



Catarina Maria Campos Coelho
Licenciada em Bioquímica

Crystallographic Studies on Molybdopterin-dependent Enzymes

Dissertação para obtenção do Grau de Doutor em Bioquímica,
Especialidade Bioquímica Estrutural

Orientadora: Doutora Maria João Romão,
Professora Associada com Agregação, FCT/UNL
Co-orientador: Doutor José Trincão, Investigador Auxiliar do
Laboratório Associado Requimte, FCT/UNL

Júri

Presidente: Prof. Doutor Luís Fernando Lopes Monteiro
Arguentes: Doutora Sandra de Macedo Ribeiro
Doutora Inês Antunes Cardoso Pereira

Vogais: Prof. Doutora Isabel Maria Andrade Martins Galhardas de Moura
Prof. Doutor Carlos Frederico de Gusmão Campos Geraldes
Prof. Doutora Margarida Maria Portela Correia dos Santos Romão
Prof. Doutora Ana Paula Costa dos Santos Peralta Leandro
Doutor Shabir Hasein Najmudin
Doutor Pablo Javier González

 **FACULDADE DE
CIÊNCIAS E TECNOLOGIA**
UNIVERSIDADE NOVA DE LISBOA

Novembro 2011

Catarina Maria Campos Coelho

Crystallographic Studies on Molybdopterin-dependent Enzymes

DISSERTAÇÃO APRESENTADA PARA A OBTENÇÃO DO GRAU DE DOUTOR EM BIOQUÍMICA,
ESPECIALIDADE BIOQUÍMICA ESTRUTURAL, PELA FACULDADE DE CIÊNCIAS E
TECNOLOGIA DA UNIVERSIDADE NOVA DE LISBOA

Caparica, 3 de Novembro de 2011

Título: “Crystallographic Studies on Molybdopterin-dependent Enzymes”

Copyright em nome de Catarina Maria Campos Coelho.

“A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa tem 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.”

O trabalho apresentado nesta Dissertação foi realizado com o apoio financeiro da Fundação para a Ciência e Tecnologia (FCT–MCTES), através da atribuição da Bolsa de Doutoramento individual SFRH/BD/37948/2007, e dos projectos POCI/QUI/57641/2004 e PTDC/QUI/64733/2006, financiados pelo programa POCI2010 e co-financiados pelo Fundo Europeu de Desenvolvimento Regional.

“Não sou nada.
Nunca serei nada.
Não posso querer ser nada.
À parte isso, tenho em mim todos os sonhos do mundo.”

Álvaro de Campos: "Tabacaria"

AGRADECIMENTOS

(Acknowledgements)

Dedico esta Tese aos meus pais, Ana Bela e Paulo, esperando deste modo mostrar uma pequena parte da gratidão que sinto por tudo o que me deram ao longo desta vida.

Aos meus pais devo aquilo que sou.

O espírito crítico que me levou a seguir a carreira científica, e o tentar ir sempre mais além foi com eles que aprendi, e jamais esquecerei.

Obrigada Queridos Pais pela fantástica vida que me proporcionaram, por todo o apoio que me deram e por estarem sempre lá para mim.

Vocês serão sempre, sem qualquer dúvida o meu porto de abrigo!

Uma palavra especial ao meu Bruno,

que me compreende como ninguém, e me faz ser Feliz!

Que o futuro se mostre risonho para todos os planos de vida que temos em conjunto.

Um agradecimento sincero à minha orientadora, Prof. Maria João Romão.

Desde o tenro começo, a João sempre acreditou em mim e auxiliou a minha formação.

Também nunca deixou de me ajudar, mesmo com um mundo de coisas em mãos. Obrigada.

Existe ainda toda uma estrutura de apoio, sem a qual a existência desta Tese (e de mim própria tal como me conheço) não seria possível. Gostaria de mencionar com carinho as minhas avós Carmo e Elisa, a minha Tia Lucinda e o meu companheiro de sempre, o meu irmão Pedro. Há amigos sempre presentes, como a Irina, o Tagarela ... e outros que trago no coração, e a quem a humildade da amizade permite saberem quem são.

Também o Zé, o Pablo, a Mineko e o Enrico, o Martin e a Silke, a Guida e a Patrícia, a Ana Luísa, o Shabir, a Teresa a Cecília e o Filipe, muito ajudaram à existência desta Tese.

Bem-haja a todos!

RESUMO

Na presente Tese é descrita a determinação da estrutura tridimensional de duas enzimas dependentes de molibdénio, assim como a sua interpretação funcional. No Capítulo 1 é dada uma introdução geral acerca da utilização do molibdénio pelos sistemas biológicos, em particular a sua incorporação no centro activo de diversas enzimas. No mesmo capítulo, é também apresentada uma visão geral sobre a técnica de cristalografia de raios-X de proteínas, assim como alguns princípios básicos da mesma.

As aldeído oxidases são proteínas homodiméricas pertencentes à família da xantina oxidase (XO) de enzimas de molibdénio. A estrutura tridimensional da proteína aldeído oxidase homóloga1 (mAOH1) de rato encontra-se descrita no Capítulo 2, constituindo a primeira estrutura existente para uma aldeído oxidase. A proteína foi simultaneamente extraída de fígado de rato e expressa num sistema bacteriano (*E.coli*). A proteína recombinante permitiu determinar condições favoráveis de cristalização. Estas condições foram utilizadas na proteína nativa e levaram à obtenção de bons cristais, que possibilitaram a determinação da estrutura a uma resolução de 2.9Å. A estrutura foi resolvida por substituição molecular, usando como modelo a xantina oxidase de leite bovino, pertencente à mesma família que a mAOH1. No geral as duas proteínas são muito semelhantes, mas no entanto, as reações que catalisam são bastante diferentes. A análise da estrutura da mAOH1 e a comparação com a estrutura da XO permitiram chegar a importantes correlações entre a estrutura e a função, que explicam algumas das diferentes especificidades das duas enzimas. Estes estudos contribuem ainda para melhor compreender o papel das aldeído oxidases na saúde humana. Esta enzima tem sido objecto de estudo de diversas empresas farmacêuticas, na medida em que está envolvida em processos de detoxificação de várias drogas e xenobióticos, assumindo particular importância em estudos de “drug design”.

A nitrato reductase periplasmática da bactéria *Cupriavidus necator* (*Cn* NapAB) é uma proteína heterodimérica, pertencente à família das proteínas de molibdénio DMSO reductase. A estrutura tridimensional da *Cn* NapAB foi obtida a 1.5Å de resolução, a partir de cristais obtidos num robot de cristalização. Estudos estruturais, funcionais e espectroscópicos da *Cn* NapAB encontram-se descritos no Capítulo 3. A elevada resolução do modelo permitiu identificar a verdadeira natureza de todos os ligandos do molibdénio. Comprovou-se que o 6º elemento de ligação ao metal é um átomo de enxofre, e não oxigénio como estava estabelecido desde a elucidação da primeira estrutura de uma nitrato reductase (NapA de *Desulfovibrio desulfuricans*). Pensa-se que esta seja uma característica comum a todas as nitrato reductases, o que conduziu a diversas revisões do mecanismo reacional.

Para melhor caracterizar a proteína, foram ainda efectuados estudos espectroscópicos e electroquímicos, que demonstraram algumas características inesperadas, tal como o potencial redox dos dois hemos do tipo-*c* presentes na enzima.

Uma forma parcialmente reduzida da enzima NapAB de *C.necator* foi também obtida, e encontra-se descrita nos Capítulos 3 e 4. Esta forma foi obtida através da utilização de compostos recentemente descritos, e denominados Líquidos Iónicos (IL). Diversos estudos cristalográficos realizados utilizando IL como agentes de cristalização, têm-se revelado uma alternativa muito atraente para optimização de cristais de proteínas. No caso específico da proteína NapAB de *C.necator*, estes compostos revelaram-se fundamentais para reproduzir em maior escala os micro cristais inicialmente obtidos pelo robot de cristalização, e também para a obtenção de cristais da enzima parcialmente reduzida, que apresentou características inesperadas no centro activo.

Algumas conclusões gerais do presente trabalho, bem como perspectivas futuras encontram-se descritas no Capítulo 5.

ABSTRACT

This Thesis reports the determination of the crystal structure of two molybdenum-dependent enzymes, as well as its functional interpretation. In Chapter 1 is given a general introduction on the use of molybdenum in biological systems, particularly its incorporation into the active site of several enzymes. In the same chapter is also included an overview on X-ray protein crystallography, briefly describing its main basic principles.

Aldehyde oxidases are homodimeric proteins belonging to the xanthine oxidase (XO) family of molybdenum containing enzymes. The three-dimensional structure of mouse aldehyde oxidase homologue1 (mAOH1) is here reported and described (Chapter 2). This constitutes the first crystal structure ever obtained for an aldehyde oxidase. The mAOH1 protein was extracted from rat liver, and heterologously expressed in *E.coli*. The recombinant protein allowed determining suitable crystallization conditions, which were reproduced using the native enzyme from mouse liver. Suitable crystals were obtained, allowing to solve the protein structure at 2.9Å resolution, using bovine milk xanthine oxidase as a search model. Both proteins belong to the XO family of Mo proteins and are very similar, although catalyzing different reactions. The structure of mAOH1 and its comparison with the XO structure allowed drawing important structure and function correlations, and to explain the different enzyme specificities. These studies have also contributed to better understand the role of aldehyde oxidase in human health. The enzyme has received considerable attention from several pharmaceutical companies, as it is involved in the detoxification of several drugs and xenobiotics, assuming particular relevance in human health and drug design studies.

Periplasmic nitrate reductase from the *Cupriavidus necator* bacterium (*Cn* NapAB) is a heterodimeric protein, and belongs to the DMSO reductase family of molybdenum containing enzymes. The three-dimensional structure of the *C.necator* NapAB was solved at 1.5Å resolution using crystals obtained from a crystallization robot. Structural, spectroscopic and functional studies of this protein are reported in Chapter 3. The high resolution of the model, allowed identifying the true nature of all Mo ligands. In the first reported nitrate reductase crystal structure (NapA from *Desulfovibrio desulfuricans*), the 6th Mo ligand had been identified as an oxygen, but in *Cn* NapAB, a sulfur atom could be unambiguously assigned to this position. It is believed that this is a general feature of all nitrate reductases, which has led to the necessary revisions on the reaction mechanism for this family of enzymes. To further characterize *C.necator* NapAB, spectroscopic and

electrochemical studies have also been performed, and have shown unexpected features, particularly regarding the potential of the two *c*-type hemes.

A partially reduced form of *C.necator* NapAB was also obtained, and is described in Chapters 3 and 4. This form was obtained using some recently described compounds named Ionic Liquids (IL). Crystallographic studies performed using Ionic Liquids as crystallization agents, have proved to be an attractive alternative for optimizing protein crystals. In the specific case of *C.necator* NapAB, these compounds were fundamental to scale up the initially nano-crystals obtained from the crystallization robot, and also in obtaining crystals of the partially reduced enzyme that provided unexpected structural features on the protein active site.

Some general conclusions from this work are summarized, and some future perspectives are outlined in Chapter 5.

Abbreviations

AO	aldehyde oxidase
AOH1	aldehyde oxidase homologue1
AOH2	aldehyde oxidase homologue2
AOH3	aldehyde oxidase homologue3
Cn	<i>Cupriavidus necator</i>
COdh	carbon monoxide dehydrogenase
Dd	<i>Desulfovibrio desulfuricans</i>
Dg	<i>Desulfovibrio gigas</i>
DMSO	dimethylsulfoxide
DTT	Dithiothreitol
Ec	<i>Escherichia coli</i>
EDTA	ethylenediamine tetra-acetic acid
EPR	electron paramagnetic resonance
FAD	flavin adenine dinucleotide
Fdh	formate dehydrogenase
Hi	<i>Haemophilus influenza</i>
IL	Ionic liquids
mAOH1	mouse aldehyde oxidase homologue1
mARC	mitochondrial amidoxime reducing component
MCD	molybdopterin cytosine dinucleotide
MCSF	moco sulfurase
MFes	molybdoflavoenzymes
MGD	molybdopterin guanine dinucleotide
MoCo	molybdenum cofactor
MOP	<i>Desulfovibrio gigas</i> aldehyde oxidoreductase
MPT	molybdopterin monophosphate form
NAD⁺	nicotinamide adenine dinucleotide
Nap	periplasmic nitrate reductase
Nar	respiratory nitrate reductase
Nas	assimilatory nitrate reductase
NO₂⁻	nitrite
NO₃⁻	nitrate
PDB	Protein Data Bank
PEG	polyethylene glycol
PFV	protein film voltammetry
PGEh	homemade pyrolytic graphite electrode
rmsd	root mean square deviation
Rs	<i>Rhodobacter sphaeroides</i>
SDH	sulfite dehydrogenase
SO	sulfite oxidase
UTGEh	commercial ultra-trace graphite electrode
XDH	xanthine dehydrogenase
XO	xanthine oxidase
XOR	xanthine oxidoreductases

INDEX

INDEX.....	IX
FIGURE INDEX.....	XI
TABLE INDEX.....	XIV
CHAPTER 1.....	17
1.1 MOLYBDENUM ENZYMES.....	3
1.1.1 <i>Why Molybdenum?</i>	3
1.1.2 <i>Molybdenum Cofactor</i>	3
1.1.3 <i>Molybdenum Containing Enzymes</i>	6
1.1.4 <i>Xanthine Oxidase Family</i>	7
1.1.5 <i>Sulfite Oxidase Family</i>	9
1.1.6 <i>DMSO Reductase Family</i>	10
1.2 BIOMOLECULAR CRYSTALLOGRAPHY.....	14
1.2.1 <i>From X-rays to Crystal Structure</i>	14
1.2.2 <i>Protein Crystallization and Crystal Symmetry</i>	15
1.2.3 <i>How to obtain Data from Crystals?</i>	18
1.2.4 <i>Resolution, Refinement and Validation</i>	23
1.2.5 <i>Synchrotron Radiation</i>	24
CHAPTER 2.....	27
2.1 INTRODUCTION.....	29
2.2 EXPERIMENTAL PROCEDURE.....	34
2.2.1 <i>Purification of AOHI from Mouse Liver</i>	34
2.2.2 <i>Expression and Purification of mAOHI from E.coli</i>	34
2.2.3 <i>Protein Crystallization</i>	35
2.2.4 <i>Data Collection, Processing and Structure Solution</i>	37
2.2.5 <i>Model Building and Refinement</i>	39
2.3 RESULTS AND DISCUSSION.....	41
2.3.1 <i>Overall Structure Description</i>	41
2.3.2 <i>The Homodimer</i>	45
2.3.3 <i>Comparison with Structurally Related Proteins</i>	45
2.3.4 <i>mAOHI Catalytic Active Site and Substrate Funnel</i>	50
2.3.5 <i>The Reaction Mechanism</i>	55
CHAPTER 3.....	58
3.1 INTRODUCTION.....	61
3.1.1 <i>Cupriavidus necator</i>	61
3.1.2 <i>Heterodimeric Periplasmic Nitrate Reductase (NapAB)</i>	62
3.2 EXPERIMENTAL PROCEDURE.....	64
3.2.1 <i>Bacterial Expression and Growth Conditions</i>	64
3.2.2 <i>Periplasm Isolation and Protein Purification</i>	65
3.2.3 <i>Protein Crystallization and Data Collection</i>	65
3.2.4 <i>NapAB Crystals and Soaking Experiments</i>	69
3.2.5 <i>Processing</i>	69
3.2.6 <i>Structure Solution, Model Building and Refinement</i>	70
3.2.7 <i>EPR Spectroscopy</i>	71
3.2.8 <i>Spectropotentiometric Redox Titrations</i>	72
3.2.9 <i>Electrochemistry</i>	73
3.3 RESULTS AND DISCUSSION.....	74
3.3.1 <i>Overall Structure Description</i>	74
3.3.2 <i>The NapAB Heterodimer</i>	77
3.3.3 <i>Comparison with Homologous Crystal Structures</i>	79
3.3.4 <i>The Molybdenum Catalytic Site</i>	81
3.3.5 <i>The Catalytic Site in Dithionite Reacted Crystals</i>	83

3.3.6 Reaction Mechanism	84
3.3.7 The Presence of Oxidized Cysteines.....	85
3.3.8 Spectroscopic Characterization	85
3.3.9 Electrochemical Characterization	89
CHAPTER 4.....	95
4.1 INTRODUCTION.....	97
4.1.1 Limitations in Protein Crystallization.....	97
4.1.2 Limitations to Scale up Robot Conditions.....	98
4.1.3 Ionic Liquids and Proteins	99
4.1.4 Protein Crystallization and Ionic Liquids.....	100
4.2 C.NECATOR NAPAB: A CASE STUDY	102
4.2.1 Protein Expression and Purification	102
4.2.2 Crystallization	102
4.2.3 Data Collection and Analysis.....	103
4.3 RESULTS AND DISCUSSION.....	104
4.4 CONCLUSIONS	109
CHAPTER 5.....	111
5.1 GENERAL CONCLUSIONS.....	113
5.2 FUTURE PERSPECTIVES	114
REFERENCES 6.....	117
APPENDIX.....	127

FIGURE INDEX

FIGURE I. 1 - THE MONONUCLEAR MO AND W ENZYMES PYRANOPTERIN COFACTOR (MoCo) STRUCTURE: IN BLUE, THE MONOPHOSPHATE FORM (MPT); IN YELLOW AND GREEN, RESPECTIVELY, GUANINE (MGD) AND CYTOSINE (MCD) DINUCLEOTIDE FORM.....	4
FIGURE I. 2 - GENERAL SCHEME OF THE THREE MAIN STEPS INVOLVED IN THE BIOSYNTHETIC PATHWAY OF THE PYRANOPTERIN-BASED MO COFACTORS. (ADAPTED FROM [6])	5
FIGURE I. 3 - NITRATE REDUCTASES LOCALIZATION: NARGHI IS MEMBRANE ANCHORED, NAPAB IN THE PERIPLASM, AND NASA IN THE CYTOPLASM. (ADAPTED FROM [27, 34])	12
FIGURE I. 4 - THE CRYSTAL TO STRUCTURE PIPELINE: MOST IMPORTANT STEPS IN X-RAY PROTEIN CRYSTALLOGRAPHY.	15
FIGURE I. 5 - (A) VAPOR DIFFUSION TECHNIQUES: LEFT HANGING-DROP, RIGHT SITTING-DROP. HANGING-DROP IS COMMONLY USED IN MANUAL SETUPS, WHILE SITTING-DROP IS PREFERRED FOR ROBOTIC SETUPS; (B) SOLUBILITY PHASE DIAGRAM: BETWEEN THE SOLUBILITY LINE (BLUE) AND THE DECOMPOSITION LINE (RED), LIES THE METASTABLE REGION REPRESENTING THE SUPERSATURATED PROTEIN SOLUTION, WHERE THE FORMATION OF CRYSTALS WILL EVENTUALLY OCCUR.	16
FIGURE I. 6 - BRAGG DIFFRACTION EQUATION (LEFT) AND GEOMETRIC CONSTRUCTION (RIGHT): IF TWO BEAMS WITH IDENTICAL WAVELENGTH AND PHASE APPROACH A CRYSTALLINE SOLID AND ARE SCATTERED, CONSTRUCTIVE INTERFERENCE OCCURS WHEN THE LENGTH $2D_{hkl} \sin\theta$ IS EQUAL TO AN INTEGER MULTIPLE OF THE WAVELENGTH OF THE RADIATION (ADAPTED FROM [37]).	18
FIGURE I. 7 - REPRESENTATION OF A BASIC X-RAY DIFFRACTION EXPERIMENT: THE CRYSTAL WHEN EXPOSED TO X-RAYS PRODUCES INDIVIDUAL REFLECTIONS IN THE DETECTOR. DIFFRACTION IMAGES ARE A TRANSFORMATION (FT) OF THE CRYSTAL REAL SPACE (REAL LATTICE) INTO RECIPROCAL SPACE (RECIPROCAL LATTICE) AND VICE-VERSA.	19
FIGURE I. 8 - THE PHASE PROBLEM IN CRYSTALLOGRAPHY: TO RECONSTRUCT THE ELECTRON DENSITY MAP, THE STRUCTURE FACTOR AMPLITUDE AND THE PHASE ANGLE FOR EACH REFLECTION, MUST BE SUPPLIED. F_{hkl} IS OBTAINED EXPERIMENTALLY BUT A_{hkl} IS NOT, AND MUST BE ACQUIRED FROM ADDITIONAL PHASING EXPERIMENTS. V IS THE VOLUME OF THE UNIT CELL. (ADAPTED FROM [38]).	21
FIGURE I. 9 - ESRF GENERAL OVERVIEW (LEFT), AND SCHEMATIC REPRESENTATION (RIGHT).	25
FIGURE II. 1 - GENERAL SCHEMATIC REPRESENTATION OF MOLYBDO-FLAVOENZYME GENES IN VERTEBRATES (ADAPTED FROM [50]).	31
FIGURE II. 2 - CRYSTALS OF MAOH1 PROTEIN: (A) NEEDLES FROM THE RECOMBINANT PROTEIN; (B) CRYSTALS FROM THE NATIVE MOUSE LIVER PROTEIN; AND (C) CRYSTALS FROM THE Y885M VARIANT.	37
FIGURE II. 3 - DIFFRACTION PATTERN OF THE NATIVE MOUSE LIVER AOH1: (A) OBTAINED AT ID14-1 (IN THE UPPER LEFT QUADRANT AFTER RE-ANNEALING); (B) AND AT ID23-1 (ESRF). RESOLUTION AT THE EDGE IS 3.0\AA	38
FIGURE II. 4 - (A) RIBBON REPRESENTATION OF MAOH1 CRYSTAL STRUCTURE. MONOMER B IN GREY (LEFT), AND MONOMER A SHOWING THE DIFFERENT DOMAINS (RIGHT), COLORED AS: DOMAIN I IN RED (RESIDUES MET1-PRO169); DOMAIN II IN GREEN (RESIDUES THR232-LEU534); AND DOMAIN III IN BLUE (RESIDUES LEU576-VAL1335). THE LINKER REGIONS ARE REPRESENTED IN GREY (LINKER I AND II, RESPECTIVELY SER170-ASN231, AND LYS535-PRO575). THE MOLYBDOPTERIN MONOPHOSPHATE (MPT), $[2\text{Fe-2S}]$ AND FAD COFACTORS ARE SHOWN AS STICKS AND COLORED BY ATOM TYPE. (B) ARRANGEMENT AND DISTANCES BETWEEN THE DIFFERENT PROTEIN COFACTORS: MPT, THE TWO DISTINCT $[2\text{Fe-2S}]$ CENTERS, AND FAD.	42
FIGURE II. 5 - COMPARISON OF THE AMINO ACID SEQUENCE OF MAOH1 WITH XOR FROM BOVINE MILK, WITH XDH FROM <i>R. CAPSULATUS</i> AND MOP FROM <i>D. GIGAS</i> . THE MAOH1 $[2\text{Fe-2S}]$ I AND II BINDING CYS RESIDUES ARE MARKED AS (*) AND (+) RESPECTIVELY. BLACK BOXES INDICATE IDENTICAL RESIDUES, AND GREY SIMILAR ONES. ALIGNMENT RESULTS OBTAINED WITH CLUSTALW MULTIPLE SEQUENCE ALIGNMENT PROGRAM AT THE EMBL-EBI WEB SERVER, AND THE FIGURE PREPARED USING BOXSHADE FROM EMBNET SERVER.....	43
FIGURE II. 6 - COMPARISON BETWEEN MAOH1 (BLUE) AND THE DIFFERENT CRYSTAL STRUCTURES OBTAINED FOR THE XO FAMILY: (A) ALDEHYDE OXIDOREDUCTASE FROM <i>D. GIGAS</i> (MOP MONOMER - YELLOW, PDB CODE: 1VLB), (B) XANTHINE DEHYDROGENASE FROM <i>R. CAPSULATUS</i> (<i>Rc</i> XDH - RED, PDB CODE: 1JRO), (C) XANTHINE DEHYDROGENASE FROM BOVINE MILK (XDH MILK - GREEN, PDB CODE: 1FO4). NEXT TO EACH STRUCTURE COMPARISON IS THE SUPERPOSITION OF THE COFACTORS INVOLVED IN ELECTRON TRANSFER (MAOH1 - ATOMS COLOR CODED, MOP IN YELLOW, <i>Rc</i> XDH IN RED, AND XDH MILK IN GREEN).	46
FIGURE II. 7 - REPRESENTATION OF THE "MOVING LOOP" (RESIDUES GLN423-LYS433, IN BOVINE XDH NUMBERING) AND IMPORTANT RESIDUES INVOLVED IN THE XDH/XO CONVERSION. (A) RESIDUES PHE549, TRP336 AND THE ORIGINAL LOOP POSITION IN BOVINE XDH (GREEN). (B) COMPARISON BETWEEN MAOH1 (BLUE) AND BOVINE XDH (GREEN), CONCERNING THE "MOVING LOOP" AND TRP336 FROM XDH (WHICH CORRESPONDS TO THR343 IN MAOH1). THE "MOVING LOOP" IS MAINLY CONSTITUTED BY HIGHLY CHARGED RESIDUES (QASRREDDIAK) IN XDH, WHICH ARE REPLACED BY LESSER CHARGED ONES IN MAOH1 (QAPRQQNAFAT) (THE 4 CONSERVED RESIDUES ARE UNDERLINED). (C) DIFFERENCES OBSERVED IN THE LOOP AND IN RESIDUE TRP336 AFTER THE XDH \rightarrow XO CONVERSION (XDH IN GREEN, AND XO IN ORANGE). (D) SAME AS IN (C) BUT FROM A DIFFERENT PERSPECTIVE, OUTLINING THE DIFFERENCES BETWEEN THE "MOVING LOOP" FROM XO (IN ORANGE), AND THE MAOH1 AND XDH STRUCTURES (BLUE AND GREEN, RESPECTIVELY). ALL THE FAD COFACTORS ARE REPRESENTED COLOR CODED, AND CORRESPOND TO THE MAOH1 STRUCTURE.	49
FIGURE II. 8 - CLOSE-UP OF THE MPT AND MO COORDINATION SPHERE FOR THE MAOH1 MONOMER A.	51
($2M_{F_o} - DF_c$ ELECTRON DENSITY MAP CONTOURED AT 1σ , BLUE).	51

FIGURE II. 9 – ACTIVE SITE COMPARISON BETWEEN MAOH1 (YELLOW) AND BOVINE XDH (GRAY). OUTLINED ARE THE MOST IMPORTANT NON-CONSERVED RESIDUES, IN MAOH1 NUMBERING. ALSO PRESENT IS THE XDH INHIBITOR OXIPURINOL MOLECULE (OXI), AND THE CONSERVED XDH AND MAOH1 RESIDUES PHE919, PHE1014 AND GLU1266.	52
FIGURE II. 10 - SURFACE VIEW OF THE FUNNEL LEADING TO THE ACTIVE SITE IN MAOH1 (LEFT) AND BOVINE XDH (RIGHT). THE MO ATOM IS SEEN AT THE END OF THE FUNNEL AS A GREEN SPHERE. THE ENTRANCE IS MUCH WIDER IN THE CASE OF ALDEHYDE OXIDASE, WITH A NARROW CONSTRICTION CLOSER TO THE MOLYBDENUM ACTIVE SITE. ELECTROSTATIC SURFACE POTENTIALS WERE CALCULATED USING THE PROGRAM DELPHY [71] AND REPRESENTED IN PYMOL [72], WITH THE COLOR OF SURFACE POTENTIALS IN THE SCALE RANGE FROM -0.10V (NEGATIVELY CHARGED, IN RED) TO +0.10V (POSITIVELY CHARGED, IN BLUE).....	55
FIGURE II. 11 – PROPOSED REACTION MECHANISM FOR ALDEHYDE OXIDASE (AO). REPRESENTED IS THE BASE-ASSISTED NUCLEOPHILIC ATTACK OF THE MO-OH ON THE SUBSTRATE CARBONYL, WITH CONCOMITANT HYDRIDE TRANSFER TO THE MO=S. (ADAPTED FROM [16, 74]).	56
FIGURE III. 1 - CRYSTALS OBTAINED USING THE CRYSTALLIZATION ROBOT AT EMBL, GRENOBLE. (A) MICROCRYSTALS OBTAINED FOR CONDITIONS NUMBER: 41, 65, 89, 82, 90, 43, 84, 93, 42, 86, 47 AND 71 FROM HAMPTON INDEX SCREEN (B) ENLARGEMENT OF CRYSTALLIZATION CONDITION NUMBER 89; CRYSTAL DIMENSIONS 0.13 x 0.09 x 0.04 MM ³	66
FIGURE III. 2 - DIFFRACTION PATTERN OBTAINED AT BEAMLINE ID14-1 (ESRF) FOR A NAPAB CRYSTAL. RESOLUTION AT THE EDGE IS 1.5Å. ...	67
FIGURE III. 3 - A) CRYSTAL STRUCTURE OF THE <i>C.NECATOR</i> NAPAB (RIBBON REPRESENTATION) SHOWING THE DOMAINS COLORED AS FOLLOWS: DOMAIN I IN RED (RESIDUES 4-68, 507-531 AND 600-641); DOMAIN II IN GREEN (RESIDUES 69-148, 386-506 AND 532-599); DOMAIN III IN YELLOW (RESIDUES 149-385); AND DOMAIN IV IN BLUE (RESIDUES 642-802). THE MO-BIS-MGD COFACTOR IS SHOWN IN STICKS AND COLORED BY ATOM TYPE, AND THE [4FE-4S] AS ORANGE STICKS. NAPB IS REPRESENTED IN GREY RIBBON AND THE 2 HEMES AS STICKS COLORED BY ATOM TYPE. B) ARRANGEMENT OF THE METAL COFACTORS: MO (MGD) ₂ , [4FE-4S] AND DI-HEME C-TYPE CYTOCHROME COLOR-CODED AS ATOM TYPES. INCLUDED ARE ALSO TWO STRICTLY CONSERVED RESIDUES FROM NAPA: LYS56A THAT MEDIATES THE CONTACT BETWEEN THE PTERIN AND THE [4FE-4S] CLUSTER AS WELL AS TYR58A, INVOLVED IN INTERSUBUNIT ELECTRON TRANSFER BETWEEN THE [4FE-4S] CENTER AND HEME I.	75
FIGURE III. 4 - SUPERPOSITION OF THE NAPB SUBUNIT FROM <i>C.NECATOR</i> (RED) WITH NAPB FROM <i>R.SPHAEROIDES</i> (GREEN) AND THE SEGMENT OF RESIDUES 38 TO 105 FROM <i>H.INFLUENZAE</i> (BLUE). THE CN NAPA IS REPRESENTED AS A GRAY SURFACE.	77
FIGURE III. 5 - A) COMPARISON OF THE AMINO ACID SEQUENCE OF NAPA FROM <i>C.NECATOR</i> WITH THE OTHER NITRATE REDUCTASE CATALYTIC SUBUNITS OF KNOW STRUCTURES (<i>R.SPHAEROIDES</i> , <i>E.COLI</i> AND <i>D.DESULFURICANS</i>). THE [4FE-4S] BINDING CYS RESIDUES ARE MARKED AS (▼) AND THE MO LIGAND CYS152 AS (*). B) COMPARISON OF THE AMINO ACID SEQUENCE HOMOLOGY OF NAPB FROM <i>C.NECATOR</i> , <i>R.SPHAEROIDES</i> , <i>H.INFLUENZAE</i> AND <i>E.COLI</i> . BINDING RESIDUES (H---CXXCH) FOR HEME I ARE MARKED AS (●) AND FOR HEME II AS (○). BLACK BOXES INDICATE IDENTICAL RESIDUES, AND GREY INDICATE SIMILAR ONES. ALIGNMENT RESULTS OBTAINED WITH CLUSTALW MULTIPLE SEQUENCE ALIGNMENT PROGRAM AT THE EMBL-EBI WEB SERVER, AND THE FIGURE PREPARED USING BOXSHADE FROM EMBNET SERVER.	78
FIGURE III. 6 - ELECTROSTATIC POTENTIALS OF NAP STRUCTURES SURFACES. ELECTROSTATIC SURFACE POTENTIALS WERE CALCULATED USING THE PROGRAM DELPHY [71] AND REPRESENTED IN PYMOL [72], WITH THE COLOR OF SURFACE POTENTIALS IN THE SCALE RANGE FROM -0.10 V (NEGATIVELY CHARGED, IN RED) TO +0.10 V (POSITIVELY CHARGED, IN BLUE). A) VIEW TOWARDS THE SUBSTRATE TUNNEL LEADING TO THE ACTIVE SITE OF Cn NAPAB, Rs NAPAB, Dd NAPA AND Ec NAPA; B) NAPA SURFACE INTERACTION WITH THE SMALL SUBUNIT NAPB, FROM Cn, Rs AND Ec (NAPB FROM Cn), ALL IN THE SAME ORIENTATION.	80
FIGURE III. 7 – STEREO VIEW OF THE SUBSTRATE FUNNEL WITH 3 FORMATE IONS AND CONSERVED CHARGED RESIDUES INVOLVED IN ORIENTING THE SUBSTRATE INTO THE MO ACTIVE SITE (DIT-IL STRUCTURE).	82
FIGURE III. 8 - CLOSE-UP OF THE MO COORDINATION SPHERE FOR THE NATIVE (A AND B) AND DITHIONITE-REDUCED FORMS (C). THE 2MF _o -DF _c ELECTRON DENSITY MAPS ARE CONTOURED AT 1Σ (BLUE) AND THE MF _o -DF _c MAP AT 3Σ (RED). A) REFINING THE SIXTH MO LIGAND AS AN OXYGEN ATOM YIELDED POSITIVE ELECTRON DENSITY ON AN MF _o -DF _c MAP. THIS DENSITY COMPLETELY DISAPPEARED WHEN ASSIGNING IT AS A SULFUR ATOM (B). C) IN THE PARTIALLY DITHIONITE-REDUCED FORM (DIT-IL) THERE IS EXTRA POSITIVE DENSITY AT THE C ^B ATOM OF CYS152 THAT COULD BE MODELED AS TWO ALTERNATE CONFORMATIONS WITH 88% AND 12% OCCUPANCIES.....	83
FIGURE III. 9 - A) UV-VIS MEDIATED POTENTIOMETRIC REDOX TITRATION OF THE C-TYPE HEMES PRESENT IN <i>C.NECATOR</i> PERIPLASMIC NITRATE REDUCTASE. THE ABSORBANCE AT 553 NM WAS NORMALIZED AND PLOTTED VERSUS THE REDOX POTENTIAL OF THE SOLUTION E (AT PH 7.00). EXPERIMENTAL DATA (CIRCLES) WERE FITTED WITH A NONLINEAR REGRESSION (GREY LINE) USING THE NERNST EQUATION FOR TWO INDEPENDENT COMPONENTS WITH E _m VALUES OF 160±5 mV AND 50±2 mV (ERRORS FROM OXIDATIVE AND REDUCTIVE TITRATIONS). B) EPR SPECTROPOTENTIOMETRIC REDOX TITRATION OF THE [4FE-4S] CLUSTER PRESENT IN THE LARGE SUBUNIT OF Cn NAPAB. THE AREA OF THE FeS EPR SIGNAL WAS NORMALIZED AND PLOTTED VERSUS THE REDOX POTENTIAL OF THE SOLUTION E (AT PH 7.00). EXPERIMENTAL DATA (CIRCLES) WERE FITTED WITH A NONLINEAR REGRESSION (GRAY LINE) USING THE NERNST EQUATION FOR ONE ELECTRON WITH AN E _m VALUE OF -15 mV.....	86
FIGURE III. 10 - - ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTRA OBTAINED FOR THE PERIPLASMIC <i>CUPRIAVIDUS NECATOR</i> NITRATE REDUCTASE HETERODIMER NAPAB. A) SAMPLE AS-PREPARED UNDER AEROBIC ATMOSPHERE MEASURED AT 15 K. B) SAMPLE AS-PREPARED REDUCED IN ANAEROBIC CONDITIONS (DIOXYGEN CONCENTRATION BELOW 1 PPM) WITH 5 mM SODIUM DITHIONITE MEASURED AT 25K. INSET OF PANEL B SHOWS THE REDUCED SAMPLE MEASURED AT 25K (BLACK) AND 50K (GRAY). ARROWS INDICATE THE POSITION OF THE MAIN G-VALUES FOR THE HEME AND [4FE-4S] ⁺¹ CENTER	88

