possível estrutura da partícula - no caso, uma proteína e os átomos que a compõe (GIACOVAZZO et al., 1992; RUPP, 2009).

2.2. Difração de raios-X

Raios-X são um tipo de radiação eletromagnética de alta energia. Eles têm níveis de energia de 200 eV a 1 MeV, que os coloca entre as radiações de raios- γ e ultravioleta (UV) no espectro eletromagnético (SURYANARAYANA; NORTON, 1998).

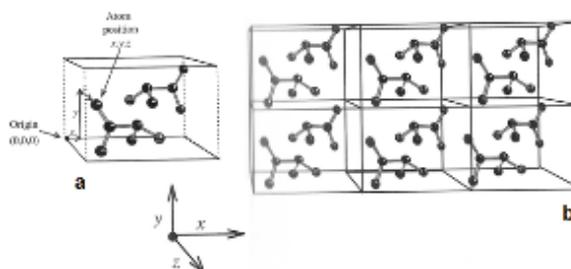
Os raios-X são produzidos quando um feixe de raios catódicos (elétrons) com alta energia é desacelerado pela colisão com um alvo metálico. Elétrons são ejetados por um filamento de tungstênio aquecido e acelerados em direção ao alvo metálico dentro de um tubo a vácuo. A radiação liberada pelo alvo consiste de uma mistura de um espectro contínuo e de diferentes linhas intensas de comprimentos de onda bem definidos que dependem do material do alvo metálico, variando de 0,5 a 3,0 Å, no intervalo de interesse da cristalografia estrutural, visto que tais comprimentos são semelhantes às distâncias interatômicas que são analisadas (STOUT; JENSEN, 1989).

Um cristal pode ser definido como um sólido poliédrico limitado por faces planas que exprimem um arranjo interno ordenado de átomos ou moléculas. No estudo da estrutura interna das substâncias pelas técnicas de raios-X, dá-se

menos ênfase às faces do cristal; considera-se um cristal como um corpo caracterizado por uma extensão de espaço tridimensional mais ou menos rígida de uma unidade característica de estrutura interna (WAHLSTROM, 1969).

Cada cristal possui uma unidade elementar, chamada de cela unitária, descrita por três vetores não coplanares x , y e z . O cristal é formado pela repetição da cela unitária através de translações nas três dimensões. O arranjo dos pontos nos vértices das celas unitárias é chamado de rede. A Figura 6 esquematiza um cristal com suas celas unitárias. A cela unitária é o elemento menor e mais simples de volume, e é totalmente representativa do cristal inteiro. Se soubermos exatamente o conteúdo de uma cela unitária, podemos imaginar o cristal todo como um arranjo empacotado de muitas celas unitárias empilhadas ao lado e em cima umas das outras (RHODES, 2006).

Figura 6. (a) Detalhe da cela unitária e (b) arranjo de seis celas unitárias na rede cristalina. Cada cela unitária contém duas moléculas de alanina em diferentes orientações.



Fonte: RHODES (1993).

Quando os raios-X incidem sobre um cristal, estes interagem com os elétrons dos átomos que o compõe. Os raios-X, como qualquer onda eletromagnética, provocam uma força que é proporcional ao campo elétrico da radiação. Essa força acelera os elétrons que por sua vez espalham a radiação em todas as direções (GIACOVAZZO et al., 1992; NONATO, 1997; RUPP, 2009).

