



Hugo David Cancela Correia

Licenciado em Bioquímica

**Estudos estruturais de interacções
proteína-ligando: Aldeído
Oxidoredutase e Tripsina**

Dissertação para obtenção do Grau de Mestre em
Bioquímica Estrutural e Funcional

Orientador: Doutora Teresa Sacadura Santos-Silva,
Investigadora Auxiliar, FCT/UNL

Júri:

Presidente: Prof. Doutor José Ricardo Ramos Franco Tavares
Arguente: Doutora Isabel Maria Travassos de Almeida de Jesus Bento
Vogal: Doutora Teresa Sacadura Santos-Silva

“Estudos estruturais de interacções proteína-ligando: Aldeído Oxidorredutase e Tripsina”

“Copyright” em nome de Hugo David Cancela Correia, da FCT/UNL e da UNL

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Agradecimentos

Durante a realização da dissertação do mestrado neste último ano, assisti a uma evolução da minha pessoa, crescendo a uma velocidade alarmante quer a nível pessoal como profissional. Gostaria, portanto, de agradecer a todas as pessoas que contribuíram directa ou indirectamente para esta experiência tão rica.

Agradeço à Professora Maria João Romão, pela oportunidade de realizar esta dissertação no Grupo de Cristalografia e pela sua opinião e orientação relativamente às experiências realizadas.

À minha orientadora, Teresa, por me aceitar como seu aluno, pelos ensinamentos desde as bases até aos conceitos mais complexos, por me mostrar a luz ao fundo do túnel nos momentos mais difíceis e por toda a confiança depositada em mim.

A todos os membros do XTAL: Ana Luísa, Abhik, Angelina, Márcia, Benedita, Catarina, Cecília, Filipe, Diana, Marino e Carmen pela ajuda, ensinamentos e por todos os bons momentos passados durante a execução desta dissertação. Agradeço também às alunas de projecto, em particular à Joana, a minha primeira “aluna”.

Ao Pablo e ao Jacopo com quem colaborei e tive a oportunidade de partilhar e discutir os vários progressos alcançados durante esta dissertação, estando sempre disponíveis quando precisei de ajuda.

À Rita, a minha “irmã” no laboratório, com quem partilhei muitos bons momentos, bem como um pouco de stress e desespero ocasional.

À Diana, a minha parceira do Mestrado, pelo bom trabalho de equipa, pela amizade demonstrada e por todas as palavras de apoio quando eu mais precisei delas.

A todos os meus amigos e colegas da Licenciatura e do Mestrado que me acompanharam durante o meu percurso académico.

Aos meus amigos do Básico e do Secundário: Ji, Martinha, Ragú, Susana, Ritinha, Sara, Sardinha, Flexa, Tiago Cardoso, Pedro, e a todos os outros, pelo apoio e amizade sempre presentes.

Ao pessoal do Meco: Piu, Slorn, Tangela, Ricky, Daniela, Toon, Bolas, Rui, Vilma, Cátia, Igor, Bibi, Angelboy, Sandra, Andreia, Mamas, Joana, Augusto e Guilherme, por todos os momentos de aventura, descontração, diversão ou pura parvoíce que me proporcionaram ao longo dos anos. Que Ocem esteja convosco!

Por último, agradeço à minha Família que sempre me apoiou incondicionalmente em tudo e me moldou na pessoa que sou hoje. Sem ela não teria sido possível a realização desta dissertação. Aos meus Pais que mesmo nos momentos mais difíceis me indicaram o melhor caminho a percorrer. Um agradecimento especial às minhas Avós que desde criança me mimaram e ajudaram, não podendo fazer melhor o papel de avó.

Um Muito Obrigado A Todos!

Resumo

O conhecimento e compreensão das interações entre proteínas e ligandos é muitas vezes a base para a resolução de questões complexas a nível biológico. Na presente dissertação são estudadas as proteínas Aldeído Oxidorredutase de *Desulfovibrio gigas* (DgAOR) e a tripsina de *Bos taurus* com vários ligandos por Cristalografia de Raios-X.

A DgAOR é uma enzima de molibdénio da família da Xantina Oxidase que apresenta a função de catalisar a hidroxilação oxidativa de aldeídos a ácidos carboxílicos. Com o objectivo de se perceber qual a diferença entre as formas activa e inactiva desta enzima, foram realizados estudos cinéticos com ditionito de sódio e sulfureto de sódio, reagentes utilizados na activação de proteínas da mesma família. Os resultados mostraram uma activação inicial seguida de uma inactivação. O modelo cristalográfico determinado nas mesmas condições sugere uma molécula de peróxido de hidrogénio ligada ao átomo de molibdénio do centro activo. Quando comparada com a estrutura obtida usando apenas peróxido de hidrogénio, as semelhanças são notórias reforçando a presença desta espécie a bloquear o centro activo. De modo a elucidar o mecanismo de reacção da enzima, foram executados estudos com diferentes aldeídos, tendo-se observado a interacção entre o cofactor e os substratos, nalguns casos originando os produtos da reacção.

A tripsina é uma protease de serina que catalisa a hidrólise de ligações peptídicas. Esta foi utilizada como modelo cristalográfico para o estudo de interações com cinco compostos desenhados para a inibição da urocinase, uma enzima muito semelhante, da mesma família. Foram obtidas estruturas com os compostos AB1 e JS62 ligados ao centro activo permitindo a visualização das interações entre estes e os resíduos chave da proteína. O composto AB11 não se ligou à proteína, enquanto que, para os restantes dois compostos, JS67 e SR5, obteve-se densidade electrónica de difícil interpretação no centro activo.

Termos chave: Cristalografia de raios-X; Cinética enzimática; Interações proteína-ligando; Aldeído Oxidorredutase; Tripsina; Espécies Reactivas de Oxigénio; Inibidores de protease; Compostos de Boro

Abstract

The knowledge and comprehension of protein-ligand interactions is often the basis for the resolution of complex biological issues. In the present dissertation the proteins Aldehyde Oxidoreductase from *Desulfovibrio gigas* (DgAOR) and trypsin from *Bos taurus* are studied with various ligands by X-ray crystallography.

DgAOR is a molybdenum enzyme of the Xanthine Oxidase family that catalyzes the oxidative hydroxylation of aldehydes to carboxylic acids. In order to distinguish the active and inactive forms of this enzyme, kinetic studies were performed with sodium dithionite and sodium sulfide, reagents used in the activation of proteins of the same family. The results showed an initial activation followed by an inactivation. The crystallographic model determined in the same conditions suggests a hydrogen peroxide molecule bound to the molybdenum atom of the active site. When compared to the structure obtained with hydrogen peroxide, the similarities are notorious, reinforcing the presence of this species blocking the active site. In order to elucidate the reaction mechanism of the enzyme, studies with different aldehydes were executed, resulting in structures where the interaction between the cofactor and the substrates, and in some cases the product, was observed.

Trypsin is a serine protease that catalyzes the hydrolysis of peptide bonds. This enzyme was used as a crystallographic model for the study of interactions with five compounds designed for urokinase inhibition, which is a very similar enzyme of the same family. Structures were obtained with the compounds AB1 and JS62 bound to the active site, allowing the characterization of the interactions between these compounds and the key residues of the protein. The compound AB11 didn't bind to the protein and some electron density of difficult interpretation was obtained for compounds JS67 and SR5.

Keywords: X-ray crystallography; Enzymatic kinetics; Protein-ligand interactions; Aldehyde Oxidoreductase; Trypsin; Reactive Oxygen Species; Protease inhibitors; Boron compounds

Índice

Agradecimentos	III
Resumo	V
<i>Abstract</i>	VII
Índice.....	IX
Índice de Figuras.....	XI
Índice de Tabelas	XV
Lista de Abreviaturas, siglas e símbolos	XVII
1. Introdução	1
1.1. Aldeído Oxidorredutase.....	3
1.1.1. Molibdénio	3
1.1.2. Cofactor do molibdénio	3
1.1.3. Enzimas de molibdénio	4
1.1.4. A família da Xantina Oxidase	6
1.1.5. Aldeído Oxidorredutase de <i>Desulfovibrio gigas</i>	7
1.1.5.1. Descrição da estrutura	8
1.1.5.2. Cofactores e centro activo	8
1.1.5.3. Cadeia de transferência electrónica	11
1.1.5.4. Inibição do centro activo	11
1.2. Tripsina.....	13
1.2.1. Proteases	13
1.2.2. Proteases de serina.....	14
1.2.3. Tripsina de <i>Bos taurus</i>	14
1.2.3.1. Descrição da estrutura	16
1.2.3.2. Mecanismo da reacção.....	18
1.2.3.3. Inibidores da tripsina	19
1.2.4. Urocinase.....	21
1.3. Cristalografia de Raios-X	21
1.3.1. Cristalização de proteínas	24
1.3.2. Experiência de difracção	28
1.3.3. Da difracção à densidade electrónica.....	30
1.3.4. Refinamento e validação da estrutura.....	32
2. Objectivos	35
3. Materiais e Métodos.....	39
3.1. Aldeído Oxidorredutase de <i>Desulfovibrio gigas</i>	41
3.1.1. Purificação.....	41

3.1.2. Ensaio cinéticos	42
3.1.3. Ensaio de cristalização	42
3.1.4. Recolha de dados e resolução da estrutura	43
3.2. Tripsina.....	44
3.2.1. Ensaio de cristalização	44
3.2.2. Recolha de dados e resolução da estrutura	46
4. Resultados e Discussão	47
4.1. Aldeído Oxidorredutase de <i>Desulfovibrio gigas</i>	49
4.1.1. Ensaio cinéticos	49
4.1.2. Ensaio de cristalização	50
4.1.3. Aldeído Oxidorredutase em complexo com ROS.....	54
4.1.3.1. Experiência de difracção.....	54
4.1.3.2. Resolução das estruturas.....	57
4.1.3.3. Refinamento das estruturas.....	58
4.1.3.4. Caracterização da estrutura.....	61
4.1.4. Aldeído Oxidorredutase em complexo com substratos	68
4.1.4.1. Experiência de difracção.....	68
4.1.4.2. Resolução das estruturas.....	69
4.1.4.3. Refinamento e validação das estruturas.....	70
4.1.4.4. Caracterização das estruturas.....	73
4.2. Tripsina de <i>Bos taurus</i>	79
4.2.1. Ensaio de cristalização	79
4.2.2. Experiência de difracção	82
4.2.3. Resolução das estruturas.....	84
4.2.4. Refinamento das estruturas.....	85
4.2.5. Caracterização das estruturas.....	87
4.2.5.1. Tripsina em complexo com AB1	89
4.2.5.2. Tripsina em complexo com JS62.....	92
4.2.5.3. Tripsina em complexo com JS67.....	93
4.2.5.4. Tripsina em complexo com SR5.....	93
5. Conclusões e Perspectivas futuras	97
6. Bibliografia	103
7. Anexos	113
7.1. Sequência peptídica da AOR de <i>D. gigas</i>	115
7.2. Sequência peptídica da tripsina de <i>B. taurus</i>	116
7.3. Estatísticas relativas à estrutura da <i>DgAOR</i> com $\text{Na}_2\text{S}_2\text{O}_4$ e Na_2S	117

Índice de Figuras

Figura 1.1. Representação esquemática da estrutura das diferentes formas da pironopterina: MPT – forma de monofosfato; MCD – <i>molybdopterin cytosine dinucleotide</i> ; MGD – <i>molybdopterin guanosine dinucleotide</i> . Adaptado de Romão 2009.....	4
Figura 1.2. Representação esquemática do centro activo para as três principais famílias de enzimas de molibdénio. Adaptado de Romão 2009.	5
Figura 1.3. Representação esquemática da estrutura do cofactor pterínico (MCD) da <i>DgAOR</i> . O ligando OH_x do molibdénio corresponde ao ligando lábil OH/OH_2	7
Figura 1.4. Representação da estrutura tridimensional da <i>DgAOR</i> (Código PDB: 1VLB). Os seus quatro domínios estão representados a amarelo (Fe-S _{distal}), laranja (Fe-S _{proximal}), vermelho (Mo1) e azul (Mo2). A cinzento estão representados os fragmentos de ligação entre os domínios. As esferas representam os cofactores: pironopterina e centros Fe-S, com o seguinte código de cores: Carbono – cinzento; Azoto – azul; Oxigénio – vermelho; Fósforo – laranja; Enxofre – amarelo; Ferro – castanho; Molibdénio – azul-esverdeado. Esta figura foi preparada utilizando o programa PyMOL.....	9
Figura 1.5. Representações tridimensional (à esquerda) e esquemática (à direita) do cofactor pterínico. S7 e S8 – grupo ditioleno da pterina; OM1 – ligando oxo apical; OR1 – ligando oxo equatorial; OM2/ OH_x – ligando lábil OH/OH_2 . Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Fósforo – laranja; Enxofre – amarelo; Molibdénio – azul-esverdeado. O programa PyMOL foi utilizado na preparação desta figura.	10
Figura 1.6. Representação esquemática do mecanismo de reacção proposto para a <i>DgAOR</i>	11
Figura 1.7. Representação esquemática do centro activo da <i>DgAOR</i> inibido com três compostos: A – arsenito; B – etilenoglicol; C – glicerol.....	12
Figura 1.8. Representação esquemática da reacção geral das proteases: hidrólise catalítica de uma ligação peptídica.	13
Figura 1.9. Representação da estrutura tridimensional da tripsina de <i>Bos taurus</i> (Código PDB: 1S0Q). A amarelo estão representados os motivos de hélice- α da sua estrutura secundária, a vermelho os motivos de folha- β e a ciano os <i>loops</i> . Esta figura foi preparada utilizando o programa PyMOL.	16
Figura 1.10. Estrutura tridimensional do centro activo da tripsina, evidenciando a tríade catalítica (Serina 195, Histidina 57 e Aspartato 102). A tracejado estão representadas as ligações de hidrogénio entre as cadeias laterais deste resíduos. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho. Esta figura foi preparada utilizando o programa PyMOL.	17
Figura 1.11. Representação esquemática do domínio de reconhecimento do substrato para as enzimas tripsina (à esquerda), quimotripsina (ao meio) e elastase (à direita). Adaptado de Voet <i>et al.</i> 2008.....	18
Figura 1.12. Representação esquemática do carácter parcial de ligação dupla de uma ligação peptídica. .	19
Figura 1.13. Representação esquemática do mecanismo de reacção da tripsina. Adaptado de Hedstrom 2002.....	19
Figura 1.14. Representação esquemática da interacção entre os ácidos borónicos e a tríade catalítica das enzimas da família da tripsina. Adaptado de Tsilikounas <i>et al</i> 1992.....	20
Figura 1.15. Representação esquemática dos principais passos necessários para a obtenção de uma estrutura tridimensional por cristalografia de raios-X.	23
Figura 1.16. Representação esquemática das técnicas de gota suspensa (à esquerda) e gota assente (à direita).	25

Figura 1.17. Diagrama de fases da cristalização. Adaptado de http://www.p212121.com/2009/11/15/the-phase-diagram-in-protein-crystallization/	25
Figura 1.18. Representação esquemática de uma célula unitária e respectivas constantes (a , b , c e α , β , γ).	26
Figura 1.19. Representação esquemática da relação entre a unidade assimétrica, a célula unitária e a rede cristalina. Adaptado de http://www.rcsb.org/pdb/101/static101.do?p=education_discussion/Looking-at-Structures/bioassembly_tutorial.html	27
Figura 1.20. Representação esquemática do arranjo experimental de uma experiência de difracção. Adaptado de http://www.esrf.eu/UsersAndScience/Publications/Highlights/2003/Materials/MAT15/	28
Figura 1.21. Representação esquemática da incidência de radiação num conjunto de planos. Adaptado de http://www-structmed.cimr.cam.ac.uk/Course/Basic_diffraction/Diffraction.html	29
Figura 3.1. Representação da estrutura dos compostos AB1, AB11, JS62, JS67 e SR5, com os quais se incubou a tripsina.	45
Figura 4.1. Representação gráfica da actividade específica da enzima <i>DgAOR</i> ao longo do tempo após incubação com ditionito de sódio e sulfureto de sódio, em condições aeróbias (círculos cheios) e anaeróbias (círculos vazios).	50
Figura 4.2. Cristal de <i>DgAOR</i> obtido em condições aeróbias a 277 K. Condições de cristalização: 30% isopropanol, 0,2 M HEPES pH 7,6 e 0,2 M MgCl ₂	51
Figura 4.3. Exemplo de um padrão de difracção do cristal <i>DgAOR</i> _H ₂ O ₂ . Resolução máxima de 1,50 Å.	54
Figura 4.4. Densidade electrónica observada para a cisteína 271, a qual foi modelada como uma hidroxicisteína, e moléculas de água na sua vizinhança. Código de cores: Carbono – ciano; Oxigénio – vermelho; Enxofre – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.	63
Figura 4.5. Estrutura cristalográfica do centro activo da amostra <i>DgAOR</i> _H ₂ O ₂ , em duas orientações diferentes. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.	64
Figura 4.6. Estrutura cristalográfica do centro activo da <i>DgAOR</i> , cristalizada na presença de Na ₂ S ₂ O ₄ e Na ₂ S, em duas orientações diferentes. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.....	65
Figura 4.7. Comparação entre o metabolismo das ROS no caso da XO (A) e os resultados aqui apresentados para a <i>DgAOR</i> (B). Código de cores: Domínio do cofactor do Mo – verde; Domínio dos centros Fe-S – vermelho; Domínio do cofactor flavínico (XO apenas) – amarelo. Representação esquemática preparada utilizando o programa Qutemol.....	67
Figura 4.8. Estabilização do anel aromático do benzaldeído pelos resíduos Phe425 e Tyr535. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Fósforo – laranja; Enxofre – amarelo; Molibdénio – azul-esverdeado. Esta figura foi preparada utilizando o programa PyMOL.	74
Figura 4.9. Estrutura cristalográfica do centro activo da amostra <i>DgAOR</i> _benz. Devido à incerteza da posição do grupo carbonilo, a molécula de benzaldeído é aqui representada em duas conformações: (A) com o grupo carbonilo orientado para o centro activo e (B) orientado numa direcção oposta. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.	75

- Figura 4.10. Densidade electrónica relativa a uma putativa segunda molécula de benzaldeído, adjacente à primeira, no canal de acesso ao centro activo. Código de cores: Carbono – ciano; Oxigénio – vermelho. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 76
- Figura 4.11. Estrutura cristalográfica do centro activo da amostra *DgAOR_form*. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 77
- Figura 4.12. Representação esquemática da reacção de hidratação do formaldeído em solução aquosa, originando metileno glicol..... 78
- Figura 4.13. Estrutura cristalográfica do centro activo da amostra *DgAOR_acet*. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 79
- Figura 4.14. Cristais de tripsina em complexo com benzamidina utilizando as condições 15 (A) e 31 (B). Estas condições apresentam PEG 3350 ou PEG 8000 como agentes precipitantes, sulfato de amónio e, no caso da condição 15, tampão cacodilato de sódio a pH = 6,5..... 80
- Figura 4.15. Cristais de tripsina incubada com os compostos AB1 (A), AB11 (B), JS62 (C) e SR5 (D), obtidos utilizando a condição de cristalização 15. Os cristais obtidos para a tripsina incubada com o composto JS67 são idênticos aos com SR5. 80
- Figura 4.16. Exemplos de dois padrões de difracção dos cristais Trip_AB1 (A) e Trip_JS62 (B). Resolução máxima de 1,42 e 1.30 Å, respectivamente. 83
- Figura 4.17. Coordenação octaédrica do ião Ca^{2+} no *Calcium-binding site* da tripsina. Este estabelece ligações de hidrogénio com os resíduos Glu70, Asn72, Val75 e Glu80 e duas moléculas de água conservadas. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Cálcio – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 87
- Figura 4.18. Conformação alternada observada para a Cys220, originada por quebra da ponte de dissulfureto. Código de cores: Carbono – ciano; Enxofre – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL. 88
- Figura 4.19. Quebra da ligação entre a Pro152 e o Asp153 no *autolysis loop* da tripsina. O grupo carbonilo da Pro152 apresenta conformação alternada. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 89
- Figura 4.20. Estrutura cristalográfica do centro activo da amostra Trip_AB1. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Boro – rosa. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 90
- Figura 4.21. Representação esquemática do composto AB1 no centro activo da tripsina. A tracejado estão representadas todas as ligações de hidrogénio entre o composto e a proteína e moléculas de água conservadas. Representados por X encontram-se átomos não identificados para os quais foi possível observar densidade electrónica..... 91
- Figura 4.22. Estrutura cristalográfica do centro activo da amostra Trip_JS62. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Boro – rosa. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL..... 92

Figura 4.23. Representação esquemática do composto JS62 no centro activo da tripsina. A tracejado estão representadas todas as ligações de hidrogénio entre o composto e a proteína e moléculas de água conservadas.93

Figura 4.24. Estrutura cristalográfica do centro activo da amostra Trip_JS67. Não foi modelado nenhum composto no centro activo, sendo possível observar alguma densidade electrónica. Foi modelado um ião sulfato perto da serina catalítica e do *oxyanion hole*. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.....94

Figura 4.25. Estrutura cristalográfica do centro activo da amostra Trip_SR5. Não foi modelado nenhum composto no centro activo, sendo possível observar alguma densidade electrónica. Foi modelado um ião sulfato perto da serina catalítica e do *oxyanion hole*. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.....94

Índice de Tabelas

Tabela 3.1. Condições de cristalização testadas para a tripsina.....	44
Tabela 4.1. Concentração final de cada composto e duração dos ensaios de <i>soaking</i> efectuados.	52
Tabela 4.2. Resumo das condições testadas e dos resultados obtidos na experiência de difracção para a DgAOR.....	53
Tabela 4.3. Estatísticas de recolha de dados para os cristais DgAOR_H ₂ O ₂ e DgAOR_Dit/S ²⁻ . Os valores entre parêntesis correspondem à última camada de resolução.....	56
Tabela 4.4. Valores de Z-score e LLG para as funções de rotação e translação obtidos utilizando o programa PHASER, para os conjuntos de dados DgAOR_H ₂ O ₂ e DgAOR_Dit/S ²⁻	58
Tabela 4.5. Factores R e R _{free} após o primeiro refinamento para as estruturas DgAOR_H ₂ O ₂ e DgAOR_Dit/S ²⁻	59
Tabela 4.6. Estatísticas de refinamento e validação para o modelo DgAOR_H ₂ O ₂	60
Tabela 4.7. Distância entre o átomo de molibdénio e os ligandos presentes na sua esfera de coordenação, no centro activo, para as duas estruturas e para a estrutura nativa.	66
Tabela 4.8. Estatísticas de recolha de dados para os cristais DgAOR_benz, DgAOR_form e DgAOR_acet. Os valores entre parêntesis correspondem à última camada de resolução.....	69
Tabela 4.9. Valores de Z-score e LLG para as funções de rotação e translação obtidos utilizando o programa PHASER, para os conjuntos de dados DgAOR_benz, DgAOR_form e DgAOR_acet.....	70
Tabela 4.10. Resultados obtidos no centro activo de todos os conjuntos de dados recolhidos, após um refinamento inicial, assim como os factores R e R _{free} iniciais.	71
Tabela 4.11. Estatísticas de refinamento actuais para os modelos DgAOR_benz, DgAOR_form e DgAOR_acet.	73
Tabela 4.12. RMSD de superposição estrutural das estruturas DgAOR_benz, DgAOR_form e DgAOR_acet com a estrutura nativa 1VLB, em relação a todos os átomos da cadeia principal da proteína.	73
Tabela 4.13. Resumo das condições testadas e dos resultados obtidos na experiência de difracção para a tripsina.	81
Tabela 4.14. Estatísticas de recolha de dados para os vários cristais de tripsina. Os valores entre parêntesis correspondem à última camada de resolução.	84
Tabela 4.15. Valores de Z-score e LLG para as funções de rotação e translação obtidos utilizando o programa PHASER, para todos os conjuntos de dados de tripsina.	85
Tabela 4.16. Estatísticas de refinamento actuais para os modelos Trip_AB1, Trip_JS62, Trip_JS67 e Trip_SR5.	86
Tabela 4.17. RMSD de superposição estrutural das estruturas Trip_AB1, Trip_JS62, Trip_JS67 e Trip_SR5 com a estrutura nativa 1S0Q, em relação a todos os átomos da cadeia principal da proteína....	87
Tabela 7.1. Estatísticas de recolha de dados para a estrutura de DgAOR obtida previamente à realização desta dissertação por <i>soaking</i> com Na ₂ S ₂ O ₄ e Na ₂ S. Os valores entre parêntesis correspondem à última camada de resolução.....	117

Tabela 7.2. Estatísticas de refinamento e validação para a estrutura de *DgAOR* obtida previamente à realização desta dissertação por *soaking* com $\text{Na}_2\text{S}_2\text{O}_4$ e Na_2S 118

Lista de Abreviaturas, siglas e símbolos

AB1 – Ácido (4-((carbamimidoiltio)metil)fenil)borónico

AB11 – Ácido (E)-(4-(((N'-metilcarbamimidoil)tio)metil)fenil)borónico

AO – Aldeído Oxidase

AOR – Aldeído Oxidorredutase

B. taurus – *Bos taurus*

BCA – Método do ácido bicinconínico

BPTI – *Basic Pancreatic Trypsin Inhibitor*

cPMP – Piranopterinina cíclica monofosfato

Crio-EM – Crio-Microscopia Electrónica

D. gigas – *Desulfovibrio gigas*

DCPIP – 2,6-diclorofenolindofenol

DgAOR – Aldeído Oxidorredutase de *Desulfovibrio gigas*

DgAOR_acet – Aldeído Oxidorredutase de *Desulfovibrio gigas* com acetaldeído

DgAOR_benz – Aldeído Oxidorredutase de *Desulfovibrio gigas* com benzaldeído

DgAOR_Dit/S²⁻ – Aldeído Oxidorredutase de *Desulfovibrio gigas* com ditionito de sódio e sulfureto de sódio

DgAOR_form – Aldeído Oxidorredutase de *Desulfovibrio gigas* com formaldeído

DgAOR_H₂O₂ – Aldeído Oxidorredutase de *Desulfovibrio gigas* com peróxido de hidrogénio

DgAOR_xant – Aldeído Oxidorredutase de *Desulfovibrio gigas* com xantina

DMSO – Dimetilsulfóxido

DMSOR – DMSO Redutase

ESRF – *European Synchrotron Radiation Facility*

FAD – Dinucleótido de flavina e adenina

GTP – Guanosina trifosfato

HEPES – Ácido N-(2-hidroxietil)piperazina-N-etanossulfónico

JS62 – 4-(4,4,5,5-tetrametil-1,3,2-dioxaborolan-2-il)benzil carbamimidotioato

JS67 – 4-(trifluorometil)benzil carbamimidotioato

KPB – Tampão de fosfato de potássio

LLG – *Log Likelihood Gain*

MAD – *Multi-wavelength anomalous dispersion*

MCD – *Molybdopterin cytosine dinucleotide*

MFEs – Molibdoflavoenzimas

MGD – *Molybdopterin guanosine dinucleotide*

MIR – *Multiple isomorphous replacement*

MoCo – *Molybdenum Cofactor*

MPT – *Molybdopterin* ou *metal-binding pterin*

MR – Substituição molecular

NAD⁺ – Nicotinamida adenina dinucleotídeo

PDB – Protein Data Bank

PEG – Polietilenoglicol

RMN – Ressonância Magnética Nuclear

RMSD – *Root-mean-square Deviation*

ROS – Espécies reactivas de oxigénio

RPE – Ressonância Paramagnética Electrónica

SAD – *single-wavelength anomalous dispersion*

SDS-PAGE – Electroforese em Gel de Poliacrilamida e Dodecil-Sulfato de Sódio

SIR – *single isomorphous replacement*

SO – Sulfito Oxidase

SR5 – Benzil carbamimidotioato

TLS – *Translation Libration Screw-motion*

Trip_AB1 – Tripsina com o composto AB1

Trip_AB11 – Tripsina com o composto AB11

Trip_JS62 – Tripsina com o composto JS62

Trip_JS67 – Tripsina com o composto JS67

Trip_SR5 – Tripsina com o composto SR5

Tris – Tris(hidroximetil)aminometano

uPA – *urokinase-type Plasminogen Activator*

uPAR – *urokinase-type Plasminogen Activator Receptor*

XDH – Xantina Desidrogenase

XO – Xantina Oxidase

Capítulo 1

Introdução

1.1. Aldeído Oxidorredutase

1.1.1. Molibdénio

Vários metais de transição são encontrados em enzimas, conferindo-lhes uma grande diversidade catalítica que não seria possível utilizando apenas os grupos funcionais das cadeias laterais dos aminoácidos. O molibdénio é um metal de transição que se encontra no 5º período e no grupo 6 da tabela periódica. Entre todos os metais de transição do 5º período da tabela periódica, este é o único que é fundamental para a grande maioria dos organismos vivos, sendo que os poucos que não necessitam deste elemento, utilizam tungsténio que se encontra imediatamente abaixo deste, no grupo 6 da tabela periódica e apresenta características muito semelhantes. É por isso considerado um microelemento para os animais, plantas e microorganismos, pois em baixas quantidades é vital, mas em quantidades mais elevadas pode ser letal. Apesar de ser muito pouco abundante na crosta terrestre, o molibdénio é o metal de transição mais abundante na água do mar, na forma de molibdato (MoO_4^{2-}). Tendo em conta que os oceanos assemelham-se à sopa primordial a partir da qual se pensa ter originado a vida, pode-se explicar a sua incorporação nos vários sistemas biológicos.¹⁻³

O molibdénio possui uma grande versatilidade, podendo apresentar estados de oxidação entre -2 e +6, sendo os estados +4 e +6 os mais estáveis, e uma coordenação entre 4 e 8 ligandos. Por estes motivos, este elemento encontra-se no centro activo de acima de 50 enzimas diferentes. Com a excepção das nitrogenases bacterianas que apresentam um centro multinuclear MoFe_7 , o molibdénio encontra-se na forma mononuclear em todos os outros casos conhecidos, coordenado por um cofactor orgânico tricíclico denominado de pterina.^{2,3}

1.1.2. Cofactor do molibdénio

Na literatura, a pterina é muitas vezes designada por MoCo (do inglês *Molybdenum Cofactor*), pois pensava-se, inicialmente, que este cofactor só existia em enzimas de molibdénio. Mais tarde foi descoberto que este também se encontra em enzimas de tungsténio, passando a designar-se por piranopterina.^{2,4}

Maioritariamente, as enzimas de molibdénio catalisam reacções redox, controlando a grande versatilidade química do metal através da piranopterina e do ambiente envolvente do centro activo. O cofactor, apesar de ser parte integral do centro activo destas enzimas, não participa directamente no mecanismo da catálise. Este coordena o metal através do seu grupo ditioleno e tem como funções principais o correcto posicionamento do metal, o controlo do seu comportamento redox e a transferência de electrões para outros grupos prostéticos.^{2,5}

A piranopterina pode estar presente nas formas de dinucleótido ou de monofosfato. Nos organismos eucarióticos esta apresenta-se na forma mais simples de monofosfato (MPT, do inglês *molybdopterin* ou *metal-binding pterin*) e nos organismos procarióticos esta encontra-se ligada a um nucleosídeo, normalmente uma citidina (MCD, do inglês *molybdopterin cytosine dinucleotide*) ou uma guanosina (MGD, do inglês *molybdopterin guanosine dinucleotide*). As diferenças entre as várias formas em que se apresenta este cofactor são evidenciadas na Figura 1.1.²

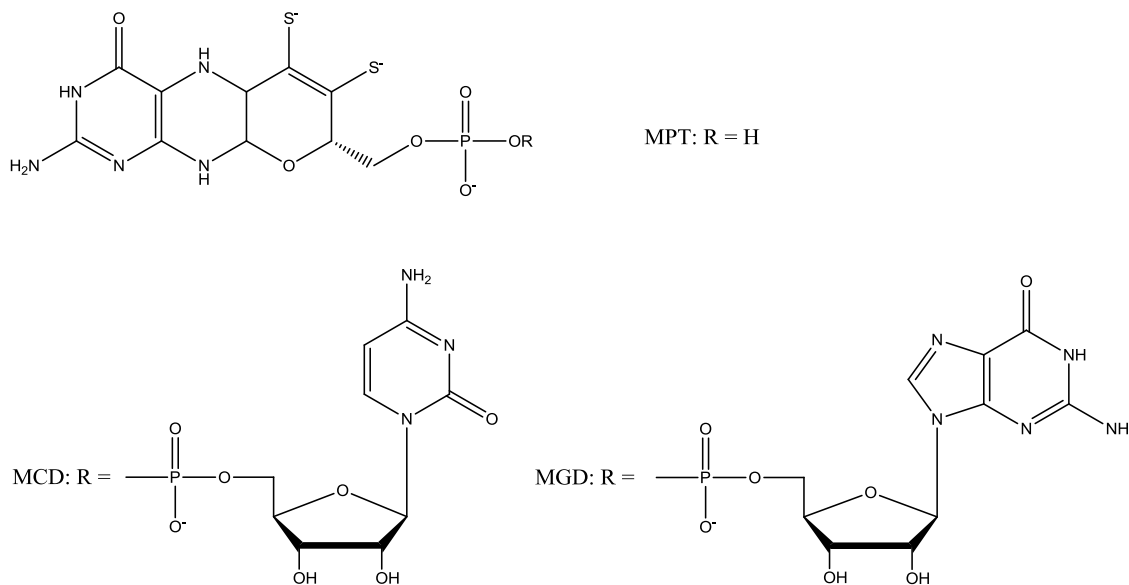


Figura 1.1. Representação esquemática da estrutura das diferentes formas da piranopterina: MPT – forma de monofosfato; MCD – *molybdopterin cytosine dinucleotide*; MGD – *molybdopterin guanosine dinucleotide*. Adaptado de Romão 2009.²

Apesar de estar presente nas mais diversas enzimas dos mais diversos organismos, a síntese da pterina é altamente conservada. Quando o molibdénio entra na célula, na forma de molibdato, este é incorporado no cofactor através da utilização de uma complexa maquinaria biossintética. Este processo ocorre em quatro passos principais: conversão do GTP a piranopterina cíclica monofosfato (cPMP, do inglês *cyclic Pyranopterin Monophosphate*), conversão do cPMP a MPT, inserção do átomo de molibdénio formando o MoCo e, finalmente, adição de diferentes nucleotídeos de modo a formar o MCD ou o MGD, no caso dos organismos procarióticos. O impedimento da biossíntese deste cofactor resulta na perda de funções metabólicas essenciais à sobrevivência dos organismos.⁵⁻⁷

1.1.3. Enzimas de molibdénio

O molibdénio é essencial para a actividade de várias enzimas de organismos procarióticos e eucarióticos. Estas estão envolvidas em diversos caminhos metabólicos, como por exemplo o catabolismo de purinas e aminoácidos sulfurados em animais ou a assimilação de nitrato e

biossíntese de hormonas em plantas. Em humanos são conhecidas a Xantina Oxidase (XO), a Sulfito Oxidase (SO) e a Aldeído Oxidase (AO) que estão envolvidas em doenças como a gota, a xantínúria ou dano provocado por radicais livres seguido de insuficiência cardíaca.²

As enzimas dependentes da pironopterina foram agrupadas em três grandes famílias, cada uma com diferentes tipos de reacções catalisadas e estruturas distintas do centro activo: a família da XO, a família da SO e a família da DMSO Redutase (DMSOR). A primeira família inclui enzimas como a XO, a AO e a Aldeído Oxidorredutase (AOR) e apresenta o átomo de molibdénio numa geometria piramidal quadrada ligado a um oxigénio na posição apical e a uma pterina (MCD ou MPT), um enxofre ou oxigénio e um grupo lábil OH/OH₂ nas posições equatoriais (Figura 1.2A). Na segunda família encontram-se enzimas como a Sulfito Oxidase e a Nitrato Redutase onde o Mo apresenta uma coordenação semelhante à família anterior mas o enxofre está ligado a uma cisteína através de uma ponte de dissulfureto (Figura 1.2B). A última família apresenta uma maior diversidade em termos de estrutura e função das enzimas que a integram, sendo que algumas podem apresentar tungsténio no centro activo em vez de molibdénio. Nesta família o metal é coordenado por duas pterinas equatoriais, um oxigénio ou enxofre apical e um sexto ligando que pode variar (Figura 1.2C). Este pode ser uma serina (no caso da DMSO Redutase⁸), uma cisteína ou selenocisteína (Formato Desidrogenase H⁹), ou um grupo oxo (Arsenito Oxidase¹⁰).^{1,2,11}

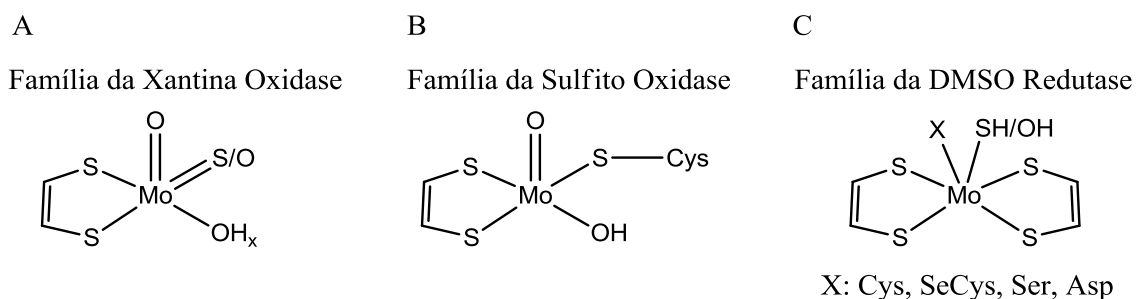


Figura 1.2. Representação esquemática do centro activo para as três principais famílias de enzimas de molibdénio. Adaptado de Romão 2009.²

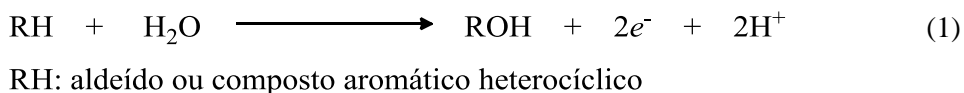
Com algumas excepções, estas enzimas catalisam reacções que envolvem a transferência de um átomo de oxigénio de uma molécula de água para o substrato (ou *vice versa*) acoplada à transferência de dois electrões do substrato para o Mo, sendo que este alterna entre o estados de oxidação VI e IV. Os electrões são depois transferidos do metal para os outros cofactores presentes nestas enzimas, como centros ferro-enxofre, hemos e flavinas, normalmente através dos anéis da pterina.^{1,2}

As molibdoflavoenzimas (MFEs) são um subgrupo homogéneo das molibdoenzimas, caracterizadas por necessitarem não só da molibdopterina mas também de um cofactor flavínico (FAD) para apresentarem actividade catalítica. Estas MFEs estão distribuídas pelas várias

famílias descritas anteriormente e pelos diferentes reinos dos organismos. Estão presentes em bactérias (p.e. xantina desidrogenase de *Rhodobacter capsulatus* – RcXDH – e a aldeído oxidorreductase periplasmática de *Escherichia coli* – PaoABC^{12,13}), fungos (p.e. purina hidroxilase I de *Aspergillus nidulans*¹⁴), plantas (p.e. aldeído oxidases 1, 2 e 3 de *Arabidopsis thaliana* – AtAO1,2,3¹⁵) e animais (p.e. aldeído oxidase – AO^{16,17}) e representam um grupo de proteínas relacionadas bioquimicamente e estruturalmente.¹⁸

1.1.4. A família da Xantina Oxidase

As enzimas da família da xantina oxidase são, entre as proteínas de molibdénio, as mais estudadas e mais caracterizadas. Estas catalisam a hidroxilação oxidativa de um diverso grupo de aldeídos e compostos aromáticos heterocíclicos em reacções que envolvem a quebra de uma ligação C–H e formação de uma ligação C–O (equação 1).²



Esta família compreende membros distribuídos entre os domínios Eucariota, Procariota e *Archaea*. Por norma estas enzimas estão organizadas em homodímeros, sendo que os cofactores podem estar localizados todos numa única subunidade (α_2) ou em subunidades diferentes ($\alpha_2\beta_2\gamma_2$), dependendo da enzima. O cofactor de molibdénio presente não se encontra ligado covalentemente à cadeia principal de aminoácidos, ao contrário do que acontece noutras famílias. Estas apresentam também centros Fe–S e poderão ou não apresentar um cofactor flavínico.²

No caso da xantina desidrogenase (XDH), esta é uma enzima chave no catabolismo de purinas, catalisando a conversão de hipoxantina a xantina e de xantina a ácido úrico. Esta está também presente em diversas situações patológicas, como a isquemia de vários órgãos, processos de inflamação e infecção.¹⁸ Trata-se de um homodímero do tipo α_2 e, para além da piranopterina, apresenta dois centros 2Fe–2S e um cofactor FAD. Esta pode ser convertida à forma de xantina oxidase, enzima que deu o nome à família. No caso destas enzimas, os electrões provenientes da hidroxilação oxidativa do substrato são transferidos para os centros Fe–S, de seguida para o FAD e finalmente para um aceitador final externo – o NAD^+ no caso da XDH e o O_2 no caso da XO (a qual forma espécies reactivas de oxigénio, ROS, como produto lateral desta reacção).¹⁹

A Aldeído Oxidase também é uma enzima de grande interesse, pois está envolvida no metabolismo de diversos aldeídos e na biotransformação de xenobióticos. Ao contrário da XDH/XO, pouco se sabe sobre a reacção catalisada por esta enzima e sobre o seu papel fisiológico. Apesar de serem muito semelhantes, estas enzimas apresentam diferentes

especificidades ao nível de substratos, inibidores e função. No entanto, tal como a XO, esta enzima também produz ROS como um produto lateral da sua actividade.^{18,20,21}

1.1.5. Aldeído Oxidorredutase de *Desulfovibrio gigas*

O objecto de estudo deste trabalho, a Aldeído Oxidorredutase, é uma molibdoenzima da família da Xantina Oxidase, extraída do citoplasma da bactéria gram-negativa e redutora de sulfato *Desulfovibrio gigas* (*D. gigas*) em condições anaeróbias. Com ligeiras diferenças em relação aos restantes membros da família, o seu centro activo possui um Mo coordenado pelo grupo ditioleno de um MCD, dois grupos oxo – um apical e outro equatorial – e um grupo OH/OH₂ lábil equatorial (Figura 1.3). Esta enzima catalisa a hidroxilação oxidativa de aldeídos aos respectivos ácidos carboxílicos.^{22,23}

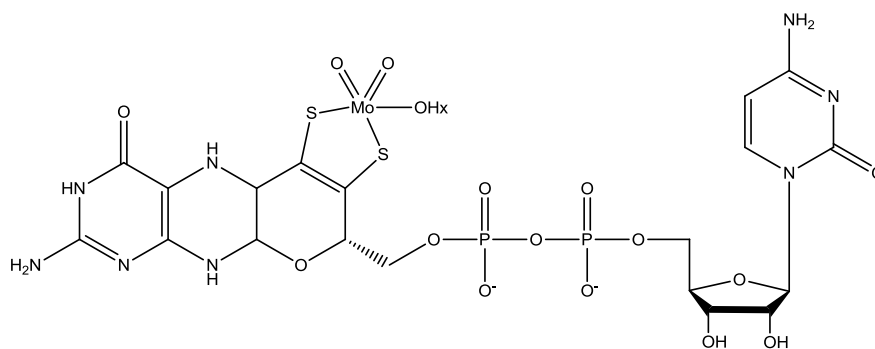


Figura 1.3. Representação esquemática da estrutura do cofactor pterínico (MCD) da *DgAOR*. O ligando OH_x do molibdénio corresponde ao ligando lábil OH/OH₂.