FIGURE III. 11 – CYCLIC VOLTAMMOGRAMS OF <i>C.NECATOR</i> NAPAB ADSORBED TO THE PYROLYTIC GRAPHITE ELECTRODE IN NON-TURNOVER CONDITIONS AND 10 mM TRIS-HCL PH 7.6 AND 0.1M NaCl (SCAN RATES: 5, 10, 15, 20, 30 AND 50 mV s^{-1}).....	89
FIGURE III. 12 - CYCLIC VOLTAMMOGRAMS OF <i>C.NECATOR</i> NAPAB ADSORBED AT A GRAPHITE ELECTRODE IN THE PRESENCE OF INCREASING CONCENTRATIONS OF H_2O_2 , IN 10 mM TRIS-HCL PH 7.6 AND 0.1 M NaCl ($v = 20 \text{ mV s}^{-1}$, 3000 RPM).	91
FIGURE IV. 1 – SOME COMMON IONIC LIQUIDS USED IN PROTEIN CRYSTALLOGRAPHY.	100
FIGURE IV. 2 - MICRODROP CRYSTALLIZATION IN THE ABSENCE OF IONIC LIQUIDS: (A) CONDITION 0.2M SODIUM FORMATE AND 15% PEG 3350, PROTEIN CONCENTRATION 5 MG/ML; (B) SEA URCHINS OBTAINED USING 0.2M SODIUM FORMATE AND 12.5% PEG 3350; (c) STREAK SEEDING FROM DROP (B); (D) LONG BUT FINE NEEDLES OBTAINED IN THE SAME CONDITION (A) BUT WITH PROTEIN CONCENTRATION 10 MG/ML.	104
FIGURE IV. 3 - MICRODROP CRYSTALLIZATION IN THE PRESENCE OF (A) 0.4M $[\text{C}_4\text{MIM}][\text{MDEGSO}_4]$ (B) 0.2M $[\text{C}_4\text{MIM}][\text{MDEGSO}_4]$ AND (C) 0.2M $[\text{C}_4\text{MIM}]\text{Cl}^-$. CRYSTALLIZATION CONDITIONS ARE 0.2M SODIUM FORMATE, 15% PEG 3350.	105
FIGURE IV. 4 - MICRODROP CRYSTALS OBTAINED IN THE PRESENCE OF 0.4M $[\text{C}_4\text{MIM}]\text{Cl}^-$ REVEALING CRYSTALLIZATION REPRODUCIBILITY. CRYSTALLIZATION CONDITIONS ARE 0.2M SODIUM FORMATE AND 15% PEG 3350. APPROXIMATE CRYSTAL DIMENSIONS 0.3 x 0.1 x 0.1 mm^3	105
FIGURE IV. 5 - <i>C.NECATOR</i> NAPAB PROTEIN DIFFRACTION PATTERN OF A CRYSTAL OBTAINED IN-HOUSE, USING 0.4M $[\text{C}_4\text{MIM}]\text{Cl}^-$. RESOLUTION AT EDGE IS 2Å AND DATASET NAME NAP_IL (2009) FROM TABLE IV.1	106
FIGURE IV. 6 - CHANGES IN THE CRYSTAL PACKING. UPPER LEFT: CRYSTAL PACKING FOR THE NATIVE PROTEIN, WITH UNIT CELL CONSTANTS $a = 142.2\text{Å}$, $b = 82.4\text{Å}$, $c = 96.8\text{Å}$ AND $\beta = 100.7^\circ$ (46.9% SOLVENT); UPPER RIGHT: PACKING OF CRYSTALS OBTAINED WITH IL, UNIT CELL CONSTANTS $a = 119.4\text{Å}$, $b = 71.4\text{Å}$, $c = 128.4\text{Å}$ AND $\beta = 121.0^\circ$ (51.5% SOLVENT CONTENT); CENTER: COMPARISON OF THE TWO UNIT CELLS CONSTANTS, PROJECTED ALONG THE b AXIS. WITH IL, a AND b AXES DECREASE RESPECTIVELY 16% AND 13%, WHILE c AXES INCREASED 33%.	107

TABLE INDEX

TABLE I. 1 - REPRESENTATIVE PYRANOPTERIN DEPENDENT ENZYMES AND RESPECTIVE FAMILIES.....	7
TABLE I. 2 - CRYSTAL SYSTEMS AND ALLOWED SPACE GROUPS FOR PROTEIN MOLECULES. FOUR TYPES OF UNIT CELL: PRIMITIVE (P); CENTERED IN THE SIDE (C), IN THE BODY (I), OR IN THE FACE (F). THE SPHERES IN THE BRAVAIS LATTICE REPRESENT THE LATTICE POINTS.	17
TABLE II. 1 – RELEVANT ALDEHYDE OXIDASE SUBSTRATES AND INHIBITORS.....	30
TABLE II. 2 - X-RAY CRYSTALLOGRAPHY DATA-COLLECTION STATISTICS. (VALUES IN PARENTHESIS CORRESPOND TO THE HIGHEST RESOLUTION SHELL)	39
TABLE II. 3 – STRUCTURE REFINEMENT DATA.	40
$^A R_{\text{WORK}} = \sum F_{\text{CALC}} - F_{\text{OBS}} / \sum F_{\text{OBS}} \times 100$, WHERE F_{CALC} AND F_{OBS} ARE THE CALCULATED AND OBSERVED STRUCTURE FACTOR AMPLITUDES, RESPECTIVELY. ($^B R_{\text{FREE}}$ IS CALCULATED FOR A RANDOMLY CHOSEN 5% OF THE REFLECTIONS).....	40
TABLE II. 4 – COMPARISON BETWEEN MAOH1 AND BOVINE XDH, CONCERNING THE MAIN RESIDUES INVOLVED IN SUBSTRATE BINDING AND CATALYSIS. NON-CONSERVED RESIDUES MARKED WITH AN ASTERISK ARE MENTIONED ABOVE, DUE TO THEIR PRIMARY IMPORTANCE.	54
TABLE III. 1 - STATISTICS ON DATA COLLECTION AND STRUCTURE REFINEMENT. VALUES IN PARENTHESIS CORRESPOND TO THE HIGHEST-RESOLUTION SHELL. $^A R_{\text{WORK}} = \sum F_{\text{CALC}} - F_{\text{OBS}} / \sum F_{\text{OBS}} \times 100$, WHERE F_{CALC} AND F_{OBS} ARE THE CALCULATED AND OBSERVED STRUCTURE FACTOR AMPLITUDES, RESPECTIVELY. (R_{FREE} IS CALCULATED FOR A RANDOMLY CHOSEN 5% OF THE REFLECTIONS FOR EACH DATASET).	68
TABLE III. 2 - THE MO ACTIVE SITE AND B FACTOR ANALYSIS.	71
TABLE III. 3 - REDOX POTENTIAL OF THE METAL COFACTORS COMPRISING THE ELECTRON PATHWAY CONNECTING NAPC TO THE MO ION AT THE ACTIVE SITE OF PERIPLASMIC NITRATE REDUCTASES. (VALUES BETWEEN PARENTHESES CORRESPOND TO REDOX POTENTIAL OBTAINED IN THE TITRATION PERFORMED ON THE MONOMER NAP A OR NAP B. VALUES ARE GIVEN IN MV WITH RESPECT TO THE STANDARD HYDROGEN ELECTRODE). <i>SE</i> , <i>SYNECHOCOCCUS ELONGATES</i> AND <i>PP</i> , <i>PARACOCCUS PANTOTROPHUS</i>	87
TABLE IV. 1 – CRYSTAL PARAMETERS AND X-RAY DIFFRACTION DATA STATISTICS FOR THE BEST DIFFRACTING CRYSTALS OBTAINED USING IONIC LIQUIDS.....	103

CHAPTER 1

General Introduction

1.1 MOLYBDENUM ENZYMES

1.1.1 Why Molybdenum?

Molybdenum is part of life, and part of our lives, and it has long been known that this element is an essential nutrient for plants, animals and microorganisms. Molybdenum is the only second row-transition metal in the periodic table, that is required by the majority of living organisms, and the few species that do not require molybdenum use tungsten, which lies immediately below Mo in the 6th group (and has similar characteristics). Molybdenum can be considered a trace element, meaning that small amounts are necessary, but higher ones can be toxic or even lethal. Both molybdenum and tungsten possess an extraordinary chemical versatility that is useful for biological systems: in the case of Mo, its oxidation state varies from +2 to +6, and it can be coordinated to 4 or up to 8 ligands. Due to molybdate anion (MoO_4^{2-}) solubility in water, molybdenum is easily available to the biological systems. In soils, for instance, the MoO_4^{2-} anion is the only form of molybdenum available to plants and bacteria [1, 2].

Although a minor constituent of the earth's crust, molybdenum is the most abundant transition metal in seawater, which resembles the primitive initial soup from where life emanated. So it is not surprising that it has been incorporated in a diverse range of biological systems such as nitrogenases and molybdopterin-dependent enzymes. Molybdenum itself is inactive, and in order to gain biological activity, it has to be complexed by a pterin compound forming the molybdenum cofactor (pyranopterin - MoCo), which is part of the catalytic center of nearly all molybdenum containing enzymes. This can be considered the most important use of molybdenum in living organisms [1, 2].

1.1.2 Molybdenum Cofactor

The majority of MoCo dependent enzymes catalyze redox reactions. These proteins take advantage of the chemical versatility of the metal, which is controlled by the cofactor itself and the enzyme environment. Two very different systems have developed in nature to control this redox state and catalytic power of molybdenum, which has the ability to function as a catalyst in oxygen transfer reactions. With the exception of the multinuclear MoFe₇ cluster present in bacterial nitrogenases, all other molybdenum dependent enzymes use the metal in a mononuclear form, with an organic tricyclic pyranopterin cofactor coordinated to it. Only in the case of carbon monoxide dehydrogenase (CODH), Mo has been found in a dinuclear Mo-S-Cu form [3]. The cofactor was originally thought

to be present only in Mo enzymes, so it was initially named MoCo. It was later discovered that this cofactor is also present in tungsten enzymes, and its designation was changed to pyranopterin to avoid confusion. The cofactor coordinates to the metal via its dithiolene function, and although representing an integral component of the active center of these enzymes, it does not seem to participate directly in catalysis. The pterin can be present in either its dinucleotide or monophosphate form. While in eukaryotes the pyranopterin is in the simplest monophosphate form (MPT), in prokaryotes it can also be conjugated to nucleosides, usually cytidine (MCD) or guanosine (MGD) (Figure I. 1). The cofactor is responsible for the correct positioning of molybdenum in the active center, for the control of its redox behavior, and also participates in the electron transfer that occurs to and from the metal. Molybdenum containing enzymes are found in all aerobic organisms, whereas tungsten enzymes occur essentially in obligate anaerobes (typically thermophilic) [1, 4-6].

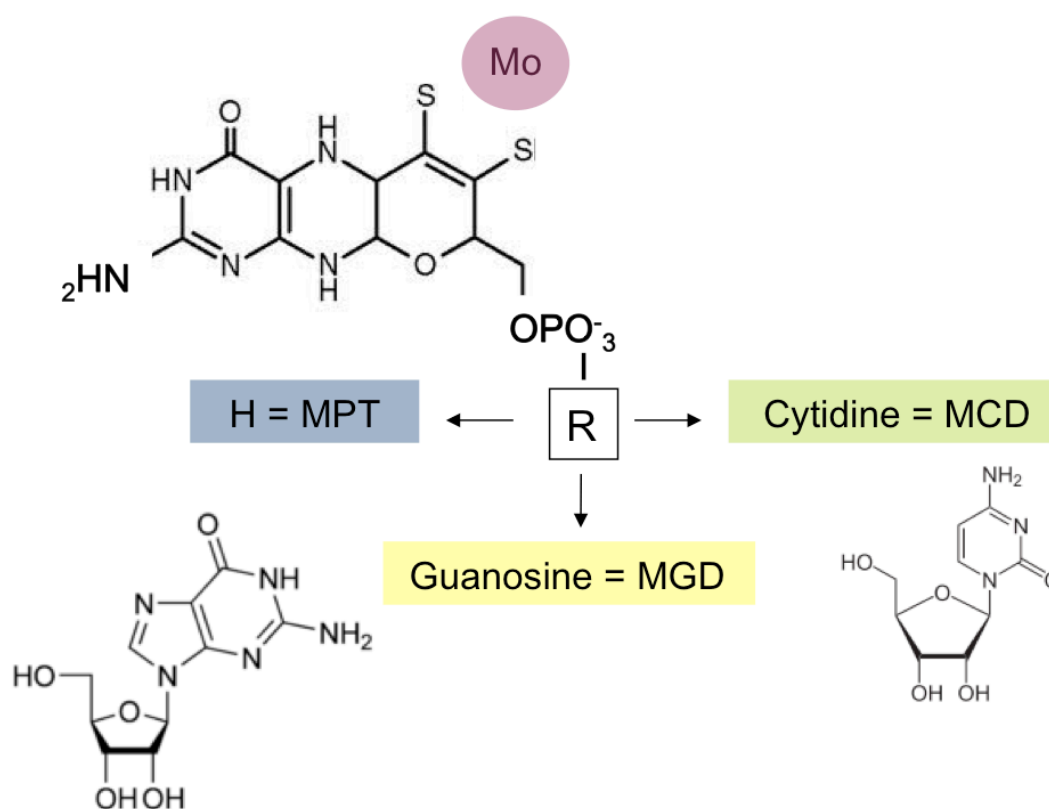


Figure I. 1 - The mononuclear Mo and W enzymes pyranopterin cofactor (MoCo) structure: in blue, the monophosphate form (MPT); in yellow and green, respectively, guanosine (MGD) and cytidine (MCD) dinucleotide form.

The mononuclear molybdenum enzymes are widely distributed in nature, dispersed in a variety of organisms from all different kingdoms. Although catalyzing a wide diversity of reactions in the metabolism of nitrogen, sulfur and carbonyl compounds, the MoCo cofactor is synthesized by a

conserved pathway. Using several biochemical, genetic, and structural approaches, MoCo biosynthesis in *E. coli* has been extensively studied. These studies have identified at least 17 different genes involved in the biosynthesis of this complex cofactor. The biosynthetic pathway can be briefly divided in four main steps, according to its intermediates: (i) formation of cyclic pyranopterin monophosphate (cPMP); (ii) conversion of cPMP to metal-binding pterin (MPT); (iii) insertion of molybdenum to form MoCo; and (iii) additional modification of MoCo by the attachment of different nucleotides to form the cytosine or the guanosine cofactors (MCD and MGD, respectively) (Figure I. 2). The first three steps are similar for all molybdenum containing proteins for all microorganisms, but the final modification step varies between different proteins and organisms, occurring only in prokaryotes. Mutational block of the cofactor biosynthesis, results in loss of essential metabolic functions [4, 6, 7].

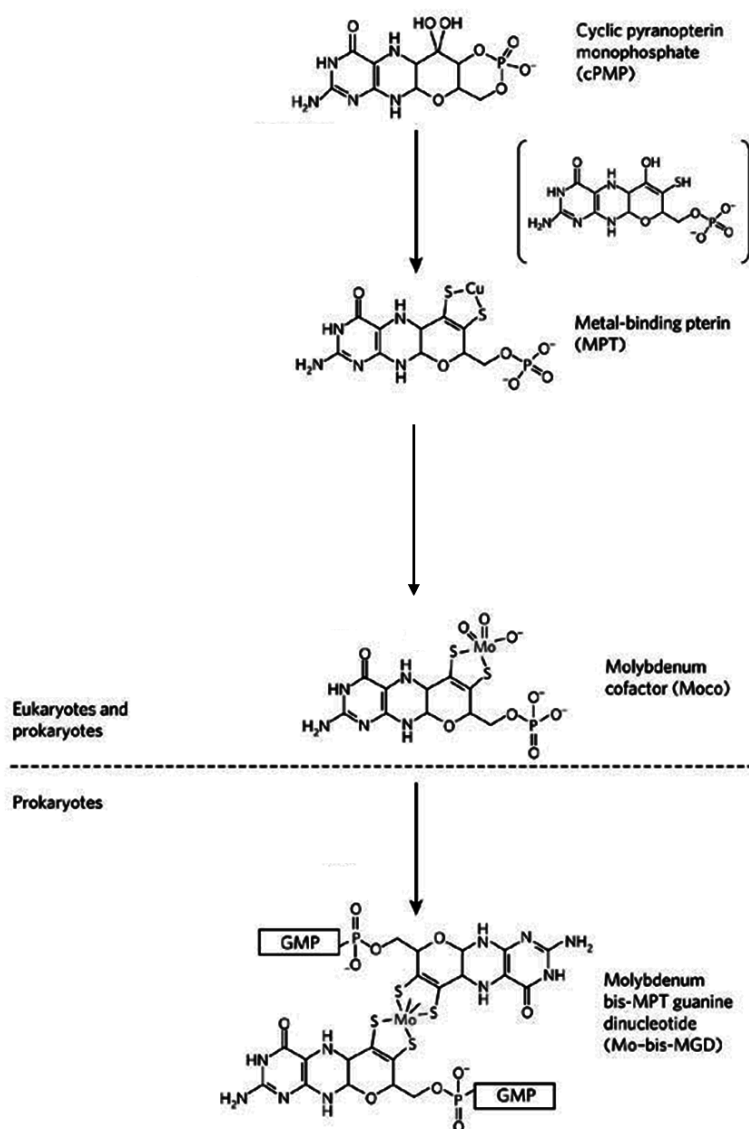


Figure I. 2 – General scheme of the three main steps involved in the biosynthetic pathway of the pyranopterin-based Mo cofactors. (Adapted from [6])

1.1.3 Molybdenum Containing Enzymes

More than 50 different MoCo containing enzymes have been described to date, which have been classified on the basis of the Mo active site coordination characteristics [5]. The majority of Mo dependent proteins catalyze oxo-transfer reactions, coupled to electron transport between substrate and other cofactors, namely iron-sulfur centers, hemes or flavins. In the catalytic center, molybdenum is coordinated to the cis-dithiolene group of one or two pyranopterins. Coordination is completed by the side chain of a cysteine, seleno-cysteine, serine or aspartate residue, and/or by coordination of a oxygen or sulfur atom in the oxo or sulfide group forms. These coordinating elements have such a large diversity that, based on X-ray structural data, primary sequence alignments and biochemical characterization, pyranopterin containing enzymes have been divided in to three large families, summarized in Table I. 1: the xanthine oxidase (XO) family, with one MCD or MPT; the sulfite oxidase (SO) (and assimilatory nitrate reductase) family, with one MPT; and the DMSO reductase (DMSOR) family, containing two MGD cofactors [5, 6, 8].

All eukaryotic molybdenum enzymes belong exclusively to either the sulfite oxidase or the xanthine oxidase family. They differ in the nature of the third Mo-S ligand, which is either provided by a terminal sulfido ligand (XO) or an enzyme-derived cysteine (SO). Members of the DMSOR family are very diverse in terms of reaction, function and structure, and the majority of them work under anaerobic conditions, whereby their respective cofactors serve as terminal electron acceptors in respiratory metabolism. Some tungsten containing enzymes, such as *Desulfovibrio gigas* formate dehydrogenase (Fdh), belong to the DMSO reductase family, and have high homology with the corresponding molybdenum enzymes. A new protein catalyzing the oxidative hydroxylation of amine substrates has been recently described, in both humans and plants. Mitochondrial amidoxime reducing component (mARC), appears to contain a distinct molybdenum coordination sphere, and could represent a new family of eukaryotic molybdenum enzymes [5, 6, 9, 10].

The study of molybdenum enzymes has become increasingly facilitated in recent years, mainly due to the possibility of heterologous protein expression in prokaryotic systems. The expression of molybdoenzymes is however an extremely complex process, controlled by the transportation of Mo into the cell, by the cofactor biosynthesis and apo-protein expression transcriptional regulation [8]. In recent decades, crystal structures for several members of the XO, SO and DMSO reductase families have provided considerable insight into how active site architecture is implicated in substrate

specificity, and enabled reaction mechanism determination. Crystallographic studies of molybdoenzymes have revealed that MoCo is deeply buried within the enzyme, and usually in close proximity to the iron/sulfur centers. A more detailed characterization of representative enzymes of each family is given below.

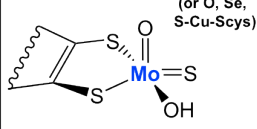
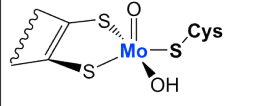
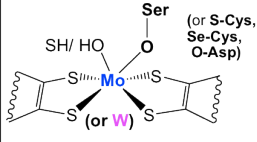
Families	Cofactor	Kingdom	Enzymes	Source	Abbrev.	PDB ID	Active site
Xanthine Oxidase	MCD / MPT	Eukaryotic Prokaryotic	Xanthine Oxidase/Dehydrogenase Xanthine Dehydrogenase Aldehyde Oxidoreductase Aldehyde Oxidase Homologue1	Bovine milk <i>R.capsulatus</i> <i>D.gigas</i> Mouse	XO / XDH <i>Rc</i> XDH MOP mAoH1	1FIQ/1FO4 1JRO 1VLB -	 (or O, Se, S-Cu-Scys)
Sulfite Oxidase	MPT	Eukaryotic Prokaryotic	Sulfite Oxidase Sulfite Oxidase Assimilatory Nitrate Reductase	Chicken liver <i>A.thaliana</i> <i>Pichia</i> <i>Angusta</i>	SO <i>At</i> SO Nas	1SOX 1OGP 2BIH/2BII	 Cys
DMSO Reductase	2 x MGD	Prokaryotic	Periplasmic Nitrate Reductase Membrane Nitrate Reductase Formate Dehydrogenase Periplasmic Nitrate Reductase	<i>D.desulfuric</i> <i>E.coli</i> <i>D.desulfuric</i> <i>C.necator</i>	NapA NarGH/GHI FDH NapAB	2NAP 1R27/1Q16 1HOH 3ML1	 Ser (or S-Cys, Se-Cys, O-Asp)

Table I. 1 - Representative pyranopterin dependent enzymes and respective families.

1.1.4 Xanthine Oxidase Family

Enzymes of the xanthine oxidase (XO) family are the best characterized mononuclear molybdenum containing enzymes, and in general catalyze the oxidative hydroxylation of a diverse range of aldehyde and aromatic heterocycles. The chemical reaction involves the cleavage of a C-H bond, and the formation of a new C-O bond:



This reaction occurs after substrate interaction with the Mo center, which is later reduced from Mo (VI) to Mo (IV). The two resulting reducing equivalents are transferred to a protein external acceptor: NAD^+ in the case of xanthine dehydrogenase (XDH), or O_2 in xanthine oxidase, along an electron transfer pathway. This process occurs through the protein cofactors, which mediates the electronic exchange between substrate and the electron acceptor. Molybdenum enzymes of the xanthine oxidase family catalyze hydroxylation of carbon centers, and use water as oxygen atom source, rather than O_2 that is later incorporated into the product [11, 12].

Members of this family are broadly distributed in eukaryotes, prokaryotes and archaea, and comprise xanthine oxidoreductases (XOR), aldehyde oxidases (AO) and carbon monoxide dehydrogenases (CODH). Xanthine oxidoreductases have been isolated from a wide range of organisms, and are the most well studied enzymes of the XO family of molybdenum proteins, since its crystal structure was solved in 2000, by Nishino and co-workers [13]. These mammalian proteins constitute the key enzymes in the catabolism of purines, oxidizing hypoxanthine to xanthine, and xanthine into the terminal catabolite uric acid. They are synthesized in the dehydrogenase form (XDH) and exist mostly as such in the cell. By oxidation of sulfhydryl residues or by proteolysis, XDH can be converted into oxidase form (XO), which fails to react with NAD^+ and exclusively uses dioxygen as electron acceptor. This results in the formation of reactive oxygen species, namely superoxide anion and hydrogen peroxide [5, 13].

Eukaryotic xanthine oxidoreductases usually exist as homodimers, and each monomer can be divided into three distinct domains, according to cofactor localization. The N-terminal domain harbors a pair of spectroscopic distinct [2Fe-2S] centers, followed by a FAD containing domain, and a C-terminal domain containing the molybdenum center and cofactor. In the case of prokaryotic enzymes a different and more complex structure can be found. XDH from *Rhodobacter capsulatus* is a dimer of dimers, combining a fusion of the FAD and [2Fe-2S] domains into one subunit (XdhA), and the molybdenum-binding portion into another (XdhB). The catalytic Mo active site possesses a distorted square pyramidal coordination geometry. The apical position is occupied by a Mo=O group, and the four equatorial ligands are: two sulfurs from the pyranopterin cofactor, a terminal Mo=S group, and a Mo-OH ligand. Generally, electron transfer pathway proceeds from Mo to the nearest iron/sulfur center, and from this (usually Fe/SII) to the next iron/sulfur center, until the electrons reach FAD. In a different variation, prokaryotic aldehyde oxidoreductase from *Desulfovibrio gigas* (MOP) does not contain the FAD cofactor, and electrons are passed through a complex of several subunits and redox centers [11, 12, 14].

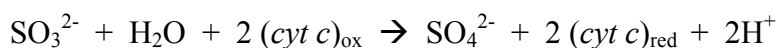
MoCo-dependent enzymes from the XO family that also contain the FAD cofactor have been classified in a sub-family named molybdo-flavoenzymes (MFEs) [15]. The most representative enzymes are the eukaryotic aldehyde oxidases, and the xanthine oxidoreductases.

Unlike xanthine oxidoreductases, which have been studied for more than a century, information on the physiological role of aldehyde oxidase in vertebrates, mammals and humans in particular, is still vague, in spite of the emerging amount of data on their primary structure. Despite the tremendous level of similarity, aldehyde oxidase and xanthine oxidoreductases have different substrate and inhibitor specificities, as well as biochemical functions. XOR recognize xanthine and hypoxanthine as substrates, but neither of them is a good substrate for the aldehyde oxidase proteins purified so far (including the homologues, explained in detail in Chapter 2). AO has relaxed substrate specificity, meaning that the protein can accommodate various types of compounds, generally characterized by aldehyde functionality (an aromatic or heterocyclic structure). It is also interesting to point out that allopurinol and oxipurinol, strong inhibitors of XOR, can inhibit AO to a lesser extent [13, 16, 17].

There is an urgent need to determine the “missing” three-dimensional structure of an aldehyde oxidase. This will allow for the identification of the structural determinants that form the basis of the different substrate specificities between AO and XOR. New structural and spectroscopic studies can also help to elucidate the exact nature of the molecular processes that occurs during catalysis. This would also improve our understanding in the evolution of these proteins [11, 12, 16, 18].

1.1.5 Sulfite Oxidase Family

The sulfite oxidase family comprises plant assimilatory nitrate reductases and sulfite oxidizing enzymes. The sulfite oxidizing enzymes can be separated in two groups according to their ability to transfer electrons to molecular oxygen: the sulfite oxidase (SO) found in animals and plants; and the sulfite dehydrogenases (SDH) encountered in bacteria. In animals, SO enzymes catalyze the oxidation of sulfite to sulfate, using ferricytochrome *c* as the physiological electron acceptor:



This constitutes the final step in the oxidative degradation of sulfur containing amino acid residues cysteine and methionine. SO also plays an important role in detoxifying the excess of sulfite and sulfur dioxide, exogenously supplied [5, 19].

In the oxidized (Mo VI) state of the enzyme, the molybdenum atom is in a distorted square pyramidal coordination geometry connected by five ligands: two from the dithiolene group of molybdopterin (MPT), the side chain of a strictly conserved cysteine, and two non-peptide oxygen

atoms (Table I. 1). Animal sulfite oxidase is a homodimer, located in the mitochondrial intermembrane space. Each subunit contains three distinct domains: a Mo cofactor domain, a dimerization domain and a smaller heme containing domain. The X-ray structure of human SO is not yet available, but the structure of the highly homologous chicken liver was solved in 1997 by Kisker and co-workers [20]. Also the crystal structures from plant *Arabidopsis thaliana* sulfite oxidase, which is the smallest eukaryotic Mo enzyme (lacking the heme domain), and the soil bacteria *Starkeya novella* sulfite dehydrogenase (SDH) were determined some years later [21, 22]. Unlike animal SO, the plant enzyme is localized in peroxisomes, and does not react with cyt *c*, but uses oxygen instead as the terminal electron acceptor. It is also important to note that, all the other families of molybdenum containing enzymes contain [Fe-S] clusters, or interact with iron/sulfur subunits, while members of the SO family do not [8, 19].

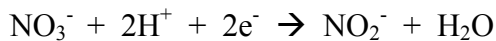
In humans, sulfite oxidase deficiency is an inherited sulfur metabolic disorder, resulting in deep birth defects, severe neonatal neurological problems and early death. Schwarz and co-workers may be close to a treatment (unpublished results), but to date no effective therapies are known. This deficiency can be caused by mutations in the SO gene, or due to defects in the MoCo biosynthetic pathway, as mentioned in section 1.1.2. Mutations in the genes can be divided into three groups: mutations causing conformational changes in the protein active site; mutations affecting dimerization contacts; and mutation of residues coordinating MoCo [19, 23].

1.1.6 DMSO Reductase Family

The dimethyl sulfoxide (DMSO) reductase family of molybdenum enzymes is the largest and displays a greater diversity. With the exception of formate dehydrogenases (Fdh), the members of this family catalyze in general the transfer of an oxygen atom, to or from the substrate. The first crystal structure reported for this family was the DMSO reductase from *Rhodobacter sphaeroides*, by Schindelin and co-workers [24]. Nevertheless, our knowledge about this family was greatly increased with further studies on nitrate reductases and formate dehydrogenases. Interestingly, all crystal structures determined to date from the DMSO reductase family of proteins possess high degree of similarity and overall fold, and considerable variations in metal coordination and active site amino acid residues. In the active site, the metal is coordinated by two pterin cofactors in the dinucleotide form (MGD) and different types of ligands. These differences account for the high diversity of functions performed by the enzymes of this family [5, 25].

Nitrate Reductases

Nitrogen is essential for biomolecules, such as proteins and nucleic acids. The nitrogen cycle involves a number of redox reactions in which prokaryotes have an important role, since only they possess the enzymatic machinery necessary to carry out the process. The dissimilatory reactions involve conversion of nitrate into dinitrogen (respiration), or into ammonia (respiration /ammonification). The assimilatory ones (which also involve conversion of nitrate to ammonia), start with the reduction of nitrate in the cytoplasm. All the reductive steps of the nitrogen cycle are started by nitrate reductases (NR), with release of one water molecule:



With the exception of eukaryotic assimilatory nitrate reductase from the SO family, all other nitrate reductases belong to the DMSO reductase family. There are three kinds of nitrate reductase that can be used to initiate the respiratory denitrification, or ammonification processes in prokaryotes. All of them bind a Mo-bis-molybdopterin guanine dinucleotide cofactor and at least one [4Fe-4S] center. Nitrate reductases are divided according to their cellular localization: Nar is a complex protein anchored to the membrane, Nap is located in the periplasmic compartment and Nas can be found in the cytoplasm [26, 27] (Figure I. 3).