O evento mais básico associado com a dispersão de raios-X à partir de uma molécula é a absorção inicial da radiação eletromagnética por átomos individuais, que conduz à excitação dos elétrons das camadas internas. O decaimento destes elétrons de volta ao estado fundamental é acompanhado pela dispersão de raios-X a partir de cada átomo individualmente (atuando como

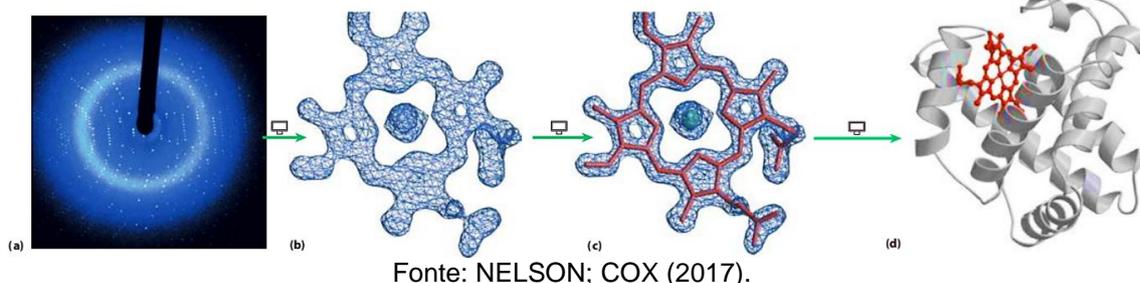
fontes pontuais) (RUPP, 2009), ambas a diferentes comprimentos de onda (espalhamento incoerente), e no mesmo comprimento de onda (espalhamento coerente). É o último efeito que é posto em prática para estudos estruturais (STODDARD, 1996).

Difração é a mudança na direção de propagação de um feixe de radiação incidente decorrente da presença de obstáculo no caminho óptico. Esta mudança de direção depende das dimensões físicas do obstáculo, do comprimento de onda da radiação incidente e do ângulo de incidência (SURYANARAYANA; NORTON, 1998).

As propriedades da difração e da interferência que decorrem do espalhamento dos raios-X por monocristais são utilizadas para se obter informações sobre a estrutura dos sólidos visto que a simetria de um padrão de difração está relacionada à simetria da estrutura que o produziu, permitindo assim, determinar a forma do objeto que gerou tal padrão e, a partir das intensidades das ondas difratadas obter a natureza do objeto (STOUT; JENSEN, 1989).

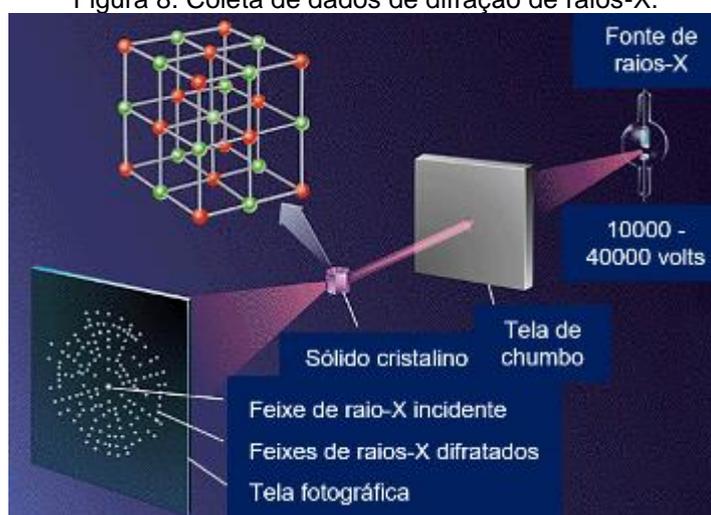
Assim, a estrutura cristalográfica de uma macromolécula pode ser obtida através da análise do padrão de difração produzida por seus cristais. Essa análise é constituída basicamente por quatro passos (Figura 7): purificação e cristalização (1); coleta e processamento de dados (2); determinação da estrutura através da densidade eletrônica (3) e refinamento (4) (NONATO, 1997; RUPP, 2009).

Figura 7. Etapas da determinação da estrutura da mioglobina de cachalote via cristalografia por raios-X. (a) O padrão da difração por raios-X é gerado a partir de um cristal da proteína. (b) Os dados obtidos pelo padrão de difração são utilizados para calcular um mapa de densidade eletrônica tridimensional. Somente a densidade eletrônica do grupo heme, uma pequena parte da estrutura, é mostrada aqui. (c) A estrutura do heme é modelada em seu mapa de densidade eletrônica. (d) A estrutura completa da mioglobina de cachalote, incluindo o heme (Protein Data Bank (PDB) ID: 2MBW).