Ainda não se conhece com certeza o papel fisiológico desta enzima em bactérias redutoras de sulfato, sabendo-se apenas que esta hidroxila um grande número de aldeídos com muito baixa especificidade. Devido ao facto do gliceraldeído ser um dos seus substratos, foi proposto que a AOR possa estar envolvida na degradação da poliglucose, a qual é acumulada por *D. gigas* em condições normais de crescimento.²³⁻²⁵

A *DgAOR* é uma proteína de 907 aminoácidos que apresenta uma massa molecular de aproximadamente 90 kDa e em solução existe na forma de homodímero com uma massa molecular de aproximadamente 210 kDa. A sua sequência encontra-se em anexo (Anexo 7.1).^{26,27} Cada monómero alberga um átomo de molibdénio, assim como quatro átomos de ferro, correspondentes a dois centros 2Fe-2S distinguíveis espectroscopicamente.²⁸⁻³⁰

Os primeiros cristais de *DgAOR* foram obtidos em 1993 por Romão *et al.*,²⁶ apresentando a forma de uma bipirâmide hexagonal pertencente ao grupo espacial *P6₁22*. A sua estrutura tridimensional foi resolvida pelo método de substituição isomórfica múltipla (MIR, do inglês *multiple isomorphous replacement*), refinada até uma resolução de 2,25 Å e publicada em

1995.³¹ Esta foi a primeira estrutura cristalográfica de uma proteína de molibdénio. Mais tarde, a qualidade dos cristais foi aperfeiçoada e a resolução foi melhorada até 1,28 Å utilizando radiação de Sincrotrão.³²

1.1.5.1. Descrição da estrutura

Na unidade assimétrica existe apenas uma molécula, constituída por uma cadeia polipeptídica única. A proteína é aproximadamente globular, com um diâmetro aproximado de 75 Å. A sua estrutura secundária apresenta 28% de hélices alfa e 21% de folhas beta. Uma representação da sua estrutura tridimensional é apresentada na Figura 1.4. A DgAOR apresenta quatro domínios discretos (Fe–S_{distal}, Fe–S_{proximal}, Mo1 e Mo2). O primeiro (resíduos 1 até 76) liga o centro ferro-enxofre distal, ou seja, o que se encontra à superfície da proteína e mais afastado do Mo. O segundo (resíduos 84 até 157) liga o centro ferro-enxofre proximal que se encontra mais próximo do Mo. Os domínios Mo1 (resíduos 196 até 581) e Mo2 (resíduos 582 até 907) estabilizam o cofactor pterínico na sua interface através de interações do tipo ligação de hidrogénio. O domínio Mo1 contribui com dois segmentos que ligam a pterina, enquanto o domínio Mo2 liga o outro lado da pterina e estabelece todas as interações que ligam o dinucleótido. Estes dois últimos domínios não só determinam a vizinhança química do centro activo, mas também formam um canal hidrofóbico com cerca de 15 Å de profundidade na sua interface, largo à superfície da proteína e estreito no interior, que conduz as moléculas de substrato em solução até ao centro activo.^{31,32}

O domínio Fe–S_{proximal} está ligado ao domínio Mo1 através de um segmento longo com estrutura secundária irregular que atravessa uma região concava da proteína. Esta região aparenta ser o sítio mais provável para a interacção com proteínas aceitadoras de electrões, como a flavodoxina que aceita electrões da DgAOR/aldeídos *in vitro*.^{23,24}

1.1.5.2. Cofactores e centro activo

Os centros 2Fe-2S encontram-se planares, estando cada átomo de ferro coordenado tetraedricamente por quatro átomos de enxofre, dois do próprio cofactor e dois das cadeias laterais de duas cisteínas. No centro Fe–S_{distal} o Fe1 encontra-se coordenado pelas cisteínas Cys40 e Cys45 e o Fe2 pelas Cys48 e Cys60, enquanto que no centro Fe–S_{proximal} o Fe1 está coordenado pelas Cys100 e Cys139 e o Fe2 pelas Cys103 e Cys137. Cada um dos centros forma um plano perpendicular com o plano formado pelas quatro cisteínas que o coordenam.³¹

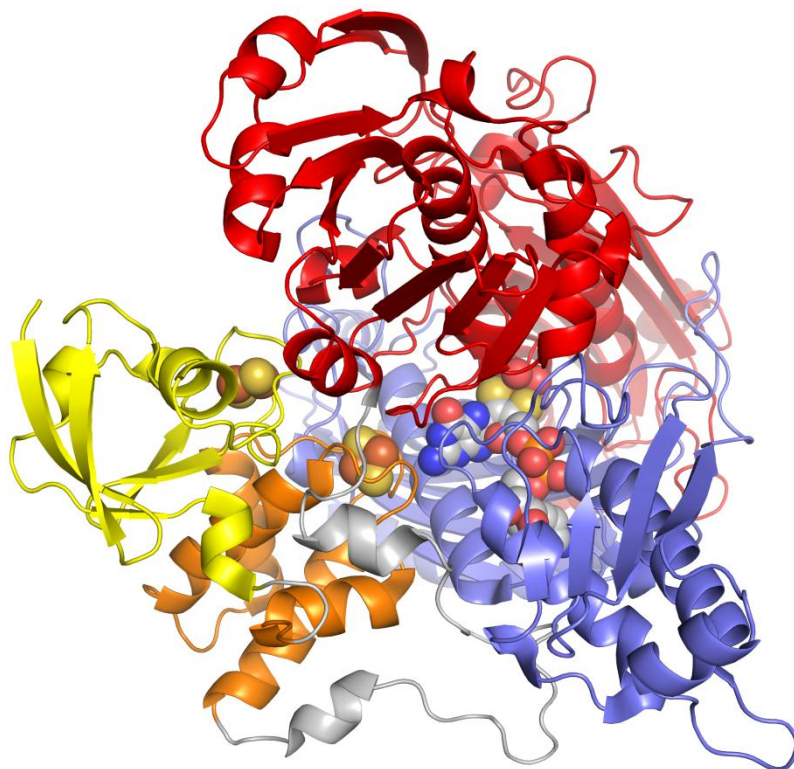


Figura 1.4. Representação da estrutura tridimensional da *DgAOR* (Código PDB: 1VLB³²). Os seus quatro domínios estão representados a amarelo (Fe-S_{distal}), laranja (Fe-S_{proximal}), vermelho (Mo1) e azul (Mo2). A cinzento estão representados os fragmentos de ligação entre os domínios. As esferas representam os cofactores: piranopterina e centros Fe-S, com o seguinte código de cores: Carbono – cinzento; Azoto – azul; Oxigênio – vermelho; Fósforo – laranja; Enxofre – amarelo; Ferro – castanho; Molibdênio – azul-esverdeado. Esta figura foi preparada utilizando o programa PyMOL.

O cofactor pterínico trata-se de um MCD situado na interface dos domínios Mo1 e Mo2. A estrutura deste cofactor só foi conseguida com a resolução da primeira estrutura tridimensional da *DgAOR*.³¹ Este apresenta um anel de pterina fundido com um anel de pirano com dois grupos tiolato a partir dos quais este se liga ao molibdênio, assim como uma citidina difosfato ligada ao anel de pirano (Figura 1.1 – R = citidina monofosfato). Na sua forma oxidada, o molibdênio está coordenado pelo grupo ditioleno da pterina, dois grupos oxo e um grupo lábil OH/OH₂, formando uma pirâmide quadrangular, de acordo com a Figura 1.5. A estrutura a 1,28 Å permitiu esclarecer com maior rigor a esfera de coordenação do molibdênio e respectivas distâncias entre o metal e os ligandos, através da análise da densidade electrónica e das distâncias entre os ligandos e o molibdênio.³² Enquanto os oxigénios dos grupos oxo se encontram a cerca de 1,6-1,8 Å do Mo, o ligando lábil OH/OH₂ encontra-se a cerca de 2-2,2 Å de distância, o que corresponde a uma ligação simples.

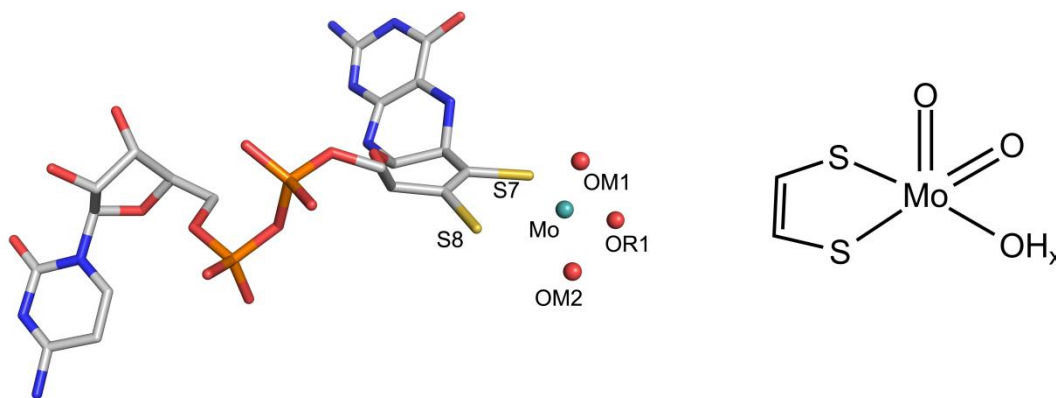


Figura 1.5. Representações tridimensional (à esquerda) e esquemática (à direita) do cofactor pterínico. S7 e S8 – grupo ditioleno da pterina; OM1 – ligando oxo apical; OR1 – ligando oxo equatorial; OM2/OH_x – ligando lábil OH/OH₂. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Fósforo – laranja; Enxofre – amarelo; Molibdénio – azul-esverdeado. O programa PyMOL foi utilizado na preparação desta figura.

Na tentativa de estabelecer qual o ligando essencial para a actividade catalítica da *DgAOR*, Huber *et al.*³³ aplicou o protocolo de ressulfuração descrito anteriormente para a XO, após o qual há uma recuperação de actividade por parte desta enzima.³⁴ Huber incubou a *DgAOR* com ião sulfureto e sulfito de sódio e cristalizou a proteína. Na estrutura resultante foi possível observar uma substituição do grupo oxo apical por um átomo de enxofre, indicando que este poderá ser essencial para a actividade da enzima.³³ No entanto, estudos cinéticos efectuados mais tarde por Santos-Silva *et al.* revelaram que a enzima apresenta actividade na sua forma não sulfurada, provando que, ao contrário da XO, esta não necessita do enxofre para ter actividade.³⁵

Muito perto do centro activo, a cerca de 3,5 Å do átomo de molibdénio, encontra-se um glutamato (Glu869), o qual é conservado nas enzimas da família da XO.^{27,31} Este pensa-se estar envolvido no mecanismo de reacção das enzimas desta família. Entre todos os ligandos do molibdénio, o que se encontra mais acessível é o OH/OH₂, que se encontra no final do canal hidrofóbico pelo qual acedem os substratos. Na vizinhança do Glu869 e deste ligando também se encontra uma cadeia de três moléculas de água conservadas, as quais participam na reacção.³¹

Existem vários mecanismos propostos para a reacção catalisada pela *DgAOR*. O mecanismo mais aceite consiste num ataque nucleofílico ao carbono da função aldeído do substrato pelo grupo lábil OH/OH₂ presente no centro activo (Figura 1.6). O glutamato 869 presente na vizinhança actua como uma base para o protão do grupo lábil, aumentando o seu carácter nucleofílico. Este ataca o carbono do grupo carbonilo do substrato originando o ácido carboxílico, o qual se mantém ligado ao molibdénio. Ao mesmo tempo dá-se a transferência do hidreto do aldeído para o ligando oxo/sulfureto equatorial do molibdénio. De seguida, o produto é libertado, o que se pensa ser facilitado pela ligação transiente do Glu869 ao molibdénio para manter a sua esfera de coordenação. O centro activo é finalmente regenerado pela ligação de

uma molécula de água, da cadeia de águas conservadas que é possível observar na estrutura cristalográfica da enzima.^{23,35}

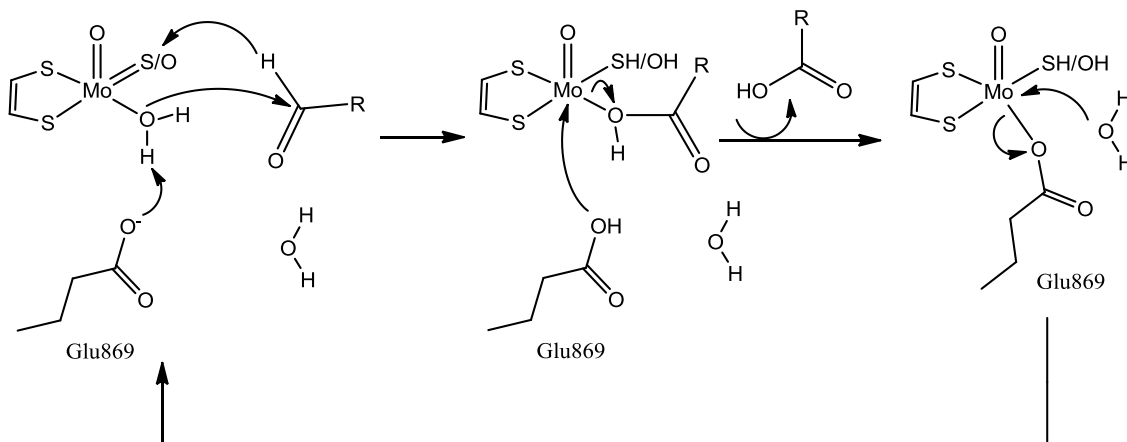


Figura 1.6. Representação esquemática do mecanismo de reação proposto para a DgAOR.^{23,35}

1.1.5.3. Cadeia de transferência electrónica

Após a oxidação dos substratos ao nível do molibdénio, este é reduzido passando do estado de oxidação VI para o estado de oxidação IV. O centro Fe-S_{proximal} encontra-se a 15 Å de distância de átomo de molibdénio e a 5.2 Å do grupo amina exocíclico da parte pterínica do cofactor. Este grupo também se encontra à distância de ligação de hidrogénio com a cisteína 139, a qual coordena um dos átomos de Fe do centro Fe-S_{proximal}. Os electrões resultantes da reacção são transferidos do molibdénio para este centro através do sistema parcialmente conjugado dos anéis da pterina. A 12.2 Å de distância deste centro, encontra-se o centro Fe-S_{distal}, o qual está mais exposto à superfície da proteína. Os electrões são transferidos para este centro mediados pelos aminoácidos da proteína. Neste ponto os electrões estão disponíveis para serem transferidos para um aceitador electrónico fisiológico desconhecido, no caso da *DgAOR*. Em MFEs como a XO os electrões são posteriormente transferidos para o cofactor flavínico. Devido à grande semelhança entre a *DgAOR* e estas proteínas, pensa-se que o aceitador final dos electrões poderá ser uma flavoproteína.^{23,24,31}

1.1.5.4. Inibição do centro activo

Nas estruturas cristalográficas da *DgAOR*^{31,32} encontra-se uma molécula de isopropanol muito próxima do seu centro activo. Esta molécula é um conhecido inibidor da enzima, assim como a maioria dos álcoois,²⁴ e encontra-se nestas estruturas pois faz parte da solução de cristalização utilizada para a obtenção destes cristais. Utilizando um tampão de estabilização sem isopropanol, é possível retirar esta molécula do centro activo.^{35,36} O local de ligação deste inibidor foi proposto como sendo o local de ligação dos substratos ao centro activo. Na posição

do grupo hidroxilo do álcool é esperado que esteja o grupo carbonilo do aldeído, em contacto próximo com o grupo lábil OH/OH₂ do centro activo. Os grupos metilo do isopropanol estabelecem contactos hidrofóbicos com os aminoácidos Phe425, Tyr535, Leu497 e Leu626, localizados no final do canal de entrada para o centro activo. Aqui é esperado que seja estabilizada a parte hidrofóbica dos aldeídos catalisados pela *DgAOR*.³³

Para além da estrutura com isopropanol, vários estudos foram feitos sobre a interacção da *DgAOR* com vários inibidores. Em 2004 foi obtida a estrutura cristalográfica desta enzima inibida com arsenito.³⁶ Este encontra-se ligado ao átomo de molibdénio através de um dos átomos de oxigénio, substituindo o ligando lábil do centro activo (Figura 1.7A). Estes resultados foram confirmados mais tarde por cristalografia de raios-X e espectroscopia de Ressonância Paramagnética Electrónica (RPE).³⁷ Ensaio cinéticos da proteína inibida com arsenito seguidos de um passo de diálise revelaram que esta se trata de uma inibição irreversível.³⁵

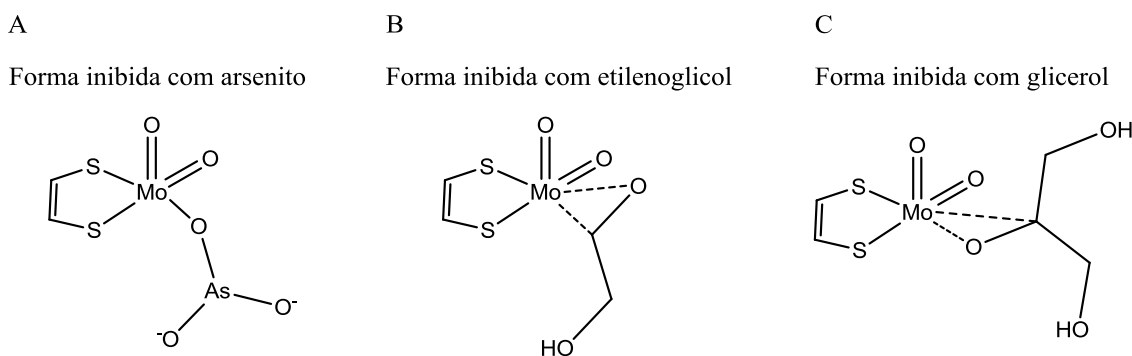


Figura 1.7. Representação esquemática do centro activo da *DgAOR* inibido com três compostos: A – arsenito; B – etilenoglicol; C – glicerol.^{35,36}

Estruturas cristalográficas da *DgAOR* inibida com etilenoglicol e glicerol também foram obtidas por Santos-Silva *et al.*³⁵ Estas moléculas aparecem numa posição semelhante à do arsenito nas estruturas supracitadas, substituindo o ligando lábil OH/OH₂ e coordenando directamente o molibdénio. No caso da estrutura com etilenoglicol, este encontra-se coordenado ao molibdénio por um dos átomos de carbono a uma distância de 2,36 Å (Figura 1.7B). Um dos grupos hidroxilo está à distância de ligação de hidrogénio do grupo carboxílico do Glu869 conservado que se encontra perto do centro activo. Em relação à estrutura com glicerol observa-se uma coordenação ao molibdénio pelo grupo hidroxilo que se encontra ligado ao carbono C2 (Figura 1.7C). Tal como o etilenoglicol, este também se encontra estabilizado na sua posição através de uma ligação de hidrogénio com o grupo hidroxilo do Glu869. Apesar de ambos estes álcoois se encontrarem na primeira esfera de coordenação do molibdénio, estudos cinéticos revelaram que, tal como o isopropanol, estes são inibidores competitivos e reversíveis desta enzima.³⁵

Este grupo de enzimas representa cerca de 2% do proteoma humano, com mais de 550 proteínas em aproximadamente 30 000 no total e são agrupadas em seis classes diferentes, consoante a composição dos seus centros activos: proteases de serina (p.e. quimotripsina, tripsina e subtilisina), proteases de treonina (p.e. subunidades catalíticas do proteassoma), proteases de cisteína (p.e. caspases e papaína), proteases de ácido aspártico (p.e. pepsina e protease do HIV), proteases de ácido glutâmico (p.e. aspergillopepsina II) e metaloproteases (p.e. colagenases e aminopeptidases).^{39,43,44}

1.2.2. Proteases de serina

As proteases de serina constituem cerca de um terço das proteases. Podem ser encontradas em vírus, bactérias e eucariotas e a origem do seu nome deve-se à presença de três resíduos Asp–His–Ser no seu centro activo. Entre os três resíduos, a serina é o agente nucleófilo que apresenta um papel muito importante na actividade catalítica. Estes resíduos são conhecidos como a “tríade catalítica” das proteases de serina. Contemplando as enzimas desta classe, é possível identificar diferentes motivos estruturais globais que permitem agrupá-las em clãs com caminhos evolutivos distintos. Consoante o autor e a data de publicação, diferentes números de clãs foram identificados, sendo que na base de dados MEROPS (<http://merops.sanger.ac.uk/>) estas encontram-se actualmente distribuídas entre quinze clãs. Os clãs representados em maior quantidade são caracterizados pelas famílias das enzimas quimotripsina, subtilisina, carboxipeptidase Y e peptidase Clp, as quais utilizam a tríade catalítica Asp–His–Ser no centro activo. O facto de quatro grandes famílias com caminhos evolutivos diferentes apresentarem a mesma maquinaria catalítica demonstra a sua eficácia.^{38,45-47}

As proteases da família da quimotripsina (família S01) são as mais numerosas e possivelmente as mais estudadas. Actualmente (18 de Setembro de 2012) existem 3203 estruturas de proteases de serina depositadas no *Protein Data Bank* (PDB – <http://www.rcsb.org/>), sendo que os membros mais representados pertencem à família S01 (trombina com 700 estruturas, tripsina com 510, factor de coagulação Xa com 251 e quimotripsina com 146). As proteínas desta família estão envolvidas em processos biológicos como a digestão, a transdução de sinal, a apoptose, a reprodução e a resposta imunitária. Para além disso, cascatas de activação sequencial de diferentes proteases de serina são responsáveis pela coagulação do sangue e fibrinólise.⁴⁵

1.2.3. Tripsina de *Bos taurus*

Neste trabalho é estudada a interacção da tripsina de pâncreas bovino (*Bos taurus*) com vários compostos. A tripsina é uma protease de serina pertencente à família S01 e é um homólogo da

quimotripsina, com uma estrutura global idêntica e cerca de 40% de homologia de sequência. Ambas as enzimas – assim como a elastase, uma protease da mesma família – são produzidas no pâncreas, tratando-se de enzimas digestivas.³⁸ No entanto, as três enzimas apresentam especificidades diferentes para os seus substratos. Enquanto a quimotripsina é específica para resíduos aromáticos, a tripsina cliva as proteínas do lado do carbono carbonilo de resíduos de arginina e lisina, os quais apresentam uma cadeia lateral longa e carregada positivamente. Por sua vez, a elastase é específica para resíduos neutros e com cadeias laterais curtas como a alanina, glicina e valina.⁴⁸⁻⁵⁰ No centro activo esta apresenta a tríade catalítica Ser195, His57 e Asp102, de acordo com o sistema de numeração da quimotripsina, o qual é utilizado como referência para comparar as proteínas desta família.^{51,52} A sequência peptídica da tripsina encontra-se em anexo (Anexo 7.2).

A tripsina tem um peso molecular de 23,3 kDa e é constituída por uma cadeia única de 223 resíduos. Esta é inicialmente produzida no pâncreas na forma de um zimogénio, ou seja, um precursor inactivo da enzima. O tripsinogénio é excretado para o duodeno, onde é clivado na ligação Lys15–Ile16 pela enteropeptidase, uma protease de serina segregada pela mucosa intestinal com especificidade muito semelhante à tripsina. Assim que algumas moléculas de tripsinogénio são clivadas, produz-se tripsina activa que, por sua vez, assume a função de clivar as restantes moléculas de tripsinogénio, considerando-se por isso como uma enzima autocatalítica. Por sua vez a tripsina também é responsável pela activação dos zimogénios das outras enzimas pancreáticas digestivas. A expressão de tripsina é regulada desta forma de modo a evitar a degradação dos tecidos onde esta é produzida. De facto, a pancreatite aguda é consequência da activação prematura das enzimas digestivas produzidas no pâncreas.^{46,50-54}

Duas isoformas aniónicas do tripsinogénio foram isoladas a partir de suco pancreático em 1969 por Figarella *et al.*⁵⁵ e foram denominadas por tripsinogénio 1 e 2. A isoforma predominante é o tripsinogénio 1 que se encontra presente numa quantidade cerca de 2 vezes superior. Devido a diferentes nomenclaturas utilizadas na literatura, vai ser utilizada a nomenclatura definida por Rinderknecht *et al.*,⁵⁶ onde o tripsinogénio 1 é considerado catiónico (por ser o menos aniónico dos dois) e o tripsinogénio 2 é considerado aniónico. Comparando a forma activa dos dois tripsinogénios, a tripsina catiónica é autoactivada mais rapidamente e é mais resistente a autólise, sendo por isso predominante. Mais tarde foi também descoberto uma terceira variante denominada por mesotripsinogénio, o qual é expresso em muito baixas quantidades. Este é resistente aos inibidores naturais da tripsina produzidos pelo próprio organismo, podendo ter a função de os degradar assim como activar os tripsinogénios mesmo na presença de inibidor em excesso. Caso seja activada prematuramente no pâncreas esta aumenta o risco de pancreatite.^{51,56,57}

A tripsina catiónica é a mais abundante *in vivo*, sendo por isso o objecto de estudo deste trabalho. A sua primeira estrutura cristalográfica foi resolvida por Stroud *et al.* em 1971, pouco tempo depois de Blow ter resolvido a estrutura da quimotripsina, a primeira das proteases de serina.^{58,59} Os cristais de tripsina podem apresentar várias formas cristalinas e grupos espaciais, sendo que a forma ortorrômbica com um grupo espacial $P2_12_12_1$ é a mais comum.

1.2.3.1. Descrição da estrutura

A tripsina é constituída por uma cadeia peptídica única e a sua estrutura terciária é mantida por seis pontes de dissulfureto. Esta enzima apresenta uma estrutura característica das enzimas da família S01 com uma forma globular e a presença de dois barris- β constituídos por folhas- β antiparalelas, estando o centro activo localizado entre eles. No interior destes barris encontram-se aminoácidos hidrofóbicos. É possível observar a sua estrutura tridimensional na Figura 1.9.^{45,50,60}

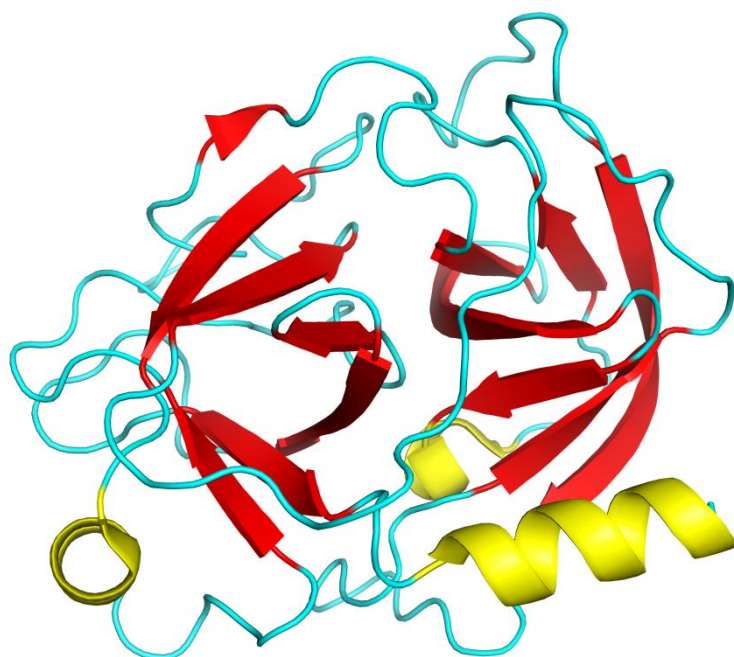


Figura 1.9. Representação da estrutura tridimensional da tripsina de *Bos taurus* (Código PDB: 1S0Q). A amarelo estão representados os motivos de hélice- α da sua estrutura secundária, a vermelho os motivos de folha- β e a ciano os *loops*. Esta figura foi preparada utilizando o programa PyMOL.

No domínio catalítico encontram-se os três resíduos que constituem a tríade catalítica, Ser195, His57 e Asp102. Nesta região a enzima apresenta uma fenda na qual se ligam os substratos, sendo que a serina catalítica, Ser195, se encontra de um lado e a His57 e o Asp102 do outro, fazendo uma rede de ligações de hidrogénio entre si e com alguns resíduos vizinhos (Figura 1.10). Aqui, a His57 assume um tautómero menos favorável, mas essencial para o mecanismo da reacção catalítica. Ambos os átomos de azoto da His57 formam ligações de hidrogénio, um

com a Ser195 e outro com o Asp102. Pensa-se que as ligações de hidrogénio entre o Asp102 e a Ser214 e os grupos amina das cadeias principais da Ala56 e His57 estejam relacionadas com o posicionamento correcto deste resíduo no centro activo. Observa-se também uma ligação de hidrogénio entre um C–H do anel da His57 e o oxigénio carboxilo da cadeia principal da Ser214, a qual poderá desempenhar um papel de distribuição da carga positiva que de outro modo ficaria acumulada no anel imidazólico.⁶¹ Esta rede de ligações de hidrogénio é encontrada não só nas enzimas desta família mas também noutras famílias, como a da subtilisina, o que evidencia a sua importância. Perto da Ser195 encontra-se o *oxyanion hole*, o qual é constituído pelos grupos amina das cadeias principais da Gly193 e da Ser195 e tem como função a formação de um local carregado positivamente que vai estabilizar o oxianião negativo do intermediário tetraédrico da reacção.⁴⁵

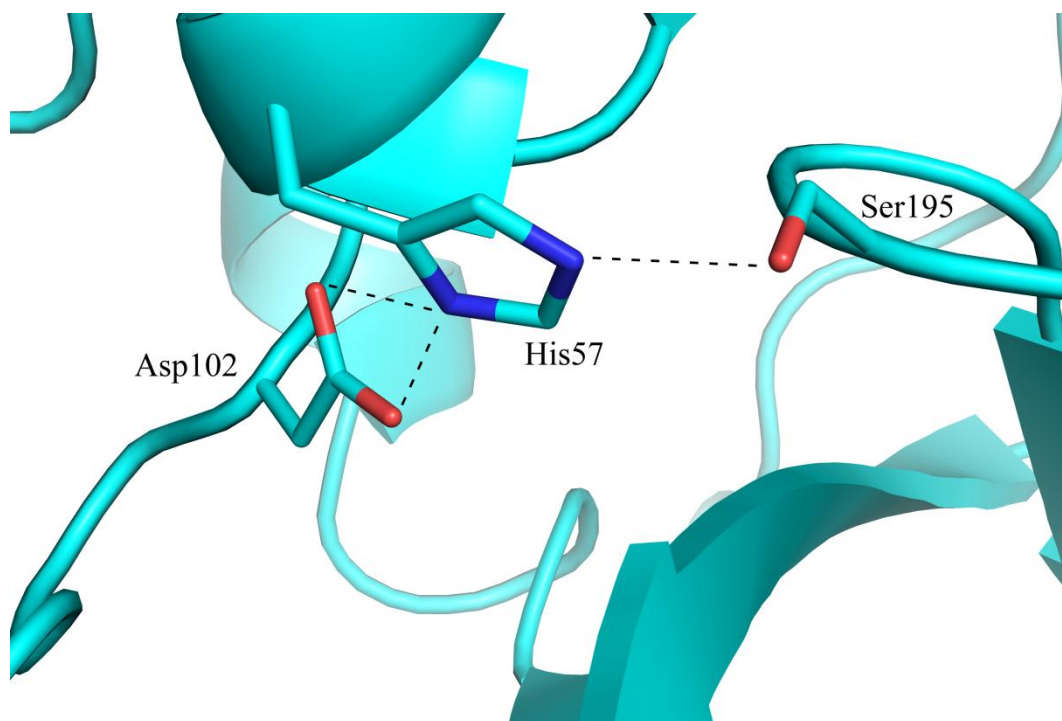


Figura 1.10. Estrutura tridimensional do centro activo da tripsina, evidenciando a tríade catalítica (Serina 195, Histidina 57 e Aspartato 102). A tracejado estão representadas as ligações de hidrogénio entre as cadeias laterais deste resíduos. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho. Esta figura foi preparada utilizando o programa PyMOL.

Na região adjacente ao centro activo encontra-se o domínio de reconhecimento e ligação do substrato. O resíduo com maior importância na ligação de substratos é o Asp189, o qual se localiza no fundo do sítio de ligação do substrato. Este estabiliza, no caso da tripsina, as cadeias laterais dos resíduos no local de corte, a lisina e a arginina, ambas carregadas positivamente. No caso da quimotripsina este resíduo é substituído por uma serina de modo a acomodar as cadeias mais volumosas dos aminoácidos aromáticos. No caso da elastase, as glicinas 216 e 226 são substituídas por uma treonina e uma valina que bloqueiam este local de ligação para

aminoácidos com cadeias laterais longas (Figura 1.11). No entanto, estudos de mutagénese foram realizados onde o Asp189 foi substituído por uma serina verificando-se que a enzima não ganhou a especificidade da quimotripsina, mas sim tornou-se numa protease não específica. Apenas mutando os resíduos de dois *loops* superficiais (resíduos 185-188 e 221-225) para os correspondentes da quimotripsina se conseguiu converter a especificidade da enzima para a da quimotripsina, o que indica que há mais resíduos envolvidos no reconhecimento e posicionamento do substrato.⁵⁰ Outras interações por ligações de hidrogénio foram estudadas entre os resíduos da tripsina e os aminoácidos adjacentes à ligação cindível do substrato, as quais não serão aqui detalhadas.⁴⁵

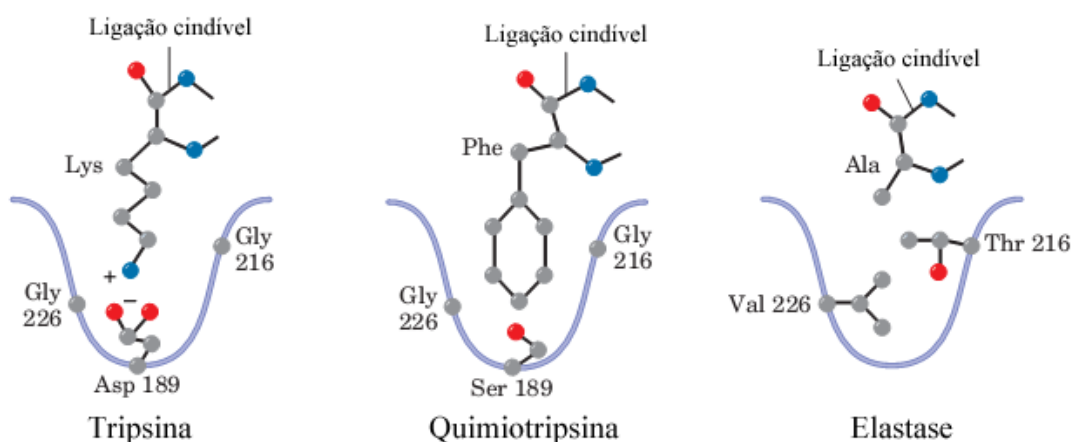


Figura 1.11. Representação esquemática do domínio de reconhecimento do substrato para as enzimas tripsina (à esquerda), quimotripsina (ao meio) e elastase (à direita). Adaptado de Voet *et al.* 2008.⁵⁰

Outro domínio importante nestas enzimas é o domínio de activação do zimogénio. Este foi descrito inicialmente por Huber por comparação das estruturas cristalográficas da tripsina e do tripsinogénio. Observando as alterações conformacionais entre as duas estruturas este domínio foi definido pelos resíduos 16-19, 142-152, 184-193 e 216-223. Esta região inclui o *oxyanion hole* e algumas partes do local de ligação do substrato. Por este motivo e por se encontrar deformada no zimogénio este não apresenta actividade catalítica. Quando o péptido de activação é clivado, a Ile16 é libertada e forma uma ponte salina com o Asp194, assim como várias ligações de hidrogénio com outros resíduos do domínio de activação, provocando uma alteração conformacional desta região e ordenando a região do centro activo.^{50,60}

1.2.3.2. Mecanismo da reacção

De modo a hidrolisar uma ligação peptídica é necessário quebrar uma ligação amida, a qual é muito estável, devido ao seu carácter parcial de ligação dupla, obtido pela doação electrónica do azoto da amida para o grupo carbonilo (Figura 1.12).³⁸

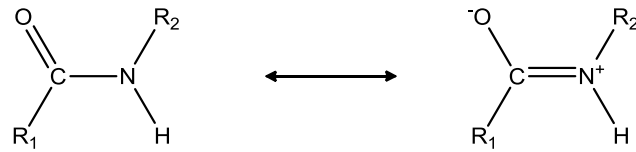


Figura 1.12. Representação esquemática do carácter parcial de ligação dupla de uma ligação peptídica.