Respiratory nitrate reductases (Nar) in bacteria have been extensively studied. All the Nar proteins isolated to date are composed of three different subunits (NarGHI): NarG which contains the MoCo active site plus a [4Fe-4S] center; NarH which has three additional [4Fe-4S] centers, and a [3Fe-4S] center; and NarI, which is an integral membrane protein with two b type hemes. NarG and NarH are in the cytoplasm, anchored to the membrane by NarI. Crystal structures for NarGHI, and NarGH were reported some years ago by Bertero and Jormakka, and co-workers respectively [28, 29]. These structures revealed interesting aspects, such as the presence of an aspartate side chain as a molybdenum ligand. Nar receives electrons from quinol, usually ubiquinol in denitrifiers, and is therefore linked to respiratory electron transfer [5, 27]. The periplasmic dissimilatory nitrate reductases (Nap) are also linked to quinol oxidation, but do not transduce the free energy from the quinol/nitrate couple into H^+ motive force. In Nap, electrons from quinol are usually transferred through one or two cytochrome containing proteins (NapC or NapB) to the catalytic subunit NapA, which contains the MoCo cofactor and a [4Fe-4S] center [27]. The first crystal structure of a dissimilatory nitrate reductase was solved in 1999, by Dias and co-workers [30]. This nitrate reductase isolated from *Desulfovibrio desulfuricans* (*Dd*) is a monomeric NapA protein, while the

majority of Nap enzymes purified from other organisms, such as *Rhodobacter sphaeroides* (*Rs*), reveals the presence of a small NapB subunit, containing two c-type hemes [31]. In the oxidized form, molybdenum is coordinated by four sulfurs from the dithiolene moiety (bis-MGD), a cysteinyl ligand from the polypeptide chain, and a recently established sulfur ligand, in a distorted hexa-coordinated Mo^{VI} geometry. The enzymatic mechanism of this group of proteins, have been recently revised and discussed, based on new structural evidences found in the *Dd* NapA, and corroborated with the recent crystal structure of *Cupriavidus necator* NapAB [32, 33]. The high resolution (1.5Å) crystal structure of *C.necator* NapAB is discussed in detail in Chapter 3 of this Thesis.

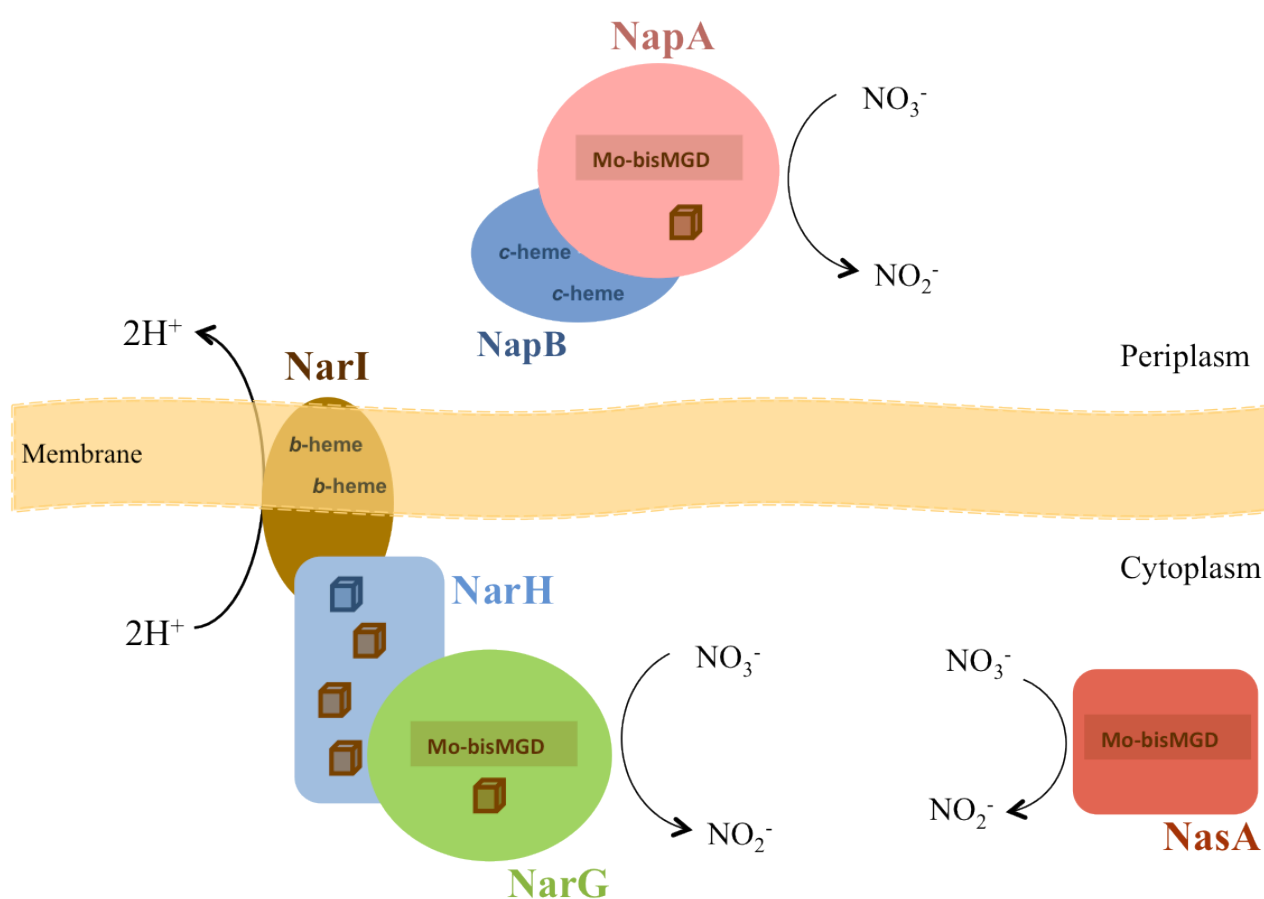


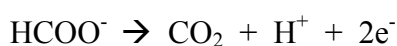
Figure I. 3 - Nitrate reductases localization: NarGHI is membrane anchored, NapAB in the periplasm, and NasA in the cytoplasm. (Adapted from [27, 34])

The assimilatory nitrate reductases (Nas) from different organisms have in common the presence of the Mo-bis-MGD cofactor, but differ largely in the number and type of electron transfer centers. No crystal structure has been reported for this sub-group of proteins, but monomeric Nas from

cyanobacteria are closely related to *Dd* NapA, indicating that Nas proteins are more related with Nap, than with Nar [35].

Formate Dehydrogenases

The few examples of enzymes capable of incorporating molybdenum or tungsten in their catalytic active site belong to the DMSO reductase family. Formate dehydrogenase (Fdh) is one of them, and it catalyzes the oxidation of formate to carbon dioxide:



The two Fdh proteins from *E.coli* (Fdh-H and Fdh-N, from the formate lyase complex and membrane bound, respectively) contain molybdenum in the catalytic center, while Fdh from *D.gigas* has tungsten. In spite of some heterogeneity in the protein assembly (*E.coli* Fdh-H is a monomer, while Fdh-N a heterotrimer and *Dg* Fdh is a heterodimer), the three known crystal structures of formate dehydrogenases exhibit very similar active site geometries. The Mo/W metal adopts a distorted trigonal prismatic geometry and is coordinated to four dithiolene sulfur atoms (from the two MGD cofactors) by a selenium atom from a conserved SeCys residue and by a sixth oxygen or sulfur ligand. Also the overall three dimensional structure and cofactor arrangement is very similar in the three enzymes [5, 25].

Given all the above and considering that the first three-dimensional crystal structure of a Mo enzyme was determined only 16 years ago (MOP protein, in 1995 by Romão and co-workers), the achievements made for this family of enzymes in such a small period of time are remarkable. This was possible mainly due to the information obtained from X-ray protein crystallography, complemented by spectroscopic and functional studies. The importance of X-ray protein crystallography is explained next.

1.2 BIOMOLECULAR CRYSTALLOGRAPHY

X-ray crystallography of important biological molecules started with Dorothy Hodgkin (1910-1994), who solved the structure of small molecules such as cholesterol, vitamin B12 and penicillin [36]. In 1950, the first protein crystal structure of sperm whale myoglobin was solved by Max Perutz (1914-2002) and Sir John Kendrew (1917-1997). D. Hodgkin was also able to solve the structure of a protein (insulin), but only in 1969, after more than thirty years of working on it. These were great achievements that were awarded the Nobel Prize in Chemistry, to D. Hodgkin in 1964, and to Perutz/Kendrew in 1962. All of this happened no more than 60 years ago, and this year (July 2011), there were over 74400 structures deposited in the protein data bank (PDB). Although some of them were obtained by other techniques, namely Nuclear Magnetic Resonance (NMR) and Cryoelectron Microscopy (cryoEM), the majority (>85%) of the structures were determined by X-ray crystallography. How can the great success of this technique be explained?

1.2.1 From X-rays to Crystal Structure

The determination of molecular structures by crystallography became possible with the discovery of X-rays, by Wilhelm Röntgen in the late 19th century. The name X-rays was due to the fact that this was an unknown type of radiation at the time, but in Germanic language it is still called Röntgen radiation. X-rays are a high-energy electromagnetic radiation and thus part of the electromagnetic spectrum, with wavelength between ultraviolet (UV) and gamma (γ) rays. The major advantages of X-rays are its high energy and short wavelength, similar to the inter-atomic bond distances, approximately 1Å. X-rays can be obtained by shooting a metal target (usually Cu or Mo, to obtain higher or shorter wavelength, respectively) with electrons produced by a heated filament and accelerated by an electric field. Due to their wave behavior, X-rays can be diffracted by the periodically arrangement of atoms in a crystal. Once obtained, a diffraction pattern from the X-ray scattering in the crystal can be used to construct electron density maps. However, additional phase information is first needed to overcome the phase problem in crystallography. It often needs acquisition of new diffraction data with added information from anomalous scatterers or heavy atoms, when estimates of the phases cannot be obtained from homologous structures. Once the phase is solved, the preliminary model can be built into the experimental electron density map, which must be progressively completed and refined. Validation is also necessary prior to deposition of the coordinates in the PDB [37-39]. X-ray crystallography can be viewed as a super microscope: we use

the “lens” to focus the diffraction pattern obtained when we pass X-rays through the orderly packed molecules in the crystal, and the similarity in magnitude between X-rays and inter-atomic bond lengths gives us the possibility to actually see the internal arrangement of atoms in our protein. This process is a pipeline, from crystal to structure, which is summarized in Figure I. 4 and will be explained below.

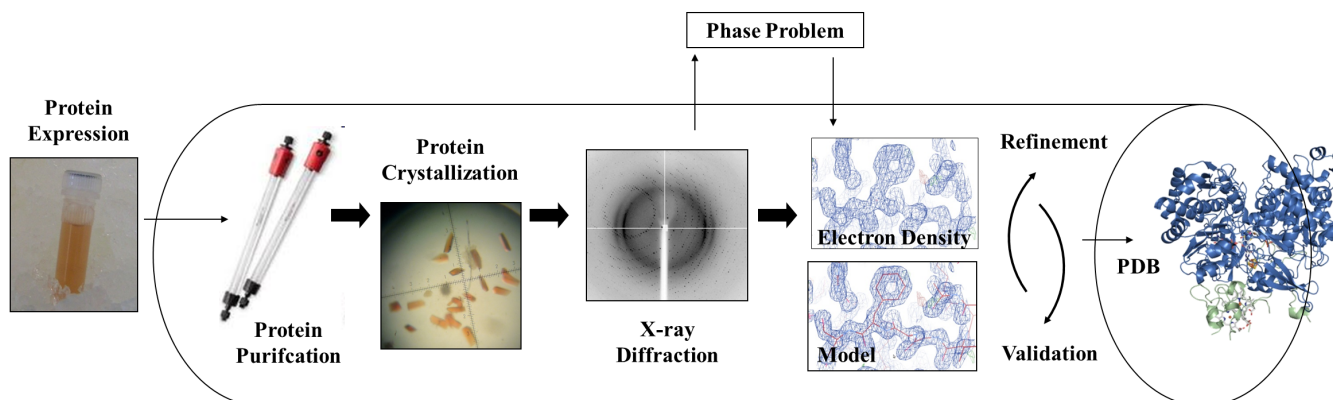


Figure I. 4 - The crystal to structure pipeline: most important steps in X-ray protein Crystallography.

1.2.2 Protein Crystallization and Crystal Symmetry

In order to determine the three-dimensional crystal structure of a macromolecule such as a protein, we must start by its expression (in the case of recombinant sources), isolation and purification. The diversity of protein purification and characterization techniques existing nowadays, such as chromatographic methodologies, has made crucial contributions in making our life easier. All crystallographers are nowadays familiar with the protein they are studying, since it’s “crystallizability” can be determined *a priori* by some intrinsic properties. Obtaining crystals from a protein is still considered the rate-limiting step in X-ray crystallography. Due to the difficulty in predicting the ideal crystallization conditions, many initial crystallization screens with different precipitants, concentrations, pHs and temperatures, must be performed. In the most popular manual way for growing protein crystals - the vapor diffusion technique - a protein/precipitant mixture is allowed to equilibrate over a reservoir, containing larger amounts of the aqueous precipitant solution. A siliconized glass slide covers and seals the reservoir, and water is slowly exchanged by controlled evaporation. To setup the crystallization process, the protein/precipitant solution can be placed in the cover slip itself, or in a support inside the crystallization reservoir - hanging and sitting-drop, respectively (Figure I. 5-a). The concentration of protein and precipitant in the drop slowly increases

by water transference to the more concentrated solution in the reservoir. When protein solubility limit is reached, the system reaches equilibrium and the exchange of water ceases. The final concentration of precipitant in the protein solution is supersaturated, and equal to the reservoir. If our crystallization conditions are within the supersaturating zone (Figure I. 5-b), nucleation can occur and the formation of crystalline nuclei may lead to a protein crystal, which should be an orderly three-dimensional array of molecules. Protein crystals have a high solvent content ($\approx 20 - 80\%$) and therefore are very fragile. They must be harvested in mother liquid solution with higher precipitant concentration, and transferred to a suitable cryoprotectant solution, usually glycerol, before data collection experiments. This procedure prevents crystal dissolution /degradation, as well as radiation damage by the intense X-rays, respectively [37, 38, 40].

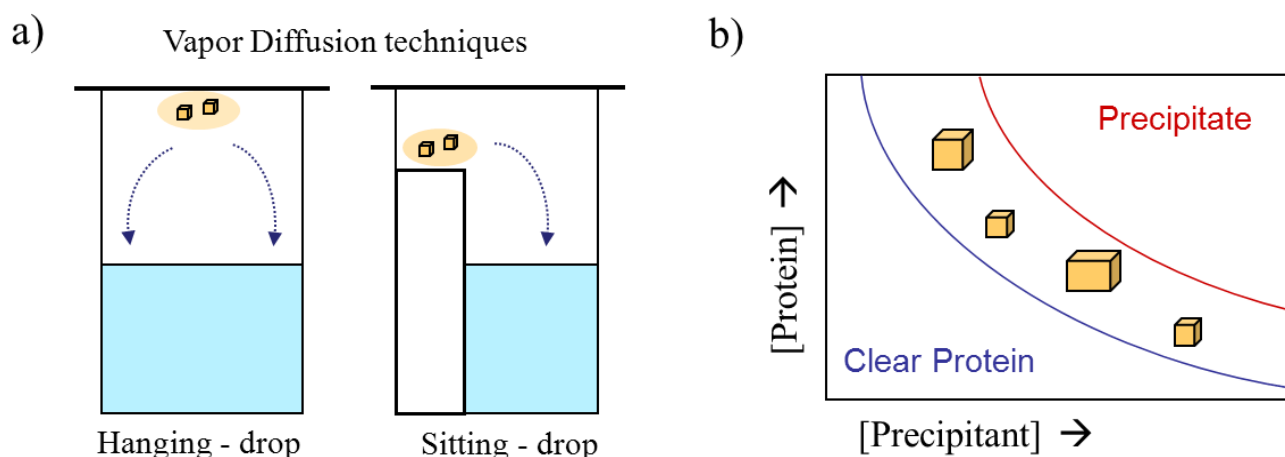


Figure I. 5 – (a) Vapor diffusion techniques: Left hanging-drop, right sitting-drop. Hanging-drop is commonly used in manual setups, while sitting-drop is preferred for robotic setups; (b) Solubility phase diagram: between the solubility line (blue) and the decomposition line (red), lies the metastable region representing the supersaturated protein solution, where the formation of crystals will eventually occur.

Nowadays the use of crystallization robots for automated crystallization setup is becoming popular. Most of the proteins that are difficult to crystallize are also those that are difficult to produce. The most significant advantage of robotic systems is the ability to miniaturize the crystallization drops (usually to the nanodrop scale), which allows a considerable increase in the number of conditions tested for the smaller amount of protein available [38].

A crystal is an orderly repetition in the three dimensional space of identical blocks, called unit cells. Each unit cell is characterized by its unit cell constants: the lengths of its three edges (a , b and c), and the angles between them (α , β and γ). The crystal unit cell contains one or more asymmetric units,

related by crystallographic symmetry. An asymmetric unit is the smallest unit of volume which contains all the structural information. Application of symmetry operations to the asymmetric unit generates the unit cell, while lattice translations of the unit cell generate the entire crystal. There are 4 different types of possible unit cells (P, C, I and F, Table I. 2), which combined with the 7 different existing crystal classes, leads to the 14 Bravais lattice. If we imagine a set of constructs dividing space in regular units, we can fit the content of our crystal in one of 230 different space groups generated by the allowed rotation and translation symmetry operations of the asymmetric unit (all of them described in *The International Tables for Crystallography, Volume A*, <http://it.iucr.org/>). However, as biological molecules are inherently chiral, neither mirror planes nor inversion centers are allowed, which reduces the number of possible space groups for proteins to 65 (Table I. 2).

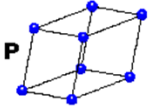
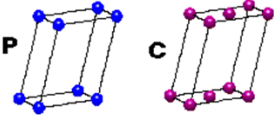
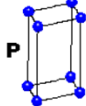
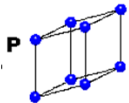
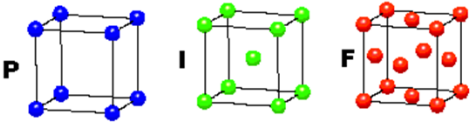
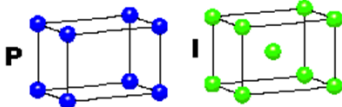
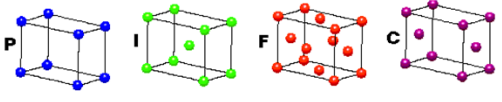
Crystal system	Cell length	Cell angles	Protein Space Groups	Bravais Lattice
<i>Triclinic</i>	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma \neq 90^\circ$	P1	
<i>Monoclinic</i>	$a \neq b \neq c$	$\alpha = \gamma = 90^\circ,$ $\beta \neq 90^\circ$	P2, P2 ₁ C2	
<i>Hexagonal</i>	$a = b \neq c$	$\alpha = \beta = 90^\circ,$ $\gamma = 120^\circ$	P6, P6 ₃ , P6 ₄ , P6 ₃ , P6 ₂ , P6 ₁ , P622, P6 ₁ 22, P6 ₂ 22, P6 ₂ 22, P6 ₃ 22, P6 ₄ 22	
<i>Trigonal</i>			P ₃ , P3 ₁ , P3 ₂ P312, P321, P3 ₁ 12 P3 ₁ 21, P3 ₂ 12, P3 ₂ 21	
<i>Cubic</i>	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$	P23, P2 ₁ 3, P432, P4 ₁ 32, P4 ₂ 32, P4 ₃ 32 I23, I2 ₁ 3, I432, I4 ₁ 32 F23, F432, F4 ₁ 32	
<i>Tetragonal</i>	$a = b \neq c$		P4, P4 ₁ , P4 ₂ , P4 ₃ , P422, P4 ₁ 2, P4 ₁ 22, P4 ₁ 2 ₁ 2, P4 ₂ 22, P4 ₂ 2 ₁ 2, P4 ₃ 2 ₁ 2, P4 ₃ 22 I4, I4 ₁ , I4 ₁ 22, I422	
<i>Orthorhombic</i>	$a \neq b \neq c$		P222, P2 ₁ 2 ₁ 2 ₁ , P2 ₁ 2 ₁ 2, P222 ₁ I222, I2 ₁ 2 ₁ 2 ₁ F222, C222, C222 ₁ ,	

Table I. 2 - Crystal systems and allowed space groups for protein molecules. Four types of unit cell: Primitive (P); Centered in the side (C), in the body (I), or in the face (F). The spheres in the Bravais Lattice represent the lattice points.

Crystals are necessary in an X-ray experiment because an amorphous aggregate of molecules would disperse the incoming X-rays, and no diffraction would be produced. The molecules need to be arranged regularly forming a crystal, in such a way that the waves diffracted by each molecule can be summed, yielding a diffraction pattern.

1.2.3 How to obtain Data from Crystals?

1.2.3.1 Bragg's Law and Diffraction

In 1913, William L. Bragg and his son, William H. Bragg, formulated a physical model to explain conditions where diffraction was observed. They considered that in a crystal lattice, regularly spaced parallel planes (characterized by h, k, l - the Miller indices) could be drawn through equivalent points. If the planes were populated by atoms or molecules, they would reflect X-rays like mirrors. Since X-rays penetrate the crystals, reflections from successive lattice planes were extinguished by interference, unless the reflected rays were in phase with each other. Considering λ as the wavelength of the X-rays, d as the perpendicular distance between the successive planes, θ as the angle between the planes and the incident or reflected X-ray, and n as an integer, then reinforcement of waves occurs only when equation in Figure I. 6 is obeyed [41]:

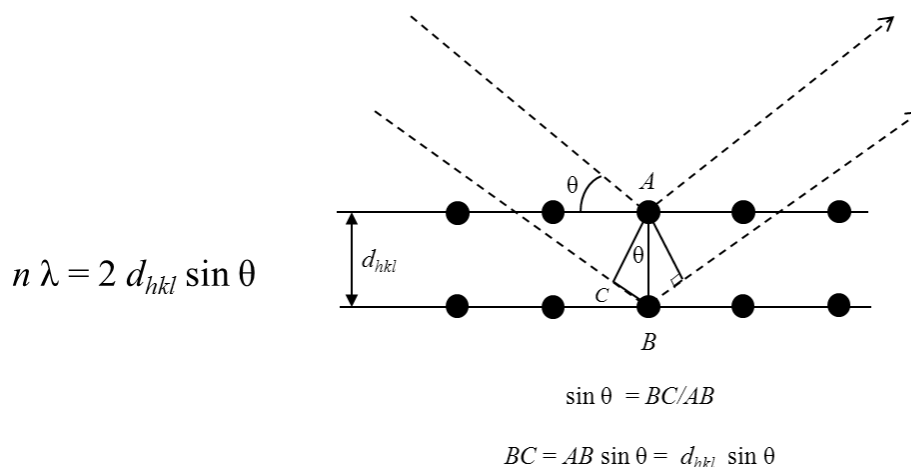


Figure I. 6 - Bragg diffraction equation (left) and geometric construction (right): If two beams with identical wavelength and phase approach a crystalline solid and are scattered, constructive interference occurs when the length $2d_{hkl} \sin\theta$ is equal to an integer multiple of the wavelength of the radiation (Adapted from [37]).

This is known as the Bragg's law. If the scattered waves from X-rays are in phase, they interfere constructively and Bragg's Law is obeyed leading to a diffraction pattern. However, if the waves are

out of phase, they are canceled out by destructive interference, and there is no diffraction. Usually, the families of planes that divide the crystal unit cell edges in integer fractions are the ones that allow constructive interference. This model also helps to have an idea about data collection geometry, since the number of measurable reflections depends on the information present in the unit cell of the crystal. Large cells contain more atoms and more information in the diffraction pattern [37].

The periodicity of molecules in the crystal amplifies the diffracted source beam into many distinct beams, each of which produces a distinct reflection (currently named spot) on the detector. With monochromatic X-rays and a stationary crystal, Bragg's law would be met for only a few sets of planes, but rotation of the crystal makes reflections from different sets of planes to become visible, one by one. Each reflection can be treated like a wave, and its phase and intensity contains the information necessary to determine the molecular structure of the protein. There is an inverse relationship between the space of unit cells in the crystal (crystal lattice), and the spacing of reflections in the detector (reciprocal lattice). In the crystal (real space), the position of an atom in the unit cell (lattice point) can be specified by a set of spatial Cartesian coordinates (x,y,z) considering one of the vertices of the cell as the origin, but in the diffraction pattern (reciprocal space), a reflection position is given by the (h,k,l) indices and the central reflection is taken as the origin (Figure I. 7) [37, 41]:

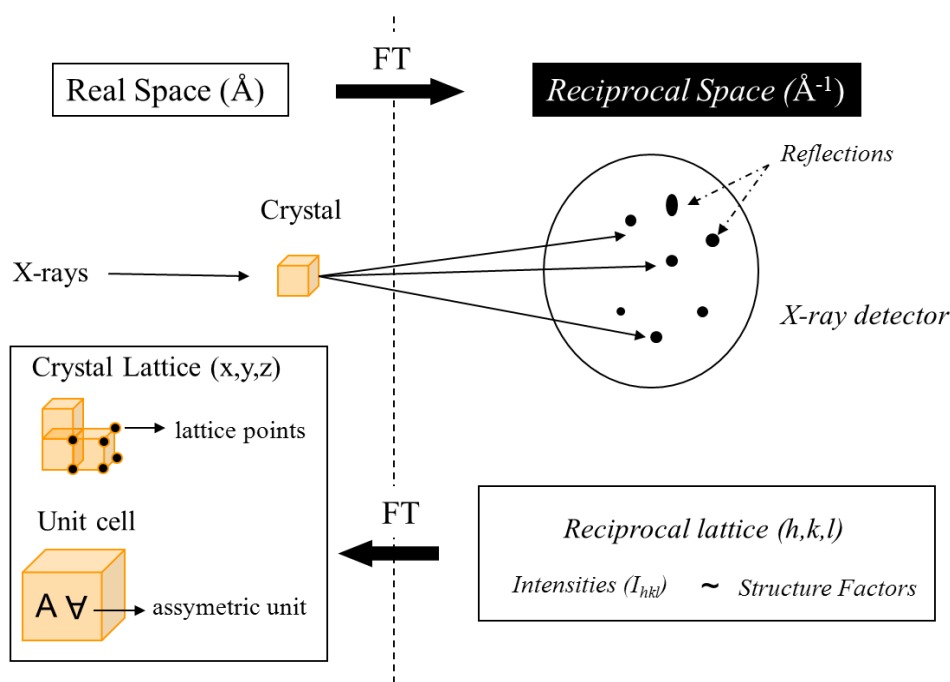


Figure I. 7 - Representation of a basic X-ray diffraction experiment: the crystal when exposed to X-rays produces individual reflections in the detector. Diffraction images are a transformation (FT) of the crystal real space (real lattice) into reciprocal space (reciprocal lattice) and vice-versa.

At the end of a diffraction experiment, or even after obtaining just a few diffraction images, we can calculate the space group and the unit cell parameters of the crystal. Also, the number of molecules present in the asymmetric unit and its volume can be determined. Protein crystals are composed not only of the repeating motif of the asymmetric unit, but also of solvent channels containing the compounds present in the crystallization solution and in the protein stock solution. Solvent content analysis was first calculated by Matthews in the 60s, [42] who plotted the volume of the asymmetric unit (V) against the molecular weight of the protein ($Z \times M_p$). Z is the number of asymmetric units present in the unit cell, and is determined by the space group. The ratio of these two values is known as the Matthews coefficient, and is defined as VM with dimensions $\text{\AA}^3\text{Da}^{-1}$. In general, protein structures have a solvent content of 30-70%, and crystals that diffract at higher resolution, usually possess lower solvent content [38]:

$$VM = V / (Z \times M_p)$$

We should be aware that when multiple copies of the same molecule are present in the asymmetric unit, they are related by local non-crystallographic symmetry (NCS). This should not be confused with the crystallographic symmetry operations of the asymmetric unit, which generates the crystal. The direct results of a diffraction experiment are now held by the crystallographer: space group and crystal unit cell parameters have been determined, solvent content was calculated, and intensity data from the individual reflection points can now be indexed, to improve consistency and maximize accuracy. However, the real purpose of the crystallographer research remains: the location of atoms inside the unit cell is still unknown.

1.2.3.2 Electron Density and the Phase Problem

To obtain the structure of the individual diffracting motif, or the distribution of electrons in the asymmetric part of the crystal, it is necessary to calculate the Fourier transform of the so-called structure factors, or F values ($F(h,k,l)$), for all Miller indices (h,k,l). A structure factor describes one diffracted X-ray, which produces one reflection received at the detector. The Fourier transform describes the mathematical relationship between an object and its inverse; it maps objects and functions from one space or domain into its reciprocal one and vice versa. Electron density (ρ) in the reciprocal space, can be given by a Fourier transform of the sum of periodic functions describing the atoms positions (x,y,z) in the real space (crystal) (Figure I. 8):

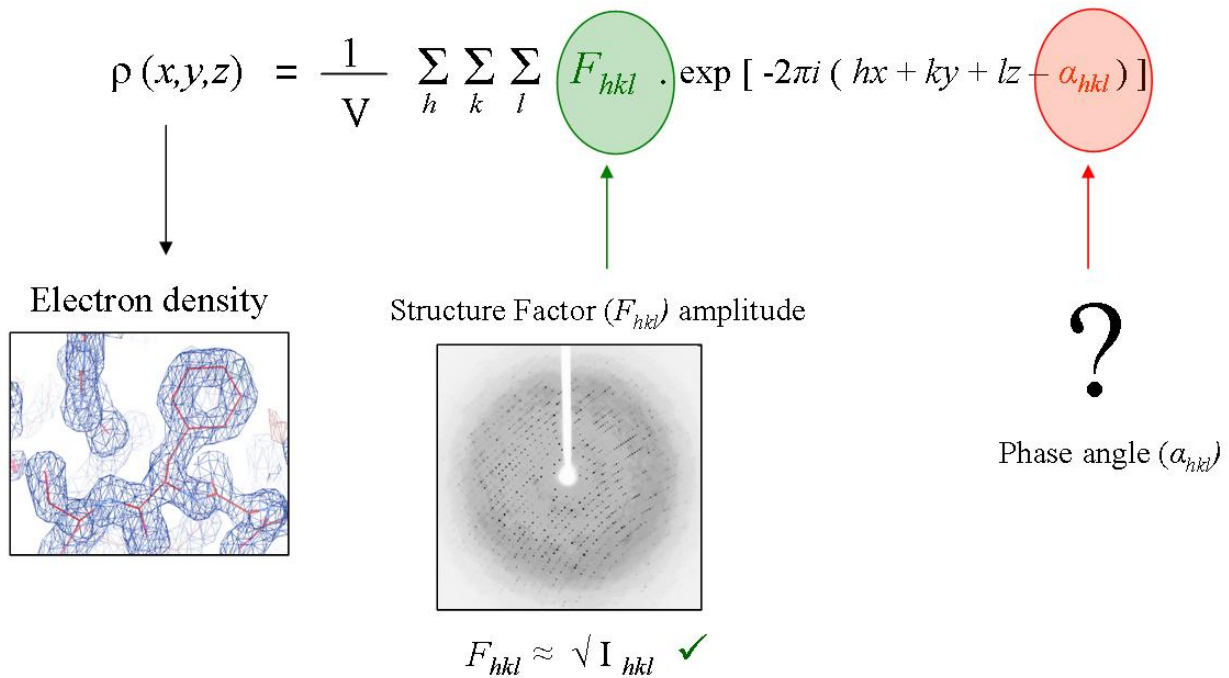


Figure I. 8 - The Phase Problem in Crystallography: to reconstruct the electron density map, the structure factor amplitude and the phase angle for each reflection, must be supplied. F_{hkl} is obtained experimentally but α_{hkl} is not, and must be acquired from additional phasing experiments. V is the volume of the unit cell. (Adapted from [38]).

This means that if we know the amplitudes (F) and the phase angle (α) of the diffracted waves, we can calculate the electron density map ($\rho(x,y,z)$). When the reflections from the diffraction experiment are indexed, a long list of intensities (I_{hkl}) and associated errors (σ_{hkl}) are obtained, for each single reflection acquired. The intensities are proportional to the square of the amplitudes, and so $F_{(h,k,l)}$ can be obtained from $\sqrt{I_{(h,k,l)}}$. However, in the diffraction experiment there is no information available regarding the phase angle that reaches the detector, which enables us to calculate the electron density map directly from the equation. To overcome this phase problem, several methods have been developed in protein crystallography to obtain the missing phases. Some of them require additional diffraction information which must be acquired from new experiments. The initial electron density map obtained will still need to be improved in an iterative way [38, 39].

1.2.3.3 Phase Determination and Model Building

Due to the lack of phase information in the diffraction patterns, direct reconstruction of the electron density map via Fourier transforms is not possible from the intensity data alone. Obtaining the phases, or solving the structure, is the most challenging part of the process following crystal growth.

In protein crystallography there are two major methods to achieve this: molecular replacement and multiple isomorphous (heavy atom or anomalous scatterer) replacement.

Molecular Replacement (MR)

If a similar model of our protein exists and has been already determined, it can be used to calculate the initial missing phases. The correct position of the search model in the crystal needs to be found, and therefore this method is called molecular replacement (MR), in the sense of repositioning but not substituting the search model. Although easily applicable, this method needs to be used carefully. Phase bias is introduced due to phase domination above the intensities during electron density reconstruction. The initial structure may largely reflect the features of the search model and not the features of the real one, therefore it is necessary to use procedures which remove phase bias to obtain a model as accurate as possible. The importance of this method is visible in the protein data bank statistics since more than 70% of the deposited structures were solved by molecular replacement. The work reported in this thesis is not an exception. The two crystal structures described here were solved using molecular replacement, with similar proteins as search models.

Multiple Isomorphous Replacement (MIR)

In the absence of a suitable similar model with known structure, the phases must be determined using a separate diffraction experiment. These alternative methods depend on the determination of a marker atom substructure, using the intensity differences between isomorphous data sets, which are usually very weak. Traditionally, these differences are found between native and derivative crystals, containing a heavy atom (*e.g.* mercury, platinum, or gold), soaked into the crystal without significantly altering its unit cell dimensions or positions of the protein atoms. From two or more such derivatives it is often possible to obtain enough phases to solve the native data set by multiple isomorphous replacement (MIR). The heavy atom substructure provides the initial phase estimates required to reconstruct the electron density. Another method now more commonly applied uses anomalous dispersion differences instead, whereby multiple data is collected from the same crystal containing an anomalous scatterer, but at different wavelengths. One example of this method is replacing the amino acid methionine, by selenomethionine through overexpressing the protein in a suitable system. The Se atom provides a proper site source for anomalous phasing signal. The choice of wavelengths above and below absorption edges is possible nowadays because of the availability of

tunable synchrotron radiation, which provides maximum differences in anomalous dispersion effects. This method is called MAD; multi wavelength anomalous dispersion method [38].