Uma fonte de raios-X emitirá radiação que atravessará a amostra de proteínas cristalizadas e sofrerá difração. Um anteparo posicionado em volta da amostra de cristais captará os raios desviados de sua trajetória (Figura 8). As emissões características de cada elemento têm uma energia associada. É possível identificar os elementos pela separação dos canais de energia. Então, medindo o ângulo de difração dos raios e conhecendo o comprimento de onda, pode-se esboçar o posicionamento dos átomos que compõem a cadeia de aminoácidos (RUPP, 2009).

Figura 8. Coleta de dados de difração de raios-X.



Fonte: Adaptado de: <https://phys4030.blogspot.com/2011/03/crystals-structure-by-x-ray-diffraction.html>

Cristais de macromoléculas são bastante sensíveis aos raios-X. Muitas vezes é necessário utilizar vários cristais para se obter um conjunto de dados necessário para prosseguir com os próximos passos. A coleta deve ser feita com o cristal resfriado instantaneamente num fluxo de nitrogênio gasoso a aproximadamente 100 K (criocristalografia) (NONATO, 1997; RUPP, 2009). Isto permite a redução dos efeitos danosos da radiação direta e aumenta o tempo de vida do cristal. Dessa forma, tem-se um melhor aproveitamento do material, pois apenas um único cristal pode fornecer dados suficientes para a resolução da estrutura e não é preciso realizar a integração com dados de cristais diferentes. Assim, cada conjunto de imagens geradas de um cristal forma um conjunto de dados a ser analisado. É possível realizar várias coletas e escolher o melhor conjunto de dados. Além disso, os dados coletados a frio apresentam melhor

qualidade, pois o resfriamento mantém o nível de organização interna do cristal, necessária para uma boa coleta de dados, preservação da rede cristalina, informação expressa nos resultados pela mosaicidade. A distância do detector ao cristal deve ser regulada de forma a chegar o mais próximo possível da melhor resolução, limitada pela qualidade do cristal.

De posse das primeiras imagens é possível realizar a indexação e determinar o grupo espacial, as dimensões e orientação da cela unitária, e estimar a mosaicidade. A partir da medida da distância e dos eixos da rede recíproca do padrão de difração é possível determinar o tamanho e os eixos do cristal no espaço real. Esse passo é necessário para planejar uma estratégia de coleta de dados eficiente, minimizando o tempo e maximizando os resultados (BARCELOS et al., 2008; KABSCH, 2010).

O método de oscilação é utilizado para a coleta, sendo o cristal rotacionado a cada 1° com o objetivo de coletar o máximo de reflexões possíveis, visando propiciar uma maior coleta de dados (FEITEN et al., 2016). A completeza expressa nos dados após a indexação e integração dos dados, representa a porcentagem do número de reflexões observadas em relação ao número de reflexões possíveis para uma determinada resolução, que está limitada pela qualidade do cristal. Uma completeza de 100% é o resultado desejável, com a maioria das reflexões medidas várias vezes (WOOLFSON, 1970; RONDEAU; SCHREUDER, 2015).

No processo de integração é quantificada a intensidade do ponto, retirando-se a contribuição proveniente da difração do solvente. Cada imagem em um mesmo conjunto de dados tem intensidade gravada em uma escala diferente umas das outras por diversos motivos, como a variação na radiação incidente, absorção dos raios difratados ou destruição do cristal. Portanto, é necessário colocar todos os pontos em uma mesma escala. Após esse passo pode ser realizada a fusão dos dados. Os pontos que foram agrupados como reflexões equivalentes simétricas de diferentes imagens terão suas intensidades somadas. Após esse processo, são obtidos parâmetros para determinar a qualidade do cristal, que é representada pelos parâmetros da rede cristalina (ESTEVES, 2010; FEITEN et al., 2016).