A tríade catalítica e as suas interações fazem com que a Ser195 seja altamente reactiva. A His57 posiciona a serina catalítica e polariza o seu grupo hidroxilo, actuando como uma base e tornando a serina num nucleófilo muito forte. Deste modo, a serina ataca o carbono do carbonilo do substrato, formando um intermediário tetragonal, estabilizado pelos grupos amina dos resíduos que formam o *oxyanion hole*. A His57 protonada é estabilizada pela ligação de hidrogénio ao Asp102. O ataque da serina enfraquece a ligação amida do substrato e a His57-H⁺ actua como um ácido que doa um protão ao grupo amina de saída, ocorrendo assim uma reacção de eliminação e formação da acil-enzima. De seguida, a His57 actua novamente como uma base tornando uma molécula de água num nucleófilo que ataca o carbono carbonilo que se encontra ligado à Ser195, formando-se um novo intermediário tetragonal. Este colapsa novamente assistido pela His57 que doa um protão à Ser195 ocorrendo expulsão do ácido carboxílico. Estes passos encontram-se resumidos na Figura 1.13.^{38,45,50,62}

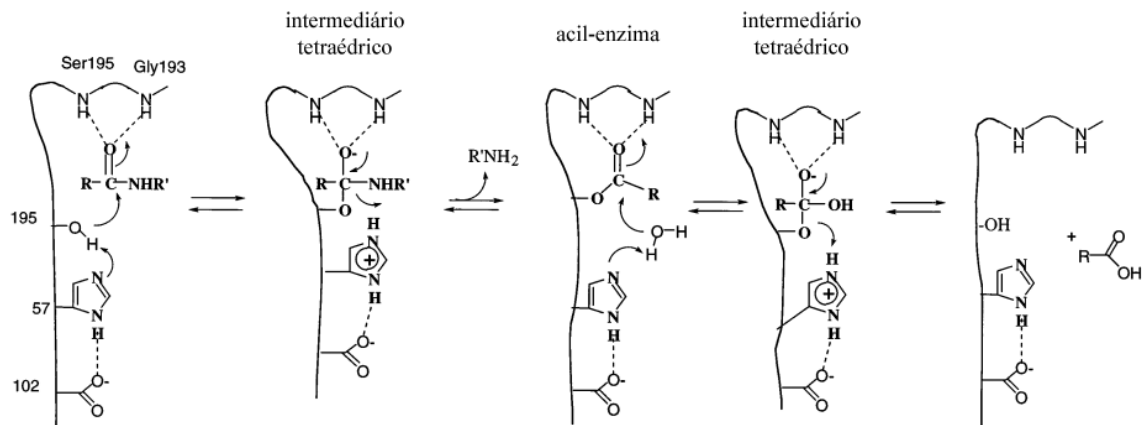


Figura 1.13. Representação esquemática do mecanismo de reacção da tripsina. Adaptado de Hedstrom 2002.⁴⁵

1.2.3.3. Inibidores da tripsina

Como descrito anteriormente, as proteases de serina desta família apresentam papéis importantes em vários processos biológicos e, conseqüentemente, em várias patologias. Por este motivo a sua regulação tem sido muito aprofundada, existindo muitos estudos farmacológicos a este nível.

O organismo regula estas enzimas através da utilização de inibidores complexos produzidos por diversos tecidos. O melhor exemplo destes inibidores é uma classe de proteínas denominadas

serpinas (do inglês *Serpin* – *Serine protease inhibitor*). A sua função principal é controlar a proteólise nos diversos tecidos e caminhos metabólicos, mas podem também ter outras funções, servindo como chaperonas, transportadoras de hormonas ou factores antiangiogénicos.⁶³ Estas ligam-se ao centro activo das proteases, pois possuem uma estrutura semelhante aos substratos respectivos de cada enzima, o que leva a uma clivagem da serpina por parte da protease. Este corte poderá levar à formação de uma ligação covalente entre a protease e a serpina ou a alterações conformacionais do complexo, inactivando a protease.⁶⁴ Um exemplo muito estudado é a aprotinina, ou inibidor básico pancreático de tripsina (BPTI, do inglês *basic pancreatic trypsin inhibitor*), o qual apresenta um mecanismo de chave-fechadura, ligando-se fortemente à tripsina através de uma extensa rede de ligações de hidrogénio. Esta proteína é encontrada no pâncreas e tem a função de inibir várias proteases de modo a que estas não exerçam a sua função até serem excretadas. O estudo deste complexo por cristalografia de raios-X permitiu obter informações preciosas sobre o mecanismo de acção destas enzimas.^{50,60}

Quando há uma falha no controlo de proteases por parte do organismo, estas podem ter um efeito devastador nos tecidos. Numerosos estudos cristalográficos foram efectuados com proteases de serina e moléculas orgânicas inibidoras, sendo as proteínas mais estudadas a trombina e a tripsina, devido à sua importância fisiológica e facilidade na cristalização. Como as enzimas dos vários clãs têm centros activos semelhantes, muitas vezes utilizam-se proteínas modelo como a tripsina de modo a estudar a orientação e interacções de moléculas inibidoras antes de se cristalizar a proteína em estudo.

Uma família de inibidores muito estudada para estas proteases é composta por ácidos borónicos. Estes podem ser inibidores muito potentes, dependendo da sua estrutura. Foi demonstrado por ¹⁵N-RMN que compostos análogos aos substratos desta família de enzimas ligam-se covalentemente à serina catalítica pelo átomo de boro, estabelecendo ligações de hidrogénio entre um dos átomos de azoto da His57 e o protão de um dos grupos hidroxilo do inibidor (Figura 1.14).⁶⁵⁻⁶⁷

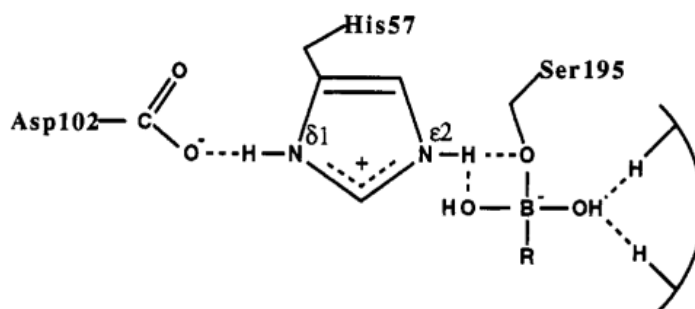


Figura 1.14. Representação esquemática da interação entre os ácidos borónicos e a tríade catalítica das enzimas da família da tripsina. Adaptado de Tsilikounas *et al* 1992.⁶⁶

Relativamente à porção das moléculas que lhes confere especificidade, estudos foram feitos onde se determinou que moléculas do tipo piridinil guanidina são inibidores potentes da urocinase, uma proteína semelhante à tripsina de elevada importância em processos cancerígenos.⁶⁸⁻⁷⁰ Estudos cristalográficos foram realizados onde se analisa a interacção de vários compostos semelhantes, observando-se a interacção dos seus grupos amina com o Asp189 da enzima, salientando a sua importância no reconhecimento de substratos específicos.⁷¹ A adição de um grupo sulfonamida a estes compostos mostrou um aumento na sua potência inibitória e selectividade para a urocinase.⁷²

1.2.4. Urocinase

Devido à necessidade de se inibir especificamente uma determinada protease, é muitas vezes necessário testar a especificidade dos compostos em estudo com outras proteases de serina, de modo a assegurar que apenas a protease alvo é inibida. Neste trabalho são testados vários compostos, desenhados para inibir a urocinase, em complexo com a tripsina.

A urocinase, também conhecida por Activador do Plasminogénio tipo urocinase (uPA, do inglês *urokinase-type Plasminogen Activator*), é uma enzima que em conjunto com o seu receptor (uPAR) tem a função de clivar o zimogénio da protease plasmina. Esta é uma protease agressiva que degrada muitas proteínas do plasma sanguíneo e está envolvida na fibrinólise. Níveis elevados de expressão da urocinase estão associados a processos como o crescimento tumoral, migração celular, metástase e angiogénese e, consequentemente, a doenças como o cancro, a esclerose múltipla e a artrite. Por este motivo esta enzima é um alvo muito atractivo neste campo de investigação.⁷³⁻⁷⁶ A cristalografia de proteínas pode ser uma mais-valia para a determinação da estrutura de complexos enzima-inibidor e esclarecer o seu modo de acção, bem como permitir o desenvolvimento de novos compostos, mais específicos e eficazes.

1.3. Cristalografia de Raios-X

A cristalografia de raios-X permite a obtenção de uma estrutura tridimensional detalhada de uma macromolécula através da interpretação da difracção de raios-X a partir de muitas moléculas idênticas ordenadas num cristal. Utiliza-se este tipo de radiação, pois apresenta um comprimento de onda compreendido entre os 0.1 e os 100 Å no espectro electromagnético. Esta gama de comprimentos de onda compreende também o comprimento da ligação atómica, podendo por este motivo ser estudada a estrutura de moléculas a nível atómico.⁷⁷ Apesar de existirem outras técnicas como a Ressonância Magnética Nuclear (RMN) e a Crio-Microscopia Electrónica (Crio-EM) que também o permitem, cada uma com as suas vantagens e

desvantagens, a cristalografia de raios-X é considerada a técnica de excelência para a determinação de estruturas tridimensionais de macromoléculas e é sem dúvida a mais utilizada. A 11 de Setembro de 2012 existem 74 250 estruturas cristalográficas de macromoléculas depositadas na base de dados PDB, num total de 84 508 estruturas, correspondendo a cerca de 89% das estruturas depositadas (<http://www.rcsb.org/>).

Os raios-X foram inicialmente descobertos em 1895 por Wilhelm Röntgen, no entanto a sua natureza (partículas ou radiação electromagnética) foi objecto de numerosos debates até 1912 quando Max von Laue postulou que os átomos numa matriz cristalina têm uma estrutura regular e periódica, com distâncias interatómicas na ordem de 1 Å, e que um cristal poderia ser usado para difractar raios-X. Laue, em conjunto com Friedrich e Knipping fotografaram pela primeira vez o padrão de difracção de um cristal de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ e provaram assim que: (1) os raios-X são radiação electromagnética, (2) a rede atómica de um cristal é periódica e (3) a distância interatómica de um cristal é da mesma ordem de grandeza dos raios-X, podendo estes ser utilizados para estudar moléculas a nível atómico. Deste modo nasceu a cristalografia de raios-X. Após esta descoberta, W. H. Bragg e o seu filho W. L. Bragg fizeram numerosas experiências de modo a estudar as características da radiação de raios-X, assim como as características das redes cristalinas. Destes estudos originaram as primeiras estruturas de moléculas pequenas (NaCl, KCl, KBr e KI) por W. L. Bragg. Nos anos seguintes, muitas estruturas de sais e minerais foram resolvidas utilizando este método. A aplicação de raios-X a moléculas biológicas começou com Dorothy Hodgkin que resolveu as estruturas do colesterol (1937), da vitamina B12 (1945) e da penicilina (1954). Em 1958 foi resolvida a primeira estrutura de uma proteína, a mioglobina de cachalote por John Kendrew e Max Perutz.⁷⁹

Nas últimas duas décadas a quantidade de estruturas resolvidas por cristalografia de raios-X aumentou exponencialmente, devido ao desenvolvimento de algoritmos e *software* informático que vieram automatizar a resolução de estruturas, assim como a utilização de radiação de Sincrotrão que veio permitir a obtenção de padrões de difracção de maior qualidade.

De modo a se conseguir obter a estrutura tridimensional de uma proteína é necessário passar por várias etapas, podendo cada uma delas ser determinante no sucesso de um estudo cristalográfico. Estas são a expressão e purificação de uma proteína, a cristalização, a experiência de difracção de raios-X, a resolução da estrutura e, finalmente, o refinamento e validação da estrutura (Figura 1.15).

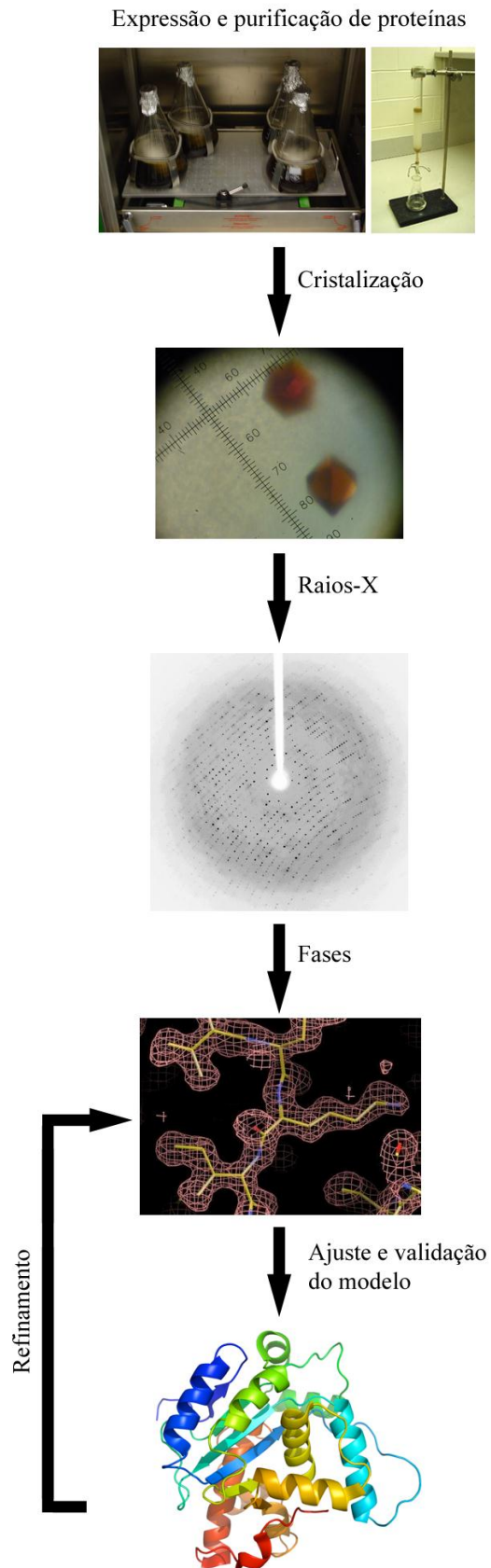


Figura 1.15. Representação esquemática dos principais passos necessários para a obtenção de uma estrutura tridimensional por cristalografia de raios-X.

1.3.1. Cristalização de proteínas

A obtenção de cristais de proteína de qualidade é muitas vezes considerada o passo mais difícil na resolução de uma estrutura cristalográfica. Inicialmente é preciso expressar/isolar e purificar a proteína em questão, sendo muitas vezes crucial a sua pureza e concentração em solução.

Um cristal de proteína é um sólido, cuja característica fundamental é a sua estrutura interna tridimensionalmente ordenada e periódica. Ao entrarem no estado cristalino, as moléculas adoptam uma, ou algumas orientações específicas e formam uma rede cristalina, que é mantida por interacções não covalentes, nomeadamente pontes de hidrogénio entre os resíduos superficiais de proteínas vizinhas mediadas por moléculas de solvente.^{77,79}

Esta rede cristalina forma-se quando as moléculas de proteína entram lentamente num estado supersaturado e agregam de uma maneira ordenada. Este processo ocorre quando se adiciona, a uma solução de proteína pura, uma solução que contém um agente precipitante numa concentração abaixo da necessária para a precipitação da proteína. De seguida, moléculas de água são retiradas por evaporação controlada até se atingir um equilíbrio, concentrando gradualmente o agente precipitante e conduzindo à precipitação da proteína. Os agentes precipitantes podem ser sais, como o sulfato de amónio, levando à precipitação da proteína pelo processo de “*salting out*” ou compostos orgânicos, como o polietilenoglicol (PEG), que interagem com os resíduos hidrofóbicos da proteína, causando por vezes desnaturação da mesma. O PEG é muito usado, pois este é um forte precipitante mas um fraco desnaturante.^{77,80,81}

O método de cristalização mais utilizado é a difusão de vapor. Este permite a evaporação controlada referida anteriormente, na qual uma gota resultante da mistura de uma solução de proteína e uma solução de precipitante (podendo-se utilizar diferentes proporções) é colocada num ambiente controlado (reservatório selado com uma lamela). No reservatório coloca-se uma quantidade elevada de solução precipitante. Deste modo a solução precipitante está mais diluída na gota, originando a evaporação de moléculas de água até se atingir um equilíbrio, reduzindo o volume da gota e levando à precipitação controlada da proteína.^{77,81,82} Dois exemplos deste método são as técnicas de gota suspensa e gota assente. Na técnica de gota suspensa a gota é colocada na lamela utilizada para selar o reservatório, enquanto na técnica de gota assente esta é colocada numa microponte presente dentro do reservatório (Figura 1.16).

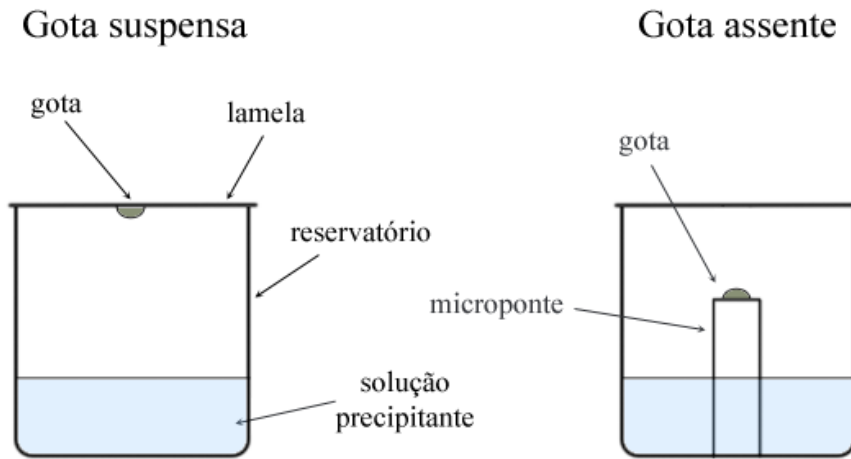


Figura 1.16. Representação esquemática das técnicas de gota suspensa (à esquerda) e gota assente (à direita).

O processo de cristalização pode ser explicado utilizando um diagrama de fases da cristalização (Figura 1.17). Este está dividido em duas grandes zonas: a zona não saturada e a zona supersaturada (onde ocorre cristalização). A zona supersaturada encontra-se subdividida em três zonas: a zona metaestável, a zona lábil e a zona precipitante. Para que ocorra cristalização tem que haver inicialmente nucleação. Esta ocorre na zona lábil e corresponde à formação de pequenos núcleos de proteína que se agrupa inicialmente, definindo à partida a estrutura do cristal. Após a fase de nucleação, a concentração de proteína diminui e ocorre a fase de

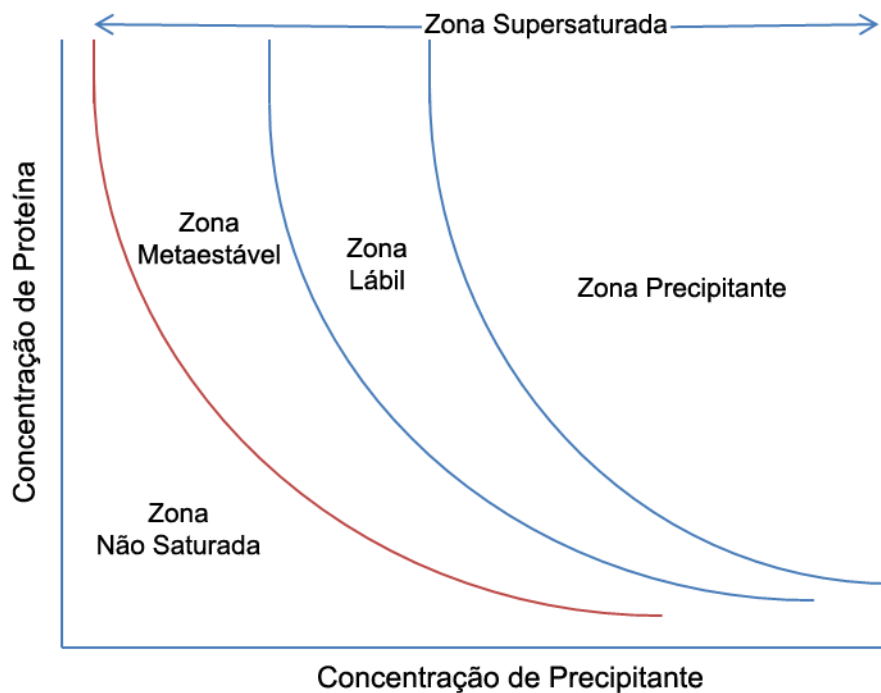


Figura 1.17. Diagrama de fases da cristalização. Adaptado de <http://www.p212121.com/2009/11/15/the-phase-diagram-in-protein-crystallization/>.

crescimento do cristal, na zona lábil ou metaestável de supersaturação. Nesta fase as moléculas de proteína precipitam lentamente e são adicionadas, numa maneira ordenada, aos núcleos anteriormente formados provocando um aumento de volume destes e originando um cristal de proteína. Na zona precipitante não ocorre cristalização pois a proteína precipita muito rapidamente, levando à formação de um sólido amorfo e não cristalino.^{81,83}

Existem vários factores que podem influenciar a cristalização de proteínas, tanto ao nível da obtenção ou não de cristais como ao nível da qualidade dos mesmos, sendo os mais importantes a pureza e concentração da proteína, a natureza e a concentração do agente precipitante, a temperatura, o pH e a força iónica. De forma a determinar as condições de cristalização de uma proteína podem ser realizados ensaios de *screening* que variam estas condições. Outras opções incluem a adição de aditivos que estabilizem a proteína, como cofactores da proteína ou inibidores em alguns casos, a adição de detergentes no caso de proteínas de membrana ou até a remoção proteolítica de uma porção mais desordenada da proteína.⁷⁷

Depois de obtido o cristal, é necessário compreender qual a organização das moléculas de proteína dentro do mesmo, para se poder interpretar o seu padrão de difracção. Ao volume mais pequeno e mais simples que é representativo da totalidade do cristal dá-se o nome de célula unitária. Por outras palavras, esta é a unidade a partir da qual, através de operações de translação, é possível gerar a totalidade de um cristal. Cada uma destas unidades está envolvida num ambiente idêntico rodeada pelas unidades vizinhas, e é a imagem tridimensional do seu conteúdo o resultado que se obtém após uma experiência de difracção. A célula unitária de um cristal é caracterizada pelas constantes da célula: três vectores (a , b , c) e três ângulos (α , β , γ) (Figura 1.18). É ainda possível definir um volume mais pequeno que se repete dentro da célula unitária: a unidade assimétrica. Esta corresponde ao volume mais pequeno a partir do qual, através de operações de simetria, é possível gerar a célula unitária (Figura 1.19).^{77,80}

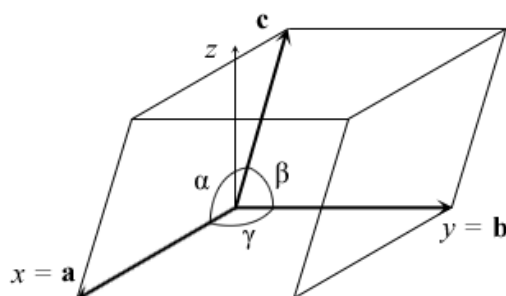


Figura 1.18. Representação esquemática de uma célula unitária e respectivas constantes (a , b , c e α , β , γ).

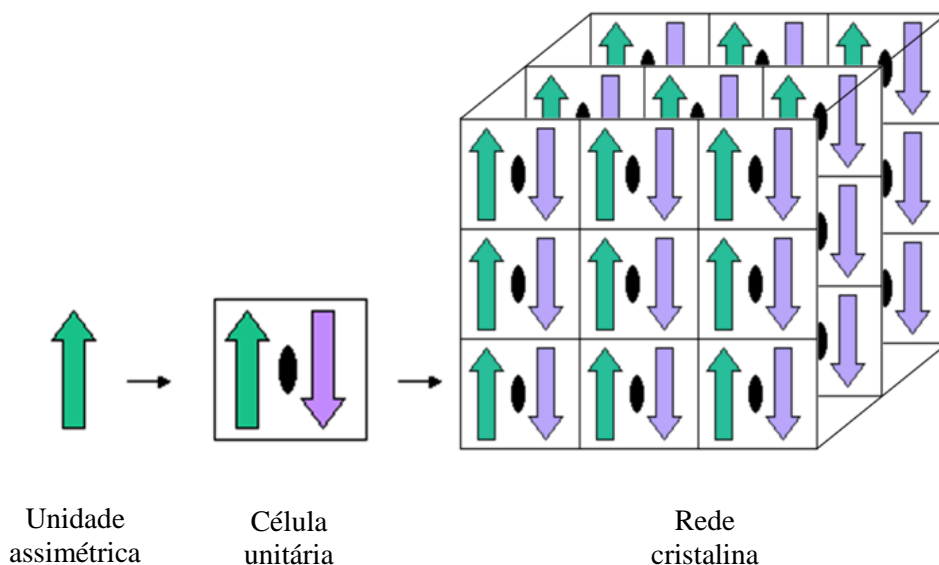


Figura 1.19. Representação esquemática da relação entre a unidade assimétrica, a célula unitária e a rede cristalina. Adaptado de http://www.rcsb.org/pdb/101/static101.do?p=education_discussion/Looking-at-Structures/bioassembly_tutorial.html.

Consoante as diferentes possibilidades de disposição espacial das partículas nas faces e no interior das células unitárias existem quatro tipos de célula unitária (P, C, I, F). Quando estas são combinadas com os sete sistemas cristalinos, resultantes das diferentes combinações das constantes das células unitárias, é possível criar 28 redes cristalinas. Na realidade, devido à simetria das soluções, existem apenas 14 configurações básicas, formando-se todas as demais a partir destas. Toda a simetria da rede cristalina é definida pelo seu grupo espacial. Se considerarmos todas as operações de simetria possíveis, obtêm-se 230 grupos espaciais. No entanto, como as proteínas são moléculas quirais, não são permitidas operações do tipo plano do espelho ou centros de inversão, reduzindo o número de grupos espaciais para 65.⁸⁰

No caso dos estudos de complexos proteína-ligando, é necessário obter cristais do mesmo e não somente da proteína. Os dois métodos mais utilizados são a cocristalização do complexo ou a cristalização da proteína seguida de *soaking* com o ligando. No primeiro método, a proteína e o ligando são incubados por um determinado período de tempo e de seguida são efectuados os ensaios de cristalização. Este método é vantajoso em casos que a adição do ligando não interfira ou até facilite a cristalização da proteína (por exemplo, adição de cofactores ou inibidores da proteína). Muitas das vezes a remoção de excesso de ligando é crucial para que ocorra cristalização. O método de *soaking* beneficia da existência de canais de solvente no cristal de proteína originalmente formado, através dos quais ocorre difusão do ligando e posterior ligação à proteína. Este método é útil em casos onde a adição do ligando interfere com a cristalização da proteína (por exemplo, indução de alterações conformacionais que dificultem a cristalização),

ou quando é necessária a medição do cristal pouco tempo depois da adição do mesmo (por exemplo, adição de substratos).⁸⁴

1.3.2. Experiência de difracção

Os raios-X são um tipo de radiação penetrante e podem ser difractados por electrões. Quando os raios passam por um átomo, apenas uma quantidade muito pequena é difractada, sendo que a restante continua o seu percurso podendo ser difractada por outros átomos, ou atravessar a totalidade do objecto sem ser difractada. Numa experiência de difracção estes raios são normalmente capturados pelo *beam stop*, pois não nos dão informações úteis sobre a estrutura da proteína.⁸⁰

Na experiência de difracção faz-se incidir radiação de raios-X a um determinado comprimento de onda (perto de 1 Å, na maioria das experiências, podendo ser alterado em alguns casos) sobre o cristal, montado num goniómetro (Figura 1.20). O cristal difracta os raios-X com ângulos específicos, produzindo reflexões que formam um padrão de difracção. Utilizando o goniómetro, o cristal é centrado no feixe, rodado em ângulos muito pequenos e são disparados vários pulsos de raios-X, recolhendo-se as imagens dos padrões de difracção correspondentes a cada rotação. A experiência pode ser feita à temperatura ambiente, mas por norma os cristais

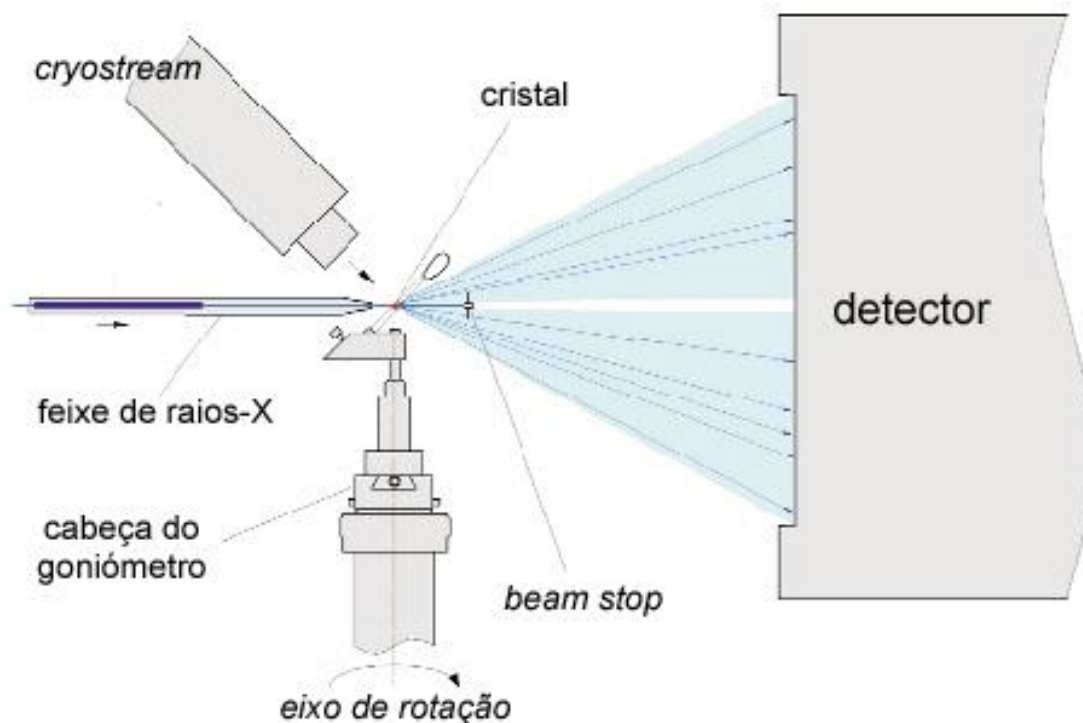


Figura 1.20. Representação esquemática do arranjo experimental de uma experiência de difracção. Adaptado de <http://www.esrf.eu/UsersAndScience/Publications/Highlights/2003/Materials/MAT15/>.

são congelados em azoto líquido e a experiência é efectuada sob um feixe criogénico (*cryostream*) a cerca de 100 K. É através da análise dos padrões de difracção que se resolve a estrutura tridimensional da molécula cristalizada.

Para se compreender a difracção dos raios-X e os padrões que se obtêm é necessário considerar a estrutura interna de uma rede cristalina. Durante os seus estudos, W. L. Bragg mostrou que os ângulos com que emergem os feixes difractados por cristais podem ser analisados como se estes fossem reflectidos por um conjunto de planos de átomos, paralelos e equivalentes, num cristal.^{77,79}

Os planos mais óbvios numa rede cristalina são aqueles determinados pelas faces das suas células unitárias. Estes e todos os outros planos espaçados regularmente que atravessem por pontos equivalentes da rede cristalina podem ser caracterizados por um conjunto de três números, designados índices de Miller e representados por hkl . Cada conjunto de planos é responsável por uma reflexão no padrão de difracção.⁷⁷

Estes conjuntos de planos apresentam diferentes distâncias interplanares. Esta distância é designada por d_{hkl} . Segundo a Lei de Bragg, um conjunto de planos com índice hkl e distância interplanar d_{hkl} produz uma reflexão quando raios-X de comprimento de onda λ o atinge com um ângulo θ e são reflectidos com o mesmo ângulo, apenas se θ satisfizer a condição

$$2d_{hkl}\sin\theta = n\lambda \quad (2)$$

sendo que n é um número inteiro. Por outras palavras, se os raios-X estiverem em fase, estes interferem construtivamente e a Lei de Bragg aplica-se, originando uma reflexão no padrão de difracção. Isto acontece quando a distância adicional percorrida pela radiação entre dois planos paralelos é igual a um número inteiro do seu comprimento de onda (Figura 1.21). Se por outro lado os raios-X atingirem o cristal a um ângulo θ' diferente, a Lei de Bragg não se aplica e os

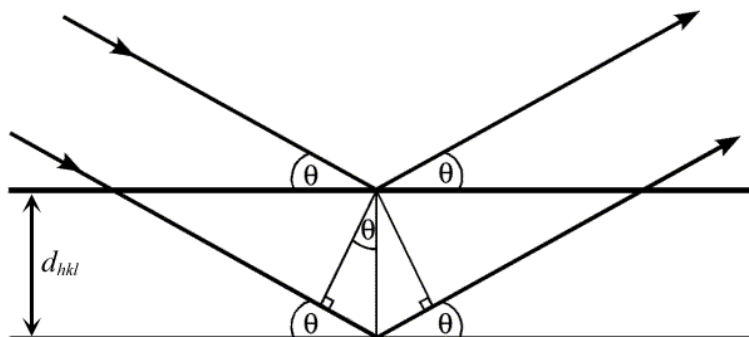


Figura 1.21. Representação esquemática da incidência de radiação num conjunto de planos. Adaptado de http://www-structmed.cimr.cam.ac.uk/Course/Basic_diffraction/Diffraction.html.

raios difractados por cada plano paralelo estão fora de fase, logo interferem destrutivamente e não se gera uma reflexão.^{77,79}

É por este motivo que é necessário rodar o cristal numa experiência de difracção de raios-X. Quando um pulso de radiação monocromática é disparado contra um cristal numa determinada orientação, apenas alguns conjuntos de planos obedecem à Lei de Bragg. Ao rodar o cristal diferentes conjuntos de planos vão produzir reflexões no padrão de difracção e só assim poderemos obter um conjunto de dados completo. Pela Lei de Bragg, também se pode observar que, para $n = 1$, quanto menor for a distância interplanar d_{hkl} maior é o ângulo de difracção θ , sendo que estas reflexões vão aparecer nos limites do padrão de difracção. Uma menor distância interplanar corresponde a uma maior resolução da estrutura da proteína. Estas reflexões dão-nos informação sobre os detalhes mais finos da estrutura. Por outro lado, reflexões situadas perto da origem do padrão de difracção correspondem a planos mais afastados entre si e transmitem informação sobre características mais abrangentes das moléculas na célula unitária, sendo também de extrema importância.^{77,79}

Enquanto a rede cristalina se encontra no espaço real (caracterizada por um conjunto de coordenadas Cartesianas x,y,z), cada reflexão (hkl) encontra-se no espaço recíproco. Estes dois espaços estão relacionados, pois é possível prever a posição da reflexão que cada conjunto de planos gera. Logo, através da análise dos padrões de difracção de um determinado cristal é possível prever qual a sua simetria e o seu grupo espacial mais prováveis, assim como a posição dos átomos da proteína no espaço real. No entanto, para descobrir estes últimos é necessário saber a intensidade relativa de cada reflexão (I_{hkl}) e o ângulo da fase da onda electromagnética que gerou essa mesma reflexão (α_{hkl}).

1.3.3. Da difracção à densidade electrónica

Como foi visto anteriormente, as reflexões representam feixes de raios-X difractados, os quais apresentam as características da radiação electromagnética: comprimento de onda, amplitude e fase. Enquanto o comprimento da onda difractada é constante, a amplitude e fase podem variar. A função periódica das ondas representadas por estas grandezas pode ser descrita como uma série de Fourier pela equação dos factores de estrutura F_{hkl} (equação 3).^{77,79}

Um factor de estrutura compreende todos os componentes da onda que originou a reflexão e representa a soma de todos os termos relativos a todos os átomos que difractam na célula unitária. Este pode ser descrito da seguinte maneira

$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad (3)$$

onde a parte exponencial corresponde às coordenadas tridimensionais dos átomos na célula unitária e f_j é o factor de dispersão atómica e representa a amplitude da contribuição do átomo j . Esta varia consoante o elemento do átomo, pois cada elemento tem um diferente número de electrões e difracta os raios-X de maneiras diferentes. Alternativamente, os átomos também podem ser descritos como densidade electrónica num determinado volume (célula unitária). Portanto, F_{hkl} pode ser descrito como um integral da contribuição de cada elemento de volume da densidade electrónica na célula unitária

$$F_{hkl} = \int_V \rho(x, y, z) e^{2\pi i(hx+ky+lz)} dV \quad (4)$$

onde $\rho(x, y, z)$ representa a função de densidade electrónica e V o volume da célula unitária. Devido ao facto da operação da transformada de Fourier ser reversível, pode-se descrever a densidade electrónica em função dos factores de estrutura

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx+ky+lz-\alpha'_{hkl})} \quad (5)$$

onde $|F_{hkl}|$ corresponde à amplitude e α'_{hkl} à fase do factor de estrutura ($\alpha' = \alpha/2\pi$).^{77,79}

Enquanto a amplitude da função dos factores de estrutura é proporcional a $\sqrt{I_{hkl}}$, sendo obtida directamente a partir do padrão de difracção e da intensidade de cada reflexão, o ângulo da fase da onda não é directamente determinado a partir dos dados experimentais, tendo que ser calculado posteriormente, o que leva ao conhecido “problema da fase” em cristalografia de raios-X.⁷⁷

Existem vários métodos de calcular uma estimativa inicial do ângulo da fase de cada reflexão. Entre eles encontram-se os métodos de substituição isomórfica (MIR/SIR, do inglês *multiple/single isomorphous replacement*) e de dispersão anómala (MAD/SAD, do inglês *multi/single-wavelength anomalous dispersion*), onde é necessária a obtenção de cristais com derivados de átomos pesados da proteína em estudo. Outro método muito utilizado é a substituição molecular (MR, do inglês *molecular replacement*), o qual requer que já seja conhecida a estrutura tridimensional da proteína em questão, como é o caso da DgAOR e da tripsina, ou de proteínas com um elevado grau de homologia (superior a 30%).⁷⁹

A substituição molecular utiliza os modelos conhecidos da proteína em questão ou de proteínas homólogas para obter informação sobre as fases. Este modelo tem de ser posicionado inicialmente na célula unitária da estrutura desconhecida, utilizando primeiro uma função de rotação seguida de uma função de translação. De seguida, podem-se calcular os factores de estrutura da molécula conhecida e já correctamente posicionada de modo a se obterem as fases iniciais da molécula desconhecida. Estas fases são então utilizadas em conjunto com as

intensidades medidas experimentalmente, obtendo-se os factores de estrutura e permitindo o cálculo da densidade electrónica inicial para a molécula desconhecida. Ao se utilizar este método é necessário ter em conta que as fases da proteína modelo têm uma grande influência na densidade electrónica resultante, correndo-se o risco desta apresentar características da proteína modelo que não correspondem à realidade para a proteína desconhecida (*model bias*). A partir da densidade electrónica é possível ajustar o modelo da proteína em estudo e melhorar as fases através de refinamentos sucessivos. Existem vários tipos de mapas de densidade electrónica que são utilizados para a avaliação e construção do modelo cristalográfico, sendo que os dois mais usuais são o mapa simples ($2F_o - F_c$) e o mapa de diferença ($F_o - F_c$). O primeiro permite a visualização da densidade contínua referente a todos os átomos na unidade assimétrica e é a partir deste que se constrói o modelo. O mapa de diferença pode apresentar picos positivos e negativos de densidade, permitindo identificar zonas ainda não construídas do modelo (picos positivos) ou zonas que apresentam demasiada população electrónica em relação ao esperado pelos dados experimentais (picos negativos).⁷⁹

1.3.4. Refinamento e validação da estrutura

Inicialmente, o modelo obtido por substituição molecular pode apresentar erros, não explicando da melhor forma a densidade electrónica obtida. Uma vez ajustado, este modelo de coordenadas x,y,z pode ser alvo de uma transformada de Fourier, obtendo-se novas fases e densidade electrónica melhoradas, as quais vão permitir que o modelo seja novamente ajustado. Este ciclo é repetido as vezes necessárias até se chegar a um modelo que esteja em conformidade com a densidade electrónica obtida, explicando adequadamente os dados experimentais.

No refinamento de uma estrutura cristalográfica, pode ser utilizado um de dois diferentes métodos: o método dos mínimos quadrados e o método da máxima verosimilhança. O objectivo é ajustar o modelo à densidade electrónica, alterando a posição dos átomos no espaço real, sendo o segundo método o mais utilizado.