The above methods allow obtaining an estimate of the phase information. In almost every case, the initial model of the protein structure obtained is further enhanced by various density modification techniques, providing substantially improved electron density maps into which the protein structure model can be built. The model building is carried out using computer graphic programs that display the electron density, allowing assignment and manipulation of protein backbone and amino acid residues. Several electron fitting and geometry refinement tools, as well as automated model building programs, greatly accelerate the process. Model building constitutes the most intense involvement of the crystallographer since chemical knowledge and previous experience on the field, are an advantage. Good phases and high resolution play a particularly decisive role in this important task [38, 43].

1.2.4 Resolution, Refinement and Validation

There are a number of reliable indicators used to determine the accuracy of the atomic coordinates prior to publication in structural databases. Some come from the process of solving the structure and reflect the experimental data itself, others are derived from the obtained coordinates. The first simple indicator is the resolution of a structure, defined as the limit in diffraction angle, up to which X-rays diffracted by a crystal can be detected. Simplifying, the resolution is a measure of how much data were collected: the more data, the greater the ratio of the number of observations to the number of atomic coordinates to be determined, and usually the more accurate the results. The higher the resolution, the more detailed the electron density map. The average number of measurements per individual, symmetrically unique reflections is called redundancy or multiplicity. Since every reflection is measured with a certain degree of error, the higher the redundancy, the more accurate the final estimation of the averaged reflection intensity. After model building, the atomic model is “refined” by varying all model parameters to achieve the best agreement between the observed reflection amplitudes (F_{obs}) and those calculated from the model (F_{calc}). Refinement is the adjustment of model parameters so that the calculated structure factors match the observations as nearly as possible, and this agreement is judged by the crystallographic “*R*-factor”. Every reported X-ray structure determination must include the statistical value *R*-factor (or *R*-work), which is simply a measure of how well the coordinates reproduce the experimental data. Also, the *R*-free parameter is

used to measure the agreement between the final model and the subset of experimental data retained for refinement. It corresponds to a little fraction of data (usually 5-10%), which is not used for the refinement process. This gives an unbiased measure of agreement. In good quality data the $R_{\text{work}}/R_{\text{free}}$ ratio should be around 20%. Another important parameter for crystallographers to judge their model is the B factor, which is related to the thermal vibration of an atom around its position. Thus, the higher the B factor the less defined (or more flexible) is that particular atom. The B factor parameter may be somewhat intuitive, but the most important criteria is that it should be within the same range of values for neighboring residues, revealing data consistency [38, 39, 43].

When analyzing a protein crystal structure, chemical parameters must also make sense, and must be in accordance with chemical constraints. Molecular chirality and stereochemistry must be obeyed, and deviations must make biological sense. The Ramachandran map provided by the graphic programs is a good guide for checking the conformations of ψ and ϕ angles, the dihedral angles for the two degrees of freedom around the $C\alpha$ atom of the amino acid residues, which define the fold for polypeptide chains.

1.2.5 Synchrotron Radiation

Although the majority of crystallographer's all around the world possess an in-house X-ray machine, the introduction of synchrotron radiation (SR) facilities for routine use, constitutes a remarkable revolution in X-ray macromolecular crystallography. The basic fundamental function of a synchrotron is that it works as a storage ring, where electrons move around in a circle, almost at the speed of light through existing magnetic fields, generating intense X-rays, which can then be used for diffraction experiments. The common use of synchrotrons helps to explain the impressive high number of structures deposited each day in the PDB. The very slow process of data-collection in house, taking usually many days to conclude has been immensely shortened to a few hours or even less. Not only the speed of data collection, but also the data quality itself has been dramatically improved. Another important feature, due to the intensity and collimation of the beam in the synchrotron, is the possible use of ever-smaller crystals in the experiments. This is particularly important in the case of crystals that are hard to improve. We have regular access to SR by using some of the macromolecular crystallography beamlines available at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, namely ID14 and ID23 (Figure I. 9). This is one of the three largest and most powerful synchrotrons in the world (the other two are APS in USA and

Spring-8 in Japan). The crystal structures described in this Thesis were obtained from data collected at ESRF.

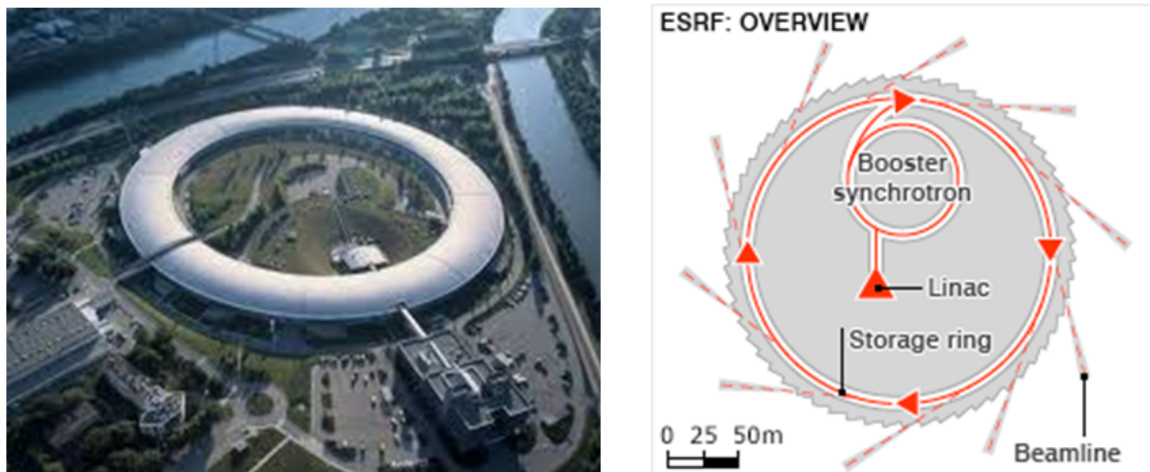


Figure I. 9 – ESRF general overview (left), and schematic representation (right).

With all the tools available nowadays, particularly the use of synchrotron radiation facilities, the available dispensing crystallizations robots for unlimited screening conditions and the emergence of more sophisticated equipment, such as fast X-ray detectors and computer processors, it is expected that protein crystallographers are tempted to “attack” proteins that have still remained a taboo for a long time. This will allow the number of macromolecular structures obtained by X-ray crystallography and deposited in the data bank to grow even more.

CHAPTER 2

mouse Aldehyde Oxidase Homologue1

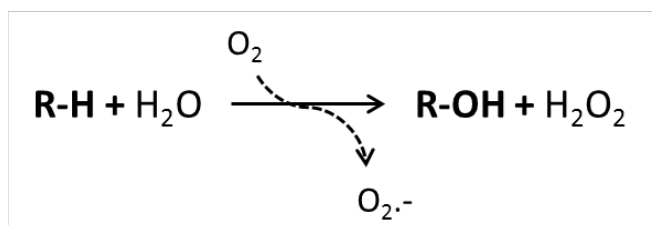
Part of the work described in this chapter, was the subject of the following publication:

M. Mahro*, C. Coelho*, J. Trincão, D. Rodrigues, M. Terao, E. Garattini, M. Saggi, F. Lenzian, P. Hildebrandt, M. J. Romão and S. Leimkühler (2011) “Characterization and crystallization of mouse Aldehyde Oxidase 3 (mAOX3): from mouse liver to *E.coli* heterologous protein expression”, *Drug Metabolism and Disposition*, published online ahead of printing.

2.1 INTRODUCTION

Aldehyde oxidases (AO or AOX1, EC 1.2.3.1) are a small group of evolutionary conserved enzymes. Contrary to xanthine oxidoreductase (XOR), which constitutes the key enzyme in the catabolism of purines, converting hypoxanthine to xanthine and xanthine to uric acid, the physiological function of aldehyde oxidases is still largely unknown. However, the enzyme is known to play an important role in the metabolism of compounds with medicinal and toxicological interest [16, 18].

Aldehyde oxidases oxidize R-H substrates using H₂O and O₂, which is the final acceptor of the reducing equivalents generated. Together with the resulting hydroxylation product (R-OH), superoxide anion radical is also produced (O₂^{•-}), which dismutates to form hydrogen peroxide (H₂O₂) (equation below). Both molecules can be potentially toxic for the cell, but can also be involved as defensive pathogenic agents.



Typical AO substrates are compounds containing aromatic heterocycles (usually N-heterocycles) or aromatic aldehydes, which are oxidized into the corresponding carboxylic acid. From the main proposed substrates [16, 44], benzaldehyde, retinaldehyde (vitamin A precursor) and pyridoxal are among the most relevant ones (Table II. 1). AO is also known to catalyze the oxidation of intermediate products of cytochrome P450 (CYP), and there is evidence in the literature that liver AO can be considered as the CYP cytosolic equivalent [45]. There is a possibility that both enzymes act in concert, activating or inactivating various types of drugs and compounds with toxicological interest. The increasing importance of AO as a drug-metabolizing enzyme has been the subject of several publications, where an extensive list of compounds (particularly substrates and inhibitors) has also been theoretically tested [45, 46]. These studies have demonstrated that the most powerful AO inhibitor is raloxifene, a selective estrogen receptor modulator. Menadione is also reported as a specific inhibitor for aldehyde oxidase, and an interesting feature is that both compounds are ineffective in terms of XOR inhibition (Table II. 1) [45, 46].

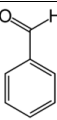
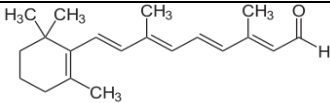
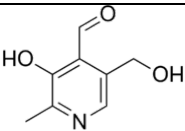
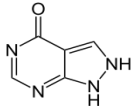
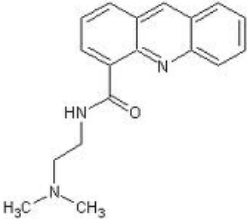
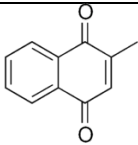
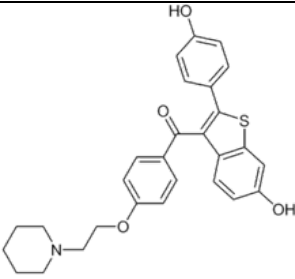
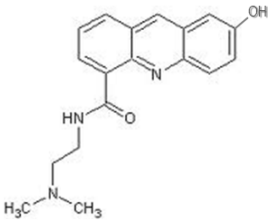
Compound	Role	Formula
Benzaldehyde	Substrate	
Retinaldehyde	Substrate	
Pyridoxal	Substrate	
Allopurinol	Substrate	
Acridine carboxamide (DACA)	Substrate	
Menadione	Inhibitor	
Raloxifene	Inhibitor	
7-hydroxy acridine carboxamide	Inhibitor	

Table II. 1 – Relevant aldehyde oxidase substrates and inhibitors.

Xanthine oxidoreductases and aldehyde oxidases possess similar primary structures in various animal species, with 40% amino acid identity. They also share similar tertiary and quaternary structures, as the catalytically active forms are homodimers composed of two identical subunits with approximately 150 kDa each. The main physiological difference between the two enzymes is that XOR can exist in two interconvertible forms (XO and XDH), while AO exists only in the oxidase form. There is another difference regarding enzymatic function, since AO utilizes only molecular oxygen as electron acceptor, while XOR are able to transfer electrons not only to oxygen but also to nicotinamide adenine dinucleotide (NAD⁺) [16, 18, 47].

It was known that in humans XOR and AO resulted from the expression of distinct genes residing on chromosome 2. A similar situation was thought to be true in the case of other vertebrates, but a few years ago the identification and characterization of novel murine genes, coding for different forms of aldehyde oxidase has changed this assumption [48]. While humans possess a single AO active gene, rodents have developed over three different genes (through a series of gene duplication events), coding for aldehyde oxidase homologue 1, 2 and 3 (mAOH1, mAOH2 and mAOH3), respectively. These aldehyde oxidase isoenzymatic forms are located on mice and rat chromosomes 1 and 9, respectively. Human and rodents aldehyde oxidases share 85% sequence identity, while 60% identity is found between the isoenzymatic forms and the human protein (Figure II. 1). One interesting feature is that the orthologous proteins from different organisms possess higher similarity than the homologous ones, present in the same organism. This may suggest that the gene duplication events occurred before species divergence [16, 17, 49].

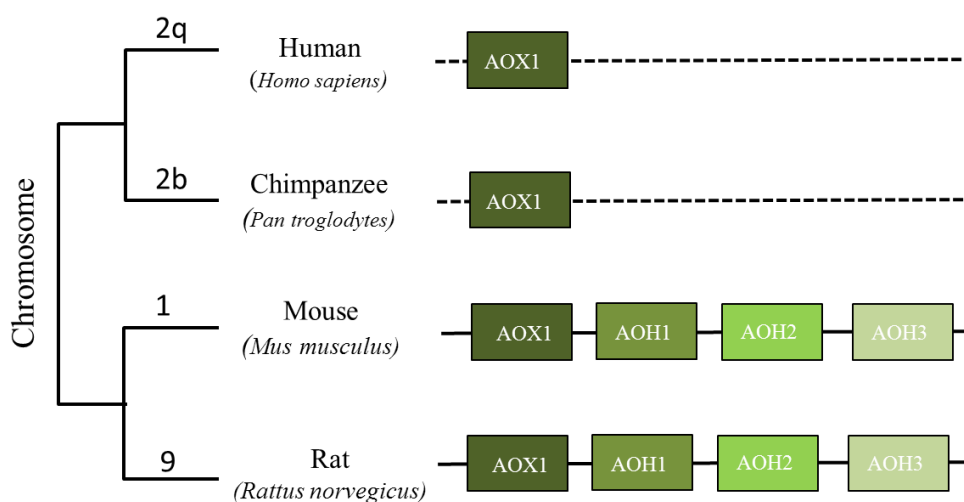


Figure II. 1 - General schematic representation of molybdo-flavoenzyme genes in vertebrates (Adapted from [50]).

Aldehyde oxidase homologues 1, 2 and 3 are expressed in different cell tissues, and the purpose of their existence is still puzzling. In rodents, the richest source of AOX1 and AOH1 is the hepatocyte component of adult liver. Expression of the two proteins is overlapping, but they differ as to the time of appearance in the liver of the developing mouse. The AOH1 transcript is readily detectable in newborn mice, while AOX1 appears later and is only measurable in the fully developed animal. Biochemical experiments demonstrated that the major source of mAOH2 is the Harderian gland, which is the most important exocrine gland located in the intra-orbital cavity of rodents and other vertebrates. The other isoform, the mAOH3 protein, could only be detected in the Bowman's gland, the principal exocrine gland located in the sub mucosal layer of the nasal cavities, representing almost 5% of total cytosolic proteins. It was also observed that in rodents the aldehyde oxidase and aldehyde oxidase homologue1 activities are influenced by the gender, since significantly higher amounts of enzyme are present in male. This suggests that estrogens may regulate AOX1 and AOH1 expression at the translational level, while androgens could exert transcriptional control. It is possible that gender specific regulation of these proteins by androgens and estrogens is an indirect effect mediated by other hormones or growth factors, as for instance the circulating growth hormone (GH) [48, 49, 51].

There is an increasing need to elucidate the main features, which dictate the specificity for the various aldehyde oxidase isoforms. These studies can only be possible with the expression of recombinant protein in heterologous systems, particularly due to the limited amounts of protein available from the native sources, and also to enable mutagenic studies. The chemical structure of MoCo in prokaryotes is different from that found in eukaryotes, preventing assembly of the holo-enzyme in engineered bacteria. With the introduction of genetic modifications in the host bacteria for eukaryotic MoCo synthesis, the authors Mahro and Schumann were able to express and purify recombinant mouse AOH1 and AOX1 proteins respectively [52, 53]. This breakthrough helped to produce sufficient amount of protein for biochemical and structural studies, which may be useful in providing some answers to the questions related with this family of proteins. Besides, it can also help to understand the possible importance and involvement of human aldehyde oxidase in the field of diseases such as obesity and cancer. Recent data support the idea that this enzyme plays a positive role in adipocyte differentiation. The specific tissue aldehyde oxidase expression could be explored to design novel anti-cancer drugs or even strategies to achieve organ and/or tumor selectivity [15, 54].

Until recently, and due to the absence of a three-dimensional structure for an aldehyde oxidase protein, *in silico* studies have been carried out based on model comparison with the XOR protein, whose crystal structure was reported in 2000 by Nishino and collaborators [13]. The system now available for producing enough quantity of the heterologous mAOH1 protein, allowed obtaining suitable crystallization conditions, which were later successfully reproduced with the native mouse protein [52]. The crystal structure of mAOH1 was solved by molecular replacement, and is the focus of this chapter. This constitutes the first three-dimensional structure for an aldehyde oxidase, and has been used for detailed comparisons with other enzymes of the xanthine oxidase family, providing the molecular details for the distinct specificities between the different enzymes (XORs and AOs). Also, comparison of the structural characteristics of the different forms within the same species (such as AOX1 and mAOH1) can help to predict if they metabolize the same or different substrates. The role of mammalian aldehyde oxidases in the activation or inactivation of foreign drug compounds can also be explained. Due to its unique distribution and substrate specificities, aldehyde oxidase has been recently recognized as an enzyme with an important role in the metabolism of drugs and xenobiotics, and has therefore high pharmacological relevance [15, 45].

2.2 EXPERIMENTAL PROCEDURE

2.2.1 Purification of AOH1 from Mouse Liver

To purify the native mouse aldehyde oxidase homologue1, 135 CD-1 mice were purchased from Charles Liver Laboratory (Como, Italy) and kept at 25°C with a 12h light/dark cycle in the “Mario Negri” Institute animal house facilities (Milan, Italy). Unless otherwise stated, all purification steps were carried out at 4°C. The mouse livers (35-40 g) were homogenized in three volumes of 100 mM sodium phosphate buffer, pH 7.5 using an Ultraturrax homogenizer (Omni 2000-Waterbury, CT). Homogenates were centrifuged for 45 min at 100,000 x g to obtain the cytosolic extracts, which were later heated for 10 minutes at 55°C, and again centrifuged to remove precipitated proteins. An equal volume of saturated ammonium sulfate solution was added to the supernatant, and incubated overnight. The resulting precipitate was collected by centrifugation and resuspended in 100 mM Tris-Glycine pH 9.0. The solubilized proteins were mixed with benzamidine Sepharose (Amersham Biosciences, Sweden) pre-equilibrated with sample buffer, and incubated for 2 hours. The resin was washed with 4 x 10 ml equilibrating buffer and the absorbed protein eluted using 5 mM benzamidine. After buffer exchange to 50 mM Tris-HCl pH 7.4, the sample was loaded on a 5/5 FPLC Mono Q column (Amersham), and mAOH1 eluted with a linear gradient 0-1M NaCl. The protein was concentrated to 10 mg/ml in 50 mM Tris-HCl, pH 7.4, and stored at -80°C.

2.2.2 Expression and Purification of mAOH1 from *E.coli*

The protein heterologous expression system was developed by Martin Mahro, at the “Institut für Biochemie and Biologie” (Potsdam, Germany). *E.coli* TP1000 ($\Delta mobAB$) cells were used for the co-expression of mAOH1 wild-type with mMCSF (pSS110), and the cultures were grown aerobically in LB medium at 30°C for 24h. The mAOH1 cDNA from mouse CD1 liver was cloned using primers designed to permit cloning into the *NdeI* and *Sall* sites of the expression vector pTrcHis. The resulting plasmid was designated pMMA1 and expresses mAOH1 as an N-terminal fusion protein with a His₆-tag. pMMA1 and pSS110 were transformed into TP1000 cells. To express the recombinant protein, cells were grown in LB medium supplemented with 150 µg/ml ampicillin, 50 µg/ml chloramphenicol, 1 mM molybdate, and 10 µM IPTG, at 30°C. After 24h the cells were harvested by centrifugation and resuspended in 50 mM sodium phosphate buffer pH 8.0, 300 mM NaCl. DNaseI and lysozyme (1 mg/L) were added to disrupt the cells at 12°C and 1.35 kbar

(Constant Cell Systems TS Benchtop Series). The cleared lysate was loaded onto a Ni²⁺ column (Ni-NTA, Qiagen), and mAOH1 eluted with 250 mM imidazole. Buffer was exchanged to 50 mM potassium phosphate, pH 7.8, 0.1 mM EDTA using PD10 columns (GE Healthcare). To increase the activity of the enzyme, chemical sulfuration was performed [55, 56]. The purified enzyme was incubated with 500 μ M sodium dithionite, 25 μ M methylviologen and 2 mM sodium sulfide, for 1 hour in an anaerobic chamber (Coy Lab Systems). After buffer exchange to 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, the sample was loaded onto a Superose 6 column (GE Healthcare), and fractions containing dimeric mAOH1 were combined and stored in 100 mM potassium phosphate pH 7.4, at -80°C.

To cleave the N-terminal His₆-tag, mAOH1 was incubated with 50 mM Tris-HCl pH 8.0 and 1mM EDTA, at 4°C for 12 hours. Thrombin cleavage site introduced by pTrc-His was used, and the 17 residues His₆-tag was cleaved after Arg17 [57]. The protein was loaded onto a HiLoad 26/60 Superdex 200 column (GE Healthcare), equilibrated with 50 mM sodium phosphate pH 8.0, 300 mM NaCl. Fractions containing dimeric Δ His₆-mAOH1 were combined and concentrated to 17.8 mg/ml, by ultrafiltration (Amicon). Protein aliquots were frozen in liquid nitrogen and stored at -80°C until usage, without loss of activity.

Using PCR mutagenesis, the amino acid exchange Y885M was introduced into the mAOH1 recombinant protein. The generated variant was expressed and purified using the same conditions as previously mentioned for the *wild type* form of the enzyme. After purification, the Y885M variant was concentrated to 17.8 mg/ml and stored in 100 mM potassium phosphate pH 7.4, at -80°C until necessary.

2.2.3 Protein Crystallization

The first crystallization trials were performed using the native mAOH1 purified from male mouse livers. Using this protein, more than 800 different conditions were tested, at three different temperatures (4, 15 and 20°C) but no successful results were obtained. In addition, the limited amount of protein obtained (9.4 mg) restricted additional crystallization trials. The recombinant mAOH1 yielded larger amounts of pure protein that allowed pursuing the crystallization studies and including more variables, such as incubating the protein with DTT and using several additives. The

first crystals appeared using the recombinant protein incubated with a freshly prepared DTT solution, polyethylene glycol (PEG) 4000 and 8000 as precipitant and Tris-HCl as buffer, at 20°C. Under these conditions long thin needles were obtained, but were unsuitable for diffraction experiments and needed to be improved (Figure II. 2-a). Several additives were tried in order to optimize the needle-shaped crystals, using commercially available additives as well as some known mAOH1 inhibitors (e.g. menadione). Another less conventional approach using the ionic liquids [C₄mim] Cl (1-butyl-3-methylimidazolium chloride) and [C₄mim] MDEGSO₄, (1-butyl-3-methylimidazolium 2(2-methoxyethoxy) ethyl sulfate) 0.2, 0.4 and 0.6 M (final concentration in the protein solution) was also attempted. In a similar case of crystal optimization, this proved to be a very attractive alternative (Chapter 4 of this Thesis). Nevertheless, this approach did not improve the mAOH1 crystals. Further extensive trials included small variations in the initial conditions, such as changing the buffer and precipitate concentration. The best protein needles of recombinant mAOH1 were obtained using 12-16% PEG 8000 as precipitant and 0.1M potassium phosphate pH 7.0 at 20°C. Using the hanging drop vapor diffusion method, equal amounts of protein (17.8 mg/ml) and precipitant solution were mixed, after incubating the protein with fresh DTT at 4°C for at least 1h (final DTT concentration in the protein solution was 8mM).

In parallel, another strategy was followed in an attempt to try to improve the crystals of the recombinant protein that consisted in cleaving the protein His₆-tag (which added 17 amino acids to the protein). His-tags usually have little effect on the native structure of a protein, but can have an impact in crystallization. We tested if this was the case for the recombinant mAOH1 [58]. Although crystals were also formed under similar conditions as the protein with the His₆-tag (14% PEG 8000, 0.1M potassium phosphate pH 7.00 and spermine tetra-HCl as additive) no significant improvement was achieved. The best diffraction obtained for crystals from the untagged recombinant protein, was around 6Å at PXI beamline at Swiss Light Source (SLS, Villigen, Switzerland), and the crystals were very difficult to reproduce.

Using a remaining amount of native mAOH1 isolated from the mouse liver, similar crystallization conditions were tested. The buffer pH for the mouse protein was optimized to 6.5 (instead of 7.0), the protein concentration used was 10 mg/ml, and 2mM EDTA was added to the crystallization solution. Larger, two dimensional rectangular-shaped crystals (0.40 x 0.15 x 0.05 mm³, Figure II. 2-b) of the native mAOH1 were reproducibly formed. These were the only crystals that allowed

collection of a usable diffraction data set. The same conditions were also used for crystallization of the variant Y885M, and nice crystals grew within 3 days (Figure II. 2-c). Although these crystals looked much better in terms of size and morphology than the native recombinant ones, diffraction was also poor.

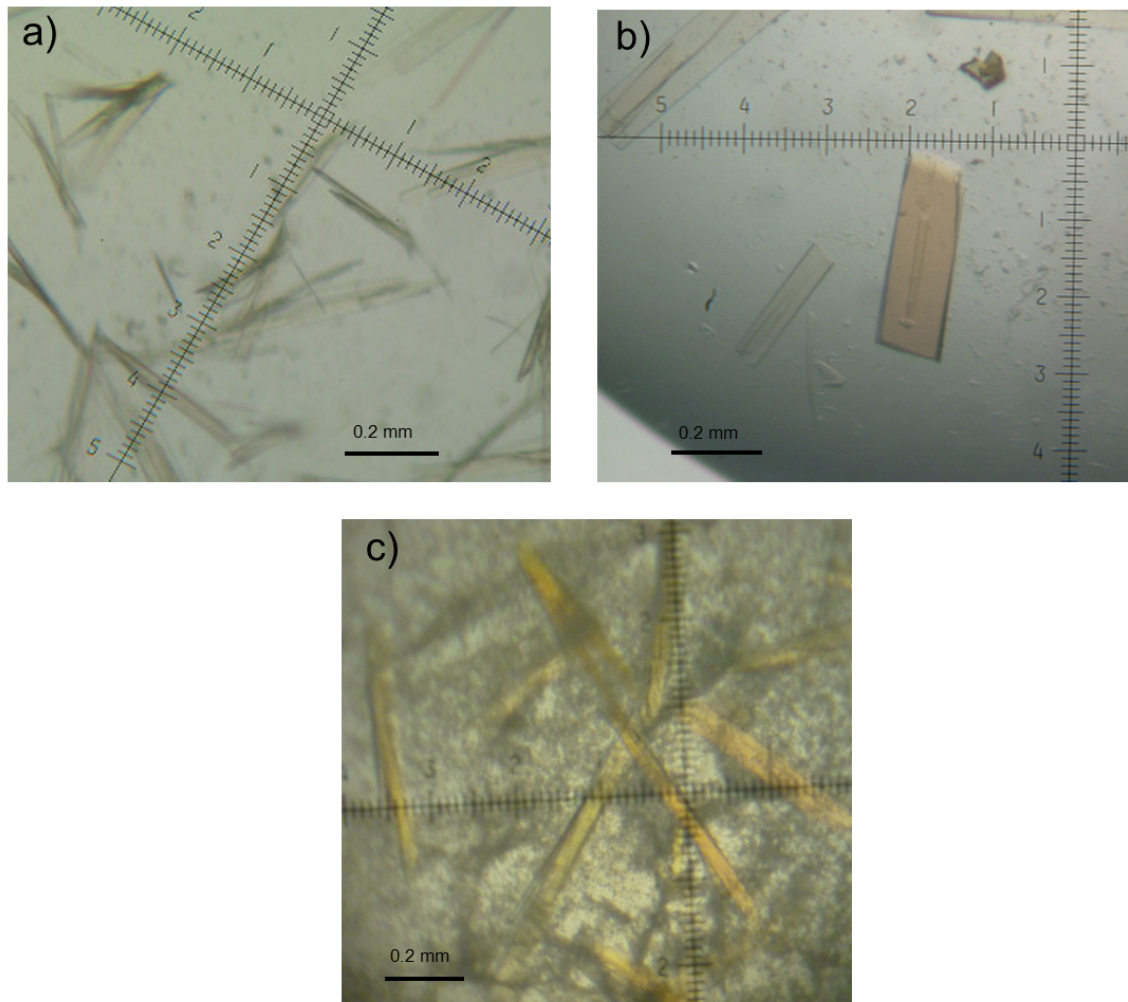


Figure II. 2 - Crystals of mAOH1 protein: (a) needles from the recombinant protein; (b) crystals from the native mouse liver protein; and (c) crystals from the Y885M variant.

2.2.4 Data Collection, Processing and Structure Solution

All tested crystals were flash-cooled directly in liquid nitrogen, using paratone oil as cryoprotectant, before transfer to a gaseous nitrogen stream for data collection. Several cryoprotectant solutions containing glycerol (30%) and higher PEG concentration (20%) were also tested. Some of the tested native mAOH1 crystals diffracted to $\sim 6\text{\AA}$ but, after annealing (using the beamline's automated annealing procedure), diffraction improved considerably to a resolution beyond 3\AA . A first data set

(mNAT-I) consisting of 180° of data was collected at ID14-1 from the European Synchrotron Radiation Facility (ESRF, Grenoble, France), at a wavelength of 0.934Å on an ADSC Q210 detector (Figure II. 3). The crystal belonged to space group P1, with unit cell dimensions $a = 91.07 \text{ \AA}$, $b = 135.02 \text{ \AA}$, $c = 147.48 \text{ \AA}$, $\alpha = 78.27^\circ$, $\beta = 77.77^\circ$ and $\gamma = 89.96^\circ$. The calculated Matthews coefficient was $2.89 \text{ \AA}^3\text{Da}^{-1}$, corresponding to two dimers in the unit cell with a solvent content of 57.4 % (Table II. 2) [42]. Because the data were very anisotropic and the mosaicity was high, the initial data processing using imosflm 1.0.4 was incomplete (~89% complete to 2.9Å) [59]. The mosaicity value had to be maintained fixed because refinement during integration was unstable and led to crashing of the program. The same crystal was later measured at ID23-1 ($\lambda=0.9748\text{\AA}$), again at the ESRF. Although a full 360° of data were collected, only about 220° were useable (mNAT-II) because the crystal was very anisotropic. This data set was worse than the first one, and even though the multiplicity was higher, the completeness was very low (~60% to 3Å) (Table II. 2). The two data sets were merged together in order to increase the completeness and improve the multiplicity. This dataset presents a very high R_{pim} and is only ~80% complete to 2.9Å (NATmerged), but the overall redundancy improved to ~3.0. The structure of the mAOH1 was solved by molecular replacement using BALBES on the mNAT-I data set [60]. The bovine milk XDH structure with a covalently bound oxipurinol inhibitor (PDB code 3BDJ) was used as a search model [61]. Four monomers were found in the unit cell, yielding a Matthews coefficient of $2.89 \text{ \AA}^3/\text{Da}^{-1}$, corresponding to a solvent content of 57.4%.

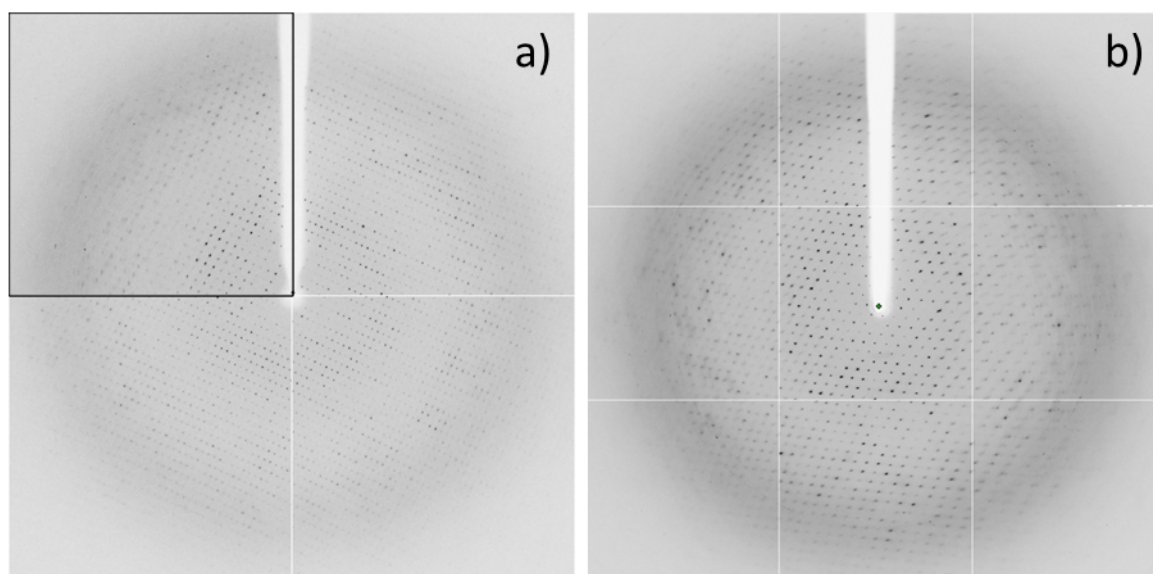


Figure II. 3 - Diffraction pattern of the native mouse liver AOH1: (a) obtained at ID14-1 (in the upper left quadrant after re-annealing); (b) and at ID23-1 (ESRF). Resolution at the edge is 3.0Å.