2.3. Determinação da estrutura tridimensional

Num experimento de difração de raios-X bem sucedido, obtemos do cristal informações sobre os parâmetros da cela unitária do cristal (parâmetros de rede e simetria) e sobre o módulo do fator de estrutura, normalmente chamado de fator de estrutura observado ($F_{\text{obs}}(\text{hkl})$), para cada reflexão de índices hkl - onde hkl são os índices de Miller para cada reflexão ocorrendo em uma dada direção. Para resolvermos a estrutura é necessário, além da informação do módulo F_{obs} , informações sobre a fase do fator de estrutura, que é perdida no experimento de difração de raios-X (BARCELOS et al., 2008; RUPP, 2009).

O principal problema na resolução das estruturas é que os dados disponíveis para a análise são somente as intensidades do espectro da difração de raios-X pelo cristal e, destes valores, obtemos apenas os módulos dos fatores de estrutura (DUCRUIX; GIEGÉ, 2000). Isso é um fator limitante para a resolução de estruturas moleculares porque impossibilita o cálculo da função de densidade eletrônica por Séries de Fourier com a ausência das fases, que são perdidas no processo de medida. Esse problema é conhecido como o “problema das fases” na cristalografia (RUPP, 2009). Um dos métodos mais utilizados para se resolver esse problema é o chamado Método Direto que, devido a sua robustez, em termos computacionais, e eficácia, é o método correntemente utilizado para pequenas moléculas (CASTRO, 2012). Como proteínas são macromoléculas, este método dificilmente pode ser aplicado.

Existem basicamente três métodos utilizados para a resolução de estruturas ou para solucionar o “problema das fases” no caso de proteínas (STOUT; JENSEN, 1989; STODDARD, 1996; NONATO, 1997; SURYANARAYANA; NORTON, 1998; BARCELOS et al., 2008; RUPP, 2009; FEITEN et al., 2016):

Método da substituição molecular (MR): Método baseado no princípio que macromoléculas apresentam alta homologia sequencial, tendendo a se enovelar de forma bastante similar, e desta forma, uma proteína cuja estrutura já tenha sido determinada serve como um modelo inicial para a obtenção de um conjunto preliminar de fases que pode ser subsequentemente refinado. O sucesso deste

método está no fato de que hoje em dia existe um grande número de macromoléculas que já tem sua estrutura conhecida, depositadas no banco de dados Protein Data Bank (PDB, 2022). O PDB foi criado em 1971 como um arquivo digital de macromoléculas biológicas (DUTTA; BERMAN, 2005), e possui atualmente em torno de 186 mil estruturas depositadas (proteínas, ácidos nucleicos e vírus).

Método da substituição isomórfica (MIR): Depende da obtenção de cristais de derivados isomorfos (com átomos pesados que tenham maior contribuição para o espalhamento). Um derivado isomorfo perfeito é aquele em que há uma única variação no mapa de densidade eletrônica e isso se deve à posição dos átomos pesados. A introdução desses átomos deve ser feita de modo a não afetar a estrutura da macromolécula original.

Método de dispersão anômala a múltiplos comprimentos de onda (MAD): Quando os elétrons se encontram em camadas mais próximas do núcleo, o tratamento de dados devido ao espalhamento desses elétrons deve ser cuidadoso, uma vez que a força de interação entre esses elétrons e o núcleo não pode mais ser desprezada e o elétron não pode mais ser tratado como livre. Devido à força de interação com o núcleo, existe a introdução de um termo imaginário ao fator de espalhamento atômico que depende do comprimento de onda da radiação.

O conhecimento das amplitudes e das fases permite a reconstituição da densidade eletrônica do cristal. Quando os raios-X incidem sobre um cristal, o mesmo gera feixes difratados. Esta radiação difratada é resultado da interação dos campos elétricos dos raios-X com a nuvem eletrônica dos átomos no cristal. Assim podemos afirmar que a interpretação do padrão de difração de raios-X fornece informação sobre a parte eletrônica do cristal, mais especificamente sobre a densidade eletrônica do cristal (GIACOVAZZO et al., 1992).