Deste modo, existem vários parâmetros que são utilizados para avaliar a qualidade do modelo construído. Entre eles, o mais importante é o factor R que verifica a concordância global entre o modelo e os dados experimentais e é dado pela equação 6.

$$R = \frac{\sum | |F_{obs}| - |F_{calc}| |}{\sum |F_{obs}|} \quad (6)$$

Quanto menor for o valor deste parâmetro, mais adequado é o modelo aos dados experimentais. Por norma, considera-se que uma estrutura está bem refinada quando este valor é inferior a

20%. Um valor de 40% pode ser aceitável para um modelo inicial, antes de qualquer refinamento, sendo que um valor de 60% corresponde à comparação entre o conjunto de amplitudes medido e um conjunto de amplitudes aleatório.^{77,85}

Outro parâmetro mais exigente é o factor R_{free} , o qual é calculado da mesma maneira que o factor R, mas utiliza um pequeno conjunto de intensidades (no mínimo 1000 reflexões) que são separadas desde o início. O R_{free} serve como um controlo de qualidade, medindo quão bem o modelo estrutural prevê um subconjunto de intensidades, as quais não foram incluídas no refinamento. Este funciona como um parâmetro de validação e não deve ser superior ao factor R em mais que 5%.^{77,85}

Enquanto os factores R e R_{free} avaliam a qualidade global do modelo, os factores de temperatura (factores B) permitem a identificação de erros locais associados ao modelo. Este parâmetro reflecte a ‘mobilidade’ de cada átomo no modelo, ou seja, a oscilação de cada átomo à volta da sua posição especificada. Estes devem apresentar valores superiores na superfície da proteína e em zonas mais móveis, enquanto no interior devem apresentar valores inferiores. De modo a identificar um erro local num determinado átomo, o seu factor B deve ser comparado com os factores B dos átomos da sua vizinhança. Se houver uma grande discrepância entre os valores é possível que haja um erro na estrutura.^{77,79,85}

Os factores de temperatura estão intimamente ligados a outro parâmetro utilizado no refinamento e validação de uma estrutura – o factor de ocupação. Este factor mede a fracção de moléculas no cristal em que determinado átomo se encontra naquela posição. Para os resíduos da proteína, esta deve ser sempre igual a 1, excepto em casos de conformação alternada da cadeia lateral de um aminoácido, podendo variar para os ligandos da proteína (cofactores, iões coordenados, inibidores, substratos, produtos). Ao se reduzir o factor de ocupação de um determinado átomo, o factor de temperatura também diminui.⁷⁷

Para além dos parâmetros que relacionam o modelo com os dados experimentais, existe uma série de protocolos de validação que possibilitam a identificação de erros relativos à geometria permitida das moléculas presentes no modelo. Entre eles encontra-se o gráfico de Ramachandran que verifica se todos os aminoácidos apresentam ângulos conformacionais permitidos na sua cadeia principal. Estes parâmetros permitem verificar se existem regiões irrealistas no modelo.⁷⁷ Quando o modelo satisfaz estes parâmetros e explica bem os dados experimentais, o refinamento dá-se por terminado e o modelo está pronto a ser depositado no PDB (<http://www.rcsb.org/>).

Capítulo 2

Objectivos

A presente dissertação tem como principais objectivos o estudo de interações de vários compostos com duas proteínas: Aldeído Oxidorredutase de *Desulfovibrio gigas* (*DgAOR*) e tripsina de *Bos taurus*. A técnica de eleição foi a cristalografia de raios-X, pois possibilita averiguar interações entre os compostos e resíduos específicos das proteínas em estudo ao nível atómico.

Seleccionou-se a *DgAOR*, que se trata de uma enzima de molibdénio da família da Xantina Oxidase, devido à facilidade de se obterem cristais de qualidade com condições de cristalização bem definidas e por se tratar de um bom modelo cristalográfico com estruturas de alta resolução depositadas no PDB. Pretende-se estudar o mecanismo de reacção desta enzima, o qual ainda não se encontra bem definido, com o objectivo de compreender melhor a sua função ao nível fisiológico, assim como a de proteínas análogas, como a Aldeído Oxidase. A *DgAOR* apresenta baixa especificidade, observando-se actividade utilizando diversos aldeídos como substrato. Por estes motivos, pretende-se determinar a estrutura desta enzima em complexo com os substratos benzaldeído, formaldeído e acetaldeído e estudar a sua interacção com o centro de molibdénio. Tenciona-se também testar a xantina, substrato da proteína Xantina Oxidase, com o objectivo de estudar a especificidade da *DgAOR*. Outro objectivo é averiguar sobre o papel desta enzima no metabolismo das espécies reactivas de oxigénio (ROS). Estruturas cristalográficas, determinadas pela Dra. Teresa Santos-Silva anteriormente à realização da presente dissertação, desta enzima em complexo com ditionito de sódio e sulfureto de sódio apontam para a ligação de uma molécula de peróxido de hidrogénio ao átomo de molibdénio no centro activo. De modo a confirmar estes resultados, pretende-se obter novas estruturas utilizando os mesmos compostos, assim como peróxido de hidrogénio.

Em relação à tripsina visa-se estudar a interacção com vários compostos, disponibilizados pelo Dr. John Spencer (University of Sussex, United Kingdom). Tendo como objectivo a obtenção e caracterização de compostos inibidores da urocinase, uma protease da mesma família da tripsina com um centro activo muito semelhante, é necessário estudar a sua especificidade, assim como as interações que estes estabelecem no centro activo destas enzimas. A tripsina é uma enzima muito estudada por cristalografia, dada a sua fácil cristalização e obtenção de dados de alta resolução. Nesta dissertação pretende-se estudar a sua interacção com os compostos AB1, AB11, JS62, JS67 e SR5, cuja estrutura se encontra representada na Figura 3.1 do subcapítulo 3.2.1. Estes compostos já foram classificados como inibidores desta proteína, apesar de estes resultados ainda se encontrarem em fase de análise e discussão.

Capítulo 3

Materiais e Métodos

3.1. Aldeído Oxidorredutase de *Desulfovibrio gigas*

3.1.1. Purificação

A DgAOR foi isolada e purificada a partir de células de *D. gigas* crescidas em meio lactato/sulfato de acordo com o que foi descrito anteriormente.^{22,29} Após o crescimento das células, estas foram recolhidas por centrifugação e ressuspensas em tampão 10 mM Tris-HCl a pH 7,6 e passadas por um homogeneizador Manton-Gaulin com uma pressão de 9 000 psi. Os restos celulares foram separados por centrifugação a 16 000 g durante 30 minutos. O sobrenadante obtido contém as proteínas solúveis e as membranas, as quais foram separadas por ultracentrifugação (Ultracentrífuga - Beckman) a 180 000 g durante uma hora.

O extracto solúvel foi submetido a uma cromatografia de troca aniónica, utilizando uma coluna dietilaminoetil-celulose (DE-52, Whatman) equilibrada com tampão 5 mM Tris-HCl a pH 7,6. A DgAOR foi eluída a uma concentração entre 250 e 300 mM do gradiente de tampão e as alíquotas recolhidas foram reunidas e dialisadas contra tampão 5 mM Tris-HCl pH 7,6. De seguida, foi realizada uma cromatografia de adsorção utilizando uma coluna de hidroxiapatite (Bio-Gel HT Hydroxyapatite, Bio Rad) equilibrada com tampão 1 mM fosfato de potássio (KPB) a pH 8,0. A eluição foi realizada através de um gradiente linear de KPB e a DgAOR foi eluída a uma molaridade de 150 mM. Após eluição, as fracções contendo DgAOR foram reunidas e a seguir dialisadas contra tampão 5 mM Tris-HCl pH 7,6. A amostra foi submetida a uma cromatografia de troca aniónica utilizando uma coluna aniónica Resource Q (GE Healthcare) equilibrada com 5 mM Tris-HCl pH 7,6. A proteína foi eluída a uma concentração de 300 mM do gradiente linear de tampão. A DgAOR pura foi dialisada contra 50 mM Tris-HCl pH 7,6 e depois concentrada.

A concentração de proteína foi determinada usando o coeficiente de extinção molar a 462 nm ($\epsilon = 24,6 \text{ mM}^{-1} \text{ cm}^{-1}$) ou com o método do BCA (Sigma).

A pureza da amostra foi verificada por SDS-PAGE a 12,5% e através do espectro de absorção UV/Visível mediante a razão de pureza estabelecida $R_{462/280} = 6$. Todos os reagentes utilizados durante a purificação apresentam grau analítico reconhecido.

Durante a realização desta dissertação, apenas o último passo da purificação foi realizado, em colaboração com o Dr. Pablo González (Grupo de química bio-inorgânica do Prof. José Moura, REQUIMTE-CQFB, FCT-UNL).

3.1.2. Ensaio cinéticos

Foram realizados estudos cinéticos de modo a determinar a actividade específica das alíquotas purificadas da *DgAOR*. Os ensaios foram realizados utilizando um espectrofotómetro *diode array* TIDAS/NMC301-MMS/16 VIS/500-1 em condições aeróbias, à temperatura ambiente (298 K). Para a determinação das velocidades iniciais de oxidação do substrato foi medida a redução do 2,6-diclorofenolindofenol (DCPIP) a 600 nm ($\epsilon = 21 \text{ mM}^{-1} \text{ cm}^{-1}$) numa célula com um percurso óptico de 1 cm, contendo a seguinte mistura: 50 mM de Tris-HCl pH 7,6, 35 μM de DCPIP como aceitador de electrões, 200 μM de benzaldeído como substrato redutor, e 450 nM de *DgAOR*.

Realizaram-se também estudos cinéticos em função do tempo de ressulfuração do centro activo da *DgAOR* segundo um protocolo previamente descrito,³⁴ o qual, segundo a literatura, lhe confere actividade. Para tal incubou-se 150 μM de *DgAOR* na presença de 7 mM de sulfureto de sódio (Na_2S) e 30 mM de ditionito de sódio ($\text{Na}_2\text{S}_2\text{O}_4$) durante diferentes tempos à temperatura ambiente e em condições aeróbias e anaeróbias e foi medida a sua actividade de acordo com o protocolo anteriormente descrito. Todos os reagentes utilizados durante estes ensaios apresentam grau analítico reconhecido.

3.1.3. Ensaio de cristalização

Foram efectuados ensaios de cristalização utilizando uma amostra de *DgAOR* a uma concentração de 10 mg/mL em 10 mM de tampão Tris-HCl a pH 7,6, procedendo de acordo com a metodologia previamente descrita.³¹ As cristalizações foram realizadas por difusão de vapor, utilizando o método de gota assente em condições aeróbias, utilizando caixas de cristalização de 24 poços (Linbro) e micropontes côncavas (Greiner). No reservatório foram colocados 700 μL de solução precipitante composta por 30% de isopropanol, 0,2 M de tampão HEPES a pH 7,6 e 0,2 M de cloreto de magnésio (MgCl_2). Na gota foram colocados 4 μL de *DgAOR* e 2 μL de solução precipitante. Os ensaios foram realizados a 277 K e a visualização das gotas foi efectuada utilizando um microscópio Olympus SZ60.

Após obtenção de cristais, foi realizado o protocolo de remoção lenta do isopropanol do centro activo da enzima descrito anteriormente.^{35,36} Foram adicionados 4 μL às gotas de uma solução HB1 (do inglês *harvesting buffer 1*) composta por 30% de isopropanol, 0,2 M de MgCl_2 , 0,2 M de tampão HEPES a pH 7,6 e 30% de PEG 3350 (PEG 3350). Após estabilização das gotas durante 48 horas, procedeu-se à adição de 2 μL de uma solução HB2 composta por 0,2 M de MgCl_2 , 0,2 M de tampão HEPES a pH 7,6 e 30% de PEG 3350 e remoção cuidadosa ao microscópio de 2 μL da solução da gota. Este procedimento foi repetido pelo menos cinco vezes

de forma a substituir por completo o conteúdo da gota pela solução HB2. Após estabilização das gotas durante 48 horas, os cristais foram transferidos para novas gotas contendo 700 μL da solução HB2 no reservatório e 5 μL na gota, onde estabilizaram durante 72 horas.

Com os cristais livres de isopropanol foram então realizados ensaios de *soakings* de vários compostos. Para isso prepararam-se diferentes soluções com a mesma composição da solução HB2 adicionando os seguintes compostos: (1) 100 μM de peróxido de hidrogénio, (2) 7 mM de sulfureto de sódio e 30 mM de ditionito de sódio, (3) 2,8 mM de benzaldeído, (4) 50 mM de acetaldeído, (5) 50 mM de formaldeído e (6) 45 μM de xantina. Adicionaram-se 5 μL das soluções 1, 2, 3 e 4 às gotas. Em relação às soluções 2 e 6 foi efectuado um processo de adição/remoção cuidadosa efectuado anteriormente, utilizando 5 μL em vez de 2 μL . Os cristais foram congelados em azoto líquido ao fim de 1 e 3 horas para todas as soluções e também 24 horas para as soluções 1, 3 e 6. Não houve necessidade de utilizar uma solução crio-protectora, pois a solução HB2 já desempenha essa função.

Foram também realizadas cristalizações em condições anaeróbias ($\text{O}_2 < 1$ ppm), sendo que todas as soluções utilizadas foram desoxigenadas utilizando árgon antes de serem utilizadas. O procedimento efectuado foi o mesmo e foram feitos ensaios de *soaking* com as soluções 3, 4, 5 e 6. Uma hora após o *soaking*, a caixa de cristalização foi removida da câmara e os cristais foram congelados em azoto líquido em condições aeróbias. Após 5 horas foram congelados novos cristais, os quais estiveram quatro horas em condições anaeróbias. Todos os reagentes utilizados durante os ensaios de cristalização apresentam grau analítico reconhecido.

3.1.4. Recolha de dados e resolução da estrutura

Os cristais congelados foram analisados por radiação de Sincrotrão no *European Synchrotron Radiation Facility* (ESRF – Grenoble, França) na linha ID29.

Para os cristais que difractaram, utilizaram-se os programas MOSFLM⁸⁶ ou XDS⁸⁷ para indexar as reflexões e determinar o grupo espacial e constantes da célula, assim como integrar as intensidades. Para processar e escalar os dados utilizou-se o programa SCALA⁸⁸ do conjunto de programas CCP4.⁸⁹ O programa PHASER⁹⁰ foi usado para determinar as fases por substituição molecular, utilizando como modelo a estrutura depositada no PDB com o código 1VLB.³² Recorreu-se ao programa DM⁹¹ para melhorar a densidade electrónica inicial. O refinamento da estrutura foi feito com o programa REFMAC5⁹² onde ciclos de refinamento com restrições foram intercalados com a visualização e ajuste dos modelos moleculares usando o programa COOT.⁹³

3.2. Tripsina

3.2.1. Ensaio de cristalização

De modo a proceder aos estudos por cristalografia de raios-X utilizou-se tripsina de pâncreas bovino comercial (Sigma). Inicialmente, testaram-se várias condições de cristalização para a tripsina incubada com benzamidina, a qual é um inibidor da tripsina, de modo a prevenir autocatálise. A tripsina foi então dissolvida a 60 mg/mL em 25 mM de HEPES a pH 7,5, 10 mM de cloreto de cálcio (CaCl_2) e 10 mg/mL de benzamidina. Utilizaram-se as condições 4, 15, 20, 30 e 31 do *screen* de cristalização *Crystal Screen* HT da Hampton Research, conhecidas por dar bons cristais desta proteína (Tabela 3.1). As cristalizações foram realizadas por difusão de vapor, utilizando o método de gota suspensa em caixas de cristalização de 24 poços (Linbro). No reservatório foram colocados 700 μL de cada solução precipitante e utilizaram-se os seguintes reagentes: tampão Tris-HCl a pH 8,5, sulfato de amónio ($(\text{NH}_4)_2\text{SO}_4$), tampão de cacodilato de sódio a pH 6,5, PEG 8000, acetato de sódio a pH 4,6 e PEG 3350. Na gota foram colocados 2 μL de tripsina e 2 μL de solução precipitante. Os ensaios foram realizados a 293 K e a visualização das gotas foi efectuada utilizando um microscópio Olympus SZH10.

Tabela 3.1. Condições de cristalização testadas para a tripsina.

*O agente crio-protector foi apenas utilizado na solução crio-protectora, e não na solução precipitante.

Condição	Sal	Tampão	Precipitante	Agente crio-protector*
4	-	Tris-HCl pH=8,5, 0,1 M	Sulfato de amónio, 2 M	Glicerol, 25%
15	Sulfato de amónio, 0,2 M	Cacodilato de sódio pH=6,5, 0,1 M	PEG 8000, 30%	Glicerol, 25%
20	Sulfato de amónio, 0,2 M	Acetato de sódio pH=4,6, 0,1 M	PEG 3350, 25%	Glicerol, 25%
30	Sulfato de amónio, 0,2 M	-	PEG 8000, 30%	Glicerol, 25%
31	Sulfato de amónio, 0,2 M	-	PEG 3350, 30%	Glicerol, 25%

Pela visualização das gotas e dos cristais correspondentes seleccionaram-se as condições 15, 30 e 31 para testar a cristalização da tripsina com vários compostos inibidores. Os cristais de tripsina inibida com benzamidina obtidos utilizando estas condições foram posteriormente

usados para testar a eficácia de soluções crio-protectoras, usando glicerol como agente crio-protector, num aparelho de raios-X *in-house* (gerador de ânodo rotativo FR591, Enraf-Nonius; detector MAR-Research 300 mm e sistema de *Criocooling* Cryostream 600, Oxford Cryosystems).

Os compostos utilizados neste estudo foram fornecidos pelo Dr. John Spencer (University of Sussex, United Kingdom) com quem temos uma colaboração. Estes são o ácido ((carbamimidoiltio)metil)fenil)borónico (AB1 – $C_8H_{11}BN_2O_2S$), o ácido (E)-4-(((N'-metilcarbamimidoil)tio)metil)fenil)borónico (AB11 – $C_9H_{13}BN_2O_2S$), o 4-(4,4,5,5-tetrametil-1,3,2-dioxaborolan-2-il)benzil carbamimidotioato (JS62 – $C_{14}H_{21}BN_2O_2S$), o 4-(trifluorometil)benzil carbamimidotioato (JS67 – $C_9H_9F_3N_2S$) e o benzil carbamimidotioato (SR5 – $C_8H_{10}N_2S$) e as suas estruturas estão representadas na Figura 3.1.

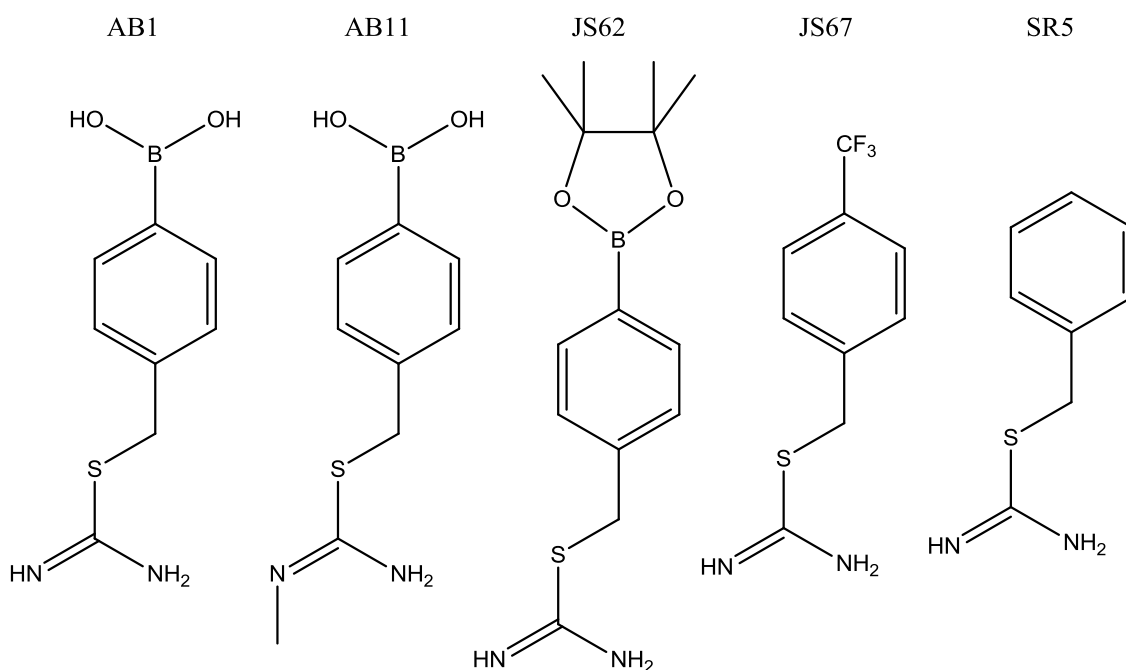


Figura 3.1. Representação da estrutura dos compostos AB1, AB11, JS62, JS67 e SR5, com os quais se incubou a tripsina.

A tripsina foi dissolvida a 60 mg/mL em soluções com 25 mM de HEPES a pH 7,5, 10 mM de $CaCl_2$ e 60 mM de cada um dos compostos. Incubou-se durante 30 minutos e foram realizados ensaios de cristalização utilizando as condições 15, 30 e 31. Após obtenção de cristais, estes foram congelados em azoto líquido utilizando tanto paratona como as soluções descritas anteriormente como solução crio-protectora.

3.2.2. Recolha de dados e resolução da estrutura

Os cristais congelados foram analisados por radiação de Sincrotrão no *European Synchrotron Radiation Facility* (ESRF – Grenoble, França) na linha ID29. O procedimento utilizado para a resolução da estrutura foi idêntico ao utilizado para a *DgAOR*. Utilizou-se o modelo depositado no PDB com o código 1S0Q para a resolução das fases com o programa PHASER. As moléculas dos compostos utilizados nos modelos finais foram criadas com o programa SKETCHER do CCP4.

Capítulo 4

Resultados e Discussão

4.1. Aldeído Oxidorredutase de *Desulfovibrio gigas*

A Aldeído Oxidorredutase de *D. gigas* é uma proteína muito estudada e caracterizada sendo um bom modelo para a realização de estudos cristalográficos, pois cristaliza com facilidade, obtendo-se cristais que difractam a alta resolução. A semelhança do seu centro activo com o de outras proteínas da família da Xantina Oxidase, como a Aldeído Oxidase humana, acrescentado ao facto de ainda existirem dúvidas em relação ao seu mecanismo de reacção, tornam a obtenção de estruturas desta proteína na presença de vários substratos um passo relevante. Por outro lado, estudos efectuados de modo a se tentar descobrir as diferenças entre as formas activa e inactiva desta proteína, discutidos no subcapítulo 4.1.1, deram azo à obtenção de estruturas desta proteína em complexo com ditionito de sódio e sulfureto de sódio, assim como espécies reactivas de oxigénio (peróxido de hidrogénio). Estes pontos serão discutidos em detalhe nos próximos subcapítulos.

4.1.1. Ensaio cinéticos

A proteína utilizada para realizar estes ensaios cinéticos foi obtida a partir de fracções previamente purificadas, tendo-se efectuado um passo de cromatografia de troca iónica de modo a assegurar a sua pureza, obtendo-se duas amostras de proteína (*DgAOR_A* a 2 mg/mL e *DgAOR_B* a 18 mg/mL). A actividade específica destas amostras foi medida, apresentando um valor de 0,42 U/mg para a amostra *DgAOR_B*, onde U corresponde a 1 μ mol de benzaldeído oxidado por minuto, sendo que a amostra *DgAOR_A* não apresenta actividade. Quando comparado com a literatura,³⁵ o valor de actividade para a amostra *DgAOR_B* é baixo, indicando que se trata de uma mistura de moléculas de *DgAOR* activa e inactiva. A diferença entre a forma activa e inactiva da enzima não é conhecida, sabendo-se, no entanto, que o protocolo de ressulfuração utilizado para recuperar a actividade da enzima XO³⁴ também confere actividade à *DgAOR*. Foram então realizados ensaios em condições aeróbias e anaeróbias onde se mediu a actividade específica da enzima ao longo do tempo após incubação com ditionito de sódio e sulfureto de sódio (Figura 4.1), utilizando a amostra inactiva *DgAOR_A*.

Quando se executa este protocolo, é possível observar um aumento de actividade da enzima numa fase inicial, seguido de um decréscimo da actividade. Em condições aeróbias, a enzima apresenta uma actividade específica máxima de 0,8 U/mg quando incubada com estes reagentes durante 50 minutos. A partir deste tempo verifica-se um decréscimo até aos 160 minutos onde a actividade da enzima é aproximadamente nula. Em condições anaeróbias, observa-se um aumento de actividade, atingindo uma actividade específica máxima de aproximadamente 1 U/mg aos 100 minutos, sendo que a partir daí se verifica uma diminuição de actividade

específica. Estes valores sugerem que, numa fase inicial, este tratamento confere actividade à enzima, mas que deve ser acompanhado pela formação de algum subproduto que se comporta como inibidor. Para explorar esta hipótese recorreu-se à cristalografia de raios-X com o objectivo de determinar a estrutura tridimensional da proteína na presença dos compostos que contribuem para a sua activação e inactivação.

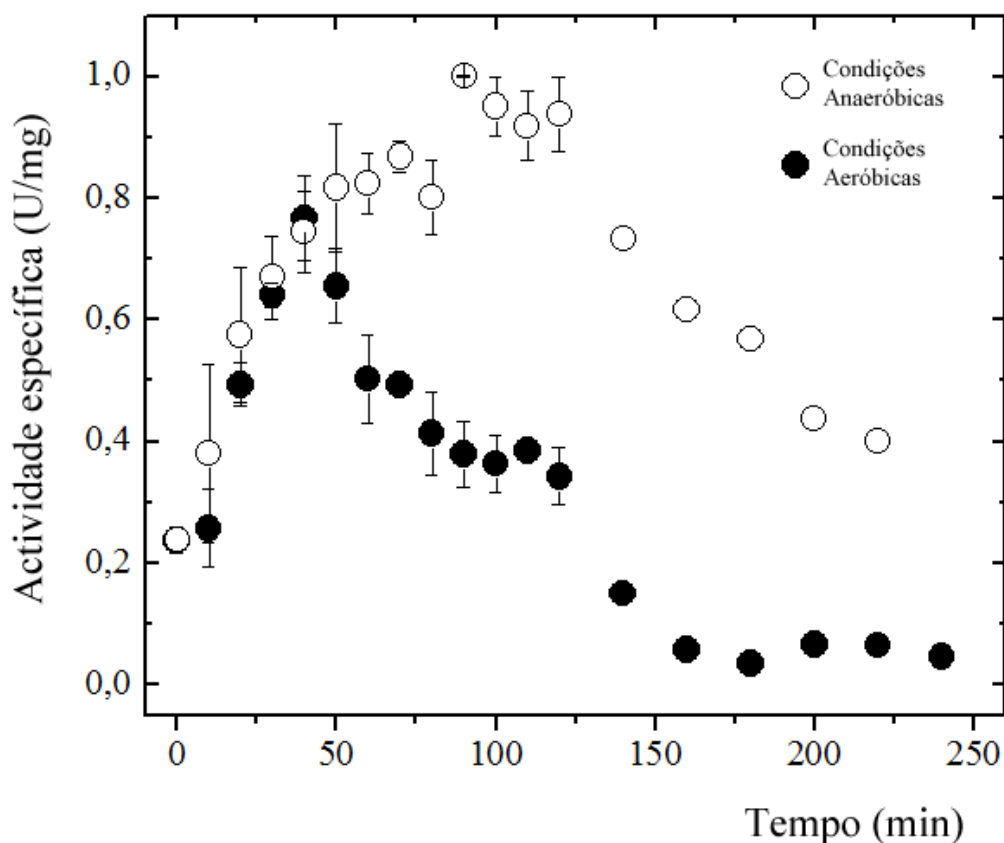


Figura 4.1. Representação gráfica da actividade específica da enzima *DgAOR* ao longo do tempo após incubação com ditionito de sódio e sulfureto de sódio, em condições aeróbicas (círculos cheios) e anaeróbicas (círculos vazios).

4.1.2. Ensaio de cristalização

Os ensaios de cristalização foram efectuados a pH 7,6 (tampão HEPES) de acordo com as condições previamente estabelecidas para esta proteína,³¹ utilizando isopropanol, um conhecido inibidor da enzima, como agente precipitante e cloreto de magnésio o qual fornece iões Mg^{2+} que estabilizam a sua estrutura cristalina. A temperatura de 277 K à qual foram realizadas as cristalizações proporciona uma evolução lenta da difusão de vapor e, conseqüentemente, do equilíbrio entre a gota e a solução do reservatório. Foi utilizada uma amostra de *DgAOR* obtida previamente à realização desta dissertação, a qual se encontrava a uma concentração de 10 mg/ml em 10 mM de tampão Tris-HCl a pH 7,5.

Nos dias seguintes à preparação das gotas foi observada a formação de bastante precipitado amorfo na base das micropontes, o que indica que as gotas se encontram na zona precipitante do diagrama de fases da cristalização (Figura 1.17). Como consequência desta precipitação, ocorre uma diminuição da concentração da proteína em solução permitindo que a proteína cristalize. Verificou-se a formação de pequenos cristais ao fim de duas semanas, os quais cresceram durante mais uma semana até atingirem um tamanho médio de 0,2 x 0,15 x 0,15 mm (Figura 4.2). Foram também obtidos cristais em condições anaeróbias com as mesmas características dos cristais obtidos em condições aeróbias.

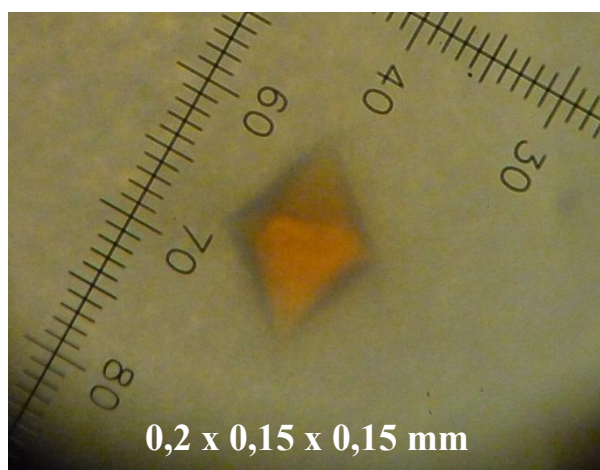


Figura 4.2. Cristal de DgAOR obtido em condições aeróbias a 277 K. Condições de cristalização: 30% isopropanol, 0,2 M HEPES pH 7,6 e 0,2 M MgCl₂.

De modo a se conseguir estudar a interação da proteína com os vários compostos, é necessário remover o isopropanol presente no centro activo da enzima. Isto é conseguido utilizando a metodologia previamente descrita no subcapítulo 3.1.3 e na literatura.^{35,36} Para além de remover o isopropanol, este método introduz PEG na gota, o que vai estabilizar os cristais de modo a que estes não se dissolvam aquando da diminuição da concentração de isopropanol, servindo também como agente crio-protector. Após o procedimento verificou-se que a maioria dos cristais não apresentaram alterações visíveis microscopicamente, enquanto alguns racharam ou apresentaram fissuras. Este resultado pode ser devido a uma alteração muito acentuada do ambiente dos cristais após transferência para novas gotas ou devido à sua manipulação aquando da transferência, indicando que este protocolo pode provocar alterações na rede cristalina, possivelmente piorando a qualidade da difracção dos cristais.

Com este procedimento os cristais encontram-se estáveis e com o centro activo livre, de modo a que se prosseguiu com a realização dos ensaios de *soaking*. Com o objectivo de não perturbar o equilíbrio estabelecido na gota, foram feitas várias soluções com a mesma composição da solução HB2 e a adição dos vários compostos. A concentração aproximada de cada composto na gota é apresentada na Tabela 4.1. Após a adição de cada composto às gotas, os cristais foram

congelados em azoto líquido ao fim de aproximadamente 1 hora, 3 horas e/ou 24 horas (Tabela 4.1), com o objectivo de se observar qual o tempo necessário para a difusão dos compostos pelos canais de solvente dos cristais e ligação ao centro activo, assim como a evolução ao longo do tempo da esfera de coordenação do molibdénio nas estruturas resultantes de cada cristal. Como a solução HB2 já é crio-protectora, o processo de congelamento dos cristais torna-se mais simples, eficiente e rápido, envolvendo menos manipulação dos cristais do que noutras situações em que uma solução crio-protectora é necessária. Os ensaios de *soaking* dos cristais obtidos em condições anaeróbias foram realizados durante 1 hora, tendo sido retiradas as gotas de cristalização para condições aeróbias para imediata congelação. Ao fim de 5 horas foram congelados novos cristais das mesmas gotas que na altura se encontravam já há 4 horas em condições aeróbias.

Tabela 4.1. Concentração final de cada composto e duração dos ensaios de *soaking* efectuados.

Composto	Condições	Concentração final na gota	Duração do ensaio de soaking (h)
Peróxido de hidrogénio	Aeróbias	50 μ M	1, 3
Sulfureto de sódio + ditionito de sódio	Aeróbias	7 mM + 30 mM	1, 3, 24
Benzaldeído	Aeróbias	1,4 mM	1, 3, 24
	Anaeróbias		1, 1+4
Acetaldeído	Aeróbias	25 mM	1, 3
	Anaeróbias		1, 1+4
Formaldeído	Aeróbias	25 mM	1, 3
	Anaeróbias		1, 1+4
Xantina	Aeróbias	45 μ M	1, 3, 24
	Anaeróbias		1

A maioria dos cristais congelados foi medida utilizando radiação de Sincrotrão na linha ID29 do ESRF. Alguns cristais não foram medidos devido a falta de tempo na linha, tendo sido guardados em azoto líquido para medições posteriores, caso seja necessário. Na Tabela 4.2 encontra-se um resumo dos resultados provenientes da experiência de difracção. Nos próximos subcapítulos serão discutidos o processamento dos dados e resolução das estruturas individualmente.

Tabela 4.2. Resumo das condições testadas e dos resultados obtidos na experiência de difração para a DgAOR.

*Designação pela qual vão ser referidas as estruturas relativas a cada composto ao longo da discussão

Designação ao longo do texto*	Composto	Condições	Duração do <i>soak</i> (h)	Experiência de difração
DgAOR_H ₂ O ₂	Peróxido de hidrogénio	Aeróbias	1	Res. máx. 1,50 Å
		Aeróbias	3	Não foi medido
		Aeróbias	24	Formação de gelo
DgAOR_Dit/S ²⁻	Sulfureto de sódio + ditionito de sódio	Aeróbias	1	Res.máx. 1,58 Å
		Aeróbias	3	Não foi medido
DgAOR_benz	Benzaldeído	Aeróbias	1	Res. máx. 1,49 Å
		Aeróbias	3	Res. máx. 1,51 Å
		Aeróbias	24	Res. máx. 1,40 Å
		Anaeróbias	1	Res. máx. 1,62 Å
		Anaeróbias	1 + 4	Res. máx. 1,75 Å
DgAOR_acet	Acetaldeído	Aeróbias	1	Res. máx. 1,50 Å
		Aeróbias	3	Res. máx. 1,66 Å
		Anaeróbias	1	Não foi medido
		Anaeróbias	1+4	Res. máx. 1,41 Å
DgAOR_form	Formaldeído	Aeróbias	1	Res. máx. 1,69 Å
		Aeróbias	3	Não foi medido
		Anaeróbias	1	Res. máx. 1,72 Å
		Anaeróbias	1+4	Não difractou
DgAOR_xant	Xantina	Aeróbias	1	Res. máx. 1,49 Å
		Aeróbias	3	Não foi medido
		Aeróbias	24	Res. máx. 1,64 Å
		Anaeróbias	1	Res. máx. 1,48 Å

4.1.3. Aldeído Oxidorredutase em complexo com ROS

Neste subcapítulo vão ser analisados os resultados referentes aos ensaios de *soaking* com peróxido de hidrogénio e com sulfureto de sódio e ditionito de sódio.

4.1.3.1. Experiência de difracção

Como se pode observar na Tabela 4.2 apenas dois conjuntos de dados foram recolhidos para estes compostos: um para a *DgAOR_H₂O₂* ao fim de uma hora e outro para a *DgAOR_Dit/S²⁻* também ao fim de uma hora. O comprimento de onda utilizado para a recolha destes conjuntos de dados, assim como todos os outros para a *DgAOR* foi de 0,976 Å e foram recolhidas 910 imagens para o *DgAOR_H₂O₂* e 1640 para a *DgAOR_Dit/S²⁻*. Na Figura 4.3 encontra-se um padrão de difracção característico desta proteína, referente ao conjunto de dados *DgAOR_H₂O₂*.

Estes dois conjuntos de dados foram tratados utilizando o programa MOSFLM no grupo espacial *P6₁22* (Tabela 4.3). De seguida as reflexões foram integradas de modo a se determinar as suas intensidades I_{hkl} . Deste processo resultou um ficheiro mtz que contém toda a informação obtida utilizando o MOSFLM.

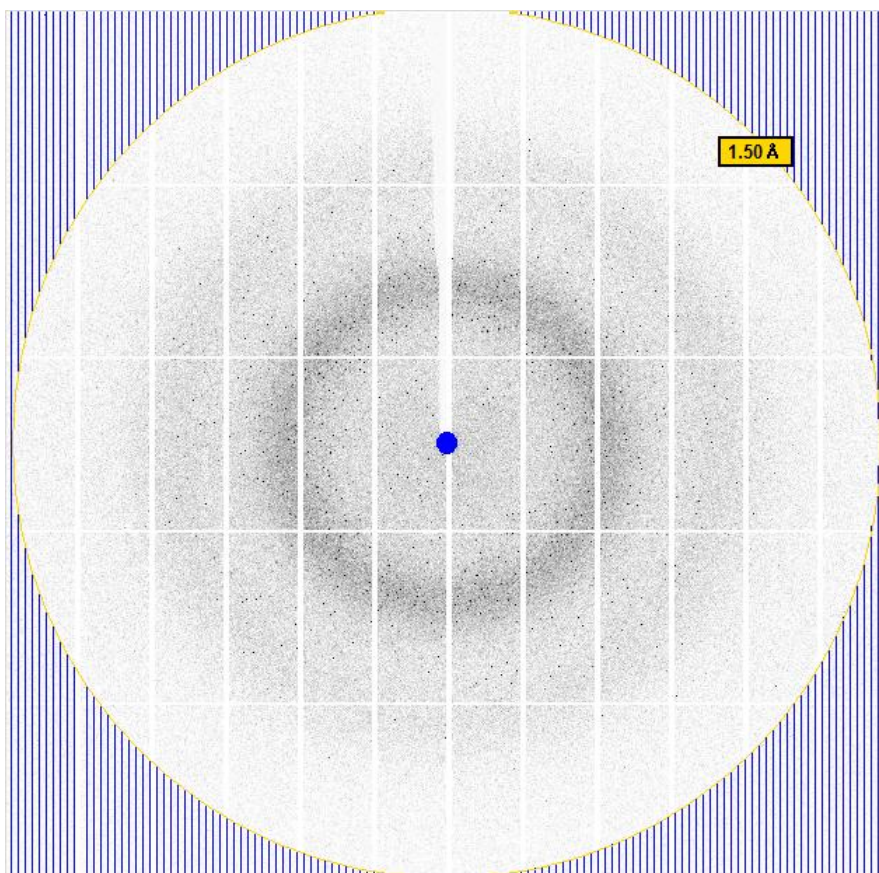


Figura 4.3. Exemplo de um padrão de difracção do cristal *DgAOR_H₂O₂*. Resolução máxima de 1,50 Å.

De modo a se determinar com maior certeza o grupo espacial, utilizou-se o programa POINTLESS⁸⁹ que analisa as intensidades das reflexões em conjunto com as constantes da célula de modo a identificar simetria no cristal, indicando os grupos espaciais com maior probabilidade. Este indicou os grupos espaciais enantiomórficos $P6_122$ e $P6_522$ com uma probabilidade de 98,9% para o $DgAOR_H_2O_2$ e 37% para o $DgAOR_Dit/S^2$. Entre os dois grupos espaciais, o $P6_122$ foi o seleccionado uma vez que, pela literatura, se sabia ser o correcto.

Procedeu-se à utilização do programa SCALA⁸⁸ o qual observa as intensidades relativas de todas as reflexões e coloca-as na mesma escala, de modo a que seja possível uma análise correcta por parte dos programas utilizados para a resolução da estrutura. As diferentes intensidades poderão estar relacionadas com variações na intensidade do feixe de radiação aquando da recolha dos dados, assim como danificação do cristal ao longo da experiência devido à radiação de Sincrotrão. Para além de escalar os dados, este programa também fornece um conjunto de estatísticas sobre a recolha de dados, o qual se encontra na Tabela 4.3. Durante este passo foram seleccionadas aleatoriamente 5% das reflexões, as quais não serão utilizadas para a resolução e refinamento da estrutura, servindo apenas para o cálculo do factor R_{free} .

Analisando a Tabela 4.3 pode-se tirar algumas conclusões sobre a qualidade dos dados de difracção. Ambos os conjuntos de dados, $DgAOR_H_2O_2$ e $DgAOR_Dit/S^2$, são de alta resolução (1,50 e 1,58 Å, respectivamente) o que indica que os cristais medidos têm uma rede cristalina ordenada e difractam bem os raios-X.