Crystal Sample	mNAT-I	mNAT-II	NATmerged
X-ray source	ID 14-1	ID 23-1	-
Crystal system	Triclinic		
Unit-cell parameters (Å, °)	$a = 91.07, b = 135.02,$ $c = 147.48$ $\alpha = 78.27, \beta = 77.77,$ $\gamma = 89.96$	$a = 91.54, b = 135.83,$ $c = 147.84$ $\alpha = 78.22, \beta = 77.89,$ $\gamma = 89.97$	-
Maximum resolution (Å)	2.9	3.0	3.0
Mosaicity (°)	1.0	1.0	1.0
Molecules per ASU	4		
Matthews coefficient (Å ³ /Da)	2.89	2.93	-
Space group	P1		
Wavelength (Å)	0.934	0.975	-
No. observed reflections	219268 (20892)	223954 (24999)	338761 (41810)
No. unique reflections	133319 (16784)	78529 (10857)	112209 (16428)
Resolution limits (Å)	51.1 – 2.9	50.0 – 3.2	50.0 – 3.20
Redundancy	1.6 (1.2)	2.8 (2.2)	3.0 (2.5)
Completeness (%)	89.8 (77.3)	69.5 (65.1)	81.8 (81.7)
R _{pim} (%)	5.6 (14.0)	14.3 (68.9)	17.5 (63.4)
I / σ (I)	9.6 (3.6)	6.3 (1.4)	7.0 (4.2)

Table II. 2 - X-ray crystallography data-collection statistics. (Values in parenthesis correspond to the highest resolution shell)

2.2.5 Model Building and Refinement

Initial protein model was manually built in Coot [62] and refined using phenix.refine [63]. Initial refinement cycles were carried out using simulated annealing. Because of the low resolution (2.9Å) and incomplete dataset (low data/parameter ratio), 4-fold NCS restraints, consisting of the 4 copies in the A.U., were imposed. Also, only one *B*-factor per residue was refined throughout the process. TLS was also used in the refinement, with the TLS groups automatically calculated by phenix. Currently, the R_{factor} and R_{free} values are 24.40% and 27.87% respectively, before addition of water molecules. Structure refinement data is present in Table II. 3. Final model included the following amino acid residues: Ser7 – Val1334. There are also some regions with no visible electron density, particularly: Pro169 - Thr199, Glu227 – Asn231, Gly399 – Ile404, Asp538 – Ile545, Leu558 – Gly563, Arg1290 – Trp1296, and Gln1321 – Pro1329.

Crystal Sample	mNAT-I
Resolution limits (Å)	51.1 – 2.9
Number of reflections	133319
R_{work} (%) ^a	24.40
R_{free} (%) ^b	27.87
RMSD from ideal geometry	
Bond lengths (Å)	0.020
Bond angles (°)	1.82
Average <i>B</i> -factor (Å ²)	41.10
Ramachandran outliers (%)	0.3%
Ramachandran favored (%)	93.6%

Table II. 3 – Structure refinement data.

^a $R_{\text{work}} = \sum ||F_{\text{calc}}| - |F_{\text{obs}}|| / \sum |F_{\text{obs}}| \times 100$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively. (^b R_{free} is calculated for a randomly chosen 5% of the reflections).

2.3 RESULTS AND DISCUSSION

The aldehyde oxidase homologue1 (mAOH1) was purified from mouse liver and from a heterologous expression system in *Escherichia coli*, and the resulting enzymes were compared. Both proteins show the same catalytic properties and characteristics, with the exception that the recombinant protein is only 30% active, while the native is 100%. Additionally, both proteins were crystallized and the best crystals were obtained from the native mAOH1 (Figure II. 2). These crystals diffracted beyond 2.9Å at the ESRF (Grenoble), and belong to space group P1 with two dimers present in the asymmetric unit (Table II. 2) [52].

2.3.1 Overall Structure Description

Mouse aldehyde oxidase homologue1 is a homodimer, composed of two identical subunits with approximately 150 kDa (Figure II. 4-a). Each monomer comprises 1335 residues and is divided in three main domains, classified according to their involvement in binding the distinct cofactors. Domain I is the N-terminal domain and the smallest one (20 kDa). It harbors the two spectroscopically different iron-sulfur clusters and comprises residues Met1-Pro169 (Figure II. 4-a, red). Domain II binds the FAD cofactor, has an approximate molecular size of 40 kDa, and includes residues Thr232-Leu534 (Figure II. 4-a, green). The C-terminal domain, is the largest one (90 kDa), encloses the MoCo (molybdopterin monophosphate, MPT) binding site and corresponds to residues Leu576-Val1335 (Domain III, Figure II. 4-a, blue). The Mo catalytic center is located at the bottom of a hydrophobic pocket accessible from the exterior through a wide funnel ~15Å long. The distance from the pterin exocyclic NH₂ to the nearest atom of [2Fe-2S] II center is 5.1Å (Figure II. 4-b). The cofactor disposition in the protein is related with the electron transfer pathway. When the oxidative hydroxylation occurs at the Mo catalytic center, the released electrons can pass consecutively through the two iron-sulfur-centers until they reach the FAD redox cofactor, and from there to the final electron acceptor, O₂. The three domains are connected by two linkers: domain I and II by a poorly conserved linkerI region (Ser170 to Asn231), and domain II and III by linkerII (Lys535 to Pro575) (Figure II. 4).

The N-terminal Iron/Sulfur Domain

The two iron–sulfur clusters are located in domain I, and are designated [2Fe-2S] I and II, according to their distinct EPR signals. Domain I can be sub-divided in two subdomains, the N-terminal subdomain (residues Met1-Val92) which harbors the [2Fe-2S] II and has a typical plant-type ferredoxin like fold, with one α -helix and five stranded β -sheets [14]. Fe/S II is bound via a Cys47-X4- Cys52-Gly-Ala-Cys55-X21-Cys77 motif. The second helical sub-domain (residues Glu93-Pro169) has a twofold symmetric arrangement, and binds the [2Fe-2S] I via the highly conserved Cys117-Gly-Phe-Cys120-X31-Cys152-Arg-Cys154 motif [13, 14]. The nearest Fe atoms from the two different Fe/S centers are 12.4Å apart from each other. The closest distance between [2Fe-2S] II and the molybdopterin cofactor is 5.1Å, while [2Fe-2S] I is approximately 9Å away from the FAD cofactor (Figure II. 4-b).

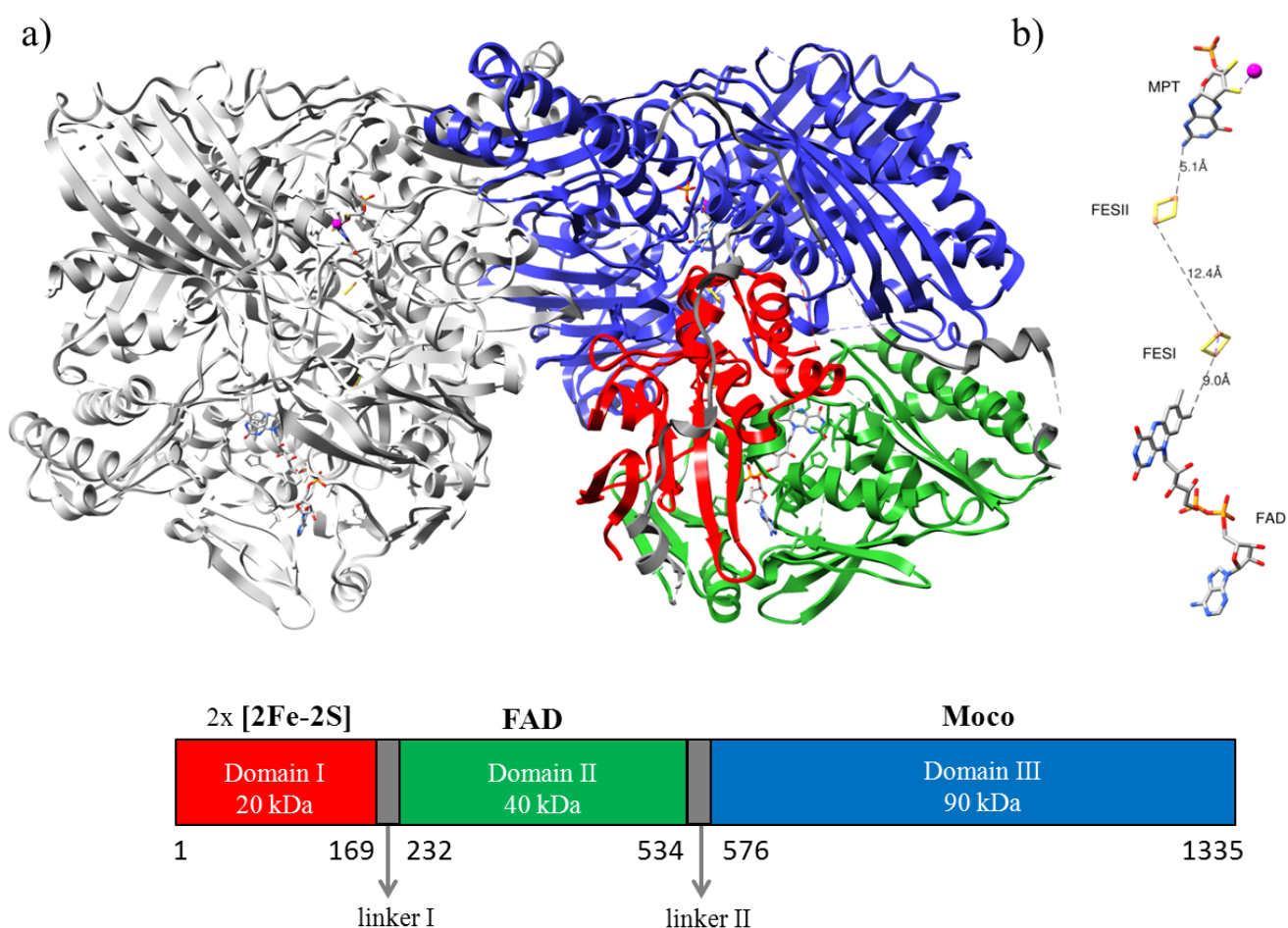


Figure II. 4 - (a) Ribbon representation of mAOH1 crystal structure. Monomer B in grey (left), and monomer A showing the different domains (right), colored as: domain I in red (residues Met1-Pro169); domain II in green (residues Thr232-Leu534); and domain III in blue (residues Leu576-Val1335). The linker regions are represented in grey (linker I and II, respectively Ser170-Asn231, and Lys535-Pro575). The molybdopterin monophosphate (MPT), [2Fe-2S] and FAD cofactors are shown as sticks and colored by atom type. (b) Arrangement and distances between the different protein cofactors: MPT, the two distinct [2Fe-2S] centers, and FAD.

The FAD Domain

The flavin adenine dinucleotide (FAD) cofactor is part of domain II, which can be subdivided in two different sub domains spanning from Thr232 to Arg419, and from Lys 420 to Lys534, respectively. The FAD cofactor is mainly coordinated by the first sub domain. Residues Pro263, Leu264, Asn268, Thr269, Tyr270, Ser354 and Leu411 make hydrogen bonds with the adenine dinucleotide region, while residues His358 and Asp367 contact the riboflavin ring. From all these residues, only the residues Thr269 and Asp367 are highly conserved among other XO family members (Figure II. 5). The second sub domain is in the MoCo domain side, and contains three long β -sheets and two α -helices. The FAD molecule occupies a vast area within the protein, with the isoalloxazine ring in close proximity to the solvent area. The riboflavin ring is pointing towards Fe/S I center (Figure II. 4-b).

The C-terminal Catalytic Domain

The C-terminal and largest domain (Domain III - 90 kDa) contains the MoCo binding site, which is located on the bottom of the substrate binding pocket. It can be structurally divided in four sub-domains, classified as for MOP and XO [13, 14]. Each sub domain is constituted by non-continuous stretches of the polypeptide chain: sub domain III.1, comprises residues 576-698 and 745-843; sub domain III.2, residues 699-744, and 844-964; sub domain III.3 possess residues 965-1009, plus 1160-1354; and finally, sub domain III.4 is constituted by residues 1010-1159. All these sub domains have a similar fold, with two central α -helices covered by three to five β -sheets, solvent exposed in the case of sub domains III.2 and III.3, and involved in dimerization contacts in sub domains III.1 and III.4. The molybdopterin monophosphate cofactor is buried in the protein, and forms hydrogen bonds with residues Ala802, Arg917 (highly conserved among XO family members, Figure II. 5), Leu1043, Ser1085 and Gly1087. The molybdenum atom is found at the bottom of a wide funnel (20Å wide at the surface), dominated by the presence of hydrophobic residues, which leads to the active site (Figure II. 10).

Figure II. 5 - Comparison of the amino acid sequence of mAOH1 with XOR from bovine milk, with XDH from *R.capsulatus* and MOP from *D.gigas*. The mAOH1 [2Fe-2S] I and II binding Cys residues are marked as (*) and (+) respectively. Black boxes indicate identical residues, and grey similar ones. Alignment results obtained with CLUSTALW multiple sequence alignment program at the EMBL-EBI web server, and the figure prepared using BOXSHADE from EMBnet server.

2.3.2 The Homodimer

The overall topology of mAHO1 is similar to the mammalian xanthine dehydrogenase protein (XDH), a typical butterfly shape structure with approximate dimensions of 150Å x 90Å x 70Å (Figure II. 4). The two monomers are tightly bound, involving 23 hydrogen bonds ($< 3.2\text{\AA}$) and 12 salt bridges. The majority of these bonds are established by residues from domain III, the MoCo binding domain. Particularly residues: Lys759, Glu761 and 764, Lys789, Asn790 and 793, Lys797, Arg798 from sub-domain III.1, and residues Gln1021, Ile1028, Thr1030, Asp1031, Ser1033, Val1034, Asp1069 and Asn1078 from sub-domain III.4. The surface contact area between the two subunits is 2500\AA^2 , approximately 5.3% of the total surface area of each monomer. The molybdenum atoms from the two different monomers are more than 50Å apart and the two subunits most likely act independent from each other.

2.3.3 Comparison with Structurally Related Proteins

The mAOH1 structure reported in this Thesis is the first crystal structure ever obtained for an aldehyde oxidase. However, several protein structures from the XO family have already been reported. The most interesting feature among these different proteins is the similar overall globular shape, while catalyzing different reactions with a wide range of substrates. The first crystal structure was reported in 1995, for the prokaryotic *D.gigas* aldehyde oxidoreductase (MOP) that was solved at 2.25Å, and later refined to almost atomic resolution (1.28Å) [14, 64]. MOP oxidizes aldehydes to carboxylic acids, with little specificity for the nature of the side group. Despite the low correspondence in sequence identity (23%, Figure II. 5), the overall fold of MOP is similar to mAOH1, in domains I and III, since MOP lacks the FAD domain and has only 907 amino acid residues (Figure II. 6). Also, the two [2Fe-2S] centers are similarly positioned in both MOP and mAOH1, and many of the residues proximal to the clusters are either identical or type conserved.

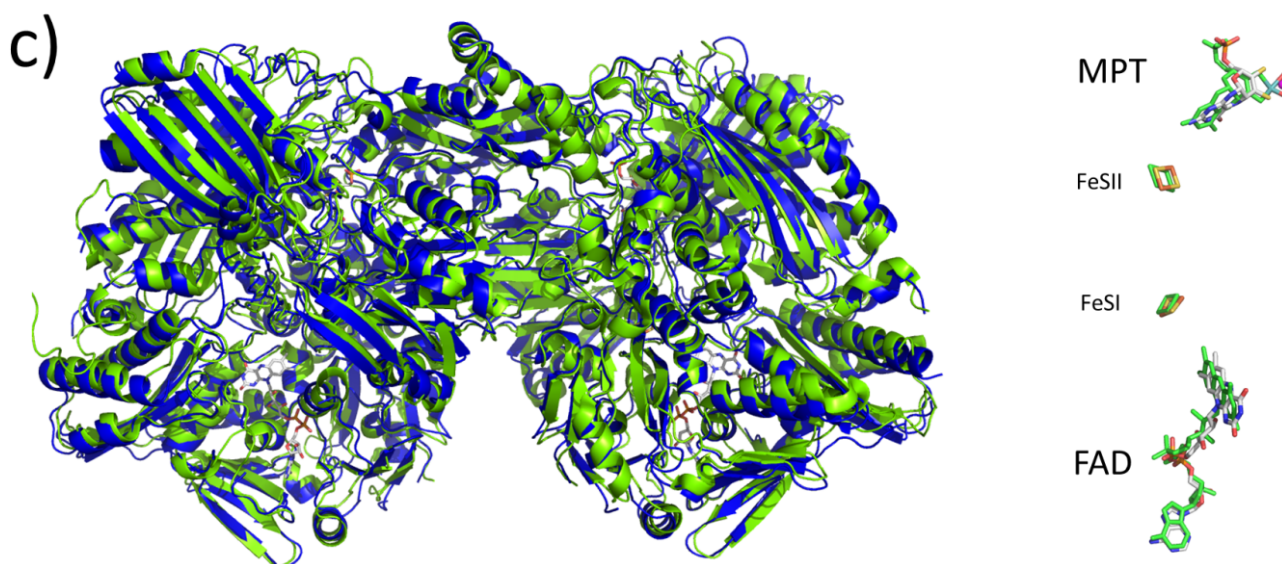
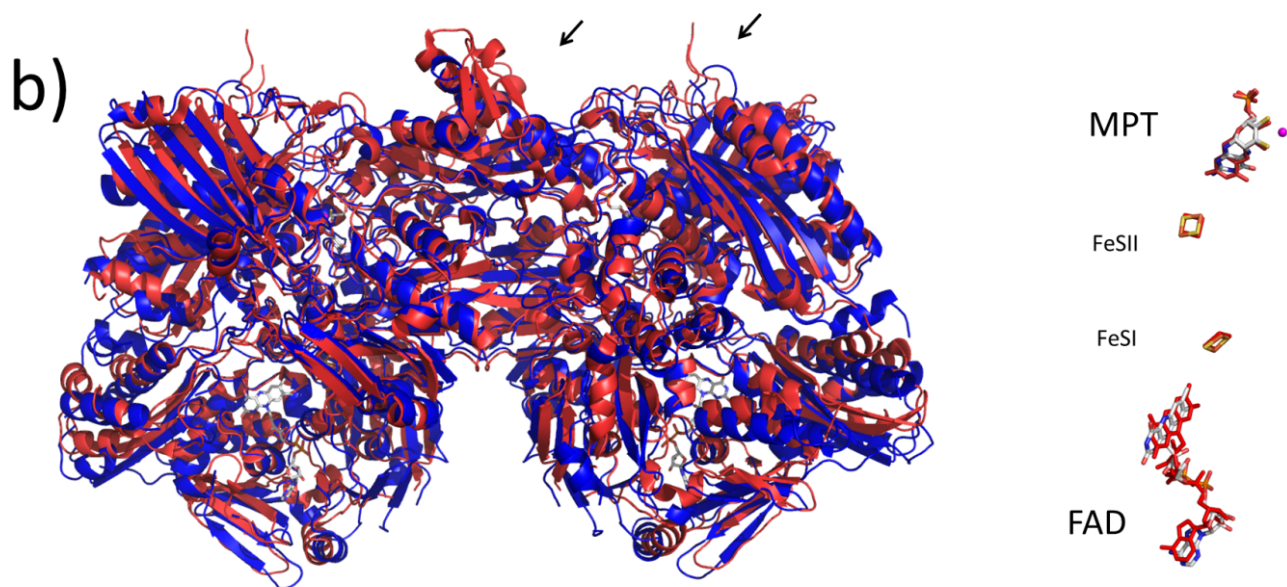
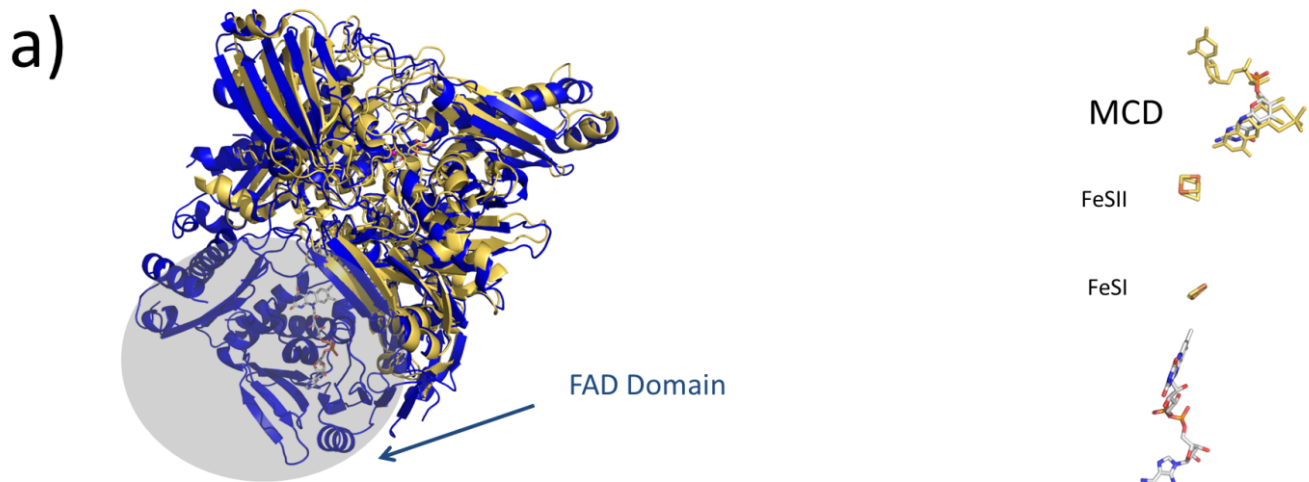
The crystal structures of bovine milk xanthine dehydrogenase (XDH) and its interconvertible form xanthine oxidase (XO) were first reported in 2000 [13]. The crystal structure of the XDH isolated from the purple bacterium *Rhodobacter capsulatus* (*Rc*) was also solved, and published in 2002 [65]. While the bovine XDH/XO is a homodimeric enzyme with ~1300 amino acid residues in a single polypeptide chain, the *R.capsulatus* XDH has an $(\alpha\beta)_2$ heterotetrameric structure, with molecular mass of 275 kDa. In the latter enzyme, the cofactors are located on different polypeptide chains, with

the two [2Fe-2S] clusters and the FAD bound to the XdhA subunit, and the MoCo cofactor located in the XdhB subunit. For an easier comparison, each mAOH1 monomer (that includes the three different cofactors) can be considered as equivalent to one *Rc* XDH heterodimer. The mAOH1 and the *Rc* XDH structures superimpose with an rms deviation of 1.09Å for 913 C_α atoms. Despite the different subunit composition, the overall fold as well as cofactors arrangement of mAOH1 and *Rc* XDH is similar, with both proteins sharing 32% sequence identity. However, there are some differences that can be noticed in the primary sequence alignment (Figure II. 5). The linker II region (residues 535 to 575, between the FAD and MoCo domains) present in AOH1 and XDH/XO is absent in the bacterial enzyme. Two other clear differences, easily seen in the structures superposition (Figure II. 6), result from the insertions of 20 and 19 amino acid residues, in the regions Arg953-Thr954 and Pro1114-Ser1115, respectively (mAOH1 numbering). These two regions are located at the MoCo binding domain and solvent exposed, and the Arg953-Thr954 sequence is at the XDH heterotrimer interface. The two insertions protrude out of the globular protein surface (marked with arrows in Figure II. 6).

The several cofactors are localized in similar positions in the four enzymes, and provide equivalent electron transfer pathways (Figure II. 6).

Two important aspects must be addressed at this point. The XDH from *R.capsulatus* possess high reactivity towards NAD⁺, as in the case of bovine XDH, and does not undergo conversion to the oxidase form, as observed for the mammalian XDH/XO (explained below). The MoCo analysis of *Rc* XDH revealed the presence of the molybdopterin cofactor monophosphate form (MPT) as found in the eukaryotic MoCo containing enzymes (Figure II. 6). This was not expected, since the majority of MoCo containing bacterial enzymes (as MOP) have the MCD form (Figure II. 6), with an additional cytosine nucleotide covalently attached to the pterin [65].

Figure II. 6 - Comparison between mAOH1 (blue) and the different crystal structures obtained for the XO family: (a) aldehyde oxidoreductase from *D.gigas* (MOP monomer – yellow, PDB code: 1VLB), (b) xanthine dehydrogenase from *R.capsulatus* (*Rc* XDH - red, PDB code: 1JRO), (c) xanthine dehydrogenase from bovine milk (XDH milk - green, PDB code: 1FO4). Next to each structure comparison is the superposition of the cofactors involved in electron transfer (mAOH1 –atoms color coded, MOP in yellow, *Rc* XDH in red, and XDH milk in green).



2.3.3.1 Mammalian XOR Inter-conversion and mAOH1

Mammalian xanthine oxidoreductases (XOR) are synthesized in the dehydrogenase form, but can be readily converted to the oxidase form by two different mechanisms: oxidation of sulfhydryl residues, and proteolysis [13]. The enzyme exists in two alternative forms (XDH and XO) of the same gene product. The global fold of both proteins remains essentially the same after the inter-conversion $XDH \rightarrow XO$, with an rms deviation in $C\alpha$ positions of 0.30\AA for 1095 atoms. The general conservation of the structure is consistent with biochemical studies, and no major differences regarding binding and catalysis of substrates are observed at the molybdenum catalytic center in both protein forms. The main difference is the binding of the cofactor NAD^+ , which occurs in the XDH but not in the XO form. This has implications in structural terms, with differences between the two forms at a loop region (see below). In XO, molecular oxygen acts as the electron acceptor, as in the case of mAOH1.

When comparing bovine milk XDH and mouse AOH1 crystal structures (Figure II. 6 - c), we can see that their global fold as well as cofactor arrangement is very similar (rms deviation of 1.26, for 2358 $C\alpha$). The major difference between the two proteins concerns important residues located not only at the active site (section 2.3.4), which explains the differences regarding substrate specificities, but also residues involved in the XOR mechanisms of inter-conversion. This kind of mechanisms has never been described for aldehyde oxidases. The reversible $XDH \rightarrow XO$ conversion mechanism involves the oxidation of two different cysteine residues (Cys535 and Cys992, bovine XDH numbering), and the formation of a new disulfide bond. These two cysteine residues are not conserved in mAOH1, corresponding to Tyr542 and Phe997 respectively. Tyr542 could not be assigned in the mAOH1 electron density map, but Phe997 locates in the protein surface, pointing towards the solvent.

XDH can also be irreversibly converted to XO by proteolysis and in both cases (oxidative and proteolytic generation of XO) there is a structural rearrangement on a loop close to the FAD ring. This “moving loop” is constituted by residues Gln423-Lys433 (bovine XDH numbering), and its displacement causes partial blockage of the NAD^+ molecule towards interaction with the FAD cofactor (Figure II. 7). As a result, the XO form does not bind NAD^+ and uses oxygen as electron acceptor. As well as the “moving loop” displacement, there is the removal of a phenylalanine residue (Phe549, bovine XDH numbering) from a tightly packed amino acid cluster, causing major

reorientation of a tryptophan residue (Trp336 in bovine XDH) (Figure II. 7– a). This residue is located at the entrance of a tightly closed cavity leading to the FAD cofactor, and its movement opens the cavity causing FAD solvent exposure, particularly in the isoalloxazine ring (around C4-C6 positions). As a result of these two different displacements (Gln423-Lys433 loop and Trp336, XDH numbering) there is a change in the FAD cofactor electrostatic environment, which accounts for the main differences between XDH and XO forms regarding nicotinamide adenine dinucleotide binding [13, 66].

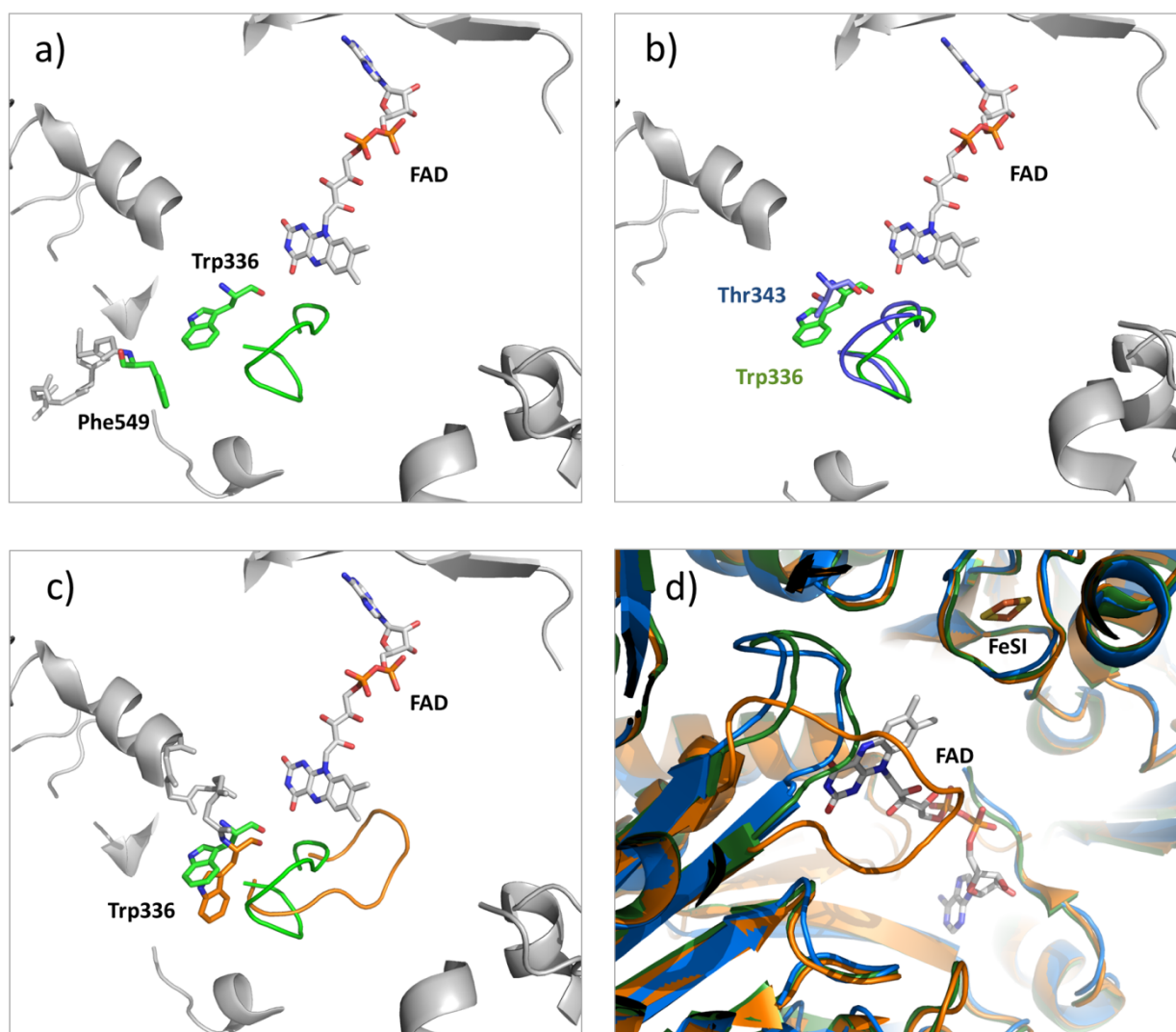


Figure II. 7 - Representation of the “moving loop” (residues Gln423-Lys433, in bovine XDH numbering) and important residues involved in the XDH/XO conversion. (a) Residues Phe549, Trp336 and the original loop position in bovine XDH (green). (b) Comparison between mAOH1 (blue) and bovine XDH (green), concerning the “moving loop” and Trp336 from XDH (which corresponds to Thr343 in mAOH1). The “moving loop” is mainly constituted by highly charged residues (QASRREDDIAK) in XDH, which are replaced by lesser-charged ones in mAOH1 (QAPRQQNAFAT) (the 4 conserved residues are underlined). (c) Differences observed in the loop and in residue Trp336 after the XDH→XO conversion (XDH in green, and XO in orange). (d) Same as in (c) but from a different perspective, outlining the differences between the “moving loop” from XO (in orange), and the mAOH1 and XDH structures (blue and green, respectively). All the FAD cofactors are represented color coded, and correspond to the mAOH1 structure.

In mAOH1, the corresponding “moving loop” is superimposable with the one observed in the XDH form, although only 4 out of 11 residues are conserved between the two proteins (Figure II. 7- b). However, as previously mentioned aldehyde oxidase does not bind the NAD^+ cofactor, using oxygen as the electron acceptor. The mAOH1 protein should therefore be more similar to the XO form than to XDH, since they possess the same electron acceptor. Nevertheless, xanthine oxidoreductase naturally occurs in the dehydrogenase form, and in the absence of an inter-conversion form described for AO, it is not surprising that the protein shows higher similarity towards the natural XDH form. The inexistence of an AO inter-conversion form could be related with some of the differences described previously.

XOR inter-conversion has been described by some authors to be connected with milk lipid secretion [67], but other opinions defend that it is involved in human ischemic-reperfusion injury defense mechanisms, due to the higher formation of radical oxygen species during the XO form catalytic reaction [67, 68].

2.3.4 mAOH1 Catalytic Active Site and Substrate Funnel

At the present resolution (2.9Å) it is not possible to identify the chemical nature of all the atoms bound to molybdenum, based on electron density alone (Figure II. 8). Two of these ligands are the sulfur atoms from the dithiolene moiety, which in mAOH1 are located in Mo-distances ranging from 2.4 to 2.7Å (for the 4 molecules in the a.u.). The exact nature of the three other ligands must be based on crystallographic data for the related XOR [13, 14, 69]. In the oxidized form of the active enzyme, the Mo coordination sphere should include an essential catalytic sulfur atom (=S), an apical oxygen atom (=O), and an equatorial OH group. This single bonded oxygen atom is pointing towards the active site funnel and is later transferred to the substrate. The five different ligands form a square pyramidal geometry.

With the mAOH1 structure firstly reported here, we are able to take a closer look at the protein active site, and draw some conclusions regarding its broad range of substrates. Several pharmaceutical companies have given their attention to aldehyde oxidases in recent years. It has been recognized that through its unique structure, distribution and substrate recognition, this protein has a vital role in drug metabolism, despite the lack of knowledge regarding biochemical and physiological

function. It is imperative now to understand all these specificities, located mainly at the protein active site.