A onda total espalhada por um pequeno volume em uma posição r terá uma amplitude proporcional à densidade eletrônica e defasada por um termo de fase. Para se obter a expressão para o espalhamento de um átomo é necessário realizar o somatório sobre toda a distribuição espacial de elétrons atômicos (nuvem eletrônica) (WOOLFSON, 1970; RONDEAU; SCHREUDER, 2015).

Assim, podemos escrever a expressão para o fator de espalhamento atômico f como a integral:

$$f(\vec{S}) = \int_V \rho(\vec{r}) \exp(2\pi i \vec{r} \cdot \vec{S}) dV \quad (1)$$

O espalhamento total da cela unitária pode ser escrito como o somatório sobre as contribuições de todos os átomos dentro da cela unitária. Esta grandeza é chamada de fator de estrutura (F), que é uma função da distribuição da densidade eletrônica na cela unitária, dado por:

$$\begin{aligned} F_{hkl} &= \sum_j f_j \exp(2\pi i (hx_j + ky_j + lz_j)) \\ &= |F_{hkl}| \exp i \alpha_{hkl} \end{aligned} \quad (2)$$

onde f_j é o fator de espalhamento atômico para o j -ésimo átomo na cela unitária, que tem coordenadas x_j , y_j , z_j e α_{hkl} é a fase do fator de estrutura $F(hkl)$. Os inteiros (hkl) são os índices de Miller para a reflexão ocorrendo em uma dada direção.

Como as equações 1 e 2 sugerem, a estrutura (eletrônica) cristalina e seu padrão de difração (densidade eletrônica) são relacionados entre si por uma Transformada de Fourier (RUPP, 2009). O padrão de difração e a transformada de Fourier da densidade eletrônica da estrutura são dados por:

$$F_{hkl} = \sum_j \rho \exp(2\pi i (hx_j + ky_j + lz_j)) \quad (3)$$

$$\begin{aligned} \rho_{xyz} &= \frac{1}{V} \sum_{hkl} |F_{hkl}| \exp(i \alpha_{hkl}) \exp(-2\pi i (hx_j \\ &\quad + ky_j + lz_j)) \end{aligned} \quad (4)$$

A equação 2 pode ser utilizada para calcular o padrão de difração esperado para a estrutura modelo e então ser usado durante o refinamento da estrutura cristalina. Esses valores encontrados são chamados de fatores de estrutura calculados, $F_{calc}(hkl)$. Essa equação descreve, matematicamente, a amplitude

observada quando os raios-X são difratados por um cristal, durante a coleta de dados (DUCRUIX; GIEGÉ, 2000).

A última etapa na determinação da estrutura tridimensional de proteínas é o refinamento, processo que visa encontrar uma melhor concordância entre o modelo proposto para a estrutura da molécula e a sua estrutura real. Esta concordância deve refletir numa igualdade entre os fatores de estrutura calculados (F_{calc}) e os fatores de estrutura observados (F_{obs}) (ARCURI, 2008; BARCELOS et al., 2008; CASTRO, 2012).

É necessária a realização de vários ciclos de refinamento para que todos os átomos do modelo representem de maneira mais fiel possível a realidade experimental medida, e, além disso, estejam no melhor posicionamento possível levando em conta a geometria das moléculas, valores dos ângulos das cadeias laterais e dimensão das ligações covalentes, que permitirão o enquadramento da cadeia principal e das cadeias laterais no mapa de densidade eletrônica e ajustes no mapa da densidade eletrônica para as densidades negativas e positivas (ESTEVES, 2010; FEITEN et al., 2016). O processo de refinamento visa trazer o valor de F_{calc} o mais próximo possível de F_{obs} , variando as coordenadas atômicas, porém sempre respeitando as restrições estereoquímicas e eletrostáticas.