A *completeness* dos dados indica a percentagem de reflexões medidas em relação ao número total de reflexões possíveis a uma dada resolução. Em ambos os casos esta encontra-se muito próxima de 100%, o que indica que a estratégia de recolha dos dados foi bem executada e que se obtiveram conjuntos de dados completos. A redundância indica o rácio entre o número de reflexões totais observadas e o número de reflexões únicas, indicando em média o número de medições independentes de cada reflexão única. Deste modo, estruturas com elevada redundância apresentam uma maior confiança nas medições das intensidades das reflexões, assim como o seu erro associado, o que é o caso das medições em causa, as quais apresentam valores elevados de redundância (9,3 e 8,1). Por outro lado, a relação sinal-ruído das reflexões é dada pelo valor de $I/\sigma(I)$, o qual apresenta valores elevados para estes dados, quer no geral (cerca de 20 para ambos os casos), quer na camada externa de resolução (cerca de 6 para ambos os casos), indicando que os dados se podem utilizar com confiança. O R_{merge} é outra medida utilizada para determinar a qualidade dos dados experimentais, a qual reflecte a concordância entre reflexões equivalentes. Este valor deve situar-se entre os 5 e 10%, o que se verifica em ambos os casos (6,4 e 6,9%), podendo aumentar até 50% na última camada de resolução, o que

também se verifica com valores de 35,1 e 35,2%. No geral pode-se concluir que os dados apresentam uma qualidade elevada e pôde-se prosseguir sem necessidade de manipulação adicional.

Tabela 4.3. Estatísticas de recolha de dados para os cristais *DgAOR_H₂O₂* e *DgAOR_Dit/S₂⁻*. Os valores entre parêntesis correspondem à última camada de resolução.

Cristal	<i>DgAOR_H₂O₂</i>	<i>DgAOR_Dit/S₂⁻</i>
Comprimento de onda (Å)	0,976	0,976
Grupo espacial	<i>P6₁22</i>	<i>P6₁22</i>
Célula unitária (Å)	a,b = 143,1 c = 162,3	a,b = 143,5 c = 162,3
Limites de resolução (Å)	49,58–1,50 (1,58–1,50)	53,76–1,58 (1,66–1,58)
Nº de reflexões observadas	1 436 348 (213 767)	1 088 202 (161 477)
Nº de reflexões únicas	155 276 (22 397)	134 408 (19 406)
<i>Completeness</i> (%)	99,9 (100,0)	99,3 (99,7)
Redundância	9,3 (9,5)	8,1 (8,3)
<i>I</i> / σ (<i>I</i>) médio	20,5 (6,0)	19,7 (5,7)
<i>R_{merge}</i> ^a (%)	6,4 (35,1)	6,9 (35,2)
<i>VM</i> ^b - Parâmetro de Matthews (Å ³ /Da)	2,47	2,48
Conteúdo de solvente ^b (%)	50,2	50,5

(a) $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, onde $I_i(hkl)$ é a intensidade integrada de uma determinada reflexão i e $\langle I(hkl) \rangle$ é a intensidade média de múltiplas reflexões relacionadas por operações de simetria.

(b) $VM = \frac{V}{Z \times Mp}$, onde V é o volume da célula unitária, Z é o número de unidades assimétricas presentes na célula unitária e Mp é a massa molecular da proteína; Conteúdo de solvente = $\left(1 - \frac{1,23}{VM}\right) \times 100$

O parâmetro de Matthews, *VM*, permite a análise do conteúdo do cristal, relacionando o volume da célula unitária com a massa molecular da proteína e o número de unidades assimétricas presentes na célula unitária. A partir deste parâmetro pode-se calcular o conteúdo de solvente presente no cristal. Sabendo que o conteúdo de solvente se deve situar entre os 30 e os 70% (salvo raras exceções) e utilizando a massa molecular da proteína é possível prever o número de moléculas presentes na célula unitária, o que é uma informação extremamente útil aquando da resolução da estrutura. No caso dos cristais de *DgAOR* estes apresentam um conteúdo de solvente de aproximadamente 50% e apenas uma molécula por unidade assimétrica.

4.1.3.2. Resolução das estruturas

Nesta fase temos apenas a informação relativa às intensidades de cada reflexão, já escaladas e prontas a ser utilizadas para a resolução da estrutura. É então necessário calcular o ângulo da fase de cada reflexão de modo a se obterem os factores de estrutura necessários para o cálculo dos mapas de densidade electrónica. Este não se obtém experimentalmente e pode ser calculado utilizando os vários métodos mencionados no capítulo 1.3.3. No caso da *DgAOR* o método de eleição é o de substituição molecular, pois a estrutura desta proteína já se encontra resolvida. Deste modo, seleccionou-se a estrutura a 1.28 Å que se encontra depositada no PDB com o código 1VLB.³² Esta apresenta alta resolução e encontra-se no mesmo grupo espacial das estruturas em estudo, sendo por isso o modelo mais adequado para este estudo. Procedeu-se à remoção de todos os ligandos, cofactores e moléculas de água deste modelo, de modo a se efectuar este passo utilizando apenas a cadeia proteica, removendo a influência (*bias*) que poderia ser introduzida por esses elementos da estrutura.

A partir do ficheiro mtz proveniente do SCALA e deste modelo, utilizou-se o programa PHASER⁹⁰ para determinar as fases e calcular a densidade electrónica. Para tal define-se o número de moléculas por unidade assimétrica e o seu grupo espacial e o programa efectua operações de rotação e translação, utilizando o modelo 1VLB, tentando orientá-lo e posicioná-lo correctamente na densidade electrónica calculada a partir das fases do modelo e das intensidades experimentais. No caso da *DgAOR* fez-se a procura utilizando os grupos espaciais enantiomórficos $P6_122$ e $P6_522$ e apenas uma molécula por unidade assimétrica.

O sucesso deste processo é avaliado através dos parâmetros LLG (do inglês *Log Likelihood Gain*) e *Z-score* para cada função (rotação e translação). O LLG mede o quão melhor os dados experimentais podem ser previstos utilizando o modelo em causa do que com uma distribuição aleatória dos mesmos átomos. O *Z-score* indica quantos desvios padrão a solução achada se encontra acima da média e é utilizado para determinar se a estrutura foi ou não resolvida. Este valor deve ser superior a 4 para a função de rotação (RFZ) e 8 para a função de translação (TFZ) para que se considerar uma solução correcta.

Para os conjuntos de dados *DgAOR_H₂O₂* e *DgAOR_Dit/S²⁻*, o PHASER encontrou soluções no grupo espacial $P6_122$. Como era expectável, os valores de LLG e *Z-score* para as funções de rotação e translação são muito elevados (Tabela 4.4), pois o modelo utilizado trata-se da própria proteína, sendo apenas expectáveis diferenças ao nível do centro activo.

Tabela 4.4. Valores de Z-score e LLG para as funções de rotação e translação obtidos utilizando o programa PHASER, para os conjuntos de dados *DgAOR_H₂O₂* e *DgAOR_Dit/S²*.

Conjunto de dados	<i>DgAOR_H₂O₂</i>	<i>DgAOR_Dit/S²</i>
Z-score (rotação)	30,1	29,1
LLG (rotação)	574,9	569,2
Z-score (translação)	91,4	83,5
LLG (translação)	8555,6	8628,9

A partir do PHASER é gerado um novo ficheiro mtz, o qual já contém informação relativa à densidade electrónica da molécula, assim como um ficheiro pdb que tem a informação relativa ao modelo ajustado à densidade electrónica experimental. O ficheiro pdb contém informação sobre a posição de todos os átomos presentes no modelo (em coordenadas cartesianas x,y,z), assim como os factores de temperatura e ocupação. Estes últimos factores não se encontram refinados e correspondem aos do modelo utilizado, neste caso o 1VLB.

Neste ponto, já é possível visualizar a densidade electrónica relativa aos dados experimentais. No entanto, antes de se prosseguir com o refinamento da estrutura, utilizou-se o programa DM,⁹¹ o qual faz um melhoramento da densidade electrónica, modificando-a tendo em conta características conhecidas do comportamento da mesma. Através da indicação do conteúdo de solvente do cristal, este programa melhora as fases obtidas utilizando técnicas de *flattening* do solvente, ou seja, dá uma maior importância à densidade electrónica contínua, correspondente à proteína, e menor ao solvente, diminuindo assim o seu ruído e obtendo-se mapas mais fáceis de interpretar.

4.1.3.3. Refinamento das estruturas

Seguidamente procedeu-se ao refinamento da estrutura. Efectuou-se um refinamento inicial, utilizando o programa REFMAC5 para as duas estruturas em causa.⁹² Este executou vários conjuntos de 10 ciclos de *restrained refinement* intercalados com a adição automática de moléculas de água em zonas de densidade não preenchida pela modalidade *findwaters* do programa COOT.⁹³ O REFMAC5 ajusta a posição dos átomos do modelo à densidade electrónica e melhora as fases empregando o método da máxima verosimilhança.

Após este primeiro refinamento, inspeccionaram-se ambos os modelos e respectivos mapas de densidade electrónica ($2F_o - F_c$ e $F_o - F_c$) utilizando o programa COOT, começando por se averiguar o centro activo. Os factores R e R_{free} obtidos são bastante razoáveis para um primeiro refinamento (Tabela 4.5), sendo possível fazer uma análise inicial dos mapas resultantes. Em

ambas as estruturas verificou-se a presença de densidade electrónica relativa a uma molécula diatómica no local de ligação do ligando lábil equatorial OH/OH₂ do molibdénio. Estes resultados mostram que os *soaks* foram efectuados com sucesso e pode-se prosseguir para o refinamento e análise das estruturas. Os dados relativos à estrutura *DgAOR_Dit/S²⁻* foram, no entanto, abandonados uma vez que são muito semelhantes a uma das estruturas previamente obtidas pela Dra. Teresa Santos-Silva (também por ensaios de *soaking* com ditionito de sódio e sulfureto de sódio), a uma resolução de 1,57 Å. A tabela com os dados estatísticos referentes à recolha de dados e refinamento da estrutura encontra-se no Anexo 7.3. Esta estrutura foi analisada e comparada com a estrutura *DgAOR_H₂O₂* aqui obtida.

Tabela 4.5. Factores R e R_{free} após o primeiro refinamento para as estruturas *DgAOR_H₂O₂* e *DgAOR_Dit/S²⁻*.

Conjunto de dados	<i>DgAOR_H₂O₂</i>	<i>DgAOR_Dit/S²⁻</i>
Factor R (%)	18,3	18,9
Factor R _{free} (%)	22,3	24,0

O refinamento foi então continuado, utilizando o programa COOT para ajustar os resíduos à densidade electrónica, assim como para adicionar os cofactores da proteína, iões provenientes do solvente e moléculas de água conservadas. Após este processo utilizou-se novamente o programa REFMAC5, fazendo 10 ciclos de *restrained refinement*. Este processo foi repetido várias vezes até se chegar a um modelo que consiga explicar satisfatoriamente a densidade electrónica. Na fase final do refinamento foram utilizados parâmetros de TLS (*Translation Libration Screw-motion*),⁹⁴ adicionalmente ao *restrained refinement*. Estes parâmetros definem grupos de átomos no modelo da proteína que se comportam como corpos rígidos que apresentam movimento anisotrópico. Ao se agrupar os átomos do modelo, é possível refinar anisotropicamente a estrutura sem ser necessária uma grande quantidade de parâmetros, o que acontece com um refinamento anisotrópico individual para cada átomo. Os domínios de TLS definidos foram os quatro conhecidos da proteína, descritos no subcapítulo 1.1.5.1. No último passo de refinamento, uma molécula de H₂O₂ foi adicionada ao centro activo do modelo na posição do ligando lábil OH/OH₂, a qual se deixou livre para evitar qualquer tipo de *bias* durante o refinamento. A caracterização do centro activo vai ser discutida no subcapítulo 4.1.4.4.

Após o refinamento do modelo, procedeu-se à sua validação de modo a assegurar que este não compreende erros graves. Para tal foram utilizadas várias ferramentas do programa COOT de modo a verificar erros a nível da geometria das ligações químicas dos aminoácidos, assim como

a análise dos seus rotâmeros permitidos e dos factores de temperatura. Utilizou-se também o programa PROCHECK⁹⁵ e o STAN,⁹⁶ tendo sido adicionados dois iões Na⁺ à estrutura de acordo com o sugerido pelo módulo WASP deste último. Na Tabela 4.6 encontram-se as estatísticas relativas ao refinamento e validação do modelo.

Tabela 4.6. Estatísticas de refinamento e validação para o modelo *DgAOR_H₂O₂*.

Conjunto de dados	<i>DgAOR_H₂O₂</i>
Limites de resolução (Å)	49,58–1,50
Factor R (%)	13,2
Nº de reflexões	147 398
Factor R _{free} (%)	15,8
Nº de reflexões (R _{free})	7 792
Nº de resíduos	907
Nº de átomos	8 397
RMSD para o comprimento de ligação (Å)	0,027
RMSD para o ângulo de ligação (°)	2,62
Factores de temperatura médios (Å ²)	
átomos da cadeia principal	11,1
átomos das cadeias laterais	13,6
moléculas de água	29,5
Gráfico de Ramachandran (%)	
resíduos nas regiões mais favoráveis	91,2
resíduos em regiões adicionalmente favoráveis	8,1
resíduos em regiões generosamente permitidas	0,4
resíduos em regiões não permitidas	0,3

Em relação ao factor R foi possível observar uma diminuição do valor inicial de 18,3 % para um valor final de 13,2 %, o que indica que o modelo foi bem construído ao longo do refinamento e está de acordo com os dados experimentais. Este valor é calculado a partir da diferença entre os factores de estrutura observados e os calculados, devendo apresentar um valor inferior a 20 %. Por sua vez, o valor de R_{free} é calculado da mesma forma que o factor R mas usando um subconjunto de reflexões separadas à partida e não utilizadas para o cálculo da estrutura. Este valor deve ser no máximo 5 a 7 % superior ao valor do factor R. A estrutura em estudo apresenta um valor de R_{free} 2,6 % superior ao factor R, o que indica que o modelo construído é válido.

Os valores de RMSD (do inglês *Root-mean-square Deviation*) para o comprimento e o ângulo de ligação indicam o desvio que a estrutura apresenta em relação aos valores expectáveis teoricamente em termos de geometria das ligações químicas. O valor de RMSD para o comprimento de ligação (0,027 Å) encontra-se um pouco mais elevado que o aceitável para estruturas de alta resolução (aproximadamente 0,02 Å). Este pode ser devido ao refinamento não ter sido suficientemente rigoroso em relação às restrições geométricas. No entanto foi efectuada validação geométrica do modelo e este não apresenta erros graves, indicando a sua validade. O valor de RMSD para o ângulo de ligação (2,62°) encontra-se dentro do expectável (inferior a 3°), o que corrobora a validade geométrica da estrutura.

Os factores de temperatura médios definem a mobilidade média das várias regiões da estrutura. De acordo com o expectável, este valor é inferior para os átomos da cadeia principal da proteína (11,07 Å²), a qual apresenta uma estrutura mais rígida, ligeiramente superior para os átomos das cadeias laterais da proteína (13,58 Å²) e muito superior para as moléculas de água (29,45 Å²), as quais apresentam uma desordem muito mais significativa.

O gráfico de Ramachandran é utilizado para a validação do modelo, analisando os resíduos individualmente em relação aos ângulos conformacionais da cadeia principal. Pode-se confirmar que a grande maioria dos resíduos (902 resíduos, 99,7 %) se encontra em regiões permitidas deste gráfico. No entanto 5 resíduos encontram-se em regiões não permitidas (Ala171, Asp172, Gly391, Arg533 e Ser874). Enquanto os primeiros dois estão numa região muito desordenada para a qual a densidade electrónica é muito fraca, não permitindo o seu correcto posicionamento, os últimos três estão posicionados de acordo com os dados experimentais, sendo que dois deles (Gly391 e Arg533) se encontram bastante próximos do centro activo. Estes resíduos adoptam esta conformação estereoquimicamente menos favorável pois estão a acomodar a unidade pterina do cofactor.

De um modo geral, o modelo explica os dados experimentais muito satisfatoriamente, não apresentando erros graves a nível das restrições geométricas que se devem ter em atenção na estrutura de uma proteína.

4.1.3.4. Caracterização da estrutura

Analisando a estrutura final da *DgAOR*_H₂O₂, conclui-se que esta apresenta uma forma globular com um diâmetro aproximado de 75 Å, apresentando dois cofactores Fe-S e um cofactor pterínico, como foi descrito anteriormente para a estrutura nativa cristalizada com isopropanol (subcapítulo 1.1.5.1). A estrutura é muito semelhante à estrutura nativa, sendo que uma superposição estrutural das duas estruturas apresenta um RMSD de 0,146 Å em relação a todos os átomos da cadeia principal da proteína. No entanto, foi possível distinguir algumas

diferenças, nomeadamente ao nível das cadeias laterais da proteína, sendo que algumas apresentam rotâmeros diferentes e foram adicionadas conformações alternadas que não existiam na estrutura nativa. Outras características adicionais à estrutura nativa foram modeladas: dois iões Na^+ , conforme sugestão do programa STAN, duas moléculas de isopropanol fora do centro activo, uma na superfície e outra no interior da proteína (substituindo um ião Cl^- na estrutura nativa), um ião Mg^{2+} e um ião Cl^- . No total a estrutura apresenta três iões Mg^{2+} , cada um coordenado por seis ligandos entre moléculas de água e átomos de oxigénio da cadeia peptídica, três iões Cl^- que interagem com regiões carregadas positivamente da superfície da proteína, dois iões Na^+ e duas moléculas de isopropanol. A presença de isopropanol na estrutura indica que o protocolo de remoção utilizado não removeu por completo o isopropanol de solução, mas foi suficiente para o remover do centro activo da enzima comprovando a sua eficácia em estudos de *soaking* utilizando a DgAOR.

Analisando o mapa de diferença $F_o - F_c$ e os factores de temperatura para o átomo de molibdénio no centro activo da enzima, verificou-se que este apresenta um pico negativo de densidade assim como um factor B elevado comparado com os factores B dos átomos vizinhos. Por estes motivos, o seu factor de ocupação foi alterado para 0,8. Esta diminuição no factor de ocupação pode ser devida ao baixo conteúdo de Mo na amostra de proteína ou ao dano provocado pela radiação ionizante dos raios-X durante a experiência de difracção. Outro efeito do dano da radiação pode ser observado na cisteína 271, a qual apresenta densidade electrónica suplementar perto do átomo de enxofre, sugerindo que este se encontra ligado a um outro átomo. Esta foi interpretada como sendo uma hidroxicisteína, formada através da reacção com radicais de oxigénio produzidos pela radiação ionizante intensa do Síncrotrão (Figura 4.4). Enquanto neste caso particular, a presença de peróxido de hidrogénio em solução pode ter influenciado a formação destes radicais, esta não é essencial, pois verificou-se o mesmo resultado noutras estruturas determinadas sem a presença de H_2O_2 . Todas as outras cisteínas na estrutura foram analisadas, nenhuma delas apresentando este comportamento.

Como referido no capítulo anterior, no centro activo da enzima é possível encontrar densidade electrónica relativa a uma molécula diatómica junto ao átomo de molibdénio, no local onde se encontra o ligando OH/OH₂ na estrutura nativa. Neste local foi modelada uma molécula de H_2O_2 coordenada ao molibdénio com uma geometria η^2 , conforme apresentado na Figura 4.5. Esta coordenação do centro activo é idêntica à obtida para a experiência de *soaking* da DgAOR com ditionito de sódio e sulfureto de sódio efectuada anteriormente, a qual se pode observar na Figura 4.6.

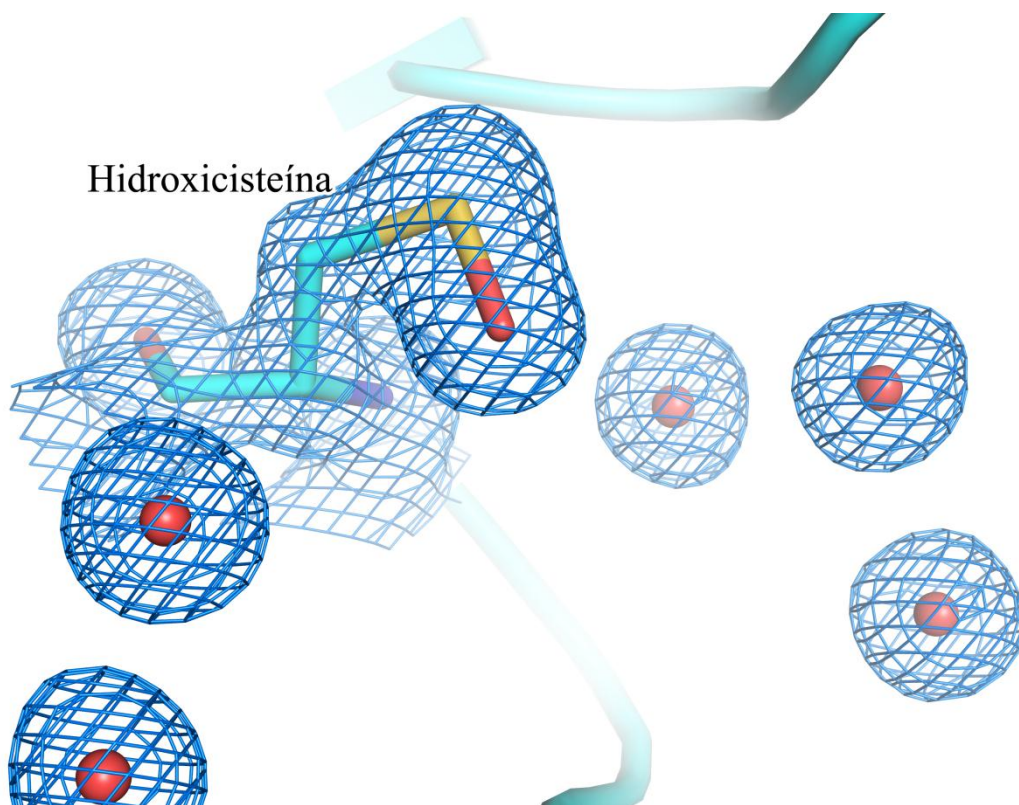


Figura 4.4. Densidade electrónica observada para a cisteína 271, a qual foi modelada como uma hidroxicisteína, e moléculas de água na sua vizinhança. Código de cores: Carbono – ciano; Oxigénio – vermelho; Enxofre – amarelo. A azul está representado o mapa $2mFo-DFc$, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

Tendo em conta as Figuras 4.5 e 4.6 e a Tabela 4.7 observa-se que o oxigénio Ox1 do peróxido de hidrogénio encontra-se mais próximo do molibdénio, a uma distância de 1,97 Å, enquanto o oxigénio Ox2 encontra-se um pouco mais afastado a 2,30 Å. Analisando os factores B da molécula de H_2O_2 em comparação com os dos átomos da vizinhança, concluiu-se que o Ox2 apresenta um factor B mais elevado que a média, tendo-se por isso diminuído o seu factor de ocupação para 0,5, enquanto o oxigénio Ox1 tem uma ocupação de 1. Isto justifica-se se em 50% das moléculas do cristal se encontrar uma molécula de peróxido e nas restantes o centro activo se manter inalterado em relação à proteína nativa, pois o Ox1 encontra-se no local do ligando lábil OH/OH₂. Este resultado pode ser devido à concentração final de H_2O_2 na gota ser baixa (50 µM).

É possível também observar uma diferença nas distâncias de ligação entre os oxigénios OR1 (equatorial) e OM1 (apical) e o molibdénio no centro activo, nas duas estruturas. Nos cristais mergulhados em agentes redutores, as distâncias são de 1,97 e 2,07 Å, respectivamente, indicando que se trata de ligações simples, enquanto na estrutura *DgAOR_H₂O₂* estas são de 1,81 e 1,85 Å, tratando-se de ligações duplas, semelhantes à proteína nativa. Assim verifica-se que na presença de agentes redutores o molibdénio se encontra num estado de oxidação mais baixo, ligado a dois grupos hidroxilo e ao peróxido de hidrogénio.

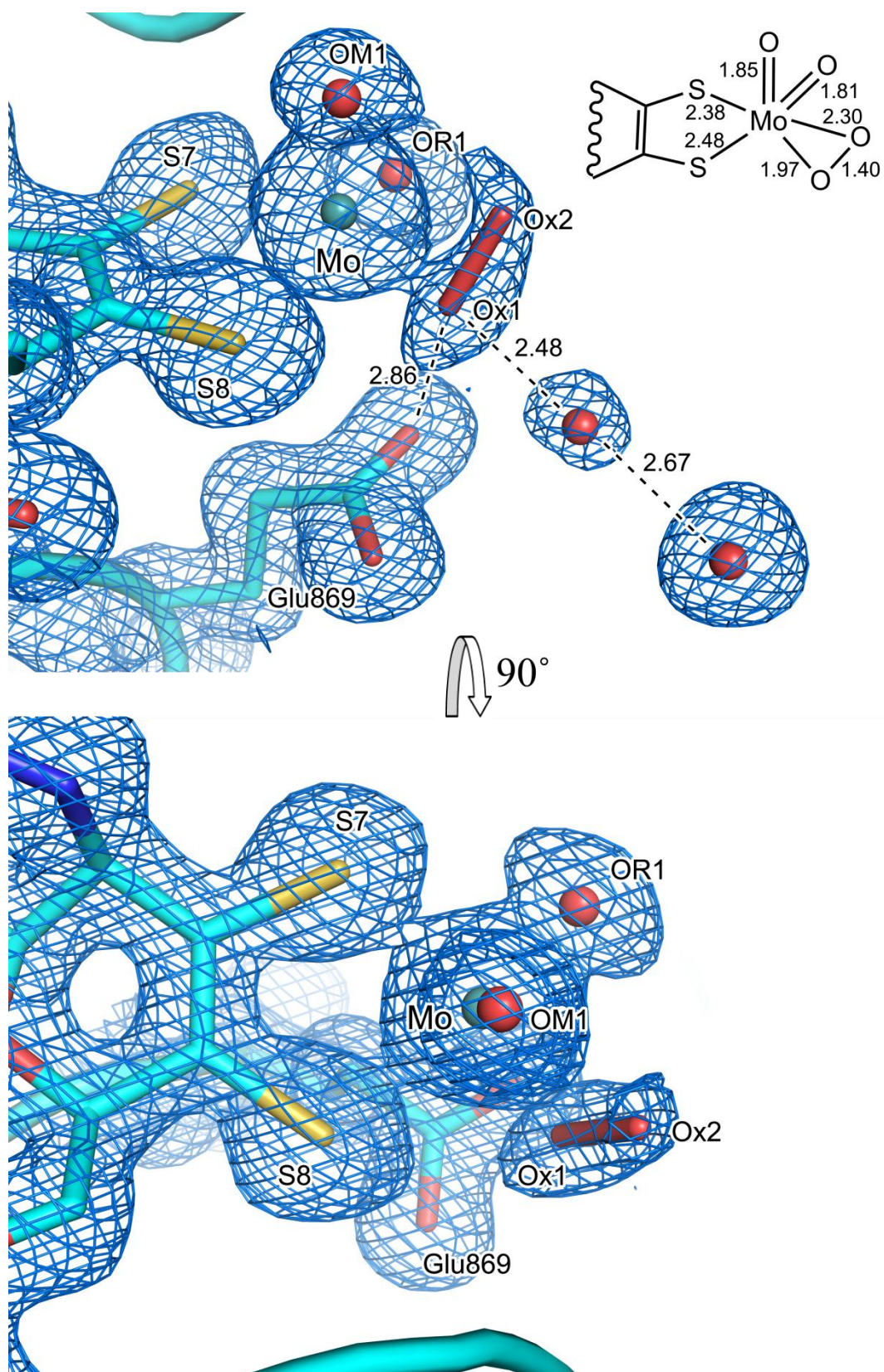


Figura 4.5. Estrutura cristalográfica do centro activo da amostra *DgAOR_H2O2*, em duas orientações diferentes. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.

Tabela 4.7. Distância entre o átomo de molibdénio e os ligandos presentes na sua esfera de coordenação, no centro activo, para as duas estruturas e para a estrutura nativa.

Distância (Å)	Nativa (1VLB)	DgAOR_H ₂ O ₂	DgAOR (<i>soaking</i> Na ₂ S ₂ O ₄ e Na ₂ S)
Mo – S7 (molibdopterina)	2,41	2,38	2,39
Mo – S8 (molibdopterina)	2,49	2,48	2,48
Mo – OM1 (apical)	1,74	1,85	2,07
Mo – OR1 (equatorial)	1,79	1,81	1,97
Mo – OM2 (equatorial)	1,99	-	-
Mo – Ox1 (H ₂ O ₂)	-	1,97	2,03
Mo – Ox2 (H ₂ O ₂)	-	2,30	2,46

A semelhança entre as duas estruturas analisadas sugere que o *soaking* com ditionito de sódio e sulfureto de sódio levou à formação de uma molécula de peróxido de hidrogénio ligada ao centro activo da enzima. Os dados obtidos com o *soak* com peróxido de hidrogénio corroboram esta hipótese, apesar de não serem suficientes para afirmar esta possibilidade. Assim, estudos por espectroscopia de RPE foram efectuados em colaboração com o Dr. Pablo González, incubando a enzima com os mesmos agentes redutores em condições aeróbias. Estes estudos, recorrendo a *spin-traps*, mostraram que não se formam radicais superóxido ou hidroxilo, o que sugere que a molécula que coordena o centro activo é mesmo o peróxido de hidrogénio.

Deste modo prova-se que a incubação da DgAOR com Na₂S₂O₄ e Na₂S gera espécies reactivas de oxigénio (ROS), as quais se ligam ao centro activo da proteína, inibindo-a. Estes resultados explicam os estudos cinéticos discutidos anteriormente, onde se concluiu que ocorre inicialmente uma activação da proteína seguida de uma inactivação. Quanto maior for o tempo de incubação da proteína com estes agentes, mais ROS se formam e maior é esta inactivação. A inactivação mais lenta em condições anaeróbias também pode ser explicada, pois a concentração de oxigénio molecular é muito menor, levando à produção mais lenta de ROS. A causa da activação inicial da proteína ainda é desconhecida, sendo que a análise das estruturas cristalográficas não revelou diferenças estruturais significativas que a possam explicar.

Os resultados obtidos são de enorme relevância quando se considera a elevada homologia entre esta proteína e outras da família como a XO e a AO. O envolvimento da Xantina Oxidase na produção de ROS é conhecido, utilizando o oxigénio como aceitador final de electrões ao nível do seu cofactor flavínico (FAD) e gerando radical superóxido.⁹⁷ As estruturas cristalográficas

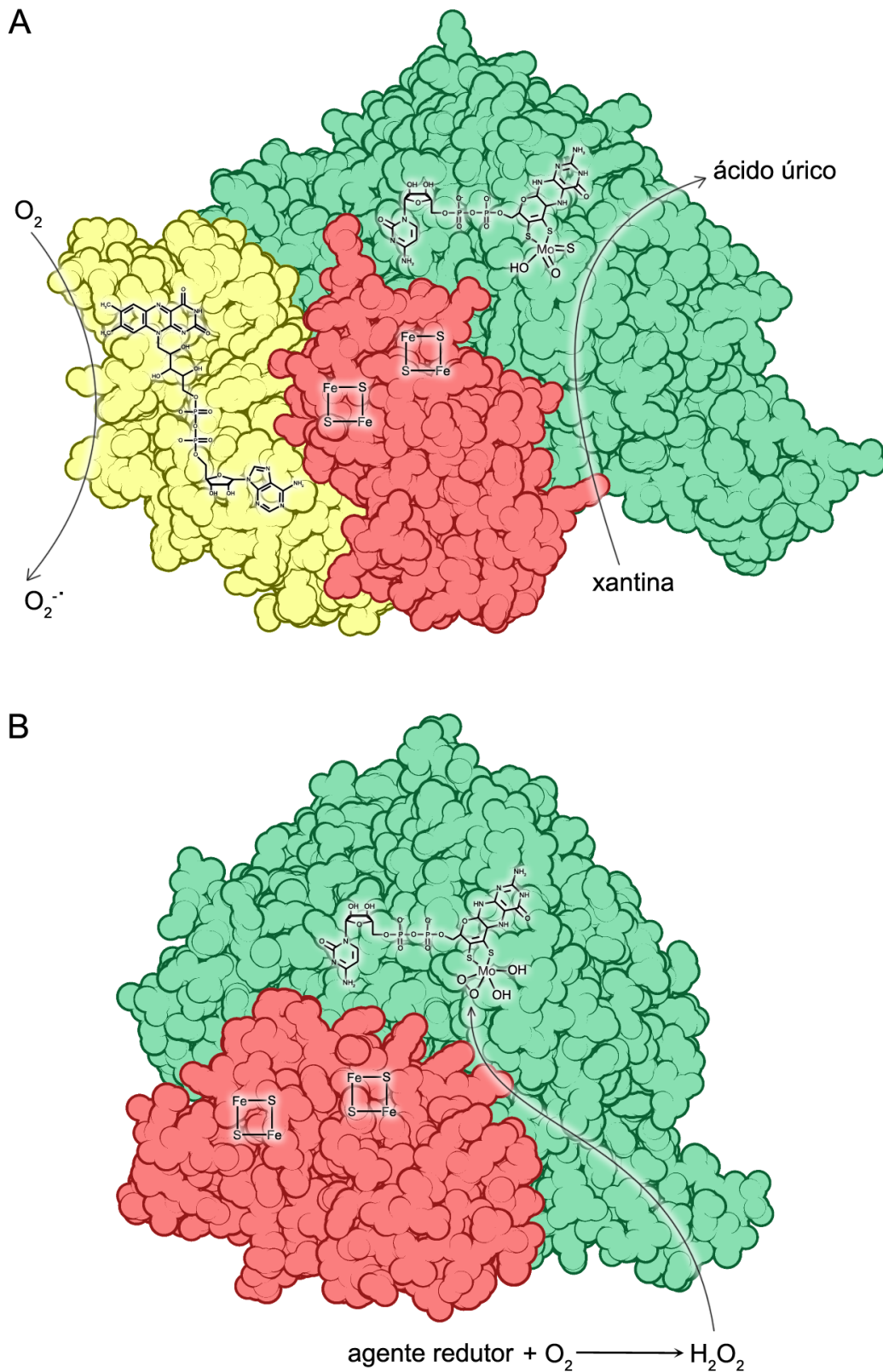


Figura 4.7. Comparação entre o metabolismo das ROS no caso da XO (A) e os resultados aqui apresentados para a *DgAOR* (B). Código de cores: Domínio do cofactor do Mo – verde; Domínio dos centros Fe-S – vermelho; Domínio do cofactor flavínico (XO apenas) – amarelo. Representação esquemática preparada utilizando o programa Qutemol.¹⁰⁰

determinadas e os dados cinéticos obtidos indicam que a *DgAOR* pode também estar envolvida no metabolismo das ROS. No entanto, ao contrário do caso da XO, o mecanismo de produção de ROS não é conhecido, sabendo-se apenas que o peróxido de hidrogénio se pode ligar ao seu centro activo (Figura 4.7). Se o mesmo se verificar em relação à Aldeído Oxidase (AO) de mamíferos, a qual apresenta um papel importante de destoxificação de aldeídos altamente reactivos, a sua inibição por ROS provenientes de processos inflamatórios ou do envelhecimento pode levar à acumulação de aldeídos e conduzir à carbonilação de proteínas.^{98,99} Estas hipóteses devem ser investigadas antes de se poderem tirar conclusões relativas ao mecanismo de acção desta família de enzimas.

4.1.4. Aldeído Oxidorredutase em complexo com substratos

Neste subcapítulo vão ser analisados os resultados referentes aos ensaios de *soaking* com benzaldeído, formaldeído, acetaldeído e xantina. O procedimento utilizado para a resolução destas estruturas foi idêntico ao descrito no subcapítulo anterior (4.1.3). Todos os conjuntos de dados recolhidos apresentados na Tabela 4.2 (subcapítulo 4.1.2) foram processados, a estrutura foi resolvida e foi efectuado um ciclo de refinamento. No entanto apenas se prosseguiu com o refinamento de três conjuntos de dados: *DgAOR_benz* ao fim de uma hora, *DgAOR_form* ao fim de uma hora e *DgAOR_acet* ao fim de uma hora, todos obtidos em condições aeróbias. Por este motivo vão ser apresentados os dados estatísticos apenas em relação a estes três conjuntos de dados.

4.1.4.1. Experiência de difracção

Os vários conjuntos de dados de *DgAOR* foram recolhidos e processados no grupo espacial $P6_122$, de acordo com os passos anteriormente descritos. As estatísticas referentes a este procedimento encontram-se na Tabela 4.8.

Analisando os valores da tabela, verifica-se que os cristais são muito equivalentes, apresentando estatísticas muito semelhantes. A maior diferença observável corresponde aos limites de resolução, sendo que os cristais *DgAOR_benz* e *DgAOR_acet* difractaram a uma resolução mais elevada ($\sim 1,5$ Å) do que o cristal *DgAOR_form* ($\sim 1,7$ Å).

Todos os parâmetros de qualidade do conjunto de dados (*completeness*, redundância, $I/\sigma(I)$ médio e R_{merge}) encontram-se dentro dos valores óptimos, indicando uma boa estratégia de recolha de dados e elevada qualidade dos cristais. Como é característico dos cristais da *DgAOR*, estes apresentam um conteúdo de solvente de 50% e uma molécula por unidade assimétrica.

Tabela 4.8. Estatísticas de recolha de dados para os cristais *DgAOR_benz*, *DgAOR_form* e *DgAOR_acet*. Os valores entre parêntesis correspondem à última camada de resolução.

Cristal	<i>DgAOR_benz</i>	<i>DgAOR_form</i>	<i>DgAOR_acet</i>
Comprimento de onda (Å)	0,976	0,976	0,976
Grupo espacial	<i>P6₁22</i>	<i>P6₁22</i>	<i>P6₁22</i>
Célula unitária (Å)	a,b = 143,2 c = 162,2	a,b = 143,2 c = 162,0	a,b = 143,4 c = 162,0
Limites de resolução (Å)	49,58–1,49 (1,57–1,49)	57,90–1,69 (1,78–1,69)	49,51–1,50 (1,58–1,50)
Nº de reflexões observadas	897 316 (146 817)	1 085 208 (165 294)	1 004 257 (153 510)
Nº de reflexões únicas	158 223 (22 898)	108 793 (15 666)	154 851 (22 429)
<i>Completeness</i> (%)	99,9 (100,0)	100,0 (100,0)	99,5 (99,9)
Redundância	5,7 (6,4)	10,0 (10,6)	6,5 (6,8)
<i>I</i> / σ (<i>I</i>) médio	20,9 (7,7)	20,7 (6,1)	20,8 (7,9)
<i>R</i> _{merge} ^a (%)	4,9 (21,0)	6,5 (29,9)	5,2 (21,2)
<i>VM</i> ^b - Parâmetro de Matthews (Å ³ /Da)	2,47	2,47	2,47
Conteúdo de solvente ^b (%)	50,3	50,2	50,3

(a) $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, onde $I_i(hkl)$ é a intensidade integrada de uma determinada reflexão i e $\langle I(hkl) \rangle$ é a intensidade média de múltiplas reflexões relacionadas por operações de simetria.

(b) $VM = \frac{V}{Z \times Mp}$, onde V é o volume da célula unitária, Z é o número de unidades assimétricas presentes na célula unitária e Mp é a massa molecular da proteína; Conteúdo de solvente = $\left(1 - \frac{1,23}{VM}\right) \times 100$

4.1.4.2. Resolução das estruturas

Com o objectivo de calcular o ângulo da fase de cada reflexão e, conseqüentemente, os factores de estrutura necessários ao cálculo da densidade electrónica, utilizou-se o método de substituição molecular, tendo-se recorrido ao programa PHASER. A estrutura utilizada como modelo foi a que se usou anteriormente (Código PDB: 1VLB³²).

Através das funções de rotação e translação, foram encontradas soluções para os três conjuntos de dados no grupo espacial *P6₁22*, apresentando valores de LLG e *Z-score* bastante elevados (Tabela 4.9). Analisando estes resultados verifica-se que as estruturas estão resolvidas, prosseguindo-se com a utilização do programa DM e respectivo melhoramento da densidade electrónica calculada anteriormente.