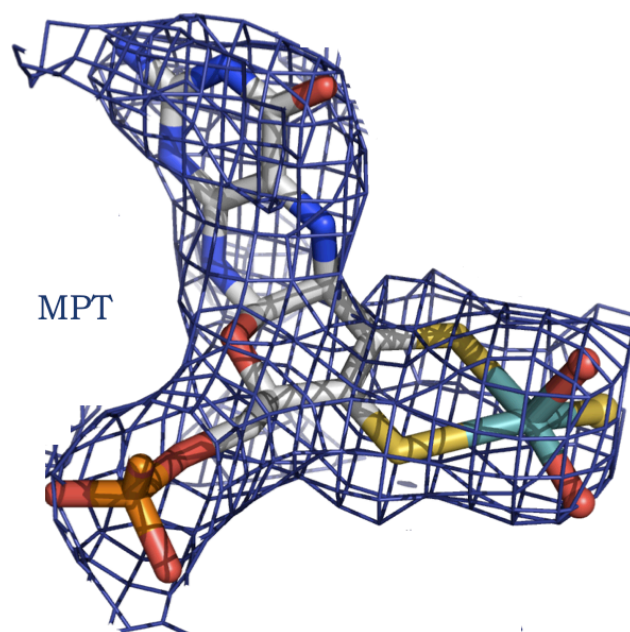


Figure II. 8 - Close-up of the MPT and Mo coordination sphere, for the mAOH1 monomer A. ($2mF_o-DF_c$ electron density map contoured at 1σ , blue).

Typical aldehyde oxidase substrates are compounds containing an aldehyde function, nitro/nitroso compounds or N-heterocycles (Table II. 1), while hypoxanthine and xanthine are XOR elected substrates. When comparing mAOH1 and bovine milk XDH structures, the most important residues conserved at the active site are the ones located up and down relative to the Mo atom (Gln772 and Glu1266) and two Phe responsible for maintaining the XDH salicylate inhibitor in place (Phe914, and Phe 1009). In XDH, the aromatic ring of Phe914 is parallel to the inhibitor molecule, while Phe1009 has an edge-on interaction. The two Phe are conserved in mAOH1 (Phe919 and Phe1014), making the model interaction of bicyclic substrate also possible. The mAOH1 Glu1266 residue (Glu1261 in bovine XDH and totally conserved in all members of the XO family) is also included in these conserved residues, and is responsible for playing a fundamental role in the reaction mechanism, as explained below.

Other important residues located at the catalytic core are in general not conserved between the two proteins, and this explains the differences regarding binding and catalysis. Particularly, residues Glu802, Arg880 and His884 (in bovine XDH numbering), which are replaced by Ala807, Tyr885

and Lys889 in mAOH1, respectively, are crucial for the different proteins specificities. Glu802 from bovine XDH is replaced by an aliphatic and smaller residue in mAOH1, Ala807 residue, which allows the active site to accommodate larger and bulkier substrates. Also the presence of a less charged residue contributes to the necessary environment for the binding of more charged compounds as seen in mAOH1. In bovine XDH another important residue, Arg880, is responsible for the correct positioning and maintenance of the salicylate and oxipurinol inhibitors (represented in Table II. 4 and Figure II. 9, respectively). This position is occupied by the aromatic ring of a tyrosine in mAOH1 (Tyr885) pointing outwards from the molybdenum atom and again increasing the mAOH1 active site availability for accommodating larger substrates. The mAOH1 Lys889 residue, although being almost 10Å away from the molybdenum center, makes also a decisive contribution in the protein catalytic core. This residue (conserved in other AOs) corresponds to His884 in bovine XDH, a bulky residue located in the active site funnel pathway, limiting the access of large substrates to the protein catalytic center. All these differences account for the binding of the distinct XOR and AOH1 substrates, particularly the substitution of smaller residues in the aldehyde oxidase protein (Figure II. 9).

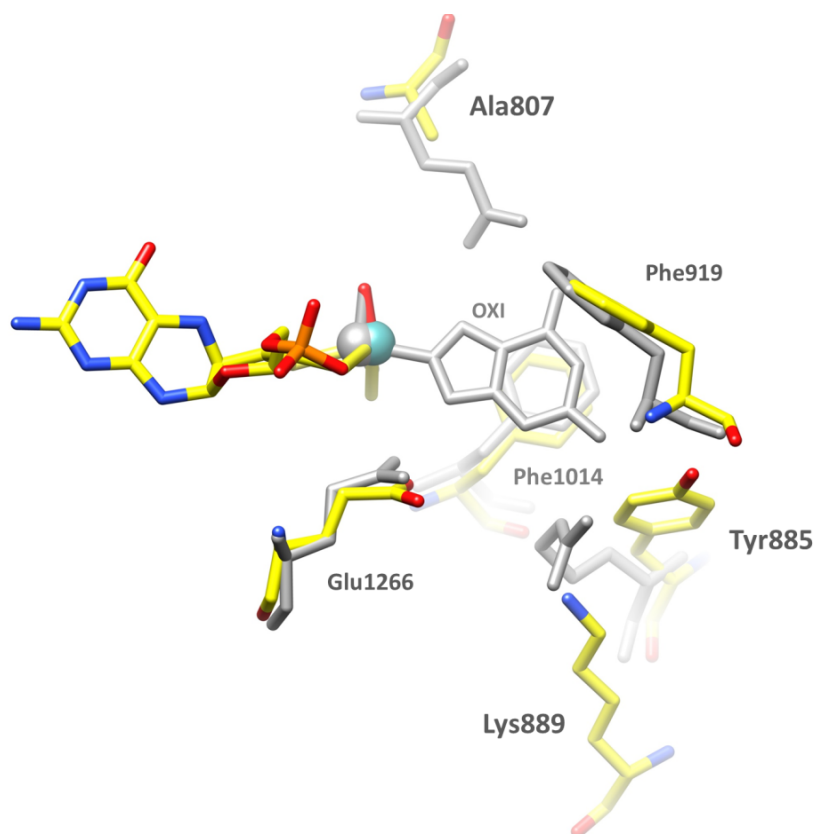


Figure II. 9 – Active site comparison between mAOH1 (yellow) and bovine XDH (gray). Outlined are the most important non-conserved residues, in mAOH1 numbering. Also present is the XDH inhibitor oxipurinol molecule (OXI), and the conserved XDH and mAOH1 residues Phe919, Phe1014 and Glu1266.

Interestingly, the Tyr885 residue from mAOH1 is also not conserved in other aldehyde oxidases, particularly in human and mouse, where a methionine can be found. Mahro and co-workers (unpublished data) have expressed and purified the mAOH1 Y885M variant, from which we were also able to obtain protein crystals. The resulting crystals had poor or non-existing diffraction. Biochemically the authors found out that the variant shows the same behavior towards benzaldehyde, but that allopurinol (the most studied mammalian XOR inhibitor) which is a substrate for the *wild type* enzyme cannot be metabolized by the variant. Two different xanthine oxidase mutational changes (E803V and R881M, XDH numbering), were also reported a few years ago by Yamaguchi and co-workers [70]. Although these residues are not conserved in mAOH1, they are conserved in AO and the aim was to understand the role of some active site residues in binding and activation of purine substrates. As a result, the two variants showed significant aldehyde oxidase activity, which also indicates that it is the chemical nature of residues responsible for the different substrate binding, and not only its location within the protein active site. After the mutations, the overall active site was similar and no structural changes could be assigned.

All the important mAOH1 residues (conserved and non-conserved), and their respective equivalents in bovine XDH are summarized in Table II. 4.

On the basis of the above discussion it is clear that mAOH1 is different from XOR, predominantly in the catalytic active site, which indicates substantial differences around the MoCo substrate-binding site. As AO and XOR are characterized by distinct substrates and inhibitors, it is also predicted that the tunnel leading to the active site is different between the two enzymes. The existence of substrates recognized by mAOH1, but not by mAOH2 (E.Garattini, unpublished data), as well the identification of better inhibitors of mAOH1 than AO, further suggests that the tunnel may vary even among the different AO isoenzymes [16].

Several residues along the funnel, which leads to the protein active site, are also not conserved between XDH and mAOH1 (Table II. 4). This explains the differences observed in the funnel shape and width, which is much wider and anionic in mAOH1 than in XDH (Figure II. 10). In mAOH1 the funnel leading to the active site is $\sim 20\text{\AA}$ wide at the surface and becomes tighter towards the molybdenum catalytic center, where it is approximately 8\AA wide. This is consistent with the entrance of larger and bulkier substrates, as the ones described to be catalyzed by aldehyde oxidases.

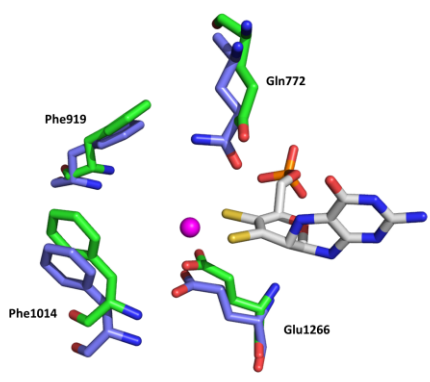
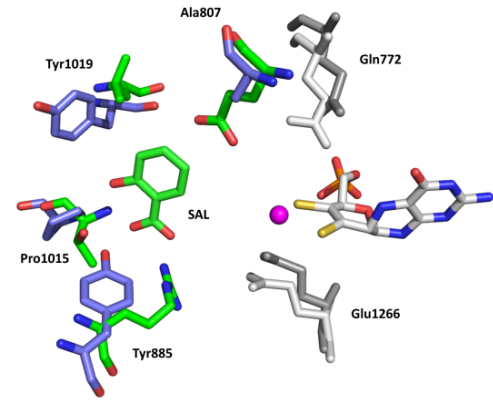
Type of residues	Location	mAOH1	XDH
Conserved	 <p>Active site</p>	Gln772	Gln767
		Phe919	Phe914
		Phe1014	Phe1009
		Glu1266	Glu1261
Non- conserved	 <p>Active Site</p>	Ala807*	Glu802
		Tyr885*	Arg880
		Lys889*	His884
		Pro1015	Thr1010
		Tyr1019	Asn1014
	Funnel	Arg717	Leu712
		Asp878	Leu873
		Glu880	His875
		Leu881	Ser876
		Thr1081	Pro1076

Table II. 4 – Comparison between mAOH1 and bovine XDH, concerning the main residues involved in substrate binding and catalysis. Non-conserved residues marked with an asterisk are mentioned above, due to their primary importance.

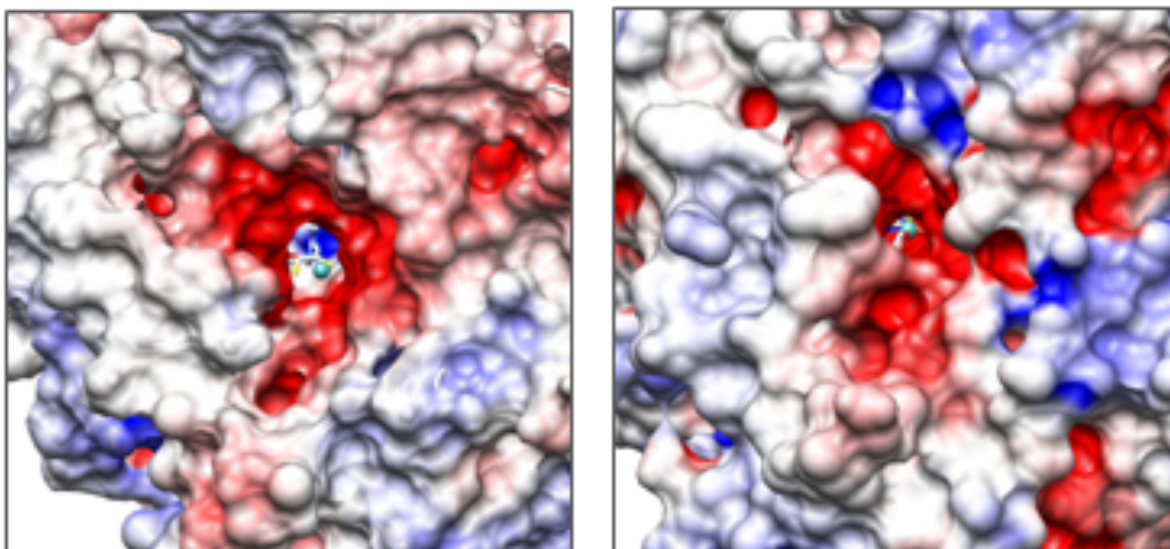


Figure II. 10 - Surface view of the funnel leading to the active site in mAOH1 (left) and bovine XDH (right). The Mo atom is seen at the end of the funnel as a green sphere. The entrance is much wider in the case of aldehyde oxidase, with a narrow constriction closer to the molybdenum active site. Electrostatic surface potentials were calculated using the program delphy [71] and represented in PyMOL [72], with the color of surface potentials in the scale range from -0.10V (negatively charged, in red) to +0.10V (positively charged, in blue).

2.3.5 The Reaction Mechanism

The initially proposed reaction mechanism for XDH/XO forms described by Nishino and co-workers has been the origin of some debate and controversy over the years regarding binding mode and substrate orientation [13, 73]. However, a similar reaction mechanism as the one described for XOR has been also proposed for aldehyde oxidases [16]. With the reported mAOH1 structure (2.9Å), it is still not possible to identify the chemical nature of the atoms bound to molybdenum based on electron density maps. Therefore, in the absence of further evidence for the oxidized form of the enzyme and owing to the presence of highly conserved residues at the catalytic center (particularly Glu1266, in mAOH1 numbering), we assume that the reaction mechanism for aldehyde oxidase is similar to the xanthine oxidase mechanism. The reaction starts with the activation of the equatorial OH group by the neighboring conserved glutamic acid residue (Glu1266). This oxygen is then transferred to the substrate and the resulting intermediate replaces the coordinated OH at the molybdenum. Simultaneously, a hydride transfer from the substrate to the equatorial sulfur atom completes the reaction intermediate. The product is afterward released from the reduced molybdenum center and the new available position is occupied by a water molecule (Figure II. 11).

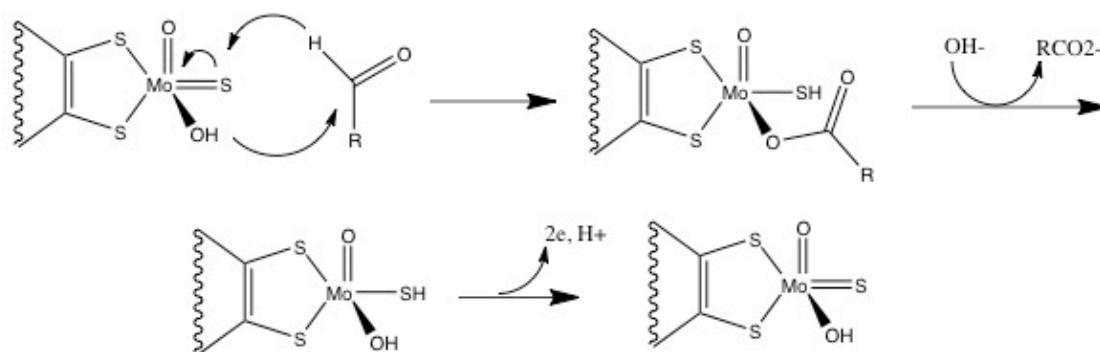


Figure II. 11 – Proposed reaction mechanism for aldehyde oxidase (AO). Represented is the base-assisted nucleophilic attack of the Mo-OH on the substrate carbonyl, with concomitant hydride transfer to the Mo=S. (Adapted from [16, 74]).

The glutamate contribution to catalysis has been confirmed by site-directed mutagenesis in the recombinant mouse aldehyde oxidase (mAOX1). The E1265Q variant was generated and purified, and was unable to metabolize any of the aldehyde or purine substrate tested. This is consistent with an essential catalytic role for this residue [53]. The Glu1261 from XDH and the Glu1266 from mAOH1 are totally superimposable in the crystal structure (Table II. 4). To further investigate the aldehyde oxidase reaction mechanism in detail and draw comparisons with XOR, two different approaches must be taken. From the structural point of view, different crystal soaking experiments with substrates and inhibitors must be performed, as soon as crystal quality and resolution can be improved. Biochemical experimental data based on mutagenic studies should also be done in parallel.

CHAPTER 3

Periplasmic Nitrate Reductase from *Cupriavidus necator*

The work described in this chapter was the subject of two publications:

C. Coelho, P.J. González, J.J.G. Moura, I. Moura, J. Trincão and M.J. Romão, (2011) “The crystal structure of *Cupriavidus necator* Nitrate Reductase in oxidized and partially reduced states” *Journal of Molecular Biology*, 408, 932-948;

C. Coelho, P.J. González, J. Trincão, A.L. Carvalho, S. Najmudin, I. Moura, J.J.G. Moura, T. Hettman, S. Dieckman and M.J. Romão, (2007) “Heterodimeric nitrate reductase (NapAB) from *Cupriavidus necator* H16: purification, crystallization and preliminary X-ray analysis” *Acta Cryst.*, F63,516-519.

3.1 INTRODUCTION

3.1.1 *Cupriavidus necator*

Cupriavidus necator (*Cn*, formerly known as *Ralstonia eutropha* or *Alcaligenes eutrophus*), is a respiratory facultative β -proteobacterium. The *Cn* H16 strain is one of the best-studied model organisms for lithoautotrophic growth on molecular hydrogen and carbon dioxide, which are used as sole energy and carbon sources in the absence of organic substrates [75, 76]. Physiological and genetic studies revealed the presence of three different nitrate reductase activities in this strain [77]. The first activity corresponds to the cytoplasmic nitrate reductase protein (Nas), and was both repressed by ammonia and not sensitive to oxygen, suggesting a nitrate-assimilatory function. The second activity was only observed in the absence of oxygen and independent of ammonia suppression, representative of a nitrate respiratory function performed by the membrane bound nitrate reductase (Nar). The third activity occurred in the soluble fraction of cells, which were grown aerobically in the presence of ammonia during the stationary phase. This activity corresponds to the periplasmic nitrate reductase (Nap), whose expression, in contrast to Nas and Nar, did not require nitrate for induction. The genes for the three different nitrate reductases are mapped at different loci: Nas and Nar are chromosomally encoded, whereas Nap is located in a megaplasmid (pHG1) [27]. To elucidate the structure and function of *C.necator* periplasmic nitrate reductase, a homologous expression system was developed by Hettmann and collaborators [78]. A new *C.necator* strain (HF210) was used, together with the complete nap operon, and the protein was purified to near homogeneity.

Several prokaryotic nitrate reductases from different organisms have been studied, and for some of them, the *nap* cluster has four genes in common: *napDABC*. The conservation of *napA*, *napB* and *napC* is expected, as they encode the terminal bis-MGD-containing moiety, the di-heme, and the tetra-heme components of the periplasmic electron transfer system respectively. The Tat (twin-arginine transport) system is responsible for the translocation of folded proteins, across the cytoplasmic membrane [79]. In *E.coli*, NapD is known to be a Tat signal peptide-binding chaperone involved in the biosynthesis of the Tat-dependent NapA. After recognition of the twin arginine-containing motif in the signal sequences of the protein cofactor, insertion in NapA occurs in the cytoplasm, and the folded protein is then exported across the cytoplasmic membrane into the periplasmic compartment [27, 80].

3.1.2 Heterodimeric Periplasmic Nitrate Reductase (NapAB)

Nitrate reductases depend on the bis-molybdopterin guanine dinucleotide (MGD) cofactor located in the active site, and belong to the DMSO reductase family of Mo proteins, described in section 1.1.6 of this Thesis. The homologous expression of the periplasmic nitrate reductase from *C.necator* allowed biochemical and structural studies to be carried out, as well as elucidating its three dimensional architecture. *Cn* NapAB is a heterodimer composed of a large catalytic subunit (90 kDa) containing the molybdenum ion bound to two MGD cofactors plus a [4Fe-4S] cluster, and by a small di-heme c-type cytochrome subunit (17 kDa), involved in electron transfer. Different roles such as redox balancing and nitrate scavenging have been proposed for this protein.

The first crystal structure for a nitrate reductase was solved by multiple-wavelength anomalous dispersion (MAD) in 1999 for the monomeric NapA from *Desulfovibrio desulfuricans* (*Dd*) at 1.9Å resolution [30]. In the following years, despite the extensive biochemical and spectroscopic characterization of several periplasmic nitrate reductases, few novel crystal structures were reported. In 2002, Brigé and co-workers published the structure of a proteolysed fragment of NapB subunit from *Haemophilus influenzae* (*Hi*) at atomic resolution (1.25Å), also with the help from MAD methods for deriving phase information [81]. One year later, the first crystal structure for a heterodimeric NapAB was obtained for the *Rhodobacter sphaeroides* (*Rs*) protein at 3.2Å resolution by Arnoux and collaborators [31]. The heterodimeric NapAB from *E.coli* was also studied in an attempt to determine its three dimensional structure, but due to heterodimer separation during purification protocols, the crystal structure could only be determined for the catalytic subunit at 2.5Å resolution [82]. In the oxidized form of these enzymes, the molybdenum atom is coordinated by two dithiolene ligands from the two MGD molecules, a S γ atom from a Cys residue and a sixth ligand which was originally proposed to be a hydroxyl/water molecule. The sixth molybdenum ligand in the first crystal structure of NapA from *D. desulfuricans* was assigned as an OH/OH₂ molecule (with Mo-O distance of approximately 2.1Å), but in recent studies obtained at higher resolution (1.8Å), and using improved refinement protocols, there was strong evidence that a heavier element (such as a non-protein sulfur ligand) is present at the sixth coordination position [33]. At the same time, crystals of the oxidized form of *C.necator* NapAB were obtained at high resolution (1.5Å) and the structure solved. The high resolution of the data provided atomic details of the protein active center and on molybdenum coordination sphere. The other Nap structures previously reported were solved to a resolution that was not enough to unequivocally identify the Mo ligands. The presence of sulfur as

the sixth molybdenum ligand implies that the formerly proposed reaction mechanisms simply based on the redox chemistry of Mo, could no longer be valid, and should also involve a sulfur-based redox mechanism. Three alternative reaction paths were proposed on the basis of the re-refined *Dd* NapA structure, and those have been the subject of theoretical and computational studies reported by different authors that provided similar results on the most favorable pathway of a putative mechanism [33, 83, 84].

3.2 EXPERIMENTAL PROCEDURE

3.2.1 Bacterial Expression and Growth Conditions

Periplasmic nitrate reductase genes are not part of the *Cupriavidus necator* chromosome, but located in the 450 kb pHG1 megaplasmid. The protein was expressed in a megaplasmid-free strain (HF210) into which a broad host-range vector (pCM62) containing the complete Nap cluster was inserted [78]. The resulting strain was grown in mineral medium (Appendix – Table 1) at 30°C until late death-phase [85]. The pre-culture was started by adding 20 µl of a *C.necator* (Δ pHG1-H210/pNAP62-2 strain) glycerol stock solution into a nutrient broth agar plate and incubated overnight at 30°C. Several colonies were afterwards chosen and resuspended, and again incubated overnight at 30°C, 190 rpm. The resulting saturated cultures were used to inoculate 18 x 1L of mineral medium and incubated at 30°C and 200 rpm until late death phase for almost 45 hours.

This yielded enough protein for the initial robot crystallization trials (16 µl of 10 mg/ml NapAB), but not for the attempts to reproduce the robot conditions in-house. To overcome this problem, the bacteria were grown in a 10L pre-equilibrated reactor (BioStat B-Plus) inoculated with 200 ml of saturated cultures. The temperature was kept constant at 30°C and the culture was continuously stirred between 200 and 400 rpm in order to maintain 20 % minimum oxygen level. After 49 h, the culture was centrifuged to isolate the periplasm. The protein obtained from the BioStat B-Plus reactor was used to scale-up the robot conditions, with the help from ionic liquids (Chapter 4), but was again insufficient to carry out the spectroscopic and electrochemical studies. To get such an enormous amount of protein, the bacteria were grown for the first time in a UD 300L Bioreactor at IBET (Oeiras, Portugal). To inoculate the industrial reactor, a pre-culture from a 10L ED 10 Bioreactor grown at 30°C for 24 h was used. All the conditions had to be optimized to such a large-scale growth (Appendix - Table1). Temperature and pressure were kept constant at 30°C and 100 mbar, agitation varied between 100 rpm – 500 rpm in order to maintain 30% oxygen. Several aliquots were retrieved from the sample and immediate measurement of optical density at 600 nm was performed in order to monitor the growth rate. The fermentation was finished 48 hours after inoculation (Appendix – Figure 1).

3.2.2 Periplasm Isolation and Protein Purification

For all the above growth experiments, periplasm had to be immediately isolated following the end of bacterial growth. In order to do this, the cultures were centrifuged at 5000g for 20 min and with a flow of 80L/h in a tubular centrifuge, respectively for the regular (10L) and industrial (300L) growths. The resulting pellets (16 g and 455 g respectively) were washed in 50 mM Tris-HCl pH 7.6 and resuspended in 10 mM Tris-HCl pH 7.6 with 0.5 M sucrose (using 5 ml per gram of pellet). EDTA was added to a final concentration of 1 mM, and the cells were slowly agitated at 30°C for 10 min. Lysozyme was added to the sample (20 mg per gram of pellet) and the cells were vigorously stirred during 30 min at room temperature. The resulting suspension was centrifuged at 5000g for 20 min and the supernatant which contained the periplasmic extract was clear, with a soft pink color very pronounced in the periplasm obtained from the cells grown in the 300L Bioreactor.

The chromatographic purification protocol for the Nap protein comprised three main steps [86]. The sample was first dialyzed overnight against 10 mM MES pH 5.5 and loaded onto an S-Sepharose (GE-Amersham) column equilibrated with the same buffer. *Cn* NapAB was eluted between 200 and 250 mM NaCl, and all the fractions containing the protein were pooled and concentrated by ultrafiltration in an Amicon system. The sample was afterward injected into a gel filtration Superdex 200 column, equilibrated with 300 mM Tris-HCl pH 7.6. After buffer exchange by dialysis, the protein in 20 mM MES pH 5.5 was loaded onto a Mono-S (GE-Amersham) column and eluted with a NaCl gradient between 200 – 250 mM ionic strength. This final step was unnecessary in the 300L protein batch growth, possibly due to the higher amount of protein available. The protein was concentrated to 10 mg/ml (using the BCA procedure from Sigma) and stored in aliquots at -80°C in 10 mM Tris-HCl pH 7.6.

3.2.3 Protein Crystallization and Data Collection

Due to the initial limited amount of protein, the crystallization conditions were screened at the High Throughput Crystallization Laboratory in EMBL, Grenoble. The experiments were set up at 293K in a Cartesian PixSys 4200 crystallization robot (Genomic Solutions, UK) using the Index Screen from Hampton Research (96 conditions, Table 2-Appendix). Microcrystals were obtained under several conditions, using the sitting drop vapor diffusion method (Figure III. 1-a). The best crystals grew within one week to approximate dimensions of 0.13 x 0.09 x 0.04 mm³ using PEG 3350 as

precipitant, respectively at 25% (w/v) with 0.1 M bis-Tris propane pH 5.5; at 20%, with 0.2 M sodium formate pH 7.0, and at 15% with 0.1 M succinic acid pH 7.0 (conditions 42, 90 and 89). Equal volume of precipitant and protein solution (100 nl) were mixed and equilibrated over 90 μ l reservoir solution. Crystals were flash-cooled directly in liquid nitrogen prior to transfer to a gaseous nitrogen stream (100K) using Paratone oil as a cryoprotectant. A 1.5 \AA high-resolution data set was obtained at beamline ID14-1 of the European Synchrotron Radiation Facility (ESRF - Grenoble, France) using an ADSC Quantum-4R CCD detector and a wavelength of 0.934 \AA (NATI dataset) for crystals from condition 42. They belong to space group C2, with unit cell parameters $a = 142.2 \text{ \AA}$, $b = 82.4 \text{ \AA}$, $c = 96.8 \text{ \AA}$, $\beta = 100.7^\circ$ and one NapAB heterodimer is present per asymmetric unit.

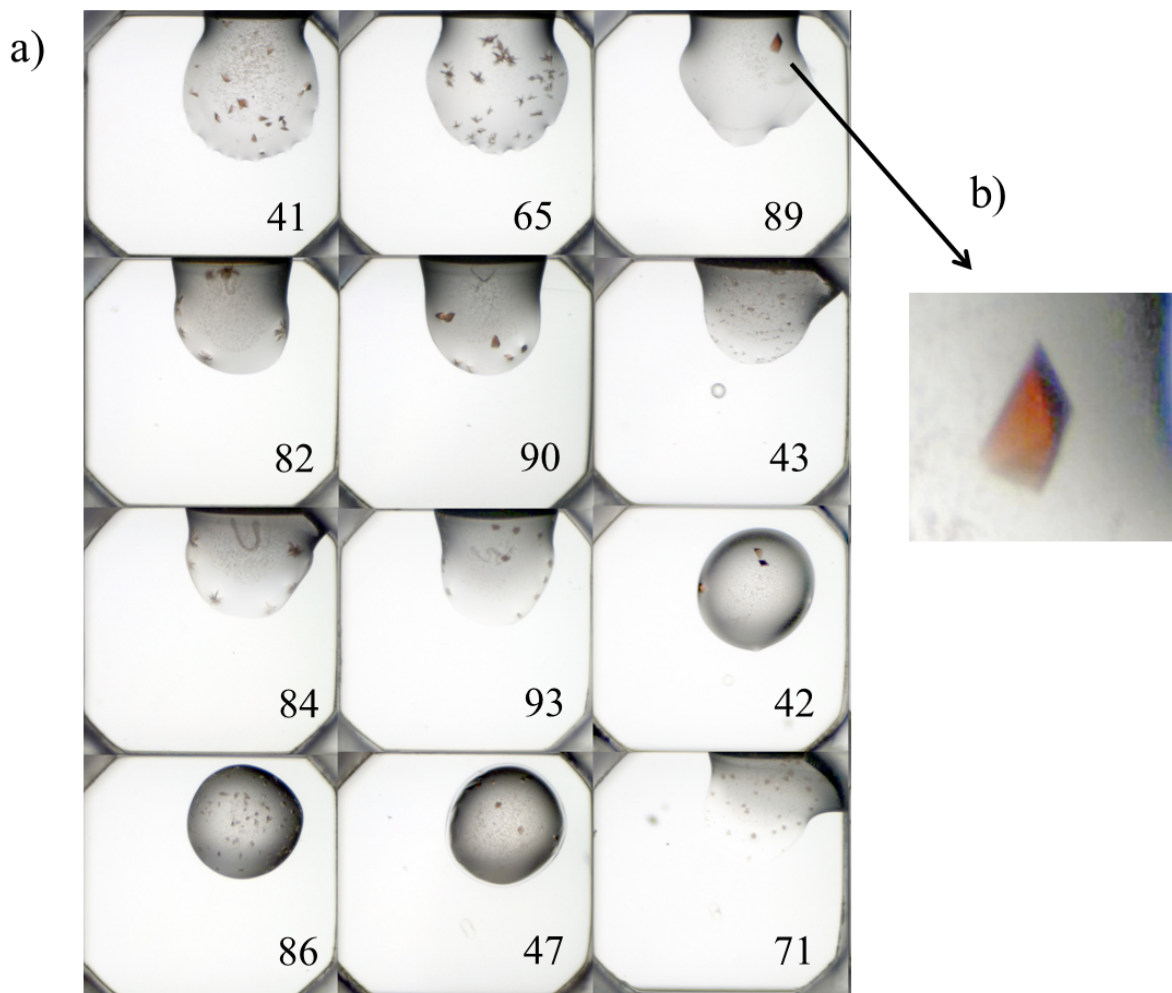


Figure III. 1 - Crystals obtained using the crystallization robot at EMBL, Grenoble. (a) Microcrystals obtained for conditions number: 41, 65, 89, 82, 90, 43, 84, 93, 42, 86, 47 and 71 from Hampton Index Screen (b) Enlargement of crystallization condition number 89; crystal dimensions 0.13 x 0.09 x 0.04 mm^3 .

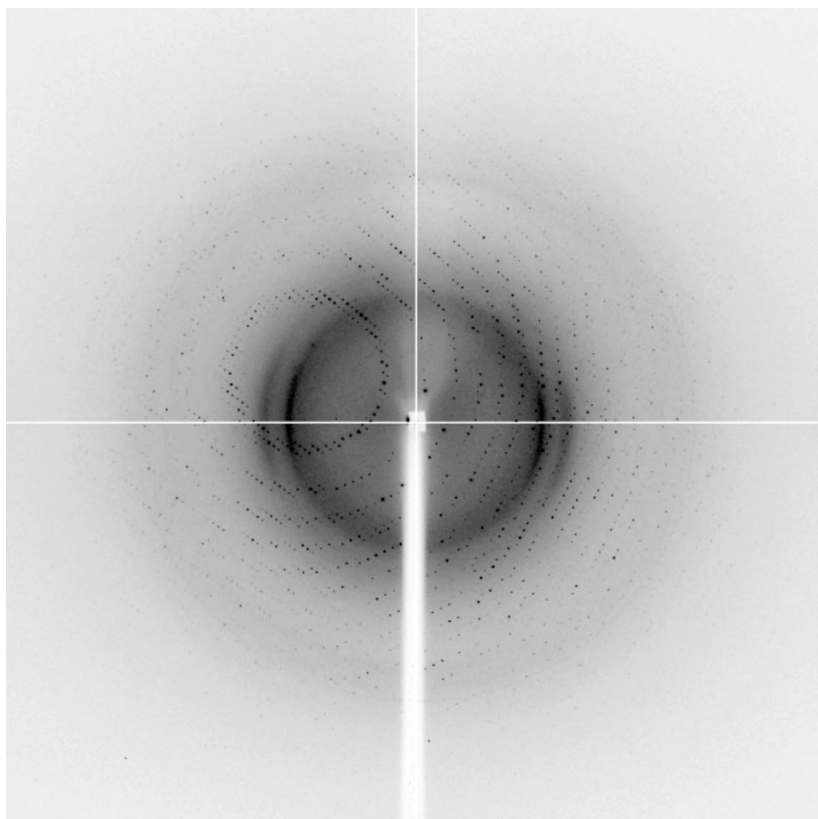


Figure III. 2 - Diffraction pattern obtained at beamline ID14-1 (ESRF) for a NapAB crystal. Resolution at the edge is 1.5Å.

Several approaches were taken in order to reproduce the robot crystallization conditions in-house, including the use of water-soluble ionic liquids (IL) as additives (as explained in Chapter 4 of this Thesis). Another effective approach for crystallizing NapAB, after the 300L bacterial growth, was the streak seeding from a new robot plate setup using 25% PEG 3350 and 0.1 M bis-Tris pH 5.5 (Hampton Research) in order to ensure reproducibility. This strategy also allowed growing suitable native crystals that diffracted beyond 1.7Å, which were used for soaking experiments. Statistics for data diffraction, processing and model refinement are presented in Table III. 1.