O acompanhamento da qualidade deste processo é feito por meio do cálculo do fator R (R_{fator}) durante o processo. O R_{fator} é uma medida da similaridade entre o modelo cristalográfico utilizado e os dados experimentais de difração de raios-X. Em outras palavras, é uma medida de quão bem a estrutura refinada prevê os dados observados no experimento de difração de raios-X (MORRIS et al., 1992).

Para estruturas cristalográficas de proteínas resolvidas à resolução melhor de 2,0 Å, espera-se que o fator R atinja valores menores que 20%. Como R pode ser minimizado artificialmente e levando em consideração a redundância dos dados de raios-X, uma pequena porcentagem das reflexões são excluídas do refinamento e utilizadas como um conjunto de teste no cálculo de um novo fator R chamado R_{livre} , que é isento de minimização artificial. Normalmente o R_{livre} possui um valor um pouco mais alto que o R, normalmente não excedendo os

10% para estruturas refinadas a alta resolução (RHODES, 1993; 2006; BURLEY et al., 2019).

São vários os algoritmos e programas computacionais necessários para a obtenção da estrutura tridimensional de enzimas. Os dados de difração de raios-X podem, por exemplo, ser processados com *XDS* (KABSCH, 2010). As fases iniciais calculadas com *Phaser* (McCOY et al., 2007) e a estrutura modelada utilizando MR e *COOT* (EMSLEY; COWTAN, 2004). Finalmente, o refinamento da estrutura tridimensional pode ser realizado com *REFMAC5* (VAGIN et al., 2004) e/ou *Phenix* (ADAMS et al., 2010).

3. Conclusão

Com todo o exposto, verifica-se que a cristalografia por difração de raios-X é uma técnica bem consolidada para a determinação da estrutura tridimensional de macromoléculas biológicas há mais de meio século. Esta técnica vem sendo apurada com o passar dos anos e é responsável pela determinação das estruturas tridimensionais de uma grande gama de enzimas já depositadas no PDB (DUTTA; BERMAN, 2005), com fácil acesso à comunidade acadêmica (BURLEY et al., 2019). Isso se deve às vantagens da cristalografia, que não apresenta limite de tamanho da molécula a ser analisada e tem se tornado um método cada vez mais rápido e automatizado. Síncrotrons fornecem fontes intensas de raios-X, enquanto que os novos detectores permitem uma rápida coleta de dados. Softwares novos e mais automatizados para determinação da estrutura de cristais, bem como o acesso relativamente fácil à computação de alta velocidade também têm reduzido muito o tempo de processamento e análise de dados. Porém, apresenta algumas etapas-chave na aplicação em proteínas de maneira geral, uma vez que pequenos detalhes no desenvolvimento da técnica podem acarretar em dificuldades na obtenção dos dados experimentais de difração ou em interpretações errôneas dos mesmos.

Mais recentemente, a cristalografia de difração de raios-X tem sido aplicada para a elucidação dos mecanismos catalíticos (BLUNDELL, 2017) e análise de alterações na estrutura conformacional (secundária, terciária e/ou quaternária)

ou no sítio ativo de enzimas, alteradas por engenharia molecular ou que foram submetidas a algum tratamento químico ou físico e tiveram sua estrutura e atividade melhoradas (DOBBEK, 2011; DUCKWORTH et al., 2013). Portanto, esta é uma importante e bem documentada técnica para a determinação de estruturas tridimensionais de macromoléculas, especialmente de enzimas; e pode ser aplicada para a elucidação das estruturas das enzimas que tiveram sua atividade catalítica alterada pelo emprego de algum método alternativo, tais como emprego de altas pressões, radiação ultrassônica ou campo magnético (KURPIEWSKA; LEWINSKI, 2010; FEITEN et al., 2016).

Referências

ADAMS, P. D. et al. *PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D*, v. D66, p. 213–221, 2010.