Tabela 4.9. Valores de Z-score e LLG para as funções de rotação e translação obtidos utilizando o programa PHASER, para os conjuntos de dados *DgAOR_benz*, *DgAOR_form* e *DgAOR_acet*.

Conjunto de dados	<i>DgAOR_benz</i>	<i>DgAOR_form</i>	<i>DgAOR_acet</i>
Z-score (rotação)	29,6	26,5	29,4
LLG (rotação)	579,8	558,0	577,4
Z-score (translação)	86,5	87,1	84,7
LLG (translação)	8 766,0	8 664,5	8 769,0

4.1.4.3. Refinamento e validação das estruturas

O refinamento das estruturas foi efectuado recorrendo ao programa REFMAC5. Foi realizado um refinamento inicial com adição automática de moléculas de água pelo programa COOT para todos os conjuntos de dados obtidos com benzaldeído, formaldeído, acetaldeído e xantina em condições aeróbias e anaeróbias. Após refinamento, a densidade electrónica no centro activo de todos os modelos foi inspeccionada. Os resultados obtidos, assim como os factores R e R_{free} e a existência de densidade electrónica extra perto do centro activo são apresentados na Tabela 4.10.

Os factores R e R_{free} apresentados após este refinamento inicial são bastante satisfatórios, tendo possibilitado a interpretação dos mapas de densidade electrónica resultantes com alguma confiança. Em condições aeróbias verificou-se a presença dos substratos no centro activo da enzima em cristais com diferentes tempos de incubação, tendo-se seleccionado apenas uma estrutura de cada para se prosseguir com o refinamento, como referido anteriormente. Nas estruturas referentes aos cristais com xantina não foi possível observar densidade electrónica extraordinária no centro activo da enzima, confirmando a especificidade da *DgAOR* para a ligação de aldeídos. Este resultado evidencia que, apesar de esta enzima apresentar o cofactor pterínico e o centro de molibdénio muito semelhante à XO, o ambiente envolvente do centro activo apresenta um papel crucial para a estabilização das moléculas de substrato.

Para o benzaldeído confirmou-se que os ensaios de *soaking* ao fim de 1 e 3 horas resultaram, tendo-se seleccionado o primeiro para se prosseguir o refinamento. Ao fim de 24 horas não se verificou a presença de densidade electrónica extraordinária no centro activo, o que pode ser devido ao substrato já não se encontrar no centro activo da enzima ao fim deste tempo, tendo sido totalmente convertido em produto e libertado para fora do interior da proteína.

Tabela 4.10. Resultados obtidos no centro activo de todos os conjuntos de dados recolhidos, após um refinamento inicial, assim como os factores R e R_{free} iniciais.

Composto	Condições	Duração do soaking (h)	Factor R (%)	Factor R_{free} (%)	Centro activo
Benzaldeído	Aeróbias	1	18,54	22,42	Anel aromático na proximidade do Mo
	Aeróbias	3	18,71	23,03	Anel aromático na proximidade do Mo
	Aeróbias	24	19,07	22,27	Vazio
	Anaeróbias	1	19,39	24,51	Anel aromático na proximidade do Mo
	Anaeróbias	1 + 4	15,90	20,46	Vazio
Formaldeído	Aeróbias	1	19,45	24,34	Densidade electrónica atípica ligada ao Mo
	Anaeróbias	1	16,77	21,33	Densidade electrónica atípica ligada ao Mo
Acetaldeído	Aeróbias	1	18,72	23,27	Densidade electrónica atípica ligada ao Mo
	Aeróbias	3	18,73	23,52	Densidade electrónica atípica ligada ao Mo
	Anaeróbias	1 + 4	19,72	23,64	Densidade electrónica atípica ligada ao Mo
Xantina	Aeróbias	1	18,38	22,14	Vazio
	Aeróbias	24	19,05	23,34	Vazio
	Anaeróbias	1	16,41	20,09	Vazio

Relativamente ao formaldeído, apenas se obteve um conjunto de dados (duração de 1 hora) o qual apresenta densidade electrónica perto do centro activo, sugerindo a presença de uma pequena molécula ligada ao molibdénio.

Tal como no caso do formaldeído, os resultados relativos ao acetaldeído ao fim de 1 e 3 horas sugerem a presença de outra pequena molécula no centro activo da enzima, observando-se

densidade relativa a quatro átomos. A densidade electrónica é idêntica para as duas estruturas, tendo-se seleccionado o ensaio com duração de 1 hora para se prosseguir com o refinamento, uma vez que este apresenta melhor resolução.

Em relação aos ensaios realizados em condições anaeróbias verificou-se que se obtiveram resultados muito semelhantes aos obtidos para os ensaios aeróbios, observando-se densidade electrónica idêntica no centro activo. Estes resultados não eram esperados, pois na ausência de oxigénio ou de outro aceitador de electrões a enzima não consegue completar o seu ciclo catalítico, sendo mais provável obter uma estrutura cristalográfica onde se captura o seu substrato interagindo com o centro activo. Tal pode ter ocorrido devido à ineficaz desoxigenação das soluções utilizadas para a realização destes ensaios ou devido ao facto dos cristais terem sido expostos ao oxigénio aquando da sua congelação. Por este motivo, não se avançou com o refinamento destas estruturas, dando-se prioridade às estruturas aeróbias. Novas experiências serão efectuadas dentro da câmara anaeróbia mas, nessa altura, o congelamento dos cristais também será efectuado na ausência de oxigénio. É de notar que a manipulação de cristais dentro da câmara é uma tarefa arduosa que, para o caso da *DgAOR* levou em experiências anteriores a uma diminuição significativa da resolução. Mesmo assim, novos ensaios serão realizados na tentativa de ultrapassar estas dificuldades.

Proseguiu-se com o refinamento das três estruturas mencionadas utilizando os programas COOT, para a visualização e ajuste dos resíduos, e REFMAC5, para o refinamento. Foram efectuados ciclos de TLS e *restrained refinement*, tal como se procedeu para a estrutura descrita anteriormente (*DgAOR_H₂O₂*). Este processo ainda está em execução atendendo que os resultados obtidos ainda estão a ser avaliados. As estatísticas relativas ao refinamento actual destas estruturas encontram-se na Tabela 4.11.

Observando os factores globais de avaliação de um modelo cristalográfico, R e R_{free} , verifica-se que estes são muito semelhantes entre as três estruturas, sendo ligeiramente mais elevados para a estrutura *DgAOR_form*, diferença esta devida à pior resolução apresentada. Tendo este factor em consideração, todas as estruturas apresentam um factor R baixo (~13%), indicando que estão de acordo com os dados experimentais, assim como um factor R_{free} também baixo (15,60% e 15,61% para *DgAOR_benz* e *DgAOR_acet* e 16,87% para *DgAOR_form*) que indica a validade do modelo em relação às reflexões independentes não utilizadas no cálculo da estrutura.

Tal como no caso da estrutura *DgAOR_H₂O₂*, os valores de RMSD para o comprimento de ligação encontram-se superiores ao ideal (0,025 Å, 0,023 Å e 0,027 Å), o que significa que o refinamento não foi muito rigoroso em relação às restrições geométricas do modelo. Os valores de RMSD para o ângulo de ligação (2,68°, 2,31° e 2,66°) encontram-se dentro dos valores aceites para estruturas de alta resolução.

Tabela 4.11. Estatísticas de refinamento actuais para os modelos *DgAOR_benz*, *DgAOR_form* e *DgAOR_acet*.

Conjunto de dados	<i>DgAOR_benz</i>	<i>DgAOR_form</i>	<i>DgAOR_acet</i>
Limites de resolução (Å)	49,58–1,49	57,90–1,69	49,51–1,50
Factor R (%)	12,96	13,45	12,85
Nº de reflexões	150 228	103 277	146 999
Factor R_{free} (%)	15,60	16,87	15,61
Nº de reflexões (R_{free})	7 930	5 428	7762
Nº de resíduos	907	907	907
Nº de átomos	8 479	8 126	8 534
RMSD para o comprimento de ligação (Å)	0,025	0,023	0,027
RMSD para o ângulo de ligação (°)	2,68	2,31	2,66

Apesar do refinamento não estar concluído para estas estruturas, os valores dos factores globais da qualidade dos modelos indicam que se pode fazer uma avaliação do seu centro activo com alguma confiança.

4.1.4.4. Caracterização das estruturas

As três estruturas obtidas são globalmente muito semelhantes entre si e à estrutura nativa, apresentando as características já discutidas para a estrutura *DgAOR_H₂O₂* no subcapítulo 4.1.4.4. Tal pode ser confirmado pelos valores baixos de RMSD da sua superposição com a estrutura nativa apresentados na Tabela 4.12.

Tabela 4.12. RMSD de superposição estrutural das estruturas *DgAOR_benz*, *DgAOR_form* e *DgAOR_acet* com a estrutura nativa 1VLB, em relação a todos os átomos da cadeia principal da proteína.

Estruturas	<i>DgAOR_benz</i>	<i>DgAOR_form</i>	<i>DgAOR_acet</i>
RMSD (Å)	0,148	0,148	0,149

Como verificado para a estrutura *DgAOR_H₂O₂* foram modelados iões Mg^{2+} e Cl^- e moléculas de isopropanol nas mesmas posições, assim como uma hidroxicisteína na posição 271 da cadeia polipeptídica. A principal diferença nestas estruturas encontra-se no centro activo da enzima e sua vizinhança.

Para o caso da estrutura *DgAOR_benz* identificou-se densidade electrónica que corresponde claramente a um anel aromático, estabilizado pelos resíduos aromáticos Phe425 e Tyr535 que se

encontram muito perto do centro activo, no local onde se encontra a molécula de isopropanol na estrutura nativa 1VLB. A cadeia lateral destes resíduos aromáticos estabelece interacções π - π com o anel aromático do benzaldeído, um efeito conhecido como empilhamento ou π -stacking, numa conformação paralela deslocada (Figura 4.8). Apesar da alta resolução dos dados experimentais, é impossível identificar densidade electrónica correspondente ao grupo carbonilo da função aldeído da molécula, reconhecendo-se apenas uma porção de densidade electrónica extraordinária na direcção do centro activo a qual poderá pertencer ao átomo de carbono deste grupo. Na Figura 4.9A apresenta-se uma hipótese de orientação do benzaldeído com o átomo de oxigénio numa posição onde não existe densidade electrónica bem definida, na direcção do centro activo. Esta orientação, no entanto, implica uma distância de apenas 1,44 Å entre os átomos de oxigénio do benzaldeído e do ligando equatorial lábil do molibdénio. Esta distância tão curta entre os átomos é muito pouco provável, devendo ser analisada com cuidado. Outra possibilidade é ter o anel numa segunda orientação onde o grupo carbonilo se encontra mais afastado do centro activo (Figura 4.9B). Uma explicação para a fraca densidade do grupo funcional aldeído nestes dados é a molécula apresentar-se em diferentes posições nas várias moléculas de proteína do cristal. Estudos por ^{13}C -RMN foram efectuados em colaboração com o

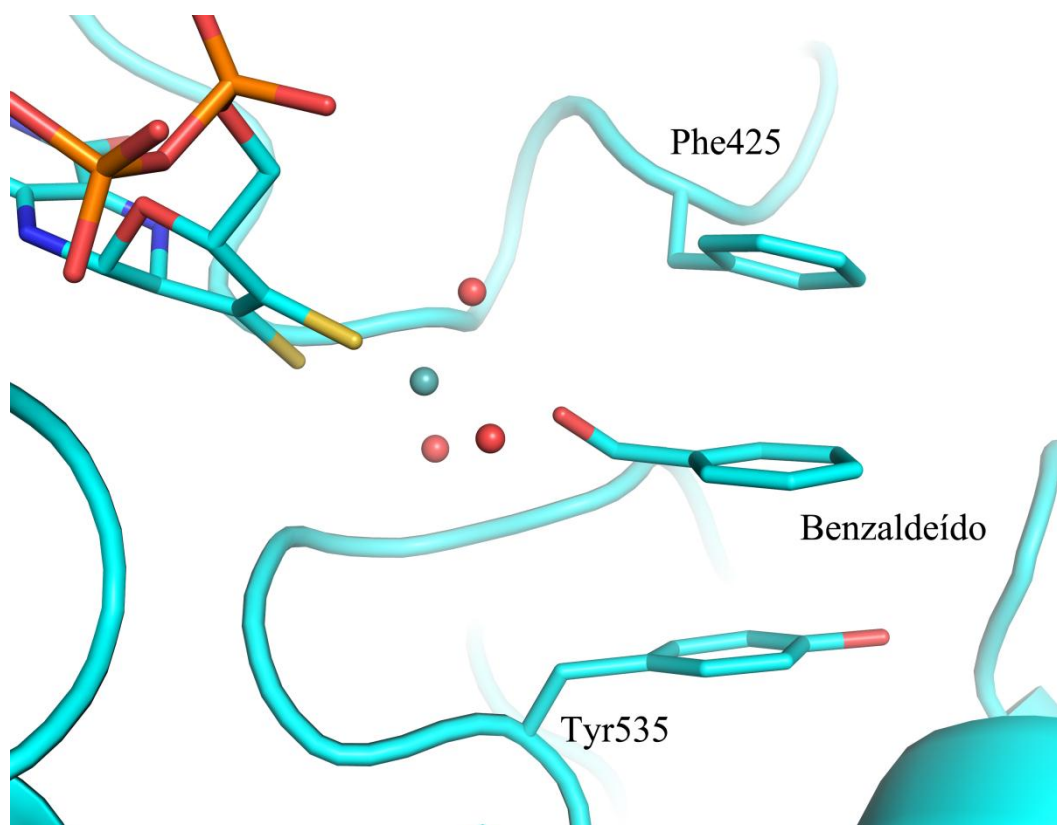


Figura 4.8. Estabilização do anel aromático do benzaldeído pelos resíduos Phe425 e Tyr535. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Fósforo – laranja; Enxofre – amarelo; Molibdénio – azul-esverdeado. Esta figura foi preparada utilizando o programa PyMOL.

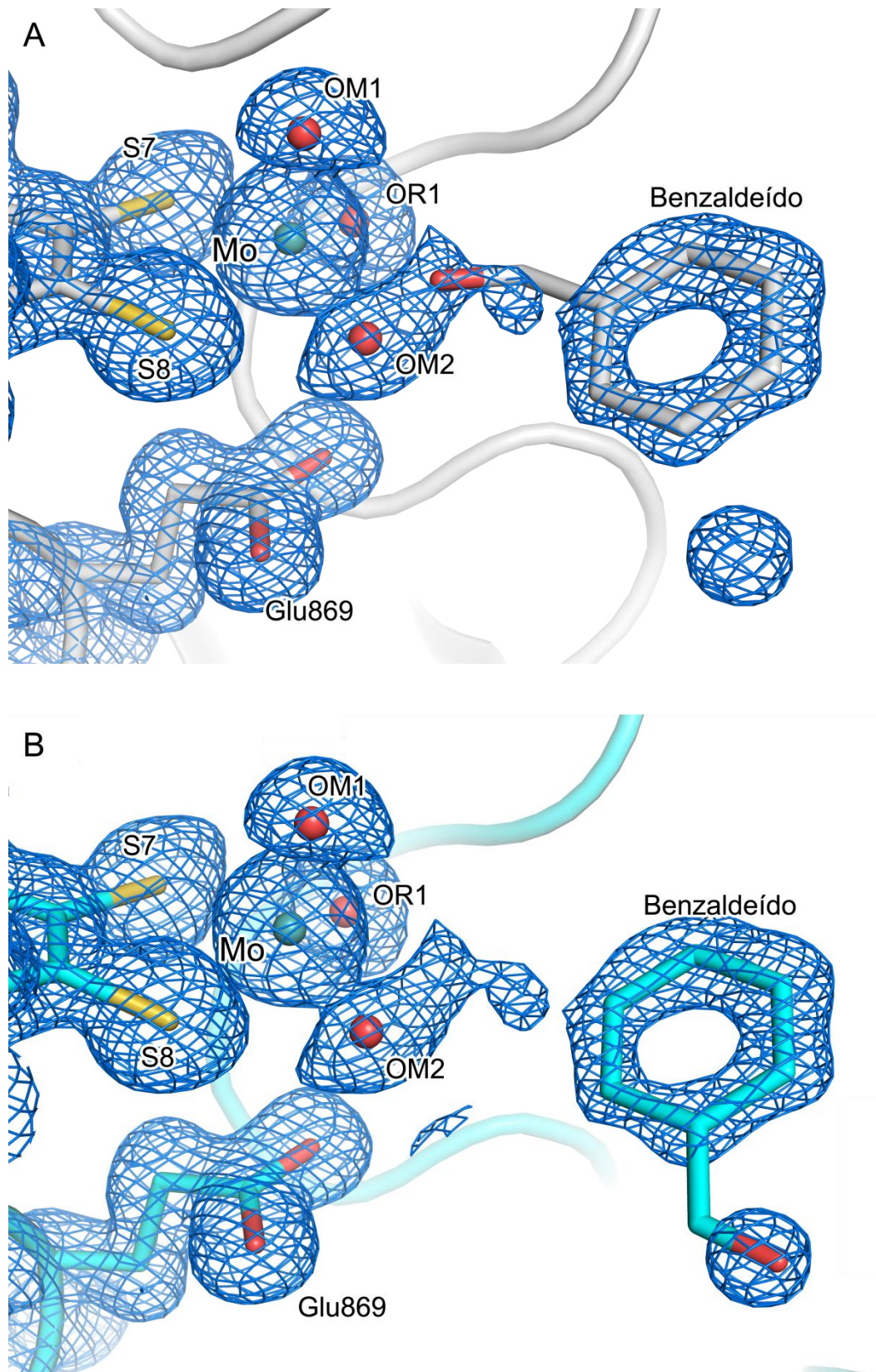


Figura 4.9. Estrutura cristalográfica do centro activo da amostra *DgAOR_benz*. Devido à incerteza da posição do grupo carbonílico, a molécula de benzaldeído é aqui representada em duas conformações: (A) com o grupo carbonílico orientado para o centro activo e (B) orientado numa direcção oposta. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ. Esta figura foi preparada utilizando o programa PyMOL.

Dr. Pablo González onde se confirmou que o único produto da reacção da *DgAOR* com o benzaldeído é o ácido benzóico. Assim, nesta estrutura, apenas uma destas duas moléculas poderão corresponder à densidade electrónica obtida. De modo a tentar esclarecer esta questão vai-se proceder ao refinamento de todas as outras estruturas obtidas com benzaldeído até ao momento, na tentativa de encontrar dados onde a visualização e interpretação deste pormenor seja mais fácil. Também se irão realizar novos *soaks* com diferentes concentrações de benzaldeído e outros análogos de substratos para recolher novos dados cristalográficos.

Outra característica interessante desta estrutura é a densidade electrónica à qual se pensa corresponder a uma segunda molécula de benzaldeído presente no canal de entrada dos substratos, muito próxima da primeira molécula de benzaldeído (Figura 4.10). Provavelmente relacionado com desordem, esta densidade electrónica não é tão clara como a primeira, sendo difícil assegurar que a posição da molécula é a correcta, ou que se trata realmente de uma molécula de benzaldeído. Mais ciclos de refinamento poderão vir a melhorar esta zona da proteína e elucidar este ponto. Caso venha a ser comprovado, este resultado dá fortes evidências sobre qual o caminho percorrido pelo substrato e produto de reacção aquando de um ciclo catalítico.

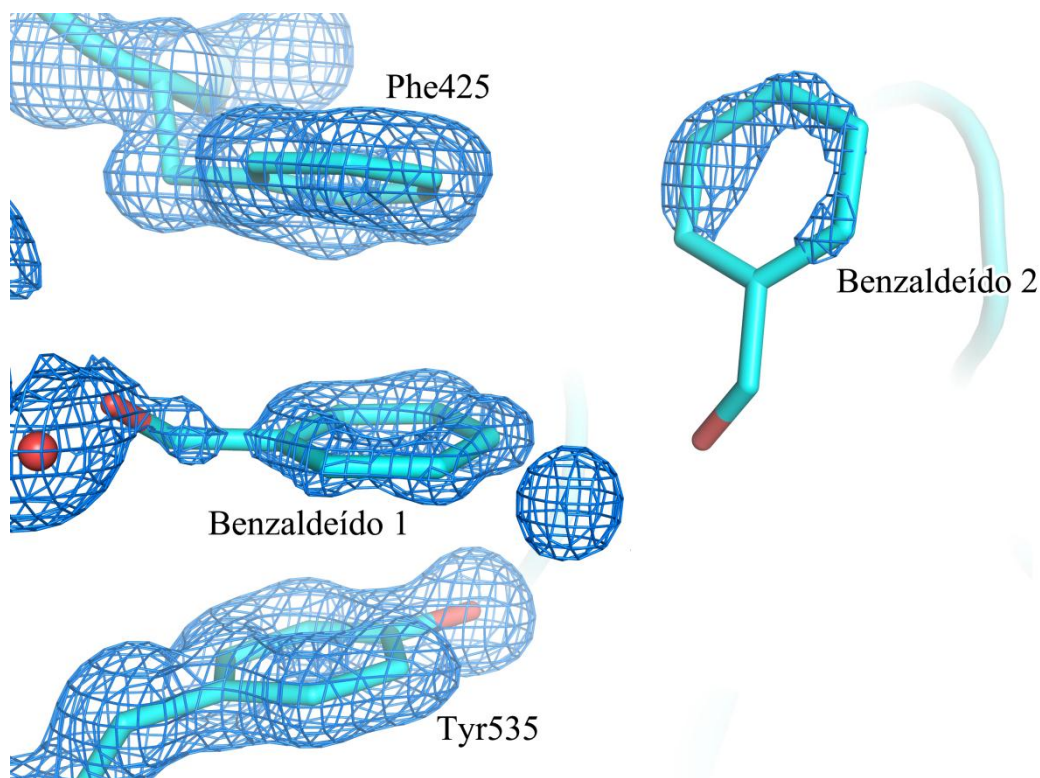


Figura 4.10. Densidade electrónica relativa a uma putativa segunda molécula de benzaldeído, adjacente à primeira, no canal de acesso ao centro activo. Código de cores: Carbono – ciano; Oxigénio – vermelho. A azul está representado o mapa $2mFo-DFc$, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

Em relação à estrutura *DgAOR_form* identificou-se, no centro activo da enzima, densidade electrónica correspondente a três átomos ligados directamente ao átomo de molibdénio no local de ligação do ligando lábil OH/OH₂, no mesmo local onde se liga a molécula de arsenito em estudos realizados anteriormente.³⁶ Foi então modelada uma molécula de formato, produto da reacção entre a *DgAOR* e o formaldeído, com uma coordenação ao molibdénio através de um átomo de oxigénio, a 2,09 Å de distância (Figura 4.11). Após um ciclo de refinamento, verificou-se que o factor B para este átomo de oxigénio é semelhante aos átomos vizinhos (~18 Å²), enquanto que, para os átomos de carbono e oxigénio mais afastados do centro, é um pouco mais elevado (aproximadamente 27 e 21 Å², respectivamente). Estes valores sugerem que esta molécula pode não estar presente em todas as moléculas de proteína no cristal, sendo que nas restantes se encontra o ligando OH/OH₂, na posição mais próxima do molibdénio. O oxigénio mais afastado do molibdénio encontra-se a uma distância de 2,75 Å de uma molécula de água da cadeia de águas conservadas que se encontra perto do centro activo e se pensam estar envolvidas no mecanismo de reacção da *DgAOR* e com o Glu869 a 2,74 Å.

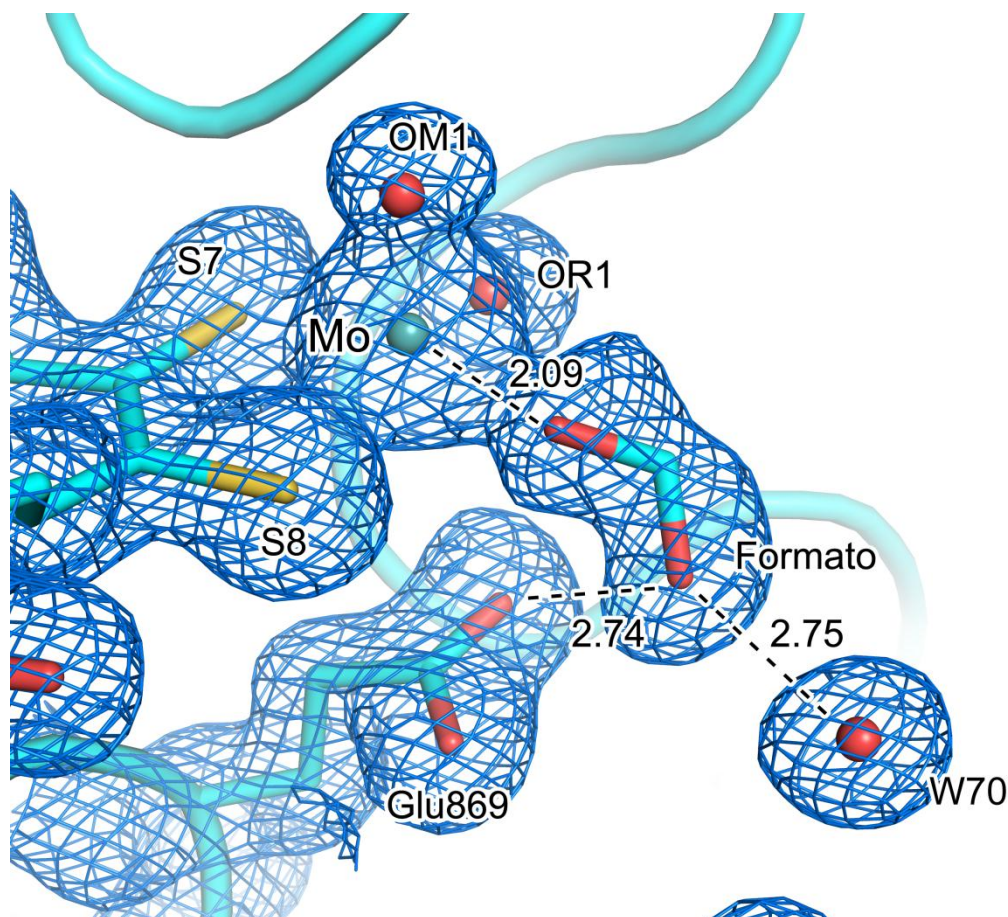


Figura 4.11. Estrutura cristalográfica do centro activo da amostra *DgAOR_form*. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Molibdénio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.

Apesar dos resultados obtidos serem bastante convincentes em relação à orientação deste substrato no centro activo da enzima, tem que se ter em consideração o facto do formaldeído em solução aquosa formar metileno glicol com uma constante de equilíbrio elevada, sendo que cerca de 80% do formaldeído em solução se apresenta nesta forma (Figura 4.12). Como cristalograficamente não se consegue distinguir os átomos de hidrogénio a esta resolução, é difícil diferenciar entre o formato e o metileno glicol, sendo que as diferenças mais importantes entre as duas moléculas correspondem ao ângulo e comprimento de ligação. Enquanto que, para o formato, o ângulo formado pelos três átomos deve ser de 120° e o comprimento das ligações de cerca de $1,2 \text{ \AA}$, para o metileno glicol estes devem ser de 104° e $1,4 \text{ \AA}$, respectivamente. Segundo uma pesquisa na base de dados de ligandos *Cambridge Structural Database* (CSD), é mais provável que exista uma ligação entre o átomo de molibdénio e um grupo carboxílico do que com um grupo oxo. Esta informação corrobora a hipótese de estarmos na presença de formato não se podendo, no entanto, descartar a outra possibilidade para a interpretação da densidade electrónica.

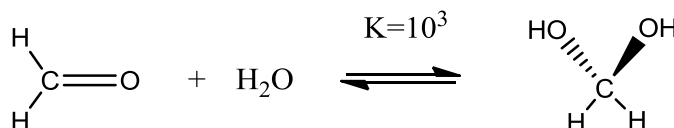


Figura 4.12. Representação esquemática da reacção de hidratação do formaldeído em solução aquosa, originando metileno glicol.

Relativamente à estrutura *DgAOR_acet* obtiveram-se resultados muito semelhantes à estrutura *DgAOR_form*, nomeadamente densidade electrónica relativa à presença de quatro átomos no mesmo local de ligação ao molibdénio. Neste local foi modelada uma molécula de acetato, produto da reacção entre a *DgAOR* e o acetaldeído, coordenada ao Mo através de um dos átomos de oxigénio a uma distância de $1,94 \text{ \AA}$ (Figura 4.13). Os factores B desta molécula após refinamento estão semelhantes aos átomos vizinhos ($\sim 9 \text{ \AA}^2$), indicando que a molécula está presente com um factor de ocupação de 1. O átomo de oxigénio que não se encontra coordenado com o molibdénio encontra-se à distância de ligação de hidrogénio ($2,85 \text{ \AA}$) com uma das águas conservadas perto do centro activo da enzima e com o Glu869 a $2,70 \text{ \AA}$.

Quando comparando os resultados das estruturas *DgAOR_form* e *DgAOR_acet* observa-se uma sobreposição quase perfeita da densidade relativa aos dois produtos, com a excepção da densidade relativa ao grupo metilo do acetato que não está presente no formato. Esta comparação permite reforçar a ideia que a molécula observada na primeira estrutura é realmente o formato, pois estas duas moléculas apresentam a mesma geometria em relação ao comprimento e ângulo das ligações.

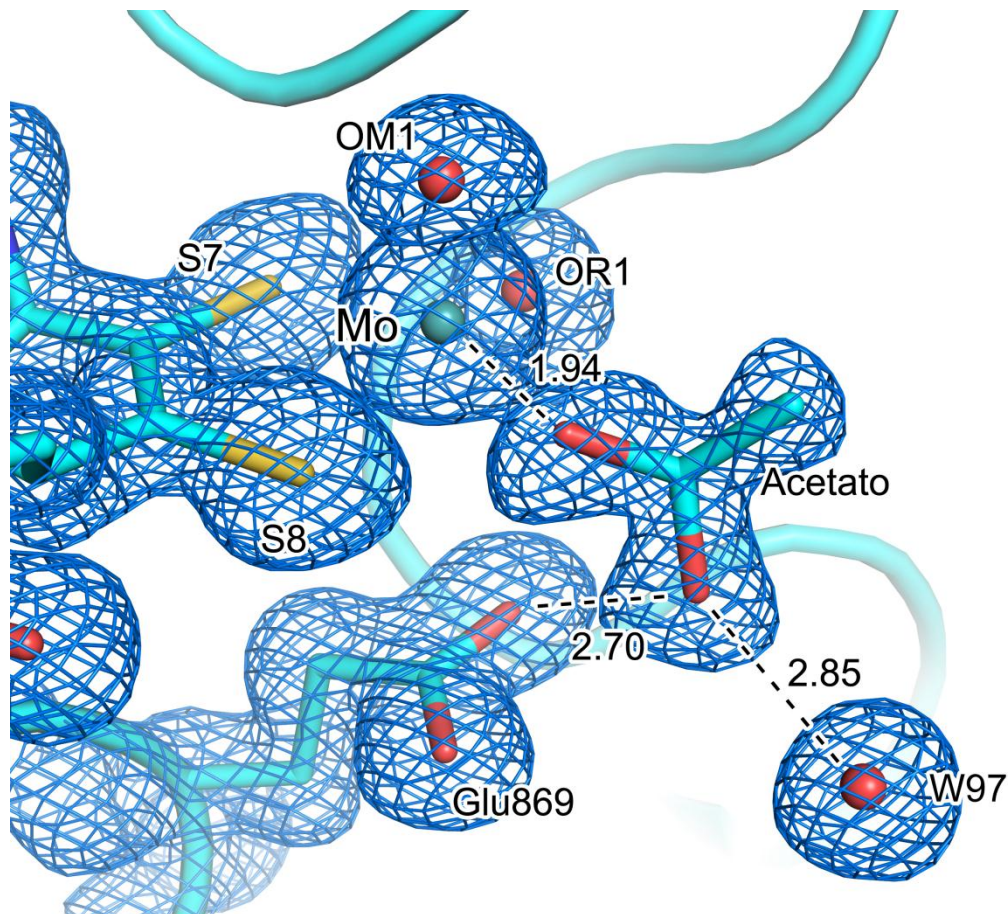


Figura 4.13. Estrutura cristalográfica do centro activo da amostra *DgAOR_acet*. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigênio – vermelho; Enxofre – amarelo; Molibdênio – azul-esverdeado. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

Uma vez refinadas, estas três estruturas vão fornecer informações importantes em relação ao mecanismo de reacção da enzima com os diversos tipos de substratos que esta catalisa.

4.2. Tripsina de *Bos taurus*

Foram realizados estudos por cristalografia de vários inibidores desenhados para a inibição da urocinase em complexo com a tripsina, pois esta é uma proteína da mesma família, sendo um modelo cristalográfico muito estudado.

4.2.1. Ensaios de cristalização

Das condições inicialmente testadas utilizando a tripsina na presença de benzamidina, obtiveram-se cristais em todas as condições (4, 15, 20, 30 e 31 do *screen* de cristalização *Crystal Screen* HT da Hampton Research). As condições 15, 30 e 31 levaram à formação de cristais em apenas 24 h e, apesar de os cristais obtidos serem todos semelhantes, os da condição 15 aparecem em maior número (Figura 4.14).

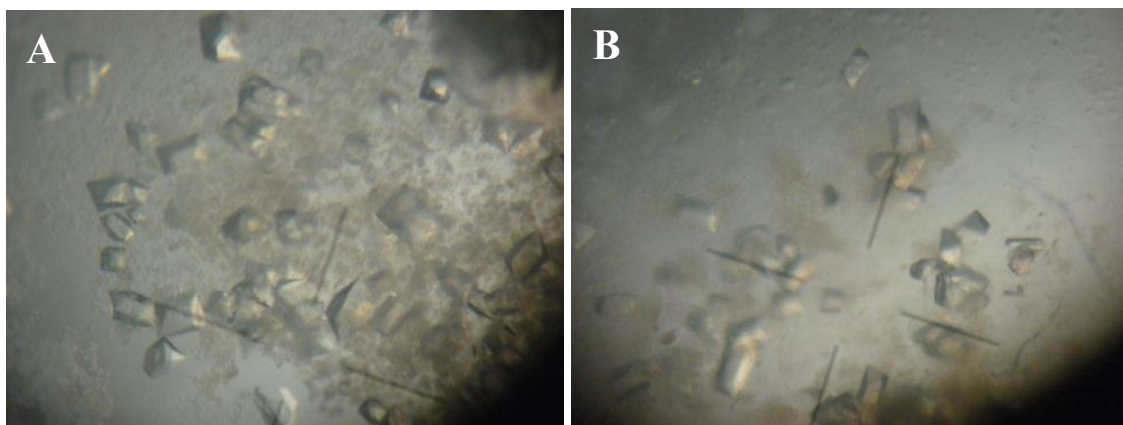


Figura 4.14. Cristais de tripsina em complexo com benzamidina utilizando as condições 15 (A) e 31 (B). Estas condições apresentam PEG 3350 ou PEG 8000 como agentes precipitantes, sulfato de amónio e, no caso da condição 15, tampão cacodilato de sódio a pH = 6,5.

Uma semana após a preparação das gotas apareceram cristais nas condições 4 e 20 em menor quantidade e com uma forma semelhante. Devido ao facto de aparecerem cristais com maior facilidade nas condições anteriores, estas foram seleccionadas para os testes de cristalização da tripsina com os compostos inibidores AB1, AB11, SR5, JS62 e JS67. Adicionalmente, o facto

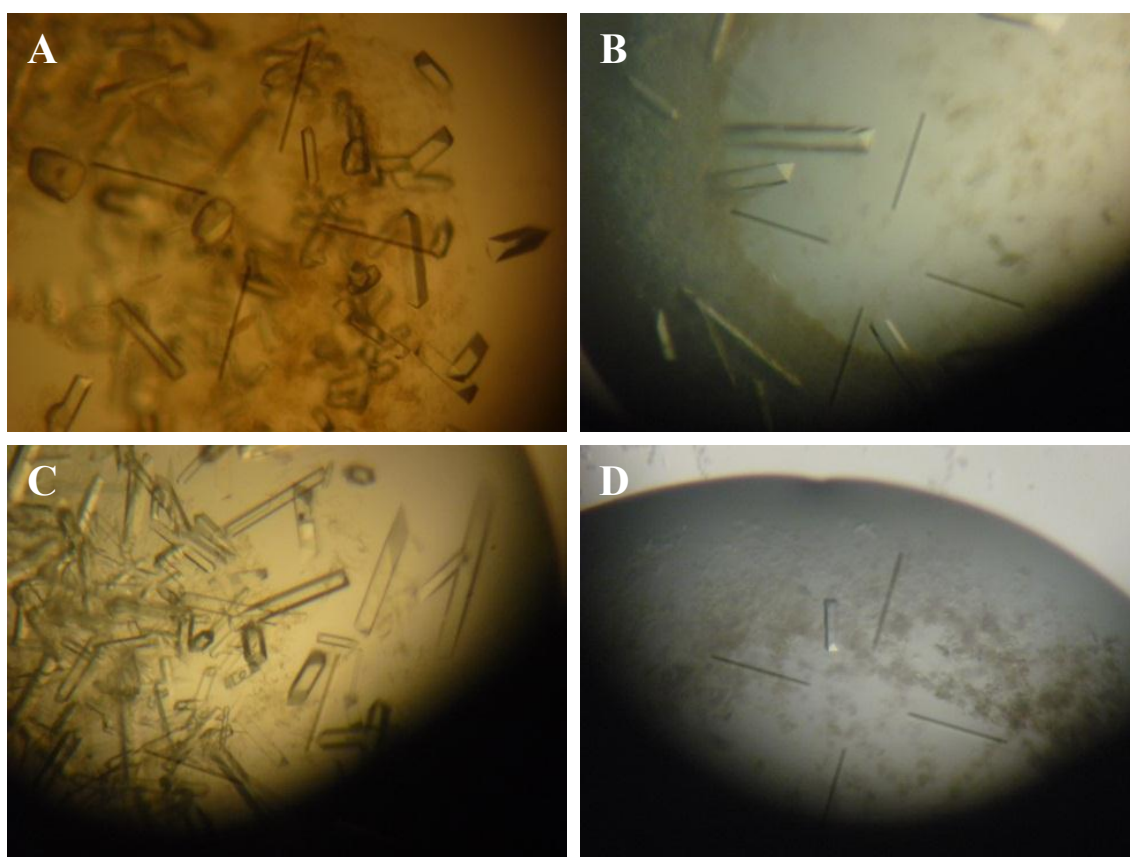


Figura 4.15. Cristais de tripsina incubada com os compostos AB1 (A), AB11 (B), JS62 (C) e SR5 (D), obtidos utilizando a condição de cristalização 15. Os cristais obtidos para a tripsina incubada com o composto JS67 são idênticos aos com SR5.

de a condição 4 apresentar uma concentração muito elevada de sulfato de amónio potencia a possibilidade de se formarem cristais de sal. Caso os cristais obtidos nas condições 15, 30 e 31 apresentem fraca difracção, podem ser testadas as condições 4 e 20. Os cristais obtidos na presença de benzamidina foram utilizados para testar soluções crio-protectoras tendo-se chegado à conclusão que os cristais difractam melhor utilizando 25% de glicerol ou paratona como agente crio-protector.

A tripsina a 60 mg/mL foi incubada com os compostos e foram realizados novos ensaios de cristalização. Após um dia apareceram numerosos cristais nas gotas com AB1, AB11 e JS62 nas condições 15 e 31. Estes cristais, apresentam-se em duas formas provavelmente devido a pertencerem a grupos espaciais diferentes (forma de prisma e forma mais arredondada) (Figura 4.15). Para os compostos SR5 e JS67 apareceram cristais ao fim de uma semana, em menor quantidade, tamanho inferior e em forma de prisma. Foram congelados cristais das duas formas

Tabela 4.13. Resumo das condições testadas e dos resultados obtidos na experiência de difracção para a tripsina.

*Designação pela qual vão ser referidas as estruturas relativas a cada composto ao longo da discussão

Designação ao longo do texto*	Composto	Forma cristalina	Agente crio-protector	Resultado da difracção
Trip_AB1	AB1	Prisma	Glicerol	Res. máx. 1,42 Å
			Paratona	Não difractou
		Arredondada	Glicerol	Res. máx. 2 Å
Trip_AB11	AB11	Prisma	Glicerol	Má difracção
			Paratona	Res. máx. 1,15 Å
Trip_SR5	SR5	Prisma	Glicerol	Res. máx. 1,17 Å
			Paratona	Má difracção
Trip_JS62	JS62	Prisma	Glicerol	Má difracção
			Paratona	Não difractou
		Arredondada	Glicerol	Res. máx. 1,3 Å
			Paratona	Não foi medido
Trip_JS67	JS67	Prisma	Glicerol	Res. máx. 1,12 Å

crystalinas para cada composto utilizando tanto solução crio-protectora com glicerol como paratona e foram medidos utilizando radiação de Sincrotrão na linha ID29 do ESRF (Tabela 4.13). Verifica-se que foi possível obter difracção a alta resolução para ambas as formas cristalinas, assim como para todos os compostos testados. Nos próximos subcapítulos serão discutidos o processamento dos dados e resolução das várias estruturas.