The maximum resolution of the data was very much dependent not only on intrinsic crystal quality but also on crystal handling and cryo-cooling. For all conditions, several crystals had to be tested until a useful data set was obtained. These difficulties may explain the isolated case of one native crystal grown with the robot that diffracted to higher resolution (1.5Å).

Crystal sample (Approximate crystal size in mm)	NATI Grown at HTX – EMBL (0.13 x 0.09 x 0.04)	DIT-IL (Grown with IL) Soak 20 mM Na dithionite (0.3 x 0.1 x 0.1)
Beamline at ESRF	ID14-1	ID14-2
Wavelength (Å)	0.934	0.933
Space group	<i>C</i> 2	<i>C</i> 2
Unit cell parameters (Å)	<i>a</i> = 142.2, <i>b</i> = 82.4, <i>c</i> = 96.8, β = 100.7°	<i>a</i> = 119.4, <i>b</i> = 71.4, <i>c</i> = 128.4, β = 121.0°
Number of molecules in the AU	1	1
Matthews coefficient (Å ³ /Da)	2.53	2.32
Solvent content (%)	51.5	46.9
Resolution limits (Å)	43.23 - 1.50	28.34 - 1.72
No. of observations	645330	363062
No. of unique observations	170014	97895
Multiplicity	3.8 (3.8)	3.7 (3.6)
Completeness (%)	97.2 (95.2)	99.7 (99.1)
<i>R</i> _{merge} (%)	8.5 (53.7)	6.8 (17.7)
<i>I</i> / σ (<i>I</i>)	8.5 (2.3)	16.8 (6.4)
Resolution (Å)	37.90 - 1.50	27.63 - 1.72
N° reflections	170003	93004
<i>R</i> _{free} (%)	19.9	18.1
<i>R</i> _{factor} (%)	16.8	14.3
Number of waters	670	787
Number of other molecules	-	19 formate ions 1 chloride ion
R.m.s. deviation from ideal geometry		
Bond lengths (Å)	0.029	0.028
Bond angles (°)	2.324	2.31
Average B factor for all atoms (Å ²)	8.54	12.38
Average B factor for waters (Å ²)	16.31	21.67
Ramachandran outliers (%)	0.67%	0.79%

Table III. 1 - Statistics on data collection and structure refinement. Values in parenthesis correspond to the highest-resolution shell. $\uparrow R_{\text{work}} = \sum ||F_{\text{calc}}| - |F_{\text{obs}}|| / \sum |F_{\text{obs}}| \times 100$, where *F*_{calc} and *F*_{obs} are the calculated and observed structure factor amplitudes, respectively. (*R*_{free} is calculated for a randomly chosen 5% of the reflections for each dataset).

3.2.4 NapAB Crystals and Soaking Experiments

In order to clarify the Nap reaction mechanism, we have tried to obtain additional structural information on different forms of the enzyme. Crystals obtained either by seeding or with IL as additives were reproducible and used for soaking experiments with a harvesting solution containing a higher PEG 3350 concentration (30%). Crystals were incubated either with 20 mM potassium nitrate, 10 and 20 mM sodium dithionite, or 20 mM potassium thiocyanate (inhibitor of nitrate reductases). Incubations with mixtures of substrate/reducing agent as well as with hydrogen peroxide were also performed. More than 20 datasets were collected at ID14-1 and 2 at the ESRF at around 2Å resolution, which after refining revealed no major differences from the native form. After extensive attempts, one unique dataset could be collected from a crystal that diffracted to 1.7Å where changes were observed in comparison to the native structure (DIT-IL). This crystal was obtained using the ionic liquid (IL) [C₄mim] Cl⁻ (Chapter 4) and soaked with 20 mM sodium dithionite (DIT-IL dataset). Data-collection and processing statistics are presented in Table III. 1.

Attempts were also made in order to crystallize *C_n* NapAB under controlled reducing conditions in an anaerobic chamber. The protein was introduced in the chamber and pre-incubated with 10 and 20 mM sodium dithionite. Crystals were obtained at 293K for the same crystallization conditions and looked very similar to the ones obtained aerobically. However, initial attempts to flash-cool the crystals in liquid nitrogen inside the chamber to maintain reducing conditions were unsuccessful. The crystals were also retrieved from the chamber and frozen outside, but were destabilized and did not diffract.

3.2.5 Processing

Diffraction data for native (NATI) and dithionite-soaked crystals (DIT-IL) were collected at the ESRF, beamlines ID 14-1 and ID 14-2, respectively. Data were processed using MOSFLM v.6.2.5 [59] and SCALA [87] from the CCP4 program package v.6.0 [88]. The calculated Matthews coefficients were 2.53 Å³Da⁻¹ and 2.32 Å³Da⁻¹, respectively, suggesting the presence of one NapAB heterodimer in the asymmetric unit for both crystal forms. The solvent content is 51.5% and 46.9% for NATI and DIT-IL data. Data-collection and processing statistics are presented in Table III. 1.

3.2.6 Structure Solution, Model Building and Refinement

The *C.necator* NapAB structure was solved by molecular replacement, using the NATI dataset and the structure of the homologous enzyme NapA from *D.desulfuricans* (PDB accession code 2NAP) as a search model, with the program Phaser [63]. The complete NapA molecule was used, excluding the two MGD cofactors, the [4Fe-4S] cluster and the water molecules. The initial phases obtained by Phaser yielded a very good initial electron density map. These were used to run automated model building in ARP/wARP [89, 90]. The small NapB was positioned by superimposing the *R.sphaeroides* NapAB model (PDB accession code 1OGY) onto the solution, and using the NapB model as a starting point for rebuilding. After rigid body refinement, the overall model presented an R_{factor} of 31.5% and an R_{free} of 33.2%. Further model building was performed using COOT, and several refinement cycles of the experimental phases, using Refmac5 from the CCP4 program package v.6.0 were also done [62, 88]. After both chains were as complete as the density allowed, all cofactors were fitted into their respective electron density. Restrained refinement with TLS, using each of the four domains from NapA subunit treated as individual TLS entities, and three domains from the NapB subunit (total of 7 TLS groups), resulted in a final model with R_{factor} and R_{free} of 16.8% and 19.9% respectively. Water molecules were automatically added with COOT.

The 6th ligand of the Mo ion was initially omitted from the electron density map and a clear positive peak was obtained in an mFo-DFc map. This density was interpreted as either a hydroxyl/water molecule or as a sulfur ligand. When refined as oxygen, the difference between the relative B factors of the 6th ligand and the other Mo ligands was significant. This could be corrected when the refinement was carried out with a terminal sulfur ligand at the sixth coordination position. The same procedure used in the case of *Dd* Nap was followed [33]. No restraints were imposed on the sixth ligand for refinement. The B factor analysis is included in Table III. 2.

The structure for the dithionite-soaked crystals (DIT-IL) was solved using the same procedure described for the native *Cn* NapAB. The native protein was used as model for molecular replacement. Water molecules were automatically added with COOT, electron density for formate molecules and one chloride ion was manually inspected and favorable contacts analyzed. The final R_{factor} and R_{free} for the DIT-IL structure were 14.3% and 18.1% respectively. No restrained refinement using TLS was used for this structure. Refinement statistics are presented in Table III. 1.

Dataset	NATI	DIT-IL	<i>Dd</i> NapA[30] (PDB: 2NAP)	<i>Ec</i> NapA* [82] (PDB: 2NYA)
Wilson B-factor (\AA^2)	11.6	11.4	39.55	-
S δ Met _{A346}	6.71	6.07	28.1	21.2/21.0
S12 MGD ₁₈₀₃	10.45	5.56	32.2	20.2/21.4
S13 MGD ₁₈₀₃	9.19	5.76	28.1	19.2/22.3
S12 MGD ₁₈₀₄	9.43	6.03	25.2	21.1/24.4
S13 MGD ₁₈₀₄	9.23	5.37	29.8	13.3/21.4
S δ Met _{A153}	7.99	7.69	30.0	20.2/25.6
S γ Cys _{A152}	9.64	4.20	33.6	20.2/25.6
Mo	9.60	5.96	29.6	20.3/24.4
S ₉₀₀	13.24	8.91	25.3	-
HOH (at S ₉₀₀ position)	2.0	2.0	7.51	5.86/2.0

*The first value is for molecule A, the second for molecule F in 2NYA

Table III. 2 - The Mo active site and B factor analysis.

After some refinement steps of the DIT-IL structure using COOT and Refmac5, unmodeled positive electron density ($>4\sigma$) was also found next to the sulfur atom of Cys784. The three clear blobs were interpreted as a modified cysteine sulfonic acid (OCS) and included in the final model. While water molecules were automatically added with COOT, electron density for formate molecules and one chloride ion was manually inspected and favorable contacts analyzed. The final R_{factor} and R_{free} for the DIT-IL structure were 14.3% and 18.1% respectively. No restrained refinement using TLS was used for this structure. The two final models included the following amino acid residues: For NATI, Val4-Val802 for the NapA subunit and Gly1-Lys124 for the NapB subunit; for DIT-IL, Lys11-Val802 for NapA and Gly1-Leu123 for NapB. In both models, the di-heme NapB subunit has two short regions with no visible electron density: residues Lys27-Met37 and Thr65-Ala72. Refinement statistics are presented in Table III. 1.

Coordinates and structure factors from both structures have been deposited in the Protein Data Bank, with the accession codes 3ML1 (NATI) and 3O5A (DIT-IL).

3.2.7 EPR Spectroscopy

In order to study the EPR properties of *C.necator* NapAB, continuous wave (CW) X-band spectra were recorded with a Bruker EMX spectrometer (model ER4116DM) equipped with a dual-mode cavity and an Oxford Instruments continuous flow cryostat. Simulations were performed using the

WIN-EPR Simfonia V1.2 software from Bruker Instruments. EPR samples were prepared in 100 mM HEPES pH 7.0, with a NapAB concentration of 200 μ M (\approx 21 mg/ml). All spectra were obtained in nonsaturating conditions: microwave power, 0.6 mW; modulation amplitude, 5 Gpp; modulation frequency, 100 kHz. Spin quantification of the Mo(V) and FeS signals were estimated by double integration and comparison with Cu-EDTA standards.

3.2.8 Spectropotentiometric Redox Titrations

UV-Vis and EPR mediated redox titrations were carried out in an anaerobic chamber at room temperature and oxygen concentration below 1 ppm. A platinum–silver/silver chloride combined electrode (Crison), calibrated with a saturated quinhydrone solution at pH 7.0 was used to determine the electrochemical potential of the solution.

For UV-Vis, a total assay volume of 3.0 ml contained 100 mM HEPES pH 7.0 buffer and final protein concentration was ca. 3 mg/ml. Mediators, at 5 μ M each, were: methyl viologen (-440 mV), neutral red (-325 mV), anthroquinone (-225 mV), fenazine (-125 mV), indigo tetrasulfonic acid (-46 mV), duroquinone (5 mV), galocyanine (30 mV), fenazine etasulfate (55 mV), fenazine metasulfate (80 mV), 2,5-dimethyl benzoquinone (180 mV), and 2,6-dichlorophenolindophenol (217 mV). Sodium dithionite and potassium ferricyanide solutions were used as reducing and oxidant agents, respectively. Optical data were collected using a TIDAS diode array spectrophotometer.

For EPR, the protein solution was incubated with the following mediator dyes, at 10 μ M: benzyl viologen (-340 mV), di-hydroxy 1,4-naphtoquinone (-145 mV), indigo disulfate (-125 mV), resorufin (-51 mV), methylene blue (11 mV), phenazine metasulfate (80 mV), 1,2-naphtoquinone (180 mV), and 1,2-naphtoquinone sulfonic acid (217 mV). The electrochemical potential was dropped using a potassium ascorbate and sodium dithionite solution, dissolved in 100 mM HEPES pH 7.0. Samples for EPR spectroscopy (100 μ l) were taken after equilibration at each potential and frozen in liquid nitrogen. Values of redox potentials are expressed relative to the standard hydrogen electrode (SHE).

3.2.9 Electrochemistry

Voltammetric measurements were performed using an Autolab PSGTAT10 potentiostat /galvanostat from ECO Chemie (Utrecht, The Netherlands). The system was controlled and data analyzed with the GPES software package from ECO Chemie. Scan rate varied between 5 and 50 mVs⁻¹. A three electrode configuration cell was utilized, with a platinum auxiliary electrode, and an Ag/AgCl reference electrode (205 mV vs standard hydrogen electrode, SHE). As working electrodes, graphite and gold disks were used. The latter was from BAS (MF-2014) with a nominal radius of 0.8 mm. As to the graphite electrodes, a homemade pyrolytic graphite (PGEh) and a commercial ultra-trace graphite (UTGEh) (Metrohm, ref. 6.1204.100) both with nominal radius of 0.1 mm, were used. Before each experiment the electrodes were polished by hand on a polishing cloth (Buehler 40-7212) using water/alumina (0.3 μm) slurry (Buehler 40-6365-006), sonicated for 1 min and then well rinsed with Milli-Q water. In most experiments with the graphite electrodes, a 2 μl drop of a mixture (1:1) of 93 μM NapAB (10 mg/ml) and 20 mM Neomycin, was deposited on the electrode surface and left to dry at room temperature for 30 min. In other experiments, a membrane electrode configuration was prepared (either graphite or gold) using a negatively charged Spectra/Por MWCO 3500 membrane.

The supporting electrolyte contained 10 mM Tris-HCl pH 7.6 and 0.1 M NaCl. In potassium nitrate assays, the final concentration in the electrolyte solution varied between 10 μM and 10 mM. In the experiments with hydrogen peroxide, final concentration varied between 50 μM and 500 μM. Oxygen was removed by the action of a nitrogen stream for 20 min. Nitrogen continued to flow on the top of the solution during all experiments in order to maintain anaerobic conditions. All measurements were performed at least in duplicate in a temperature-controlled room at 20 ± 1 °C. All potential values are referred to the standard hydrogen electrode SHE.

3.3 RESULTS AND DISCUSSION

3.3.1 Overall Structure Description

The periplasmic nitrate reductase from *Cupriavidus necator* is a heterodimer (NapAB). The large catalytic subunit (NapA) comprises 802 amino acid residues and contains the Mo ion, bound to two MGD cofactors and a [4Fe-4S] cluster. The small subunit (NapB) is 134 residues long and harbors two low-spin *c*-type hemes, with two histidines as axial ligands (Figure III. 3). One of the histidine residues is preceded by a second Cys in the canonical CXXCH binding motif. NapB is responsible for transferring electrons from the physiological partner to NapA, where nitrate is reduced to nitrite. During purification and crystallization, the two subunits remained tightly bound, providing the complete heterodimeric form of the protein. This was also reported for the heterodimeric nitrate reductase from *R.sphaeroides*. Dissociation of the two subunits was observed in the *E.coli* NapAB in which case NapA and NapB have been purified independently [26, 31, 82].

The overall topology of the *C.necator* catalytic chain is very similar to the monomeric NapA from *D.desulfuricans*, with the exception of two loops (residues 267-297 and 543-596) that protrude from the globular structure in the *Cn* NapA structure. The overall globular shape of *Cn* NapAB has approximate dimensions of 65Å x 65Å x 58Å. The catalytic core is deeply buried in the protein with a substrate channel of ~15Å leading to it. The channel is oriented opposite to the heterodimer contact interface. The distance from Mo to the nearest Fe atom of the [4Fe-4S] cluster is 12Å, with a conserved Lys (K56) mediating electron transfer between the two cofactors. *Cn* NapA is divided into four main domains classified as defined for *Dd* NapA (Figure III. 3). The catalytic subunit NapA overall fold is formed by discontinuous stretches of the polypeptide chain. The N-terminal domain I (red) is composed of three segments (residues 4-68, 507-531 and 600-641) and contains the four Cys residues (Cys19, 22, 26 and 54) that coordinate the [4Fe-4S] center. Domains II and III share an overall α/β -fold very similar to each other (represented in green and yellow respectively), and each one harbors a MGD cofactor. Domain II is also composed by three non-contiguous chains of residues (69-148, 386-506 and 532-599) while domain III comprises residues 149-385. The C-terminal domain IV (blue) presents a typical β -barrel fold and is very solvent exposed (residues 642-802) [30].

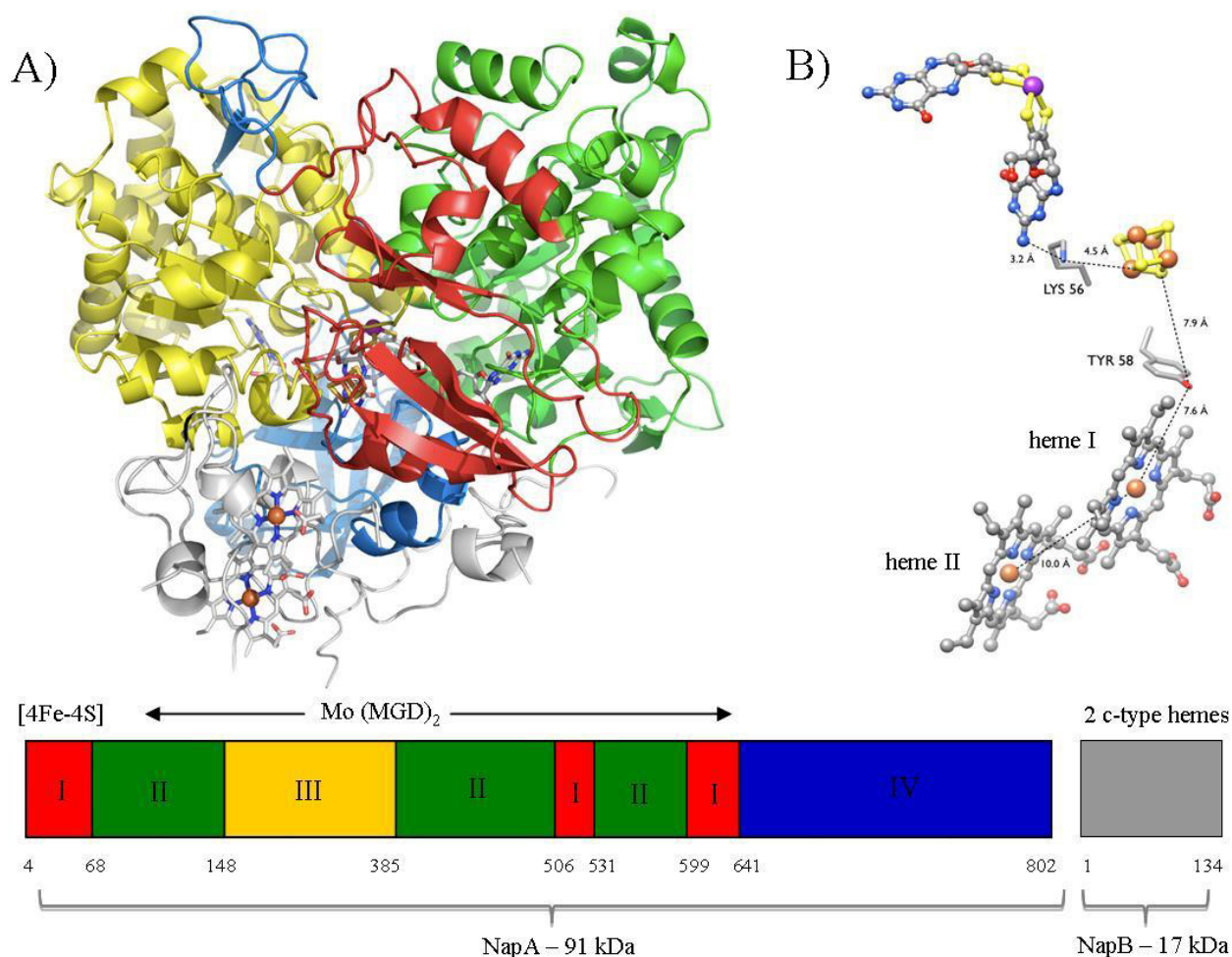


Figure III. 3 - A) Crystal structure of the *C. necator* NapAB (ribbon representation) showing the domains colored as follows: domain I in red (residues 4-68, 507-531 and 600-641); domain II in green (residues 69-148, 386-506 and 532-599); domain III in yellow (residues 149-385); and domain IV in blue (residues 642-802). The Mo-bis-MGD cofactor is shown in sticks and colored by atom type, and the [4Fe-4S] as orange sticks. NapB is represented in grey ribbon and the 2 hemes as sticks colored by atom type. B) Arrangement of the metal cofactors: Mo (MGD)₂, [4Fe-4S] and di-heme c-type cytochrome color-coded as atom types. Included are also two strictly conserved residues from NapA: Lys56A that mediates the contact between the pterin and the [4Fe-4S] cluster as well as Tyr58A, involved in intersubunit electron transfer between the [4Fe-4S] center and heme I.

The NapB subunit is characterized by a low structural complexity. The polypeptide chain has a spread-out shape with extended N- and C-terminal segments and a small globular domain in the middle (Figure III. 3 and Figure III. 4). The last 10 residues, as well as two solvent exposed regions of *Cn* NapB had no visible electron density (residues 28-36 and 66-71). The globular domain has approximate dimensions of 30Å x 20Å x 20Å and is composed of three short α -helices (Arg57-His62, Ile77-His79 and Cys98-Cys101) linked by large loops. The long N-terminus extended region (first 27 residues as random coil) in addition to the extended C-terminus (residues 103 to 124 that include a short α -helix, Val119-Leu124) resemble two arms that embrace NapA, spreading from

domain II all the way around domain IV. The stabilization of the NapAB complex is dominated by hydrogen bonds and salt bridges, involving mainly residues from both arms (Figure III. 4). The two c-type hemes in NapB are almost parallel to each other, in a stacked position with an iron-to-iron distance of 10.0Å, and at 14.7Å distance between the heme II iron atom, and the nearest iron from the [4Fe-4S] center. The side chain of Tyr58 from NapA lies between both centers at the intersubunit interface. Its hydroxyl group is 7.9Å and 7.6Å away from the nearest atom of the Fe/S cluster, and from the Fe atom of heme II, respectively (Figure III. 3). This tyrosine is conserved in all heterodimeric nitrate reductases but not in the monomeric *Dd* NapA, where is replaced by a serine residue. In the crystal structures of the heterodimeric nitrate reductases (*Cn* NapAB, *Rs* NapAB and *Ec* NapA(B)) Tyr58 adopts the same orientation, strongly suggesting its role in mediating electron transfer between NapA and NapB.

The propionate side chains from both hemes are oriented towards the solvent. The stacked arrangement of both hemes, within van der Waals contact distances, is unusual and resembles the heme arrangement first reported for the split-Soret cytochrome *c* (SSC) from *D.desulfuricans* ATCC 27774 [91]. However, there is no relationship between the fold of *Dd* SSC and NapB. The amino acid sequence of the NapB cytochrome shows two typical CXXCH motifs (Figure III. 5), with the N-terminal heme-binding site located nearly halfway along the protein sequence. The Fe atom of each heme is coordinated by two strictly conserved axial histidine residues (His44B and His62B for heme I, and His79B and His102B for heme II) (Figure III. 5). No significant sequence homology to other c-type cytochromes was found. Based on this, and on its particular structural features, NapB can be considered as representing an unusual class of c-type cytochromes with short Fe-Fe distance and parallel arrangement of the hemes located on one side of the protein. The structural features of c-type cytochrome NapB as well as its orientation in relation to NapA may favor intermolecular electron transfer between NapA, NapB and NapC.

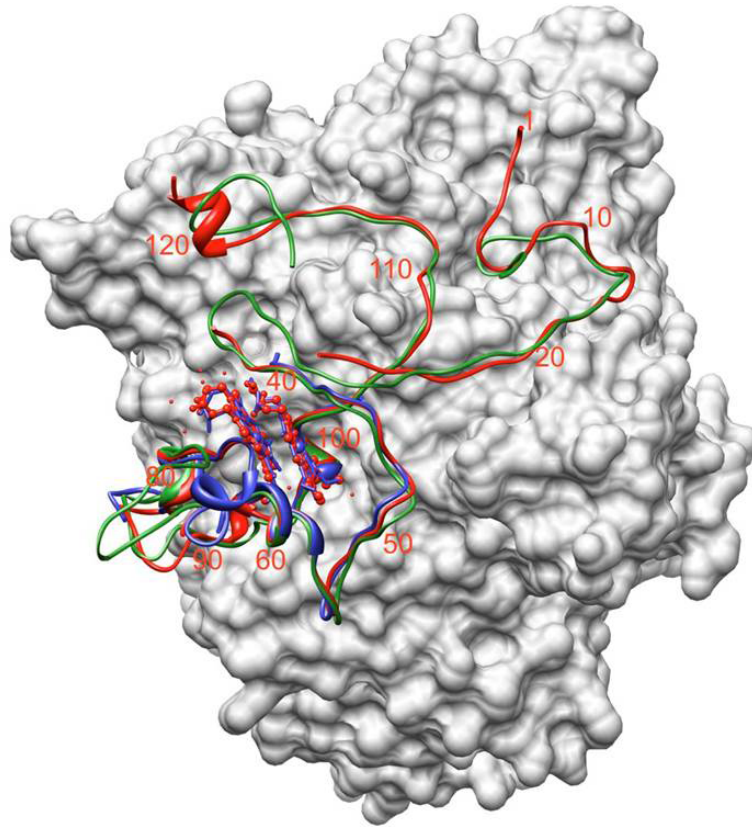
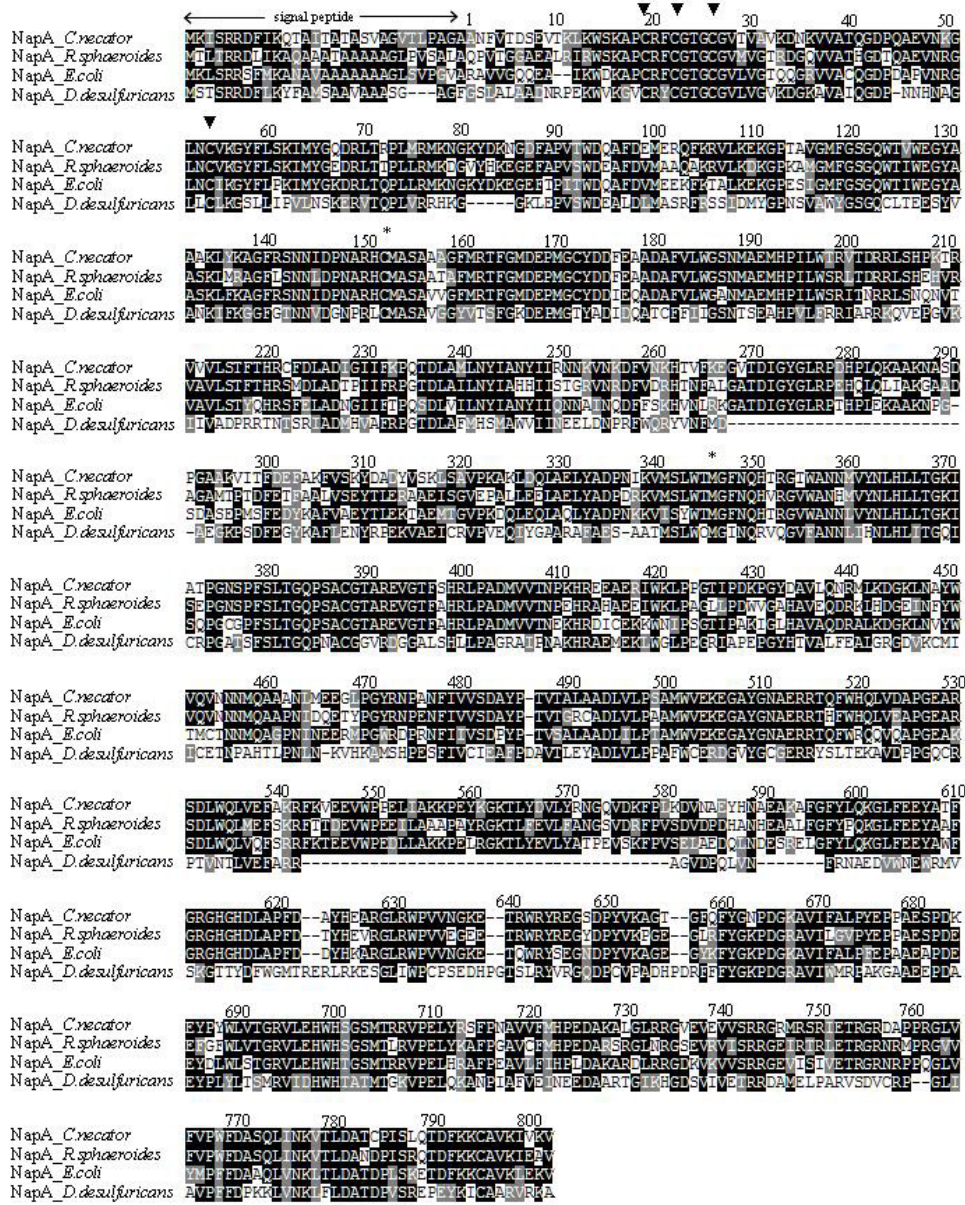


Figure III. 4 - Superposition of the NapB subunit from *C.necator* (red) with NapB from *R.sphaeroides* (green) and the segment of residues 38 to 105 from *H.influenzae* (blue). The *Cn* NapA is represented as a gray surface.

3.3.2 The NapAB Heterodimer

In *C.necator*, the catalytic subunit NapA is tightly bound to the subunit NapB, with a total buried surface area of 6005\AA^2 . This interaction involves several hydrogen bonds and salt bridges. The majority of these bonds are established by residues mostly from NapA domains I, III and IV, that interact with residues from NapB. From a total of 24 hydrogen bonds ($< 3.2\text{\AA}$) and 4 salt bridges between NapB and NapA residues, only 5 involve residues from the core domain of NapB. The other 19 involve residues from the extended N-arm (Asp4B, Met6B, Arg7B, Asn14B, Glu15B, Ala18B, Pro19B and Glu25B) and C-arm (Gln105B, Ala106B, Thr108B, Gly113B, Asn114B and Asn115B) of NapB. The 4 salt bridges that mediate the complex interface involve residues Asp4B, Glu15B, Glu25B and Arg83B. The majority of the conserved residues in the NapB proteins are located in the proximity of the two heme groups in the small globular domain (residues 27 to 103) (Figure III. 5). The less conserved regions correspond to the two long arms and the differences may explain why in some organisms, such as *E.coli* or *H.influenzae*, dissociation between NapA and NapB occurs.

A)



B)

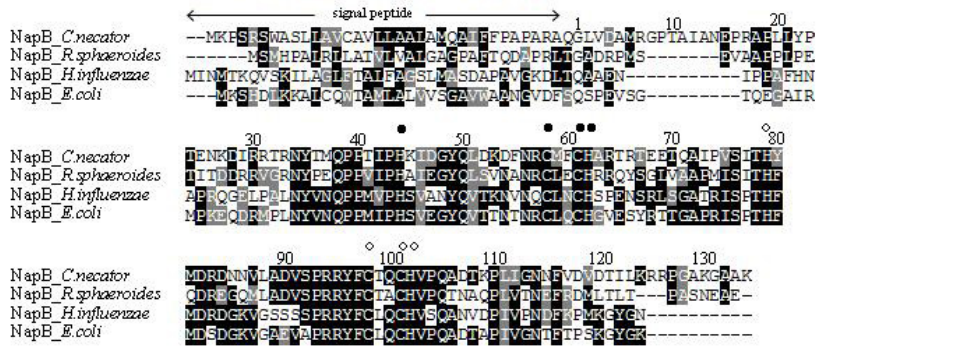


Figure III. 5 - a) Comparison of the amino acid sequence of NapA from *C.necator* with the other nitrate reductase catalytic subunits of known structures (*R.sphaeroides*, *E.coli* and *D.desulfuricans*). The [4Fe-4S] binding Cys residues are marked as (▼) and the Mo ligand Cys152 as (*). b) Comparison of the amino acid sequence homology of NapB from *C.necator*, *R.sphaeroides*, *H.influenzae* and *E.coli*. Binding residues (H---CXXCH) for heme I are marked as (●) and for heme II as (○). Black boxes indicate identical residues, and grey indicate similar ones. Alignment results obtained with CLUSTALW multiple sequence alignment program at the EMBL-EBI web server, and the figure prepared using BOXSHADE from EMBnet server.

3.3.3 Comparison with Homologous Crystal Structures

Several crystal structures of periplasmic nitrate reductases have been reported in the past few years. As mentioned before, the first was the monomeric NapA from *D.desulfuricans* at 1.9Å [30], while the NapAB from *R.sphaeroides* was the first structure of a heterodimeric Nap protein [31]. Crystal structures of the individual subunit NapA from *E.coli* [82], as well as of a proteolysed fragment of *H.influenzae* NapB were also reported, the later corresponding to a globular domain of *Hi* NapB [81] obtained by proteolysis of recombinant protein (a 8.5 kDa fragment, instead of the complete 15 kDa subunit).

Comparing the two heterodimeric NapAB crystal structures (*C.necator* and *R.sphaeroides*) we can see that both proteins share a similar overall globular shape, folding and cofactor localization. The buried surface area as well as the number and type of intersubunit contacts found in both complexes are also very similar. The *Cn* NapA and *Rs* NapA subunits share 71% sequence identity and the same number of amino acid residues (802). Both structures superimpose with an rms deviation of 0.67Å for 780 $C\alpha$ atoms. The monomeric NapA from *D.desulfuricans* shares only 36% sequence identity with *Cn* NapA but the two structures are also quite similar (rmsd of 0.9Å over 590 $C\alpha$ superimposed). Structural comparisons with *E.coli* NapA also revealed no major differences. Both share 69% sequence identity with an rmsd of 0.51Å for 781 superimposed $C\alpha$ positions. The main differences among all four NapA structures are the two exposed loops conserved in *Cn*, *Rs* and *Ec* NapA but absent in the *Dd* NapA structure (residues 267-297 and 543-596, *Cn* NapA numbering).