ARCURI, Helen Andrade. *Bioinformática estrutural aplicada ao estudo de enzimas da via metabólica do ácido chiquímico*. Tese [Doutorado em Biofísica Molecular]. Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual de São Paulo. São José do Rio Preto, 113 p. 2008.

BARCELOS, G. B. et al. Molecular Modeling as a Tool for Drug Discovery. *Current Drug Targets*, v. 9, p. 1084-1091, 2008.

BLUNDELL, Tom L. Protein crystallography and drug discovery: recollections of knowledge exchange between academia and industry. *International Union of Crystallography Journal*, v. 4, n. 4, p. 308-321, 2017.

BORZANI, W. et al. *Biotecnologia industrial. Volume I – Fundamentos*. São Paulo: Edgard Blüncher Ltda., 2001.

BURLEY, S. K. et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Research*, v. 47 (D1): 464-474, 2019.

CASTRO, Rosane de Paula. *Análise estrutural cristalográfica de protótipos de fármacos derivados de N-fenilpiperazina e de N-acilidrazonas*. Dissertação [Mestrado em Física]. Instituto de Física, Universidade Federal de Goiás. Goiânia, 78 p. 2012.

COLNAGO, L. A.; ALMEIDA, F. C. L.; VALENTE, A. P. Espectrometria de massa e RMN no estudo de macromoléculas biológicas. *Química Nova na Escola*, n.

16, 2002. Disponível em: http://qnesc.sbg.org.br/online/qnesc16/v16_A04.pdf
Acesso em: 28 Jan. 2022.

DOBBEK, Holger. Structural aspects of mononuclear Mo/W-enzymes. *Coordination Chemistry Reviews*, v. 255, p. 1104 – 1116, 2011.

DUCKWORTH, H. W. et al. Enzyme-substrate complexes of allosteric citrate synthase: Evidence for a novel intermediate in substrate binding. *Biochimica et Biophysica Acta*, v. 1834, p. 2546 – 2553, 2013.

DUCRUIX, A.; GIEGÉ, R. *Crystallization of nucleic acids and proteins: A practical approach*. 2. ed. New York: Oxford University Press, 2000.

DUTTA, S.; BERMAN, H. M. Large macromolecular complexes in the Protein Data Bank: A status report. *Structure*, v. 13, p. 381 – 388, 2005.

EMSLEY, P.; COWTAN, K. D. *Coot*: model-building tools for molecular graphics. *Acta Crystallographica Section D*, v. D60, p. 2126 - 2132, 2004.

ESTEVEZ, Gisele Ferreira. *Resolução da estrutura tridimensional do inibidor trípico e quimotríptico de *Vigna unguiculata* em complexos binário e ternário*. Tese [Doutorado em Biologia Molecular]. Instituto de Ciências Biológicas, Departamento de Biologia Molecular, Universidade de Brasília. Brasília, 163 p. 2010.

FEITEN, M. C. et al. X-ray crystallography as a tool to determine three-dimensional structures of commercial enzymes subjected to treatment in pressurized fluids. *Applied Biochemistry and Biotechnology*, v. 182, n. 2, p. 429 – 451, 2016.

GIACOVAZZO, C. et al. *Fundamentals of crystallography*. New York: Oxford University Press, 1992.

KABSCH, Wolfgang. XDS. *Acta Crystallographica Section D*, v. D66, p. 125 – 132, 2010.

KURPIEWSKA K.; LEWINSKI K. High pressure macromolecular crystallography for structural biology: a review. *Central European Journal of Biology*, v. 5, n. 5, p. 531 – 542, 2010.

LAI, J. et al. X-ray and NMR crystallography in an enzyme active site: The indoline quinonoid intermediate in tryptophan synthase. *Journal of the American Chemical Society*, v. 133, p. 4 – 7, 2011.

MATTHEWS, Brian W. Solvent content of protein crystals. *Journal of Molecular Biology*, v. 33, p. 491 – 497, 1968.