4.2.2. Experiência de difracção

Foi seleccionado o conjunto de dados com maior resolução para cada composto, os quais foram processados utilizando os programas MOSFLM ou XDS.⁸⁷ As reflexões foram indexadas e integradas e as constantes da célula foram refinadas. Utilizou-se o programa POINTLESS para determinar o grupo espacial, verificando-se que os cristais Trip_AB1, Trip_AB11, Trip_JS67 e Trip_SR5 com forma de prisma pertencem ao sistema cristalino ortorrômbico e ao grupo espacial $P2_12_12_1$ e que os cristais de Trip_AB1 e Trip_JS62 com forma cristalina arredondada pertencem ao sistema cristalino trigonal e ao grupo espacial $P3_121$. Entre os cristais que difractaram da condição Trip_AB1, foi seleccionado o que apresenta uma maior resolução (1,42 Å) para se resolver a estrutura. Na Figura 4.16 encontram-se dois padrões de difracção correspondentes a dois cristais com grupos espaciais diferentes (Trip_AB1 com forma de prisma, pertencente ao grupo espacial $P2_12_12_1$, e Trip_JS62 com forma arredondada, pertencente ao grupo espacial $P3_121$).

De seguida, recorreu-se ao programa SCALA para escalar os dados e analisar a qualidade dos cristais. Uma porção dos dados (5%) foi separada para cálculo do factor R_{free} aquando do refinamento. As estatísticas referentes a este procedimento encontram-se na Tabela 4.14.

Comparando os resultados obtidos para os vários cristais verifica-se que todos apresentam um valor de *completeness* bastante elevado (~98% ou superior), assim como um valor de redundância superior a 3, indicando uma boa estratégia de recolha de dados. Os valores de $I/\sigma(I)$ médio são particularmente elevados na última camada de resolução para os conjuntos de dados Trip_AB1, Trip_JS67 e Trip_SR5, sendo que estes poderiam ter sido recolhidos a uma resolução ainda maior, o que é confirmado pelos valores baixos de R_{merge} na última camada de resolução para estes conjuntos.

Em relação ao conteúdo de solvente, este é mais elevado para os cristais pertencentes ao sistema cristalino ortorrômbico (entre 43% e 45%) e mais baixo para o cristal pertencente ao sistema cristalino trigonal (37%), sendo que todos os cristais apresentam apenas uma molécula de proteína por unidade assimétrica.

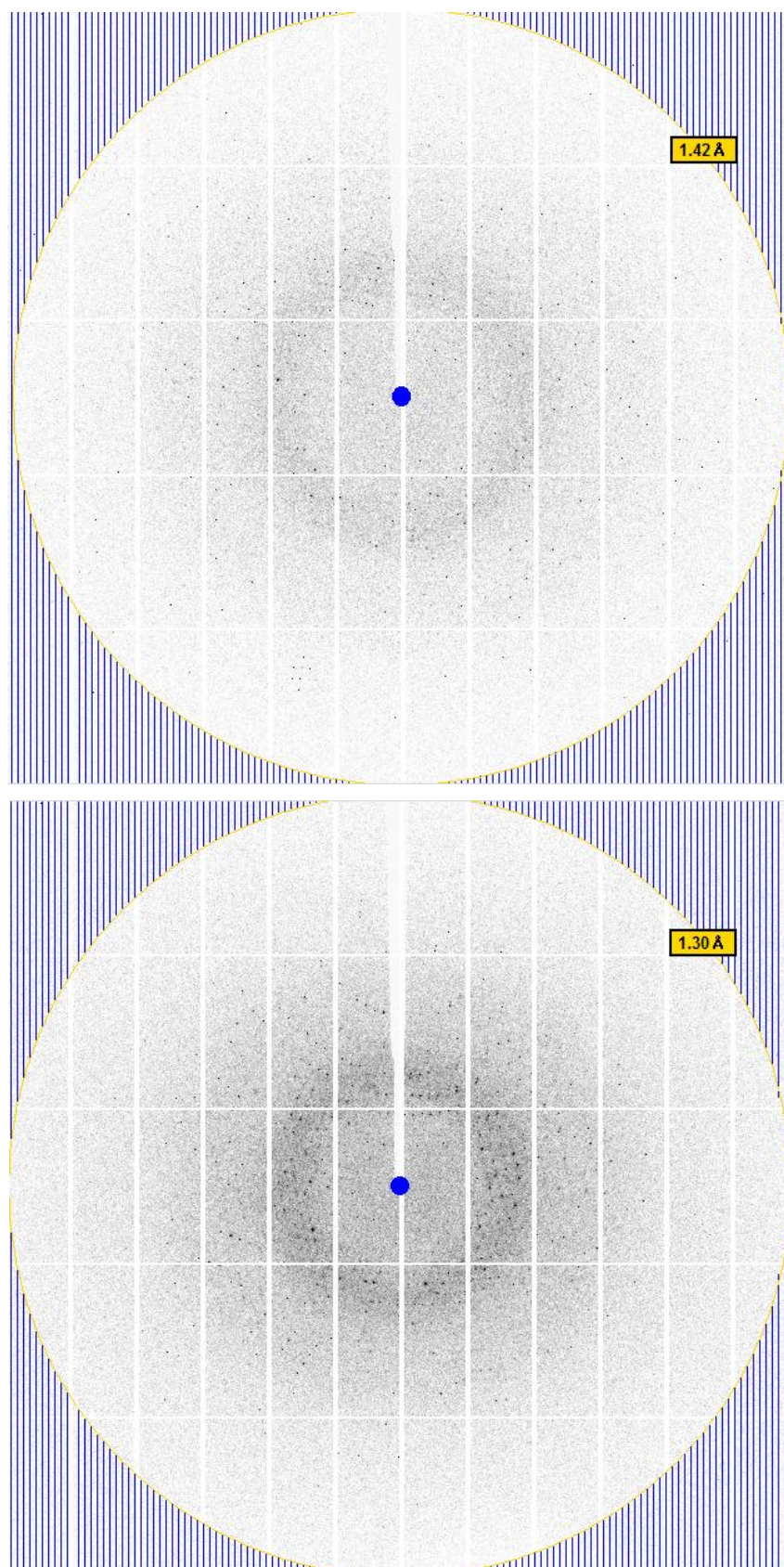


Figura 4.16. Exemplos de dois padrões de difracção dos cristais Trip_AB1 (A) e Trip_JS62 (B). Resolução máxima de 1,42 e 1,30 Å, respectivamente.

Tabela 4.14. Estatísticas de recolha de dados para os vários cristais de tripsina. Os valores entre parêntesis correspondem à última camada de resolução.

Cristal	Trip_AB1	Trip_AB11	Trip_JS62	Trip_JS67	Trip_SR5
Comprimento de onda (Å)	0,976	0,976	0,976	0,976	0,976
Grupo espacial	$P2_12_12_1$	$P2_12_12_1$	$P3_12_1$	$P2_12_12_1$	$P2_12_12_1$
Célula unitária (Å)	a = 54,5 b = 56,7 c = 66,5	a = 54,7 b = 58,4 c = 66,4	a,b = 54,4 c = 107,9	a = 54,2 b = 58,1 c = 66,9	a = 54,3 b = 58,1 c = 66,9
Limites de resolução (Å)	43,15–1,42 (1,50–1,42)	34,21–1,15 (1,22–1,15)	47,07–1,30 (1,37–1,30)	28,97–1,12 (1,18–1,12)	34,11–1,17 (1,24–1,17)
Nº de reflexões observadas	175 487 (26 078)	270 257 (39 246)	275 239 (39 744)	315 909 (45 678)	278 277 (40 620)
Nº de reflexões únicas	39 086 (5 637)	74 063 (10 687)	46 181 (6 618)	82 295 (11 871)	69 708 (10 109)
Completeness (%)	99,4 (99,4)	98,5 (98,3)	99,9 (99,9)	99,4 (99,5)	97,9 (98,1)
Redundância	4,5 (4,6)	3,6 (3,7)	6,0 (6,0)	3,8 (3,8)	4,0 (4,0)
$I/\sigma(I)$ médio	22,3 (11,9)	11,2 (3,7)	19,2 (4,8)	16,1 (7,2)	17,2 (9,5)
R_{merge}^a (%)	3,8 (9,6)	5,4 (31,2)	4,3 (36,3)	4,4 (16,3)	4,4 (13,6)
VM^b - Parâmetro de Matthews (Å ³ /Da)	2,17	2,24	1,94	2,22	2,23
Conteúdo de solvente ^b (%)	43,4	45,1	36,7	44,7	44,8

(a) $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, onde $I_i(hkl)$ é a intensidade integrada de uma determinada reflexão i e $\langle I(hkl) \rangle$ é a intensidade média de múltiplas reflexões relacionadas por operações de simetria.

(b) $VM = \frac{V}{Z \times Mp}$, onde V é o volume da célula unitária, Z é o número de unidades assimétricas presentes na célula unitária e Mp é a massa molecular da proteína; Conteúdo de solvente = $\left(1 - \frac{1,23}{VM}\right) \times 100$

4.2.3. Resolução das estruturas

Para o cálculo dos ângulos da fase necessários para a resolução da estrutura, foi utilizado o método de substituição molecular, pois a estrutura desta proteína já é conhecida. Assim, seleccionou-se a estrutura com o código 1SOQ, que se encontra depositada no PDB, como modelo. Esta estrutura apresenta uma resolução bastante elevada de 1,02 Å, encontra-se no grupo espacial $P2_12_12_1$ e apresenta apenas um ião cálcio como ligando. Procedeu-se à remoção

deste ligando e de todas as moléculas de água da estrutura, utilizando-se apenas a cadeia peptídica da proteína para a resolução das estruturas.

Foi então usado o programa PHASER para calcular os ângulos da fase necessários para o cálculo dos factores de estrutura, usando as funções de rotação e translação para procurar uma molécula de proteína na unidade assimétrica. Foram encontradas soluções nos grupos espaciais previamente determinados, apresentando valores de LLG e *Z-score* que indicam que as estruturas dos 5 cristais foram resolvidas (Tabela 4.15). De seguida, recorreu-se ao programa DM para melhorar a densidade electrónica calculada anteriormente, através de métodos de *flattening* de solvente.

Tabela 4.15. Valores de *Z-score* e LLG para as funções de rotação e translação obtidos utilizando o programa PHASER, para todos os conjuntos de dados de tripsina.

Conjunto de dados	Trip_AB1	Trip_AB11	Trip_JS62	Trip_JS67	Trip_SR5
<i>Z-score</i> (rotação)	24,3	27,4	19,4	26,1	26,7
LLG (rotação)	367,5	435,9	209,8	407,7	409,8
<i>Z-score</i> (translação)	39,5	32,3	39,7	29,6	30,0
LLG (translação)	3040	3630	2979	3550	3565

4.2.4. Refinamento das estruturas

As estruturas foram refinadas utilizando o programa REFMAC5, tendo sido feito um refinamento inicial com adição de moléculas de água pelo programa COOT. Após este refinamento, os mapas de densidade electrónica foram inspeccionados tomando especial atenção ao centro activo da enzima. Foi possível observar densidade electrónica correspondente aos respectivos compostos para os conjuntos de dados Trip_AB1 e Trip_JS62, alguma densidade electrónica extraordinária para os conjuntos de dados Trip_SR5 e Trip_JS67 e nenhuma densidade electrónica extra no centro activo do conjunto de dados Trip_AB11. O facto de não haver densidade electrónica referente ao inibidor para este conjunto de dados pode significar que este não se liga à tripsina ou que a incubação seguida de cristalização do complexo não foi eficaz. Por este motivo abandonou-se o refinamento da estrutura Trip_AB11 e prosseguiu-se para as restantes quatro estruturas. Este refinamento foi efectuado utilizando o programa COOT para ajustar o modelo à densidade electrónica e o programa REFMAC5 para recalculer os ângulos da fase, repetindo-se este processo até se obterem modelos que expliquem da melhor forma a densidade electrónica. O passo de refinamento das estruturas ainda não se encontra

concluído, pois o seu centro activo ainda está sob avaliação. As estatísticas relativas a este refinamento encontram-se na Tabela 4.16.

Tabela 4.16. Estatísticas de refinamento actuais para os modelos Trip_AB1, Trip_JS62, Trip_JS67 e Trip_SR5.

Conjunto de dados	Trip_AB1	Trip_JS62	Trip_JS67	Trip_SR5
Limites de resolução (Å)	43,15–1,42	47,07–1,30	43,84–1,12	43,84–1,17
Factor R (%)	14,8	14,93	14,01	14,02
Nº de reflexões	37 078	43 793	78 107	66 124
Factor R _{free} (%)	17,4	17,71	15,50	15,94
Nº de reflexões (R _{free})	1 955	2 341	4 116	3 527
Nº de resíduos	223	223	223	223
Nº de átomos	2 139	2 052	2 226	2 199
RMSD para o comprimento de ligação (Å)	0,028	0,029	0,34	0,033
RMSD para o ângulo de ligação (°)	2,67	2,96	2,73	2,74

Analisando os factores R e R_{free} das quatro estruturas verifica-se que estes estão dentro dos valores aceites, o que significa que as estruturas estão bem resolvidas, explicando satisfatoriamente as reflexões obtidas. Os modelos Trip_JS67 e Trip_SR5 apresentam valores mais baixos destes parâmetros, pois têm uma maior resolução, sendo de esperar mais facilidades na modelação dos detalhes mais finos das estruturas devido à qualidade dos mapas de densidade electrónica.

Os valores de RMSD para o comprimento de ligação são superiores ao valor ideal indicando um baixo rigor no refinamento em relação às restrições geométricas. No entanto, como visto pela análise dos factores R e R_{free}, os modelos explicam os dados experimentais, sendo difícil melhorar os valores de RMSD. No que toca ao ângulo de ligação os valores encontram-se dentro dos valores ideais, não excedendo os 3°.

Tendo em mente que o refinamento das estruturas não se encontra concluído, no próximo subcapítulo é feita uma caracterização preliminar da sua estrutura global do seu centro activo.

4.2.5. Caracterização das estruturas

A tripsina apresenta uma estrutura globular com um diâmetro aproximado de 40 Å. A sua estrutura secundária apresenta dois barris- β , conforme descrito anteriormente (subcapítulo 1.2.3.1). As estruturas são semelhantes à estrutura nativa, apresentando valores baixos de RMSD para a superposição das estruturas (Tabela 4.17), sendo que as estruturas Trip_AB1 e Trip_JS62 apresentam valores mais elevados comparando com as estruturas Trip_JS67 e Trip_SR5. Este resultado indica que ocorreram alterações conformacionais mais acentuadas nas duas primeiras estruturas. Assim, foi possível encontrar diferenças em relação a algumas cadeias laterais que se encontram em posições diferentes, ou apresentam conformações alternadas, não existentes na estrutura nativa. Foram também modeladas moléculas de glicerol, PEG e iões Cl^- e SO_4^{2-} em quantidades variáveis nas diferentes estruturas.

Tabela 4.17. RMSD de superposição estrutural das estruturas Trip_AB1, Trip_JS62, Trip_JS67 e Trip_SR5 com a estrutura nativa 1S0Q, em relação a todos os átomos da cadeia principal da proteína.

Estruturas	Trip_AB1	Trip_JS62	Trip_JS67	Trip_SR5
RMSD (Å)	0,227	0,235	0,168	0,155

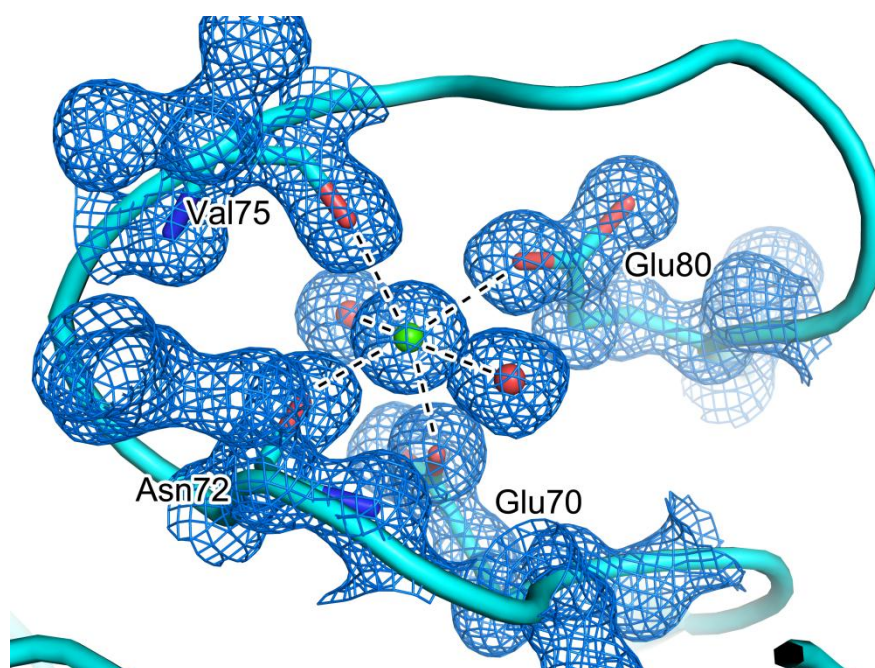


Figura 4.17. Coordenação octaédrica do íon Ca^{2+} no *Calcium-binding site* da tripsina. Este estabelece ligações de hidrogénio com os resíduos Glu70, Asn72, Val75 e Glu80 e duas moléculas de água conservadas. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Cálcio – amarelo. A azul está representado o mapa $2m\text{Fo}-\text{DFc}$, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

Todas as estruturas apresentam um ião Ca^{2+} coordenado por átomos de oxigénio dos resíduos Glu70, Asn72, Val75 e Glu80 e por duas moléculas de água conservadas (Figura 4.17), no denominado *Calcium-binding site*. Este ião confere estabilidade à proteína, evitando a sua auto-degradação.

Nas estruturas obtidas verifica-se a presença de seis pontes de dissulfureto, no entanto é possível observar conformações alternadas para a cadeia lateral de algumas cisteínas, indicando que estas pontes foram quebradas numa percentagem das moléculas de proteína presentes nos cristais. A quebra destas ligações (Figura 4.18) deve ser resultado da radiação ionizante dos raios-X de alta intensidade, como já foi descrito na literatura para esta proteína.¹⁰¹

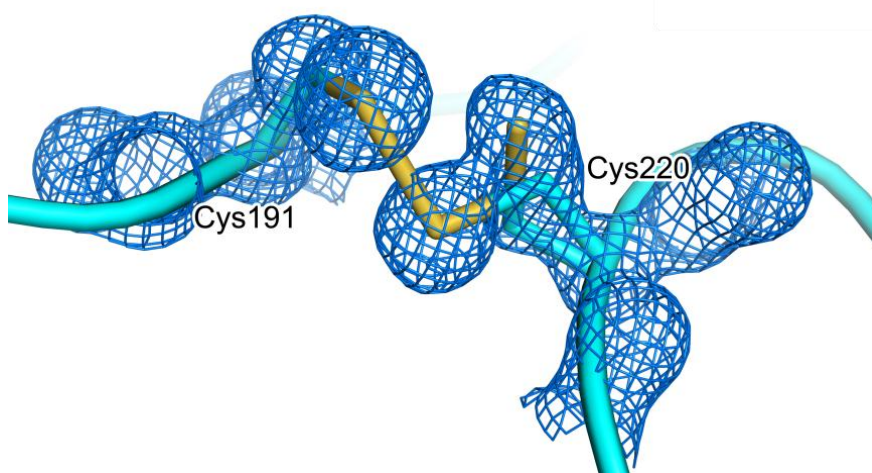


Figura 4.18. Conformação alternada observada para a Cys220, originada por quebra da ponte de dissulfureto. Código de cores: Carbono – ciano; Enxofre – amarelo. A azul está representado o mapa $2m\text{Fo-DFc}$, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

É também possível observar, em todas as estruturas resolvidas, uma quebra da cadeia polipeptídica principal, entre a Pro152 e o Asp153 (Figura 4.19). Estes aminoácidos estão situados no *autolysis loop* da tripsina, o que indica que ocorreu auto-degradação das moléculas de proteína presentes no cristal. Este fenómeno pode ser indicativo da fraca potência dos inibidores testados relativamente à tripsina, pois indica que existe actividade catalítica na mistura. Este resultado pode também ser devido à ineficácia do procedimento de incubação com os vários inibidores, o qual pode não ter sido feito de modo a evitar a autólise.

Nos próximos subcapítulos vão ser analisados os centros activos das quatro estruturas individualmente.

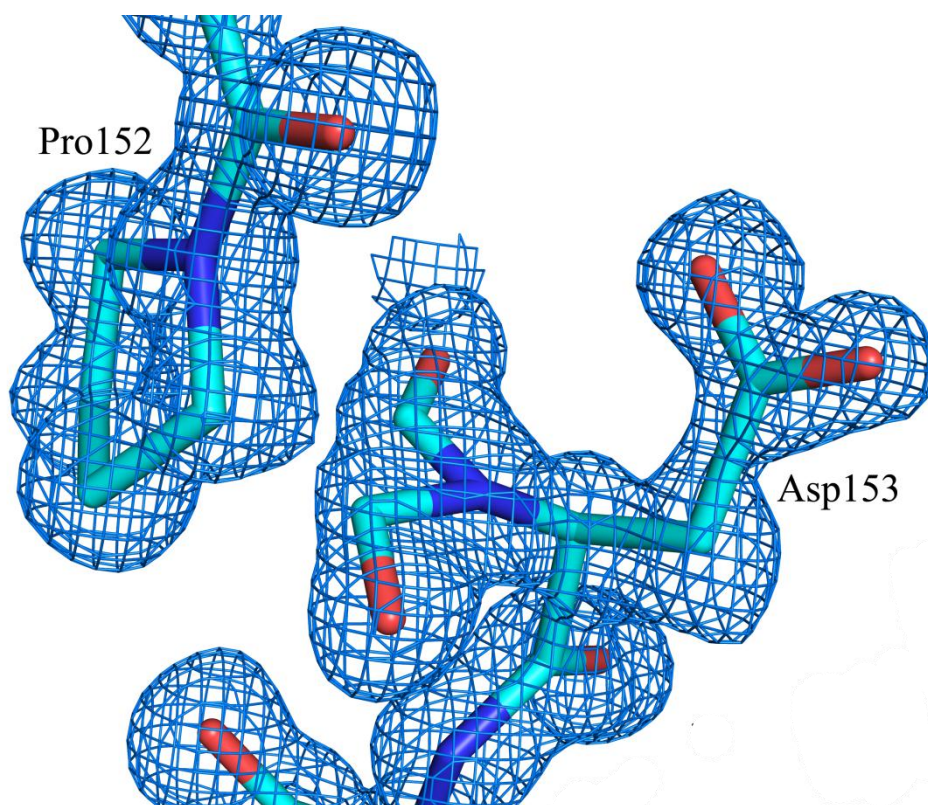


Figura 4.19. Quebra da ligação entre a Pro152 e o Asp153 no *autolysis loop* da tripsina. O grupo carbonilo da Pro152 apresenta conformação alternada. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho. A azul está representado o mapa $2mF_o-DF_c$, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

4.2.5.1. Tripsina em complexo com AB1

Analisando o centro activo da estrutura Trip_AB1 foi possível identificar densidade electrónica correspondente ao composto AB1. No entanto, ao se introduzir este composto e refinar a estrutura foi possível visualizar, por inspecção dos mapas $2F_o-F_c$ e F_o-F_c , densidade electrónica extraordinária perto dos grupos hidroxilo ligados ao átomo de boro do composto. Este resultado é inesperado pois mostra que a estrutura inicial do composto não corresponde à considerada. No entanto também pode indicar que o composto inicial reagiu com a proteína ou algum componente da solução de cristalização tendo-se formado um sub-produto que ficou a interagir com o centro activo da enzima. De modo a tentar explicar a densidade electrónica encontrada, foram adicionados dois grupos CH_2 ligados aos grupos hidroxilo e entre si, formando um anel. A estrutura foi refinada, resultando em mais densidade electrónica adjacente a estes átomos. Foram adicionados mais dois grupos metilo e obteve-se a estrutura representada na Figura 4.20. Com a excepção destes átomos adicionais, os valores dos factores B dos átomos correspondentes ao composto são semelhantes entre si (entre 8 e 13 \AA^2) e em relação aos resíduos do centro activo (entre 7 e 11 \AA^2). Os quatro carbonos adicionais apresentam factores B mais elevados (entre 15 e 24 \AA^2), o que sugere que estes poderão não estar presentes em todas

as moléculas do cristal. Este modelo explica os dados experimentais apesar de não se conseguir assegurar que corresponda à realidade. De modo a se determinar qual o motivo deste resultado, pretende-se realizar estudos por espectroscopia de $^1\text{H-RMN}$ do composto antes de este ser incubado com a proteína. Assim, espera-se determinar se o composto inicial corresponde à forma com os grupos hidroxilo livres e se a densidade obtida é o produto de uma reacção inespecífica entre o inibidor e a proteína. Também pode ser o caso de, depois da síntese, este composto se ter degradado para formar outro diferente que também pode interagir com a proteína, inibindo-a. Este assunto merece uma discussão aprofundada com os nossos colaboradores que sintetizam os compostos, de modo a esclarecer estas dúvidas.

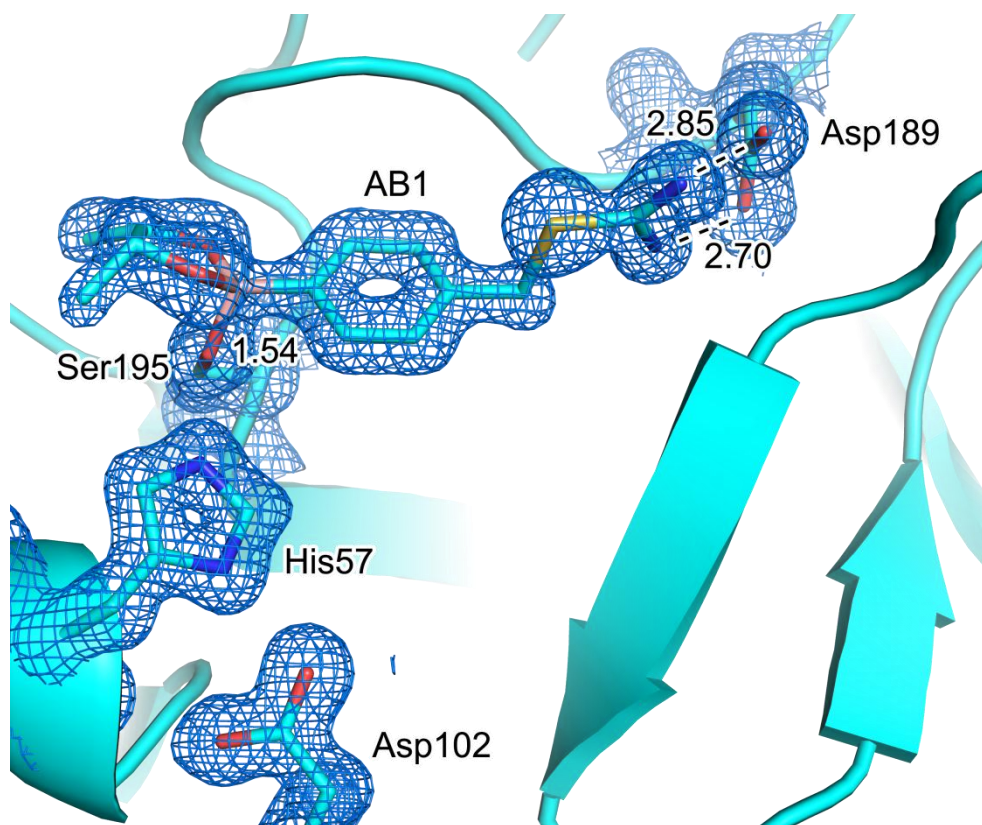


Figura 4.20. Estrutura cristalográfica do centro activo da amostra Trip_AB1. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Boro – rosa. A azul está representado o mapa $2m\text{Fo-DFc}$, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

Em todo o caso, procedeu-se à análise das interacções entre a proteína e o ligando. A serina catalítica está coordenada ao átomo de boro, como esperado, a uma distância de 1,54 Å. É possível observar várias interacções entre o composto e a proteína à distância de ligação de hidrogénio. O grupo amina da Gly193, parte do *oxyanion hole* que estabiliza o intermediário tetraédrico da reacção catalisada, está a 2,74 Å de distância de um dos oxigénios ligado ao átomo de boro. O outro oxigénio estabelece uma ligação de hidrogénio com uma das moléculas de água conservadas nas estruturas (2,71 Å). Por último, os grupos amina do composto

encontram-se à distância de ligação de hidrogénio do grupo carboxílico da cadeia lateral do Asp189, responsável pelo reconhecimento do local de corte da tripsina, assim como do grupo hidroxilo da Ser190 e do oxigénio do grupo carbonilo da cadeia principal da Gly219. Estas interações encontram-se representadas na Figura 4.21. Comparando a estrutura do composto AB1 com o AB11, verifica-se que a única diferença entre estes dois inibidores é a presença de um grupo metilo ligado a um dos grupos amina. Dado que, para o composto AB11 não se verificou ligação à proteína, é provável que esta metilação esteja a impedir a interacção dos grupos amina com o Asp189. Este resultado destaca a importância que este resíduo tem na especificidade da interacção de ligandos, quer inibidores, quer cadeias polipeptídicas, na actividade da protease. Mais ainda, estes dados cristalográficos são uma base sólida para o desenho de novos compostos com um poder inibidor mais acentuado, contribuindo para o desenvolvimento desta área de investigação.

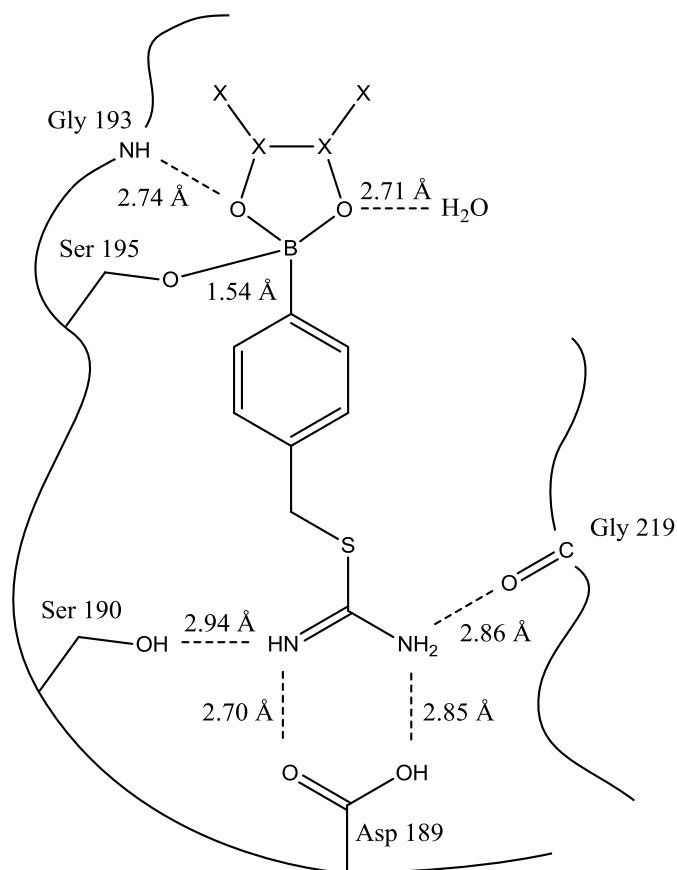


Figura 4.21. Representação esquemática do composto AB1 no centro activo da tripsina. A tracejado estão representadas todas as ligações de hidrogénio entre o composto e a proteína e moléculas de água conservadas. Representados por X encontram-se átomos não identificados para os quais foi possível observar densidade electrónica.

4.2.5.2. Tripsina em complexo com JS62

No centro activo da estrutura Trip_JS62 observou-se densidade electrónica relativa ao composto JS62, o qual foi modelado. Este encontra-se coordenado à serina catalítica pelo átomo de boro. Não é possível observar densidade electrónica relativa aos grupos metilo da unidade pinacol, devido possivelmente à sua desordem nas várias moléculas do cristal (Figura 4.22). O composto encontra-se na mesma posição do composto AB1, revelando o mesmo tipo de interacções com a proteína (Figura 4.23). O átomo de boro encontra-se a 1,50 Å de distância do grupo hidroxilo da serina catalítica. Os oxigénios ligados ao boro fazem ligações de hidrogénio com a Gly193 e uma água conservada, a 2,77 e 2,64 Å de distância, respectivamente. Um dos grupos amina do composto estabelece ligações de hidrogénio com um dos oxigénios do grupo carboxílico do Asp189 e com o grupo hidroxilo da Ser190, enquanto o outro grupo amina do composto estabelece ligações de hidrogénio com o outro oxigénio da cadeia lateral do Asp189, com o oxigénio do grupo carbonilo da cadeia principal da Gly219 e com uma molécula de água conservada. Após refinamento das estruturas, estas vão ser analisadas em comparação com estruturas de tripsina inibida com compostos de boro previamente descritas na literatura, com o objectivo de traçar conclusões em relação à especificidade e eficácia dos compostos em estudo.

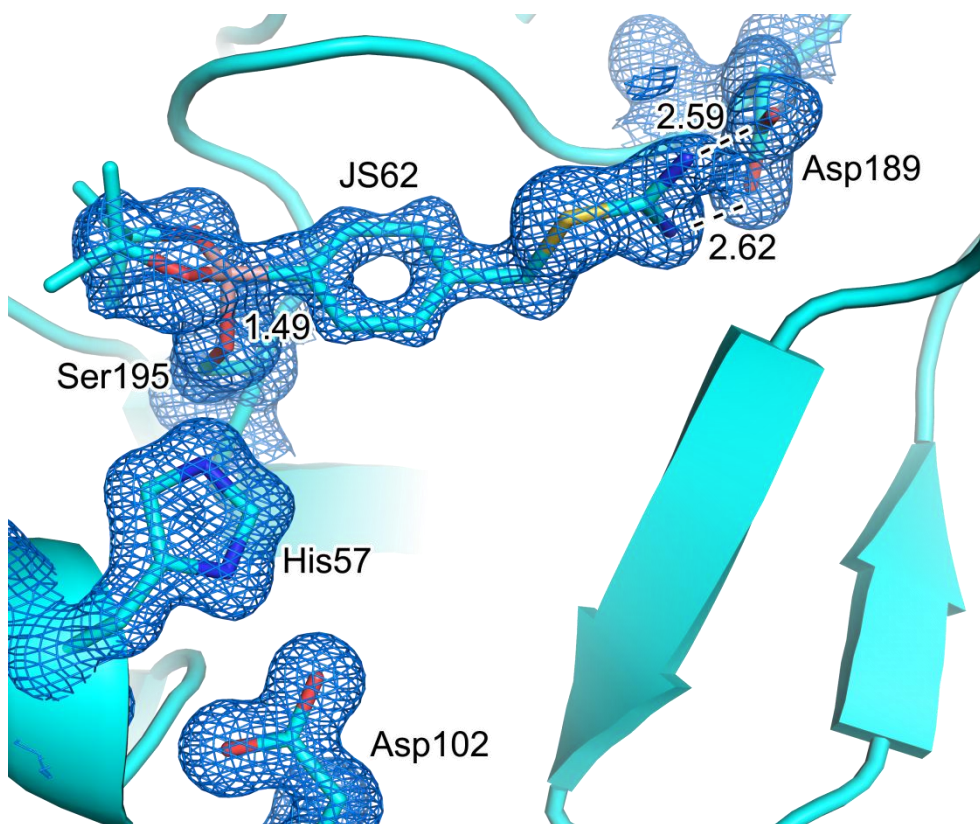


Figura 4.22. Estrutura cristalográfica do centro activo da amostra Trip_JS62. Todas as distâncias apresentadas estão em Å. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo; Boro – rosa. A azul está representado o mapa 2mFo-DFc, com um contorno de 1σ . Esta figura foi preparada utilizando o programa PyMOL.

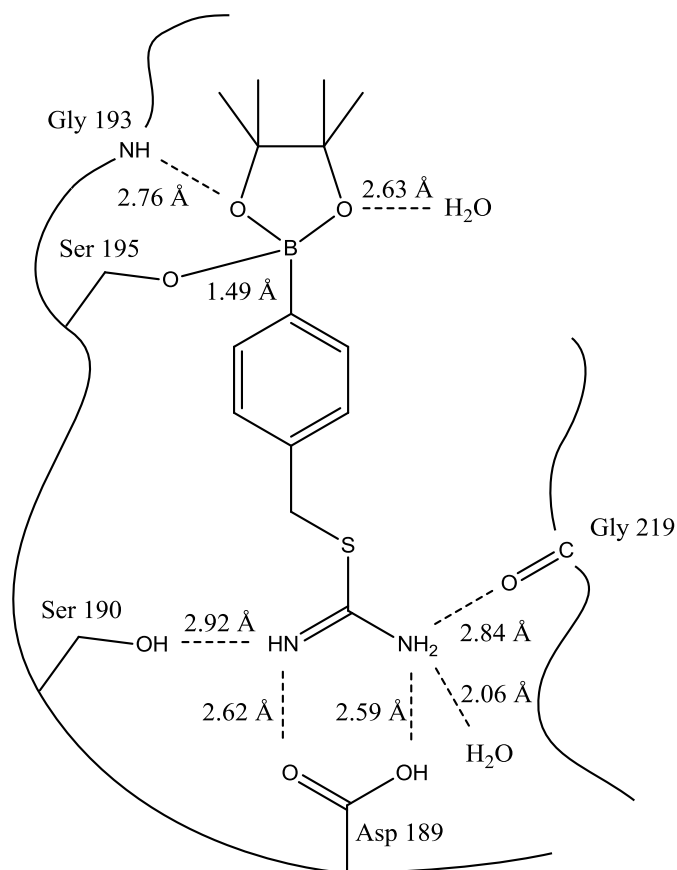


Figura 4.23. Representação esquemática do composto JS62 no centro activo da tripsina. A tracejado estão representadas todas as ligações de hidrogénio entre o composto e a proteína e moléculas de água conservadas.

4.2.5.3. Tripsina em complexo com JS67

Na estrutura Trip_JS67, apesar de se observar densidade electrónica no centro activo da enzima, esta não corresponde ao composto JS67. Após refinamento de toda a estrutura e consequente melhoramento da densidade electrónica, modelou-se um ião sulfato perto da serina catalítica, o qual está a interagir com o *oxyanion hole*. Mais afastada encontra-se densidade relativa ao que se pensa ser um produto de degradação do composto o qual ainda não foi modelado (Figura 4.24).

4.2.5.4. Tripsina em complexo com SR5

O caso da estrutura Trip_SR5 é semelhante ao anterior (Trip_JS67) onde também se observa densidade electrónica correspondente a um ião sulfato a interagir com os resíduos que formam o *oxyanion hole* (Figura 4.25), assim como densidade electrónica que se pensa corresponder a produtos de degradação do composto.

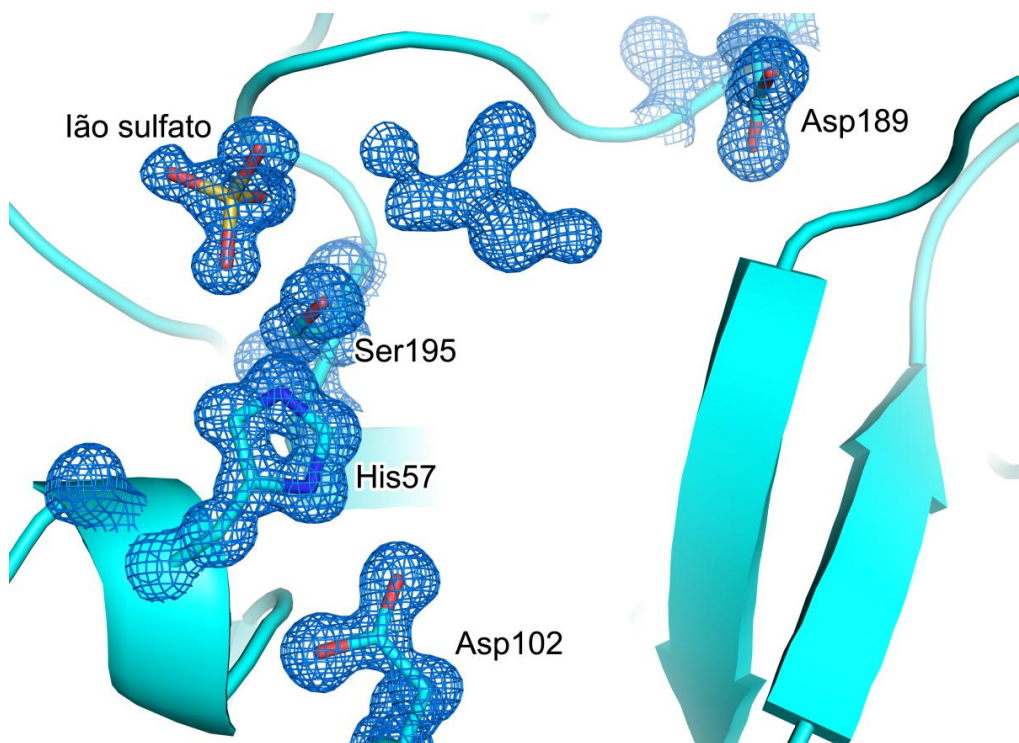


Figura 4.24. Estrutura cristalográfica do centro activo da amostra Trip_JS67. Não foi modelado nenhum composto no centro activo, sendo possível observar alguma densidade electrónica. Foi modelado um ião sulfato perto da serina catalítica e do *oxyanion hole*. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.

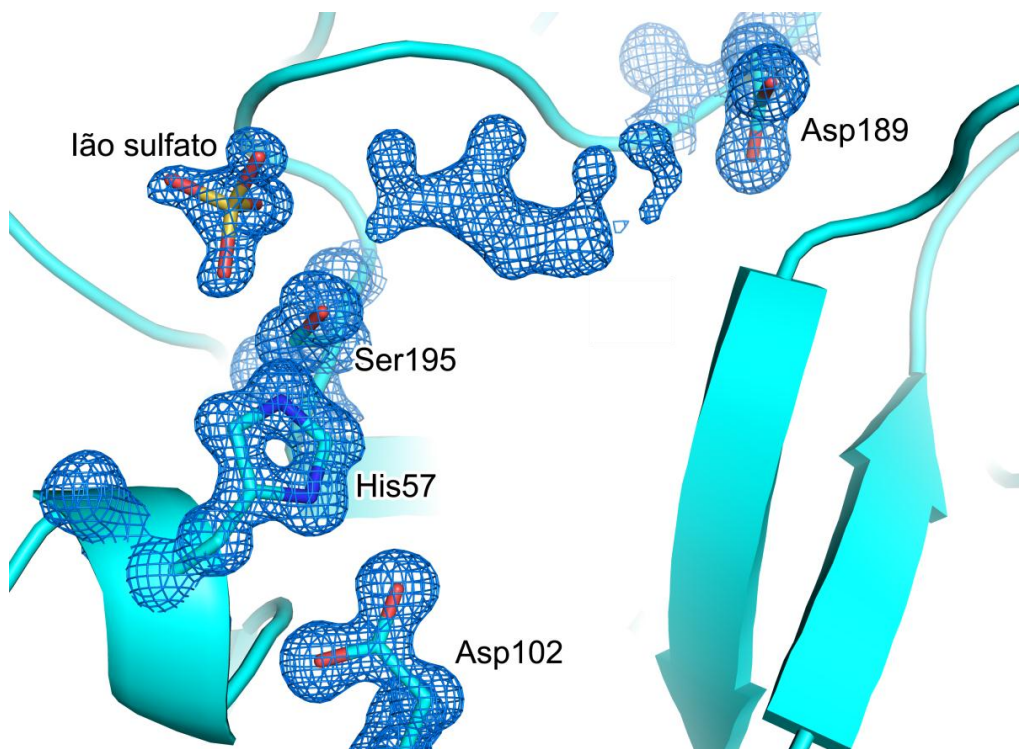


Figura 4.25. Estrutura cristalográfica do centro activo da amostra Trip_SR5. Não foi modelado nenhum composto no centro activo, sendo possível observar alguma densidade electrónica. Foi modelado um ião sulfato perto da serina catalítica e do *oxyanion hole*. Código de cores: Carbono – ciano; Azoto – azul; Oxigénio – vermelho; Enxofre – amarelo. A azul está representado o mapa 2mFo-DFc, com um contorno de 1 σ . Esta figura foi preparada utilizando o programa PyMOL.

É de notar que estas estruturas (Trip_JS67 e Trip_SR5) foram as que demoraram mais tempo a cristalizar, sendo que os cristais obtidos eram de menores dimensões e difractavam a maior resolução. Apesar de os conjuntos de dados recolhidos serem melhores do que para as outras estruturas inibidas, a interpretação do centro activo da enzima é mais difícil. A estabilidade ou capacidade inibitória dos próprios compostos devem ser responsáveis por estas diferenças sendo necessários mais estudos cinéticos e de caracterização estrutural para interpretar estes resultados. No entanto, os resultados obtidos apontam para uma especificidade maior por parte dos compostos de boro (AB1 e JS62) que se ligam á serina catalítica, em contraste com os compostos JS67 e SR5 que não têm boro na sua estrutura.

Capítulo 5

Conclusões e Perspectivas futuras

A cristalografia de raios-X é uma ferramenta poderosa na caracterização de interações entre ligandos e proteínas, bem como aspectos específicos da sua estrutura. Neste trabalho foi possível cristalizar a DgAOR e a tripsina em quantidades suficientes para se estudar a sua interação com diversos compostos, através da resolução e análise de estruturas de alta resolução.

A DgAOR é uma enzima de molibdénio que funciona como modelo para se estudarem outras proteínas da sua família com elevada homologia, como a Xantina Oxidase ou a Aldeído Oxidase. Os estudos cinéticos feitos com ditionito de sódio e sulfureto de sódio revelaram que a enzima ganha actividade numa fase inicial, a qual acaba por perder ao longo do tempo. Estudos cristalográficos usando as mesmas condições indicaram a presença de uma molécula diatómica, interpretada como peróxido de hidrogénio, na esfera de coordenação do molibdénio com uma geometria η^2 . Estes resultados foram confirmados por espectroscopia de RPE e por cristalografia de raios-X ao cristalizar a enzima na presença de peróxido de hidrogénio. Resolvendo esta estrutura obteve-se densidade electrónica, no centro activo da enzima, idêntica à obtida anteriormente, sugerindo que a molécula encontrada depois do *soak* de um cristal de DgAOR com ditionito e sulfureto é realmente peróxido de hidrogénio. Assim conclui-se que estes agentes redutores, na presença de proteína, geram a formação de peróxido de hidrogénio que inibe a enzima, causando uma diminuição da sua actividade. O aumento de actividade na fase inicial ainda não se conseguiu explicar. Estes resultados sugerem que o protocolo de ressulfuração usado para a recuperação de actividade pela Xantina Oxidase pode causar a formação de peróxido de hidrogénio que, por sua vez, também pode inibir a enzima. Este facto nunca foi descrito na literatura dado o curto espaço de tempo dos ensaios cinéticos outrora realizados (80 min).³⁴ Futuramente pretende-se averiguar se existem diferenças estruturais entre a forma activa e inactiva da DgAOR, com o objectivo de explicar os dados cinéticos obtidos. Numa fase mais avançada tenciona-se estender estes estudos a outras enzimas da mesma família, as quais apresentam uma maior relevância ao nível da patofisiologia humana, como a Aldeído Oxidase ou a Xantina Oxidase.

Em relação aos estudos efectuados com substratos da DgAOR, foram obtidas estruturas da enzima em complexo com benzaldeído, formaldeído e acetaldeído, sendo que o seu refinamento não se encontra concluído. Relativamente ao benzaldeído é possível identificar a posição do anel aromático, estabilizado por interações de π -stacking, adjacente ao centro activo. No entanto, a função carbonilo do aldeído não apresenta densidade bem definida. É portanto importante terminar o refinamento desta e de outras estruturas obtidas com este aldeído, na tentativa de se conseguir visualizar melhor este detalhe essencial para a compreensão da interação desta molécula com o molibdénio no centro activo. Novas cristalizações também serão efectuadas neste sentido. As estruturas obtidas após os ensaios de *soaking* com

formaldeído e acetaldeído apresentam o produto de reacção no centro activo da enzima (formato e acetato, respectivamente). Estes encontram-se coordenados ao molibdénio por um dos oxigénios do seu grupo carbonilo, na posição do grupo lábil OH/OH₂ na estrutura nativa. Após terminar o refinamento destas estruturas e analisar o centro activo espera-se compreender melhor o mecanismo de reacção da enzima, o qual não se encontra bem definido. De modo a auxiliar nesta tarefa e para determinar uma razão para a baixa especificidade da enzima, pretende-se obter estruturas com mais substratos, desta vez maiores e com cadeias alifáticas mais longas. O conhecimento sobre o mecanismo de acção e da especificidade da DgAOR é um passo importante para a compreensão dos mesmos em relação à Aldeído Oxidase, à qual se pretende estender estes estudos.

O outro grupo de proteínas explorado nesta dissertação pertence à família das proteases de serina. A urocinase tem o papel de activar a protease plasmina. A sua desregulação *in vivo* está envolvida em processos como o crescimento tumoral ou a metástase, sendo por isso essencial o desenvolvimento de fármacos inibidores desta proteína. Vários compostos foram testados por cristalografia de raios-X utilizando a tripsina de *Bos taurus*. Esta é uma protease da mesma família da urocinase muito estudada por esta técnica. Dos cinco compostos testados (AB1, AB11, JS62, JS67 e SR5), foi possível identificar densidade electrónica no centro activo da enzima relativa aos compostos AB1 e JS62. Em relação ao primeiro, densidade electrónica extra foi encontrada, levantando dúvidas quanto à composição inicial do composto usado ou se este foi alvo de alguma reacção antes de se ligar à enzima. Para averiguar esta situação vai ser realizada espectroscopia de ¹H-RMN do composto inicial e, possivelmente, novas cristalizações, caso se justifique. Em ambas as estruturas obtidas, verifica-se a ligação dos compostos à serina catalítica da tripsina pelo átomo de boro e estabilização dos seus grupos amina terminais pelo aspartato 189, o qual é responsável pela especificidade da enzima. Um dos oxigénios ligados ao átomo de boro dos compostos interage com o *oxyanion hole*, responsável pela estabilização do intermediário na reacção *in vivo*. Estas duas estruturas serão comparadas com outras estruturas descritas na literatura com compostos de boro, de modo a se estudar a sua eficácia de inibição. Para o caso do composto AB11, não ocorreu ligação deste à enzima, possivelmente porque um dos seus grupos amina se encontra metilado. Este resultado indica que os grupos amina terminais do inibidor e a sua interacção com o aspartato 189 são de extrema importância para a ligação do composto à proteína. Os compostos JS67 e SR5 não se ligaram à enzima, apresentando densidade electrónica no centro activo que se pensa corresponder ao produto de degradação dos mesmos. Ao contrário dos três primeiros, estes dois compostos não apresentam boro na sua estrutura, o que revela a sua importância na inactivação da enzima. As estruturas de alta resolução obtidas são uma óptima ferramenta e um bom ponto de partida para a produção de

mais compostos da mesma família, contribuindo de uma forma crucial para este projecto de *drug-design*.

A execução desta dissertação proporcionou-me explorar exaustivamente a técnica de cristalografia de raios-X, desde a cristalização de duas proteínas ate à análise das suas estruturas, passando pela recolha de dados com radiação de Sincrotrão, resolução e refinamento das estruturas. Tive ainda possibilidade de ter contacto com as técnicas de purificação de proteínas e ensaios cinéticos, em colaboração com o Dr. Pablo González.

Os resultados obtidos ao longo deste ano de trabalho vão ser publicados brevemente. A estrutura da *DgAOR* obtida com $\text{Na}_2\text{S}_2\text{O}_4$ e Na_2S e a estrutura *DgAOR_H₂O₂*, juntamente com os ensaios cinéticos apresentados nesta dissertação, foram analisados e discutidos num artigo submetido recentemente. Em relação à tripsina, a estrutura *Trip_JS62* foi descrita num artigo que se encontra de momento em preparação, onde a síntese dos inibidores e os seus dados biológicos estão também incluídos. Todas as outras estruturas de ambas as proteínas encontram-se numa fase de refinamento/análise preliminar.

Capítulo 6

Bibliografia

1. Hille, R. (2002). Molybdenum and tungsten in biology. *Trends in Biochemical Sciences* 27, 360–367.
2. Romão, M.J. (2009). Molybdenum and tungsten enzymes: a crystallographic and mechanistic overview. *Dalton Transactions* 4053.
3. Mendel, R.R., Bittner, F. (2006). Cell biology of molybdenum. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research* 1763, 621–635.
4. Johnson, M.K., Rees, D.C., Adams, M.W.W. (1996). Tungstoenzymes. *Chem. Rev.* 96, 2817–2840.
5. Schwarz, G., Mendel, R.R., Ribbe, M.W. (2009). Molybdenum cofactors, enzymes and pathways. *Nature* 460, 839–847.
6. Schwarz, G., Mendel, R.R. (2006). Molybdenum Cofactor Biosynthesis and Molybdenum Enzymes. *Annual Review of Plant Biology* 57, 623–647.
7. Leimkühler, S., Wuebbens, M.M., Rajagopalan, K.V. (2001). Characterization of *Escherichia coli* MoeB and Its Involvement in the Activation of Molybdopterin Synthase for the Biosynthesis of the Molybdenum Cofactor. *J. Biol. Chem.* 276, 34695–34701.
8. Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K.V., Rees, D.C. (1996). Crystal structure of DMSO reductase: redox-linked changes in molybdopterin coordination. *Science* 272, 1615–1621.
9. Boyington, J.C., Gladyshev, V.N., Khangulov, S.V., Stadtman, T.C., Sun, P.D. (1997). Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocysteine, and an Fe₄S₄ cluster. *Science* 275, 1305–1308.
10. Ellis, P.J., Conrads, T., Hille, R., Kuhn, P. (2001). Crystal Structure of the 100 kDa Arsenite Oxidase from *Alcaligenes faecalis* in Two Crystal Forms at 1.64 Å and 2.03 Å. *Structure* 9, 125–132.
11. Hille, R. (1996). The Mononuclear Molybdenum Enzymes. *Chem. Rev.* 96, 2757–2816.
12. Leimkühler, S., Kern, M., Solomon, P.S., McEwan, A.G., Schwarz, G., Mendel, R.R., Klipp, W. (1998). Xanthine dehydrogenase from the phototrophic purple bacterium *Rhodobacter capsulatus* is more similar to its eukaryotic counterparts than to prokaryotic molybdenum enzymes. *Molecular Microbiology* 27, 853–869.
13. Neumann, M., Mittelstädt, G., Iobbi-Nivol, C., Saggi, M., Lenzian, F., Hildebrandt, P., Leimkühler, S. (2009). A periplasmic aldehyde oxidoreductase represents the first molybdopterin cytosine dinucleotide cofactor containing molybdo-flavoenzyme from *Escherichia coli*. *FEBS Journal* 276, 2762–2774.
14. Glatigny, A., Scazzocchio, C. (1995). Cloning and Molecular Characterization of hxA, the Gene Coding for the Xanthine Dehydrogenase (Purine Hydroxylase I) of *Aspergillus nidulans*. *J. Biol. Chem.* 270, 3534–3550.
15. Hoff, T., Frandsen, G.I., Rocher, A., Mundy, J. (1998). Biochemical and genetic characterization of three molybdenum cofactor hydroxylases in *Arabidopsis thaliana*. *Biochimica Et Biophysica Acta (BBA) - Gene Structure and Expression* 1398, 397–402.

16. Mahler, H.R., Mackler, B., Green, D.E., Bock, R.M. (1954). Studies on Metalloflavoproteins III. Aldehyde Oxidase: A Molybdoflavoprotein. *J. Biol. Chem.* *210*, 465–480.
17. Mahro, M., Coelho, C., Trincão, J., Rodrigues, D., Terao, M., Garattini, E., Saggi, M., Lenzian, F., Hildebrandt, P., Romão, M.J., Leimkühler, S. (2011). Characterization and crystallization of mouse aldehyde oxidase 3: from mouse liver to *Escherichia coli* heterologous protein expression. *Drug Metab. Dispos.* *39*, 1939–1945.
18. Garattini, E., Mendel, R., Romão, M.J., Wright, R., Terao, M. (2003). Mammalian molybdo-flavoenzymes, an expanding family of proteins: structure, genetics, regulation, function and pathophysiology. *Biochem J* *372*, 15–32.
19. Nishino, T., Okamoto, K., Eger, B.T., Pai, E.F., Nishino, T. (2008). Mammalian xanthine oxidoreductase – mechanism of transition from xanthine dehydrogenase to xanthine oxidase. *FEBS Journal* *275*, 3278–3289.
20. Garattini, E., Terao, M. (2011). Increasing recognition of the importance of aldehyde oxidase in drug development and discovery. *Drug Metabolism Reviews* *43*, 374–386.
21. Pryde, D.C., Dalvie, D., Hu, Q., Jones, P., Obach, R.S., Tran, T.-D. (2010). Aldehyde Oxidase: An Enzyme of Emerging Importance in Drug Discovery. *J. Med. Chem.* *53*, 8441–8460.
22. Moura, J.J.G., Xavier, A.V., Bruschi, M., Le Gall, J., Hall, D.O., and Cammack, R. (1976). A molybdenum-containing iron-sulphur protein from *Desulfovibrio gigas*. *Biochemical and Biophysical Research Communications* *72*, 782–789.
23. Romão, M.J., Moura, J.J.G. (2001). Aldehyde Oxidoreductase (MOP). In *Handbook of Metalloproteins* (John Wiley & Sons, Ltd), 1037–1047.
24. Barata, B.A., LeGall, J., and Moura, J.J. (1993). Aldehyde oxidoreductase activity in *Desulfovibrio gigas*: in vitro reconstitution of an electron-transfer chain from aldehydes to the production of molecular hydrogen. *Biochemistry* *32*, 11559–11568.
25. Stams, F.J.M., Veenhuis, M., Weenk, G.H., Hansen, T.A. (1983). Occurrence of polyglucose as a storage polymer in *Desulfovibrio* species and *Desulfobulbus propionicus*. *Archives of Microbiology* *136*, 54–59.
26. Romão, M.J., Barata, B.A.S., Archer, M., Lobeck, K., Moura, I., Carrondo, M.A., LeGall, J., Lottspeich, F., Huber, R., Moura, J.J.G. (1993). Subunit composition, crystallization and preliminary crystallographic studies of the *Desulfovibrio gigas* aldehyde oxidoreductase containing molybdenum and [2Fe-2S] centers. *European Journal of Biochemistry* *215*, 729–732.
27. Thoenes, U., Flores, O.L., Neves, A., Devreese, B., Van Beeumen, J.J., Huber, R., Romão, M.J., LeGall, J., Moura, J.J.G., Rodrigues-Pousada, C. (1994). Molecular cloning and sequence analysis of the gene of the molybdenum-containing aldehyde oxidoreductase of *Desulfovibrio gigas*. *European Journal of Biochemistry* *220*, 901–910.
28. Moura, J.J.G., Xavier, A.V., Bruschi, M., LeGall, J., Cabral, J.M.P. (1977). A molybdenum-containing (2Fe, 2S) protein from *desulphovibrio gigas*, a sulphate reducer. *Journal of the Less Common Metals* *54*, 555–562.

29. Moura, J.J., Xavier, A.V., Cammack, R., Hall, D.O., Bruschi, M., LeGall, J. (1978). Oxidation-reduction studies of the Mo-(2Fe-2S) protein from *Desulfovibrio gigas*. *Biochem J* 173, 419–425.
30. Barata, B.A., Liang, J., Moura, I., LeGall, J., Moura, J.J., Huynh, B.H. (1992). Mössbauer study of the native, reduced and substrate-reacted *Desulfovibrio gigas* aldehyde oxidoreductase. *Eur. J. Biochem.* 204, 773–778.
31. Romão, M.J., Archer, M., Moura, I., Moura, J.J., LeGall, J., Engh, R., Schneider, M., Hof, P., Huber, R. (1995). Crystal structure of the xanthine oxidase-related aldehyde oxido-reductase from *D. gigas*. *Science* 270, 1170–1176.
32. Rebelo, J.M., Dias, J.M., Huber, R., Moura, J.J., Romão, M.J. (2001). Structure refinement of the aldehyde oxidoreductase from *Desulfovibrio gigas* (MOP) at 1.28 Å. *J. Biol. Inorg. Chem.* 6, 791–800.
33. Huber, R., Hof, P., Duarte, R.O., Moura, J.J., Moura, I., Liu, M.Y., LeGall, J., Hille, R., Archer, M., Romão, M.J. (1996). A Structure-Based Catalytic Mechanism for the Xanthine Oxidase Family of Molybdenum Enzymes. *PNAS* 93, 8846–8851.
34. Wahl, R., Rajagopalan, K. (1982). Evidence for the inorganic nature of the cyanolyzable sulfur of molybdenum hydroxylases. *J. Biol. Chem.* 257, 1354–1359.
35. Santos-Silva, T., Ferroni, F., Thapper, A., Marangon, J., González, P.J., Rizzi, A.C., Moura, I., Moura, J.J.G., Romão, M.J., Brondino, C.D. (2009). Kinetic, Structural, and EPR Studies Reveal That Aldehyde Oxidoreductase from *Desulfovibrio gigas* Does Not Need a Sulfido Ligand for Catalysis and Give Evidence for a Direct Mo–C Interaction in a Biological System. *J. Am. Chem. Soc.* 131, 7990–7998.
36. Boer, D.R., Thapper, A., Brondino, C.D., Romão, M.J., Moura, J.J.G. (2004). X-ray Crystal Structure and EPR Spectra of “Arsenite-Inhibited” *Desulfovibrio gigas* Aldehyde Dehydrogenase: A Member of the Xanthine Oxidase Family. *J. Am. Chem. Soc.* 126, 8614–8615.
37. Thapper, A., Boer, D., Brondino, C., Moura, J., Romão, M. (2007). Correlating EPR and X-ray structural analysis of arsenite-inhibited forms of aldehyde oxidoreductase. *Journal of Biological Inorganic Chemistry* 12, 353–366.
38. Berg, J.M., Tymoczko, J.L., Stryer, L. (2002). *Biochemistry*. 5ª Edição (W. H. Freeman and Company). Seção 9.1, Proteases: Facilitating a Difficult Reaction.
39. Madala, P.K., Tyndall, J.D.A., Nall, T., Fairlie, D.P. (2010). Update 1 of: Proteases Universally Recognize Beta Strands In Their Active Sites. *Chem. Rev.* 110, PR1–PR31.
40. Quesada, V., Ordóñez, G.R., Sánchez, L.M., Puente, X.S., López-Otín, C. (2009). The Degradome database: mammalian proteases and diseases of proteolysis. *Nucleic Acids Res* 37, D239–D243.
41. López-Otín, C., Bond, J.S. (2008). Proteases: Multifunctional Enzymes in Life and Disease. *J. Biol. Chem.* 283, 30433–30437.
42. Han, N., Jin, K., He, K., Cao, J., Teng, L. (2011). Protease-activated receptors in cancer: A systematic review. *Oncol Lett* 2, 599–608.

43. Base de dados MEROPS (<http://merops.sanger.ac.uk/>), último acesso a 17 de Agosto de 2012.
44. Barrett, A.J., Rawlings, N.D., O'Brien, E.A. (2001). The MEROPS Database as a Protease Information System. *Journal of Structural Biology* 134, 95–102.
45. Hedstrom, L. (2002). Serine Protease Mechanism and Specificity. *Chem. Rev.* 102, 4501–4524.
46. Di Cera, E. (2009). Serine Proteases. *IUBMB Life* 61, 510–515.
47. Rawlings, N.D., Barrett, A.J. (1994). Families of serine peptidases. *Meth. Enzymol.* 244, 19–61.
48. Kraut, J. (1977). Serine Proteases: Structure and Mechanism of Catalysis. *Annual Review of Biochemistry* 46, 331–358.
49. Evnin, L.B., Vásquez, J.R., Craik, C.S. (1990). Substrate specificity of trypsin investigated by using a genetic selection. *Proc Natl Acad Sci U S A* 87, 6659–6663.
50. Voet, D., Voet, J.G., Pratt, C.W. (2008). *Fundamentals of Biochemistry: Life at the Molecular Level*. 3ª Edição (John Wiley & Sons, Inc.). Secção 11.5, Serine Proteases, 347–360.
51. Chen, J.M., Ferec, C. (2000). Genes, cloned cDNAs, and proteins of human trypsinogens and pancreatitis-associated cationic trypsinogen mutations. *Pancreas* 21, 57–62.
52. Whitcomb, D. (1999). Hereditary pancreatitis: new insights into acute and chronic pancreatitis. *Gut* 45, 317–322.
53. Hirota, M., Ohmuraya, M., and Baba, H. (2006). The role of trypsin, trypsin inhibitor, and trypsin receptor in the onset and aggravation of pancreatitis. *Journal of Gastroenterology* 41, 832–836.
54. Neurath, H., Dreyer, W.J. (1955). Mechanism of activation of trypsinogen and chymotrypsinogen. *Discuss. Faraday Soc.* 20, 32–43.
55. Figarella, C., Clemente, F., Guy, O. (1969). On zymogens of human pancreatic juice. *FEBS Letters* 3, 351–353.
56. Rinderknecht, H., Renner, I.G., Carmack, C. (1979). Trypsinogen variants in pancreatic juice of healthy volunteers, chronic alcoholics, and patients with pancreatitis and cancer of the pancreas. *Gut* 20, 886–891.
57. Szmola, R., Kukor, Z., Sahin-Tóth, M. (2003). Human Mesotrypsin Is a Unique Digestive Protease Specialized for the Degradation of Trypsin Inhibitors. *J. Biol. Chem.* 278, 48580–48589.
58. Krieger, M., Kay, L.M., Stroud, R.M. (1974). Structure and specific binding of trypsin: comparison of inhibited derivatives and a model for substrate binding. *J. Mol. Biol.* 83, 209–230.
59. Blow, D.M. (1997). The tortuous story of Asp...His...Ser: Structural analysis of α -chymotrypsin. *Trends in Biochemical Sciences* 22, 405–408.

60. Huber, R., and Bode, W. (1978). Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* *11*, 114–122.
61. Ash, E.L., Sudmeier, J.L., Day, R.M., Vincent, M., Torchilin, E.V., Haddad, K.C., Bradshaw, E.M., Sanford, D.G., Bachovchin, W.W. (2000). Unusual ^1H NMR chemical shifts support (His) C ϵ 1—H \cdots O=C H-bond: Proposal for reaction-driven ring flip mechanism in serine protease catalysis. *Proc Natl Acad Sci U S A* *97*, 10371–10376.
62. Polgár, L. (2005). The catalytic triad of serine peptidases. *Cellular and Molecular Life Sciences* *62*, 2161–2172.
63. Silverman, G.A., Whisstock, J.C., Bottomley, S.P., Huntington, J.A., Kaiserman, D., Luke, C.J., Pak, S.C., Reichhart, J.-M., Bird, P.I. (2010). Serpins Flex Their Muscle. *J Biol Chem* *285*, 24299–24305.
64. Whisstock, J.C., Silverman, G.A., Bird, P.I., Bottomley, S.P., Kaiserman, D., Luke, C.J., Pak, S.C., Reichhart, J.-M., Huntington, J.A. (2010). Serpins Flex Their Muscle II. Structural Insights Into Target Peptidase Recognition, Polymerization, And Transport Functions. *J. Biol. Chem.* *285*, 24307–24312.
65. Bachovchin, W.W., Wong, W.Y.L., Farr-Jones, S., Shenvi, A.B., Kettner, C.A. (1988). Nitrogen-15 NMR spectroscopy of the catalytic-triad histidine of a serine protease in peptide boronic acid inhibitor complexes. *Biochemistry* *27*, 7689–7697.
66. Tsilikounas, E., Kettner, C.A., Bachovchin, W.W. (1992). Identification of serine and histidine adducts in complexes of trypsin and trypsinogen with peptide and nonpeptide boronic acid inhibitors by ^1H NMR spectroscopy. *Biochemistry* *31*, 12839–12846.
67. Kettner, C.A., Shenvi, A.B. (1984). Inhibition of the serine proteases leukocyte elastase, pancreatic elastase, cathepsin G, and chymotrypsin by peptide boronic acids. *J. Biol. Chem.* *259*, 15106–15114.
68. Barber, C.G., Dickinson, R.P., Horne, V.A. (2002). Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 1: 2-Pyridinylguanidines. *Bioorganic & Medicinal Chemistry Letters* *12*, 181–184.
69. Barber, C.G., Dickinson, R.P. (2002). Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 2: (3-Substituted-5-halo-2-pyridinyl)guanidines. *Bioorganic & Medicinal Chemistry Letters* *12*, 185–187.
70. Barber, C.G., Dickinson, R.P., Fish, P.V. (2004). Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 3: 1-Isoquinolinylguanidines. *Bioorganic & Medicinal Chemistry Letters* *14*, 3227–3230.
71. Yamane, J., Yao, M., Zhou, Y., Hiramatsu, Y., Fujiwara, K., Yamaguchi, T., Yamaguchi, H., Togame, H., Tsujishita, H., Takemoto, H., Tanaka, I. (2011). In-crystal affinity ranking of fragment hit compounds reveals a relationship with their inhibitory activities. *Journal of Applied Crystallography* *44*, 798–804.
72. Fish, P.V., Barber, C.G., Brown, D.G., Butt, R., Collis, M.G., Dickinson, R.P., Henry, B.T., Horne, V.A., Huggins, J.P., King, E., O’Gara, M., McCleverty, D., McIntosh, F., Philips, C., Webster, R. (2007). Selective Urokinase-Type Plasminogen Activator Inhibitors. 4. 1-(7-Sulfonamidoisoquinolinyl)guanidines. *J. Med. Chem.* *50*, 2341–2351.

73. Ulisse, S., Baldini, E., Sorrenti, S., D'Armiento, M. (2009). The urokinase plasminogen activator system: a target for anti-cancer therapy. *Curr Cancer Drug Targets* 9, 32–71.
74. Stillfried, G.E., Saunders, D.N., Ranson, M. (2007). Plasminogen binding and activation at the breast cancer cell surface: the integral role of urokinase activity. *Breast Cancer Res* 9, R14.
75. Plesner, T., Behrendt, N., Ploug, M. (1997). Structure, Function and Expression on Blood and Bone Marrow Cells of the Urokinase-Type Plasminogen Activator Receptor, uPAR. *Stem Cells* 15, 398–408.
76. Zhou, H.-M., Nichols, A., Meda, P., Vassalli, J.-D. (2000). Urokinase-type plasminogen activator and its receptor synergize to promote pathogenic proteolysis. *EMBO J* 19, 4817–4826.
77. Rhodes, G. (2006). *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models* (Academic Press).
78. Ewald, P.P., Editor (1962). *Fifty Years of X-ray Diffraction* (Reprinted in pdf format for the IUCr XVIII Congress, Glasgow, Scotland, Copyright © 1962, 1999 International Union of Crystallography).
79. Romão, M.J. (1996). Cristalografia de Proteínas: metodologias e aplicações em Bioquímica. *Boletim de Biotecnologia* 53, 18-36.
80. Blow, D. (2002). *Outline of Crystallography for Biologists* (Oxford University Press).
81. Li, M., Chang, W. (2009). Protein crystallization. *Photosyn. Res.* 102, 223–229.
82. Chayen, N.E., Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods* 5, 147–153.
83. Weber, P.C. (1997). [2] Overview of protein crystallization methods. In *Macromolecular Crystallography Part A*, (Academic Press), 13–22.
84. Hassell, A.M., An, G., Bledsoe, R.K., Bynum, J.M., Carter, H.L., Deng, S.-J.J., Gampe, R.T., Grisard, T.E., Madauss, K.P., Nolte, R.T., Rocque, W.J., Wang, L., Weaver, K.L., Williams, S.P., Wisely, G.B., Xu, R., Shewchuk, L.M. (2007). Crystallization of protein–ligand complexes. *Acta Crystallogr D Biol Crystallogr* 63, 72–79.
85. 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. *FEBS Journal* 275, 1–21.
86. Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G.W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D Biological Crystallography* 67, 271–281.
87. Kabsch, W. (2010). XDS. *Acta Crystallographica Section D Biological Crystallography* 66, 125–132.
88. Evans, P. (2005). Scaling and assessment of data quality. *Acta Crystallographica Section D Biological Crystallography* 62, 72–82.

89. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50, 760–763.
90. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography* 40, 658–674.
91. Cowtan K. (1994). 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, 31, p34-38.
92. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography* 67, 355–367.
93. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography* 66, 486–501.
94. Winn, M.D., Isupov, M.N., Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallographica Section D Biological Crystallography* 57, 122–133.
95. Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26, 283–291.
96. Nayal, M., Di Cera, E. (1996). Valence screening of water in protein crystals reveals potential Na⁺ binding sites. *J. Mol. Biol.* 256, 228–234.
97. Fridovich, I. (1970). Quantitative Aspects of the Production of Superoxide Anion Radical by Milk Xanthine Oxidase. *J. Biol. Chem.* 245, 4053–4057.
98. Chen, K., Kazachkov, M., Yu, P.H. (2007). Effect of aldehydes derived from oxidative deamination and oxidative stress on β -amyloid aggregation; pathological implications to Alzheimer's disease. *Journal of Neural Transmission* 114, 835–839.
99. Grimsrud, P.A., Xie, H., Griffin, T.J., Bernlohr, D.A. (2008). Oxidative Stress and Covalent Modification of Protein with Bioactive Aldehydes. *J. Biol. Chem.* 283, 21837–21841.
100. Tarini, M., Cignoni, P., Montani, C. (2006). Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 1237–1244.
101. Leiros, H.-K.S., McSweeney, S.M., Smalås, A.O. (2001). Atomic resolution structures of trypsin provide insight into structural radiation damage. *Acta Crystallographica Section D Biological Crystallography* 57, 488–497.

Capítulo 7

Anexos

7.1. Sequência peptídica da AOR de *D. gigas*

MIQKVITVNGIEQNLFVDAEALLSDVLRQQLGLTGVKVGCEQGQCACSVILDGKVVVRAC
 1 10 20 30 40 50 60
 VTKMKRVADGAQITTTIEGVGQPENLHPLQKAWVLHGGACQCGFCSPGFIVSAKGLLDTNAD
 61 70 80 90 100 110 120
 PSREDVRDWFQKHRNACRCTGYKPLVDAVMDAAAVINGKKPETDLEFKMPADGRIWGSKY
 121 130 140 150 160 170 180
 PRPTAVAKVTGTLDYGADLGLKMPAGTLHLAMVQAKVSHANIKGIDTSEALTMPGVHVSVI
 181 190 200 210 220 230 240
 THKDVKGKNRITGLITFPTNKGDGWRPILCDEKVFQYGDCIALVCADSEANARAAAEKV
 241 250 260 270 280 290 300
 KVDLEELPAYMSGPAAAAEDAIEIHPGTPNVYFEQPIVKGEDTGPIFASADVTVEGDFYV
 301 310 320 330 340 350 360
 GRQPHMPIEPDVAFAVMGDDGKCYIHSKSIGVHLHLYMIAPGVGLEPDQLVLVANPMGGT
 361 370 380 390 400 410 420
 FGYKFSPTSEALVAVAAMATGRPVHLRYNYQQQQYTGKRSPWEMNVKFAAKKDGTLLAM
 421 430 440 450 460 470 480
 ESDWLVDHGPYSEFGDLLTLRGAQFIGAGYNIPNIRGLGRTVATNHVWGSAFRGYGAPQS
 481 490 500 510 520 530 540
 MFASECLMDMLAEKLGMDPLELRYKNAYRPGDTNPTGQPEVFSLPDMIDQLRPKYQAAL
 541 550 560 570 580 590 600
 EKAQKESTATHKKGVGISIGVYGSGLDGPDAEAWAELNADGTITVHTAWEDHGQGADIG
 601 610 620 630 640 650 660
 CVGTAHEALRPMGVAPKIKFTWPNTATTPNSGPSGGSRQQVMTGNAIRVACENLLKACE
 661 670 680 690 700 710 720
 KPGGGYYTYDELKAADKPTKITGNWTASGATHCDAVTGLGKPFVVYMYGVFMAEVTVDVA
 721 730 740 750 760 770 780
 TGQTTVDGMTLMADLGLCNQLATDGGIYGGLAQQIGLALSDEFEDIKKHATLVGAGFPF
 781 790 800 810 820 830 840
 IKQIPDKLDIVYVNHPRPDGPF GASGVGELPLTSPHAAIINAIKSATGVRIRLPAYPEK
 841 850 860 870 880 890 900
 VLEALKA
 901 907

7.2. Sequência peptídica da tripsina de *B. taurus**

I VGGYTCGANTVPYQVSLNSGYHFCGGS LINSQWVVSAAHCYKSGIQVRLGEDNINVVEG
16 20 30 40 50 60 70 78
NEQFISASKSIVHPSYNSNTLNNDIMLIKLSAASLNSRVASISLPTSCASAGTQCLISG
79 90 100 110 120 130 140
WGNTKSSGTSYPDVLKCLKAPILSDSSCKSAYPGQITSNMFCAGYLEGGKDSCQGDSGGP
141 150 160 170 180 190 198
VVCSGKLQGI VSWGSGCAQKNKPGVYTKVCNYVSWIKQTIASN
199 210 220 230 240 245

*Foi utilizado o sistema de numeração da quimotripsina.

7.3. Estatísticas relativas à estrutura da DgAOR com Na₂S₂O₄ e Na₂S

Tabela 7.1. Estatísticas de recolha de dados para a estrutura de DgAOR obtida previamente à realização desta dissertação por *soaking* com Na₂S₂O₄ e Na₂S. Os valores entre parêntesis correspondem à última camada de resolução.

Cristal	DgAOR (<i>soaking</i> Na ₂ S ₂ O ₄ e Na ₂ S)
Comprimento de onda (Å)	0,931
Grupo espacial	P6 ₁ 22
Célula unitária (Å)	a,b = 143,3 c = 162,3
Limites de resolução (Å)	25,68–1,57 (1,65–1,57)
Nº de reflexões observadas	715 713 (96 433)
Nº de reflexões únicas	135 137 (19 580)
<i>Completeness</i> (%)	99,5 (99,9)
Redundância	5,3 (4,9)
I/σ(I) médio	12,5 (3,8)
R _{merge} ^a (%)	8,9 (35,4)
VM ^b - Parâmetro de Matthews (Å ³ /Da)	2,47
Conteúdo de solvente ^b (%)	50,2

(a) $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, onde $I_i(hkl)$ é a intensidade integrada de uma determinada reflexão i e $\langle I(hkl) \rangle$ é a intensidade média de múltiplas reflexões relacionadas por operações de simetria.

(b) $VM = \frac{V}{Z \times Mp}$, onde V é o volume da célula unitária, Z é o número de unidades assimétricas presentes na célula unitária e Mp é a massa molecular da proteína; Conteúdo de solvente = $\left(1 - \frac{1,23}{VM}\right) \times 100$

Tabela 7.2. Estatísticas de refinamento e validação para a estrutura de *DgAOR* obtida previamente à realização desta dissertação por *soaking* com $\text{Na}_2\text{S}_2\text{O}_4$ e Na_2S .

Conjunto de dados	<i>DgAOR</i> (<i>soaking</i> $\text{Na}_2\text{S}_2\text{O}_4$ e Na_2S)
Limites de resolução (Å)	25,68–1,57
Factor R (%)	14,1
Nº de reflexões	128 343
Factor R_{free} (%)	17,3
Nº de reflexões (R_{free})	6 793
Nº de resíduos	907
Nº de átomos	8 564
RMSD para o comprimento de ligação (Å)	0,027
RMSD para o ângulo de ligação (°)	2,51
Factores de temperatura médios (Å ²)	
átomos da cadeia principal	9,9
átomos das cadeias laterais	12,1
moléculas de água	29,5
Gráfico de Ramachandran (%)	
resíduos nas regiões mais favoráveis	91,3
resíduos em regiões adicionalmente favoráveis	8,0
resíduos em regiões generosamente permitidas	0,3
resíduos em regiões não permitidas	0,4