McCOY, A. J. et al. *Phaser* crystallographic software. *Journal of Applied Crystallography*, v. 40, p. 658 – 674, 2007.

MORRIS, A. L. et al. Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function and Bioinformatics*, v. 12, p. 345 – 364, 1992.

MUNTE, Claudia Elisabeth. *Ressonância magnética nuclear na determinação de estrutura de proteínas: Aplicação à mutante HIS15ALA de HPR de Staphylococcus Aureus, e ao peptídeo-C da proinsulina humana*. Tese [Doutorado em Física]. Instituto de Física de São Carlos, Universidade de São Paulo. São Carlos, 151 p. 2001.

NASCIMENTO, C. J.; BLOCH JR, C. Ressonância magnética nuclear: Gradus Primus*. *Biotecnologia Ciência & Desenvolvimento*, v. 21, p. 52-61, 2001.

NELSON, D. L.; COX, M. M. *Princípios de Bioquímica de Lehninger*. 6. ed. New York: Freeman, 2013.

NELSON, D. L.; COX, M. M. *Lehninger principles of biochemistry*. 7. ed. New York: Freeman, 2017.

NONATO, Maria Cistina Costa. *Estudos cristalográficos em macromoléculas biológicas: Aplicações em Calgranulina C de granulócitos porcinos, Tripanotona redutase de Trypanosoma cruzi e Fosfolipase A₂ extraída do veneno da serpente Bothrops moojeni*. Tese [Doutorado em Física]. Instituto de Física de São Carlos, Universidade de São Paulo. São Carlos, 155 p. 1997.

PDB. Worldwide Protein Data Bank. Disponível em: <http://www.pdb.org/index.html> Acesso em: 14 Jan. 2022.

RHODES, Gale. *Crystallography made crystal clear: a guide for users of macromolecular models*. London: Academic Press, 1993.

RHODES, Gale. *Crystallography Made Crystal Clear: a guide for users of macromolecular models*. 3. ed. London: Academic Press, 2006.

RONDEAU, J. M.; SCHREUDER, H. *Protein crystallography and drug discovery. The Practice of Medicinal Chemistry*. 4. ed. London: Academic Press, 2015.

RUPP, Bernhard. *Biomolecular crystallography: principles, practice, and application to structural biology*. New York: Garland Science, 2009.

SABADINI, E.; CARVALHO, L. V. C. Visual demonstration of the ionic strength effect in the classroom: the Debye-Hückel limiting law. *Química Nova*, v. 36, n. 1, p. 187-189, 2013.

SCRIBAN, René. *Biotecnologia*. São Paulo: Malone, 1985.

SILVESTRINI, L.; CIANCI, M. Principles of lipid–enzyme interactions in the limbus region of the catalytic site of *Candida antarctica* Lipase B. *International Journal of Biological Macromolecules*, v. 158, p. 358–363, 2020.

STODDARD, Barry L. Intermediate trapping and Laue X-ray diffraction: Potential for enzyme mechanism, dynamics, and inhibitor screening. *Pharmacology & Therapeutics*, v. 70, n. 3, p. 215 - 256, 1996.

STOUT, G. H.; JENSEN, L. H. *X-Ray Structure Determination: A Practical Guide*. 2. ed. New York: John Wiley & Sons, 1989.

SURYANARAYANA, C.; NORTON, M. G. *X-ray diffraction: A practical approach*. New York: Plenum Press, 1998.

VAGIN, A. A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D - Biological crystallography*, v. 60, p. 2184 – 2195, 2004.

VOET, D.; VOET, J. G. *Biochemistry*. 3. ed. New York: John Wiley & Sons, 2004.

WAHLSTROM, Ernest E. *Optical crystallography*. 3. ed. New York: John Wiley & Sons, 1969.

WOOLFSON, Michael Mark. *An introduction to X-ray crystallography*. Editora Cambridge University Press, Londres, 1970.