In contrast to the catalytic subunit *Cn* NapA, the electron transfer subunit *Cn* NapB has a lower sequence identity (52%) with *R.sphaeroides* NapB and the two structures superimpose with an rmsd of 0.96Å (86 superimposed out of 134 $C\alpha$ atoms). The structure of the *Hi* NapB proteolysed fragment superimposes onto *Cn* NapB with an rmsd of 0.78Å for 49 $C\alpha$ atoms. These two proteins share only 37% sequence identity. The core domain of the three NapB structures is very similar with the exception of a short α -helix present in the *Hi* NapB (residues 65-72) (Figure III. 4). This corresponds to an exposed polypeptide that appears disordered in the *Cn* NapB structure (residues Arg66- Gln71) showing no continuous electron density. In *Rs* NapB this same region is unstructured and protrudes towards the solvent.

Comparing the electrostatic surface potential for the different enzymes, we can see that around the solvent exposed funnel (Figure III. 6-A) *Cn* NapAB is markedly positively charged, and similar to the surfaces of NapA from *E.coli* and from *D.desulfuricans*. In the case of the *R. sphaeroides* protein, the electrostatic surface potential is clearly more anionic. The similarity between the *Cn* NapA, *Ec* NapA and *Dd* NapA funnel surfaces may be explained by common evolution of the proteins in the different bacteria, possibly due to evolutionary conditions of adaptation, while the *Rs* NapA could have a distinct evolution. The electrostatic contact surface between NapA and NapB (Figure III. 6-B) is quite similar in both *Cn* and *Rs* heterodimeric proteins. The corresponding surface on *Ec* NapA at the NapB binding interface is apparently more positive than for *Cn* and *Ec* NapA. In spite of the observed differences, there is no clear pattern and it is hard to draw conclusions simply on the basis of the surface electrostatic potentials, regarding the stability of the NapAB complex, namely the lower dimerization affinity in *Ec* NapAB.

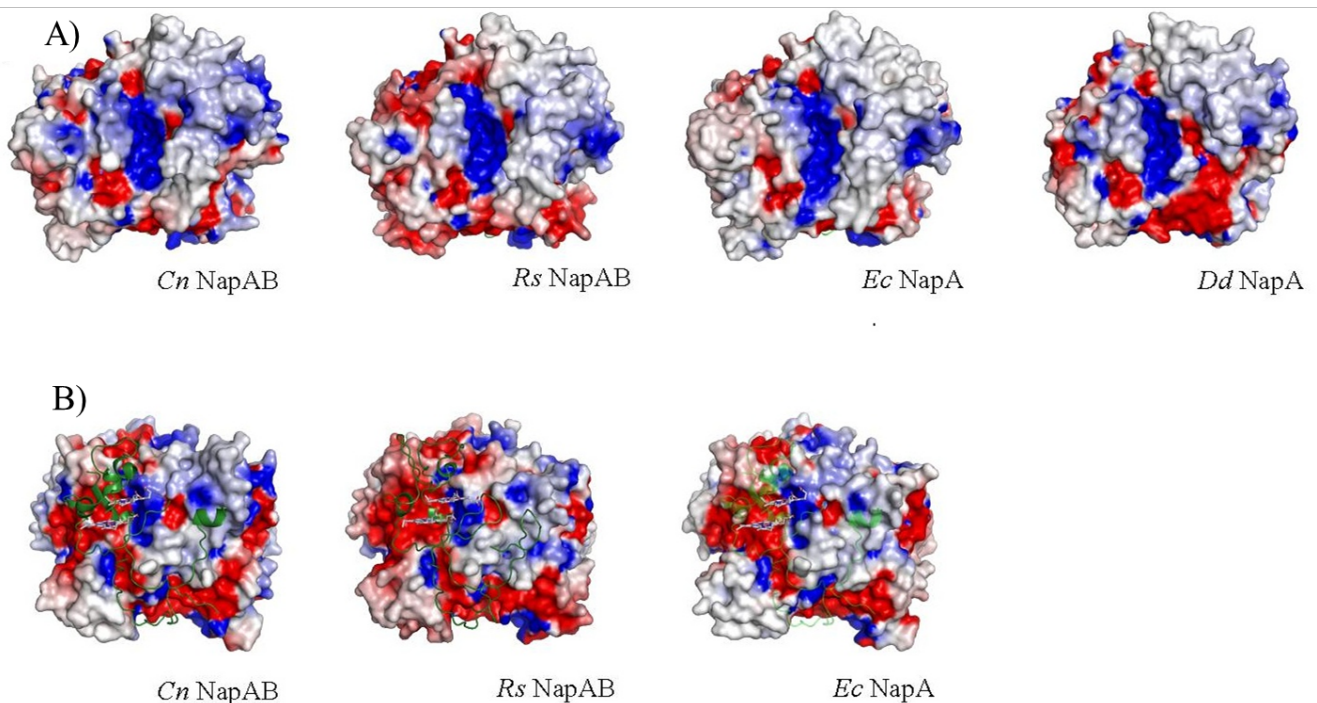


Figure III. 6 - Electrostatic potentials of Nap structures surfaces. Electrostatic surface potentials were calculated using the program delphi [71] and represented in PyMOL [72], with the color of surface potentials in the scale range from -0.10 V (negatively charged, in red) to +0.10 V (positively charged, in blue). A) View towards the substrate tunnel leading to the active site of *Cn* NapAB, *Rs* NapAB, *Dd* NapA and *Ec* NapA; B) NapA surface interaction with the small subunit NapB, from *Cn*, *Rs* and *Ec* (NapB from *Cn*), all in the same orientation.

Amino acid sequence alignment for the four NapB subunits (Figure III. 5-B) shows that the core globular domain sequence that binds the two hemes (residues 34 to 106), is highly conserved. In contrast, the N- and C-terminal arms differ in their primary structures, although these are the regions of NapB that mostly contribute to the complex stabilization. Most of the residues from NapA that form the dimer interface are from segments 47A to 67A and 706A to 777A, and are involved in hydrogen bonding interactions (direct or water mediated) and salt bridges. It is interesting to note that from these interface residues, some are conserved only in the proteins that purify as dimers (*Cn* and *Rs*) and not in *Ec* and *Dd* NapA: Glu47, Tyr712, Glu737, Arg748, Arg750, Arg750 and Ser772 (all from subunit A). In a recent study of NapA from *Shewanella gelidimarina*, the authors proposed that only two of those residues (Glu47 and Ser772 using *Cn* numbering) should be crucial for the heterodimer stabilization [92]. However, when analyzing the structure interfaces of *Cn* NapAB and of *Rs* NapAB it appears more likely that the stability of the NapAB heterodimer, and therefore the preservation of its integrity upon purification, is due to a combination of multiple intersubunit interactions. Analyzing for the presence of some of these residues might help to predict whether a certain Nap protein is expected to isolate as a NapA monomer or as a NapAB heterodimer.

3.3.4 The Molybdenum Catalytic Site

The NapA active site is accessible from the solvent area, through a funnel-shaped cavity defined by residues from domains II and III (Figure III. 7). The entrance from the wider part of the funnel is coated with highly conserved and charged residues (Arg392, 400 and 150, Asp167 and Glu168). Arg392 and Asp167 form a salt bridge and are directed towards the active site, while Glu168 is a Ramachandran outlier in all Nap structures. Asp167 and Glu168 are located in an exposed loop at the mouth of the funnel. Mutation of Asp167 to Ala (in single and double mutants, D167A and D167A-E168G, respectively) leads to total loss of nitrate reductase activity [93]. These observations suggest that the positioning of those conserved residues and the charged patch are essential for anchoring and orienting the substrate towards the Mo active site or are important for protein maturation and/or cofactor insertion [6, 94]. The NapA funnel cavity is lined with several ordered water molecules. In the DIT-IL structure, where formate is present in the crystallization buffer, three formate ions were found at mid height of the funnel. A total of 19 formate ions were modeled, with all other present on the solvent exposed surface of the heterodimer. The formate molecule (Fmt1) closest to the active site (13Å from the Mo) contacts the guanidinium group of Arg400. Fmt5 contacts Arg150, and Fmt7 is stabilized through H-bonding interactions with Asp167, Glu168 and His616 (Figure III. 7).

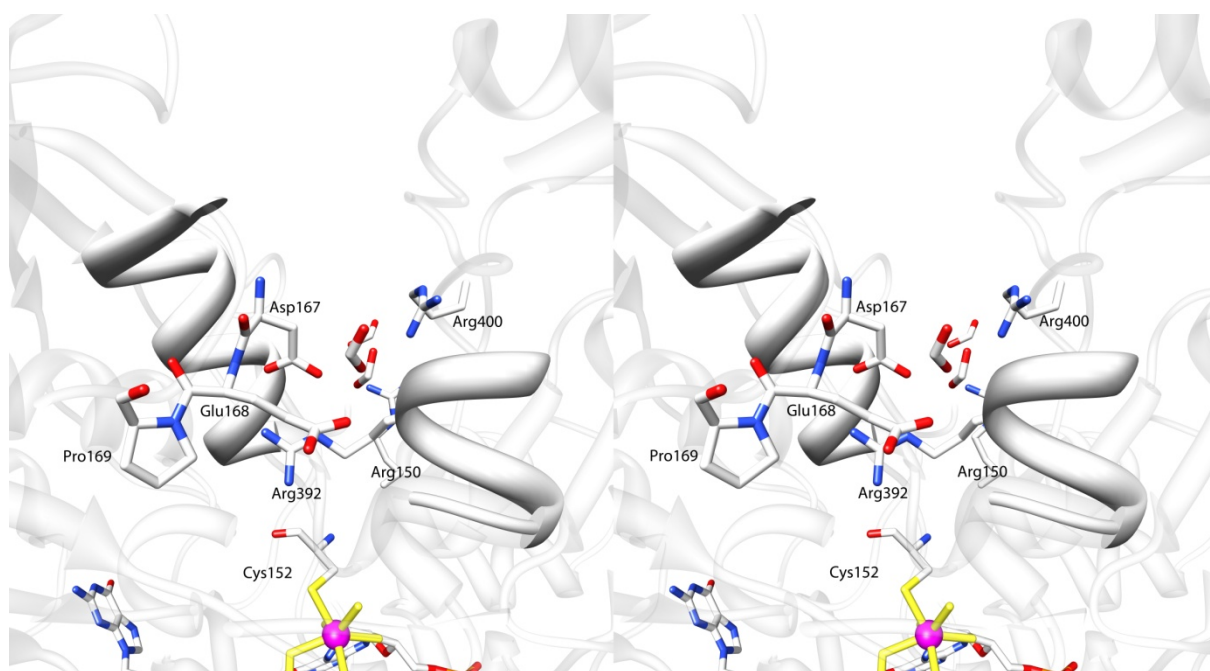


Figure III. 7 – Stereo view of the substrate funnel with 3 formate ions and conserved charged residues involved in orienting the substrate into the Mo active site (DIT-IL structure).

The molybdenum active site is found at the bottom of the solvent accessible funnel and in the native oxidized form, the Mo atom is coordinated by six sulfur ligands: four provided by the two dithiolene ligands (MGD1803 and MGD1804), one from the S γ atom of Cys152, and a sixth terminal sulfur ligand. Although in the original *Dd* NapA structure the sixth ligand was assigned as an OH ligand [30]), recent studies on the same enzyme reacted with reducing agents, substrates and inhibitors, provided new evidence on molybdenum active site details, allowing to assess the sixth Mo ligand as a sulfur atom [33]. The structure of the NapAB from *C.necator* here described constitutes the first true atomic resolution structure of a nitrate reductase. A similar analysis for *Cn* NapAB with 1.5Å data (as performed for *Dd* NapA, [33]) confirmed the true nature of the molybdenum coordination sphere (Table III. 2). When refining the 6th Mo ligand as oxygen, positive electron density was obtained on a difference density map, which completely disappeared when assigning it a sulfur atom (Figure III. 8-A and Figure III. 8-B). Besides, the B-factor of an oxygen atom at this position is lower than 2 Å², well below the average B-factor of the surrounding atoms (~9Å²). When considered as a sulfur atom, its B-factor refines to 11Å² (Table III. 2). As in the *Dd* NapA structure, the two sulfur atoms form a partial disulfide bond (S--S, 2.55Å) (Figure III. 8-C). A unique coordination sphere of six sulfur ligands bound to molybdenum suggests that this is most likely a general feature for all periplasmic nitrate reductases.

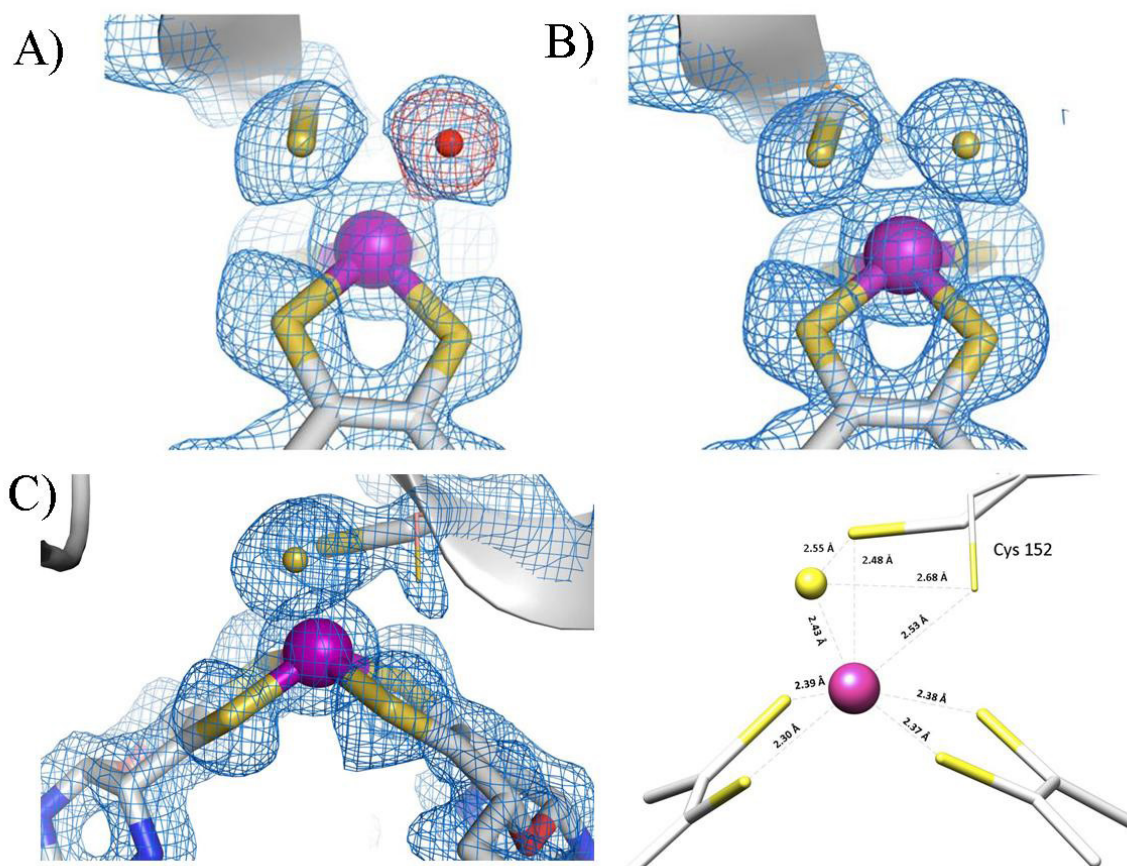


Figure III. 8 - Close-up of the Mo coordination sphere for the native (A and B) and dithionite-reduced forms (C). The $2mF_o-DF_c$ electron density maps are contoured at 1σ (blue) and the mF_o-DF_c map at 3σ (red). A) Refining the sixth Mo ligand as an oxygen atom yielded positive electron density on an mF_o-DF_c map. This density completely disappeared when assigning it as a sulfur atom (B). C) In the partially dithionite-reduced form (DIT-IL) there is extra positive density at the C^β atom of Cys152 that could be modeled as two alternate conformations with 88% and 12% occupancies.

3.3.5 The Catalytic Site in Dithionite Reacted Crystals

To further clarify Nap reaction mechanism, it was important to have additional structural information on the reduced form of the enzyme. Several strategies to prepare crystals of the reduced form of *Cn* NapAB by soaking reducing agents into previously obtained crystals were used. However, the majority of the reacted crystals was damaged and did not diffract at all or only to poor resolution. On the other hand, some dithionite treated crystals that did diffract revealed no significant changes at the active site. After extensive attempts, one unique dataset could finally be collected from a crystal that diffracted to 1.7\AA . This crystal was obtained using the ionic liquid (IL) $[\text{C}_4\text{mim}]\text{Cl}$, and soaked with 20 mM sodium dithionite (DIT-IL). Data analysis revealed interesting changes in the vicinity of the molybdenum active site. Of particular relevance was the presence of extra positive electron density close to the C^β of the coordinating Cys152, which suggested a possible movement of the corresponding side chain. Towards the end of the refinement, these modifications at the active site

were interpreted as two alternate conformations of Cys152, with 0.88 and 0.12 occupancies. In the final DIT-IL model all atoms for both conformations had B-factors within the same range as the surrounding atoms. Figure III. 8-C represents the final coordination geometry for the two conformers. However, due to the low occupancy (12%) of the displaced Cys conformation, one should be cautious when analyzing the corresponding bond distances. At this level of occupancy the uncertainty of the atomic positions is too large and one cannot claim if this conformation corresponds to an unbound form or not, although the shape of the electron density may suggest it.

3.3.6 Reaction Mechanism

The new coordination sphere of molybdenum, with a total of six sulfur ligands and no oxygen directly bound to the metal, was first revealed by the reanalysis of the *Dd* NapA crystal structure and confirmed with the *Cn* NapAB structure here described. In both active site structures, the sulfur atom from the coordinating Cys is making a partial disulfide bond with the terminal sulfido ligand (Figure III. 8-C). The new nitrate reductase active site definition required a revision of the reaction mechanism. Three alternatives were proposed, considering that substrate binding and reaction pathway must combine molybdenum and sulfur redox chemistry [33]. Two of the possible mechanisms involve direct binding of nitrate to the Mo atom (first coordination sphere) while the third one involves binding of nitrate to the terminal sulfur ligand, in the second coordination sphere of the metal. Subsequently, theoretical and computational studies were performed by different authors in order to investigate which of the alternative paths was energetically more favored [83, 84, 105]. Interestingly they all came to the same conclusion that the most viable mechanism should involve a direct binding of nitrate to the Mo atom. In order to open free access to the metal center, the Cys ligand must disconnect from Mo, but remain bound to the sulfur ligand (fragment Mo-S-S γ Cys).

The alternate conformation of Cys152 found in the *Cn* NapAB structure of the partially reduced form here reported, reflects the redox interplay between Mo and the sulfur ligands, supporting a reaction mechanism, which involves rearrangement of the Cys152 coordinating position.

3.3.7 The Presence of Oxidized Cysteines

As previously mentioned, during final refinement steps of the DIT-IL data, three positive electron density blobs were clearly visible ($> 4\sigma$), near the sulfur atom of Cys784, close to the substrate funnel. This density was modeled as three oxygen atoms of cysteine sulfonic acid ($-\text{CH}_2\text{SO}_3\text{H}$). After refining this modified Cys784, *B*-factors for the three oxygen atoms were within the range of the surrounding atoms. In addition, Cys388 also revealed extra positive electron density but not as clear as for Cys784, and therefore it was left unmodified in the deposited model. The presence of oxidized cysteine residues was totally unexpected and was observed only in the data from *Cn* NapAB crystals treated with sodium dithionite (DIT-IL). The oxidation of cysteine residues has been reported in the literature as the result of peroxidation, and has been also related with the important role of cysteines as redox-sensitive regulatory switches [94, 98, 99]. As the soak experiments were performed in the presence of O_2 , there is a possibility that the free-thiol group of cysteines had reacted with hydrogen peroxide resulting from the reaction reduction of O_2 by dithionite (according to the reaction: $\text{Na}_2\text{S}_2\text{O}_4 + \text{O}_2 + 2\text{H}_2\text{O} \rightarrow 2\text{NaHSO}_3 + \text{H}_2\text{O}_2$). Although dithionite can also react with the formed peroxide, this reaction is much slower and H_2O_2 may accumulate in solution [100]. In order to test this hypothesis, native *Cn* NapAB crystals were treated with a hydrogen peroxide solution and datasets were collected (data not deposited). As expected, similar results were obtained for this crystals and clear electron density revealed the cysteine Cys784 oxidation (and of Cys388 to lower extent). It is important to note that in these peroxide soaked crystals no changes were observed in the active site, in particular at the Mo ligand Cys152. Both cysteine residues are solvent accessible, and located at (Cys388) or close (Cys784) to the substrate channel, although only Cys388 is strictly conserved among periplasmic nitrate reductases. The presence of oxidized cysteines in *Cn* NapAB is most probably irrelevant from the functionality point of view. Nevertheless, care must be taken when interpreting structural data where dithionite has been used as a reducing agent, as is often the case in crystal structures of reduced forms of metalloproteins [101, 102].

3.3.8 Spectroscopic Characterization

The UV-visible spectrum of *Cn* NapAB exhibits typical features of a heme protein, with a Soret band centered at 409 nm ($\epsilon = 200,000 \text{ M}^{-1}\text{cm}^{-1}$) and a protein band at 275 nm ($\epsilon = 240,000 \text{ M}^{-1}\text{cm}^{-1}$). The characteristic absorption features of the iron-sulfur cluster, usually found around 350-500 nm, are covered by the heme absorption bands and cannot be observed. Upon ascorbate or dithionite addition in anaerobic atmosphere, the Soret band shifts to 419 nm ($\epsilon = 280,000 \text{ M}^{-1}\text{cm}^{-1}$) and α and β bands

can be observed at 550 nm ($\epsilon = 66,000 \text{ M}^{-1}\text{cm}^{-1}$) and 522 nm ($\epsilon = 41,000 \text{ M}^{-1}\text{cm}^{-1}$), respectively. The latter features are characteristic of *c*-type hemes with ferrous ion in a low electronic spin configuration. A UV-visible mediated redox titration in the +300 to -400 mV range, revealed that changes in the 550 nm absorption band can be observed in the range of +200 to 0 mV. Similar to what was observed for other periplasmic nitrate reductases, these changes are bi-phasic and can be fitted with a Nernstian curve of two independent hemes with $n=1$ and E_m values of +50 mV and +160 mV (Figure III. 9). As the *Cn* NapAB complex is very tight, it was impossible to dissociate both subunits to study their redox properties separately, as performed in Nap from *R.sphaeroides* and from *E.coli* [31, 82].

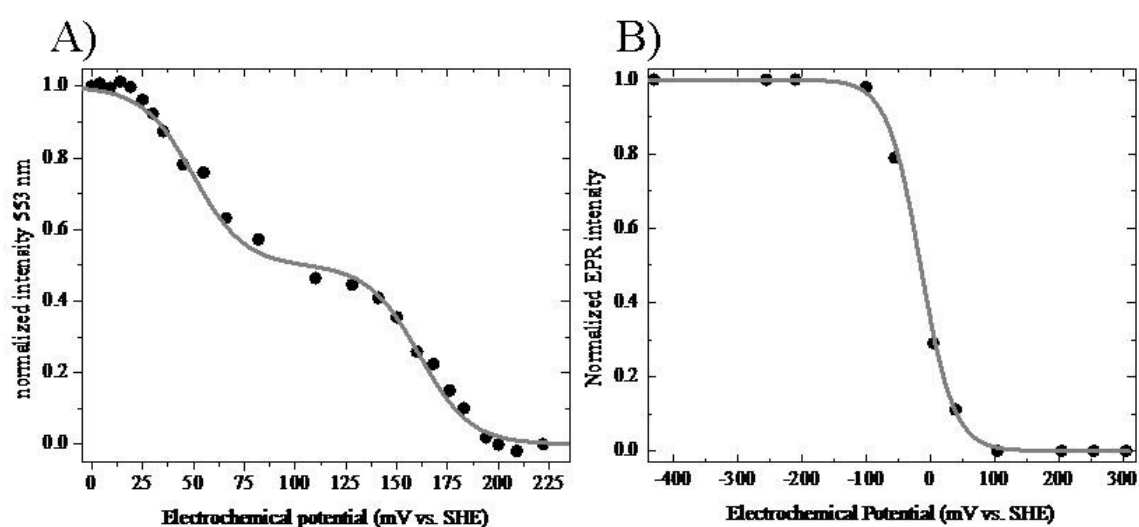


Figure III. 9 - A) UV-Vis mediated potentiometric redox titration of the *c*-type hemes present in *C.necator* periplasmic nitrate reductase. The absorbance at 553 nm was normalized and plotted *versus* the redox potential of the solution E (at pH 7.00). Experimental data (circles) were fitted with a nonlinear regression (grey line) using the Nernst equation for two independent components with E_m values of 160 ± 5 mV and 50 ± 2 mV (errors from oxidative and reductive titrations). B) EPR spectropotentiometric redox titration of the [4Fe-4S] cluster present in the large subunit of *Cn* NapAB. The area of the FeS EPR signal was normalized and plotted *versus* the redox potential of the solution E (at pH 7.00). Experimental data (circles) were fitted with a nonlinear regression (gray line) using the Nernst equation for one electron with an E_m value of -15 mV.

The midpoint redox potentials of the hemes of the *Cn* NapAB heterodimer, and those of other Nap proteins are reported in Table III. 3. The values obtained for *Cn* NapAB are higher (mainly for heme I) and in the positive limit of the redox potential range expected for *c*-type cytochromes with a bis-histidinyll coordinated heme (-350 mV to +150mV) [104]. At present, there is no established correlation between the redox potential of a heme and the factors that tune such value, and no unusual features can be observed in the crystal structure. The continuous-wave X-band electron

paramagnetic resonance (EPR) spectrum of air-oxidized *Cn* NapAB recorded at 15 K (Figure III. 10) contains a low spin ($S = \frac{1}{2}$) rhombic species with g -values of 2.935 (90 G), 2.275 (100 G) and 1.487 (400 G) (line widths in parentheses). A feature at $g \sim 2$ arising from a fraction (0.02-0.05 spin/molecule) of the Mo ion present in the paramagnetic Mo (V) state ($S = \frac{1}{2}$) is also visible. The spin integration of the rhombic signal yields near 2 spin per NapAB heterodimer, indicating that both hemes contribute to this EPR signal. The g -values used for spectrum simulation yield crystal field parameters $\Delta/\lambda = 2.93$ and $V/\Delta = 0.63$, which are characteristic of low spin hemes with bis-His coordination, in which the planes of the imidazole rings from the histidines are nearly parallel [103]. This is in agreement with the X-ray structure obtained for the oxidized crystal form.

Sample	heme I	heme II	[4Fe-4S]	Reference
<i>Cn</i> NapAB	+50	+160	-15	This work
<i>Ec</i> NapAB (NapA)	-160	+10	-250 (+20)	Jepson <i>et al</i> , 2007 [82]
<i>Rs</i> NapAB (NapA or B)	-210	-65 (-110)	-70 (-250)	Arnoux <i>et al</i> , 2003 [31]
<i>Pp</i> NapAB	-15	+80	+160	Berks <i>et al</i> , 1995 Breton <i>et al</i> , 1994
Dd NapA	-	-	-390	González <i>et al</i> , 2006 [26]
<i>Se</i> NarB	-	-	-190	Jepson <i>et al</i> , 2004
<i>Hi</i> NapB	-170	-25	-	Brigé <i>et al</i> , 2002 [81]

Table III. 3 - Redox potential of the metal cofactors comprising the electron pathway connecting NapC to the Mo ion at the active site of periplasmic nitrate reductases. (Values between parentheses correspond to redox potential obtained in the titration performed on the monomer NapA or NapB. Values are given in mV with respect to the standard hydrogen electrode). *Se*, *Synechococcus elongates* and *Pp*, *Paracoccus pantotrophus*.

When an as-prepared sample of *Cn* NapAB is incubated with 5 mM sodium dithionite anaerobically, the EPR signature of the hemes disappears, consistent with the reduction of the low spin ferric porphyrin iron to the ferrous state. Much of the Mo site is not EPR detectable (the intensity of the EPR signal at $g \sim 2$, assigned to a minor Mo(V) species accounts to less than 0.01 spin per molecule). In this reduced state, a characteristic signal from $[4Fe-4S]^{+1}$ ($S = 1/2$) cluster develops (Figure III. 10 - A). This paramagnetic species recorded at 25K yields g -values $g_{max} = 2.047$ (10), $g_{mid} = 1.951$ (8) and $g_{min} = 1.906-1.900$ (10) (line shape parameter Lorentzian/Gaussian = 0.75). At low temperatures, the g_{min} region presents a complex structure that vanishes completely at 50K (inset, Figure III. 10 - B),

yielding the g -values $g_{\max} = 2.047$ (16), $g_{\text{mid}} = 1.951$ (20) and $g_{\min} = 1.904$ (20) (line shape parameter Lorentzian/Gaussian= 0.10). The behavior of the g_{\min} feature with temperature is difficult to interpret at this moment. For instance, the Mo-containing aldehyde oxidoreductase from *Desulfovibrio gigas* (MOP) presents a split of the g_{\max} feature for one of the [2Fe-2S] clusters and in the three g -values of the Mo (V) species due to the magnetic coupling between the three paramagnetic ($S= 1/2$) redox cofactors present in the enzyme [82, 106, 107]. However, this explanation is difficult to apply in the present case since in the dithionite-reduced state, the only paramagnetic species present in the enzyme is the $[4\text{Fe-4S}]^{+1}$ cluster.

Another interesting aspect of *Cn* NapAB [4Fe-4S] cluster is that it can be partially reduced (deduced from spin quantification) by incubation with an excess of potassium ascorbate (20 mM). The FeS cluster EPR signal is identical to the one obtained with sodium dithionite. An EPR-mediated potentiometric titration of this paramagnetic species indicates that the redox potential of the metal cluster is -15 mV, which is relatively high for the $[4\text{Fe-4S}]^{2+/1+}$ couple in a periplasmic nitrate reductase. A relatively high redox potential for this couple was also observed in the FeS cluster of the monomeric NapA from *E.coli* (+20 mV), but this potential turns negative (-250 mV) when the NapAB heterodimer is formed (Table III. 3). Jepson *et al.* proposed that the reduction potentials of the FeS clusters can be affected by the electrostatic surface potential, where a more negative surface potential would make the FeS cluster more difficult to reduce, i.e. have lower redox potentials [82].

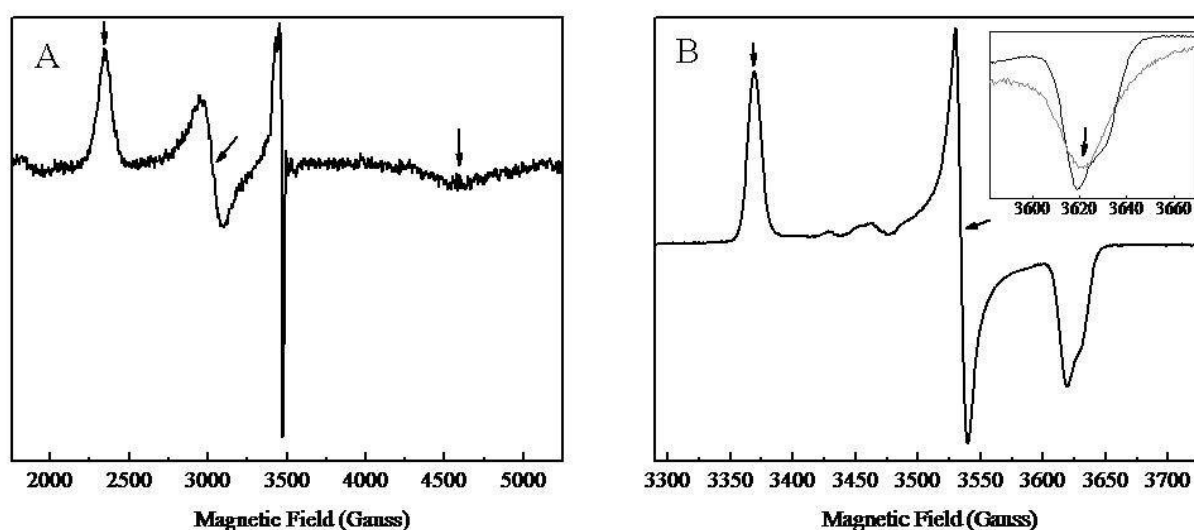


Figure III. 10 - - Electron paramagnetic resonance (EPR) spectra obtained for the periplasmic *Cupriavidus necator* nitrate reductase heterodimer NapAB. A) Sample as-prepared under aerobic atmosphere measured at 15 K. B) Sample as-prepared reduced in anaerobic conditions (dioxygen concentration below 1 ppm) with 5 mM sodium dithionite measured at 25K. Inset of panel B shows the reduced sample measured at 25K (black) and 50K (gray). Arrows indicate the position of the main g -values for the heme and $[4\text{Fe-4S}]^{+1}$ center

3.3.9 Electrochemical Characterization

Non-turnover conditions

The *C.necator* NapAB electrochemical behavior was investigated by protein film voltammetry, PFV, using graphite electrodes [108]. The heterodimeric periplasmic nitrate reductases from *P.pantotrophus* and *R.sphaeroides* form very stable films when adsorbed onto graphite electrodes [109], and this was experienced for the case of NapAB from *C.necator*. Cyclic voltammograms (CV) at the PGEh electrode in the potential range +160 mV to -600 mV vs SHE and scan rates (v) between 5 - 50 mVs^{-1} revealed the presence of one cathodic peak, with an anodic counterpart (Figure III. 11). From the difference $E_p - E_{p/2} = (65-90)$ mV, it is reasonable to consider that $n=1$. Although the separation between the cathodic, E_p^c , and the anodic, E_p^a , peak potentials ($\Delta E_p = E_p^c - E_p^a$) increased with v , the average $(E_p^c + E_p^a)/2$ remained almost constant for all scan rates. Therefore, a formal reduction potential $E^{0'} = (E_p^c + E_p^a)/2 = -161$ mV vs SHE could be estimated at pH 7.6. In the CVs at the UTGEh electrode and in the same experimental conditions, the anodic peak is not easy to measure, but E_p^c compares with the values obtained with the PGEh. From the amount of charge consumed in the reduction process and taking into account the geometric areas of the electrodes, identical surface coverage areas Γ 45 pmol/cm^2 were estimated on both electrodes. No voltammetric signal was observed with the gold electrode either with the enzyme immobilized on the electrode surface (PFV) or using the membrane electrode configuration.

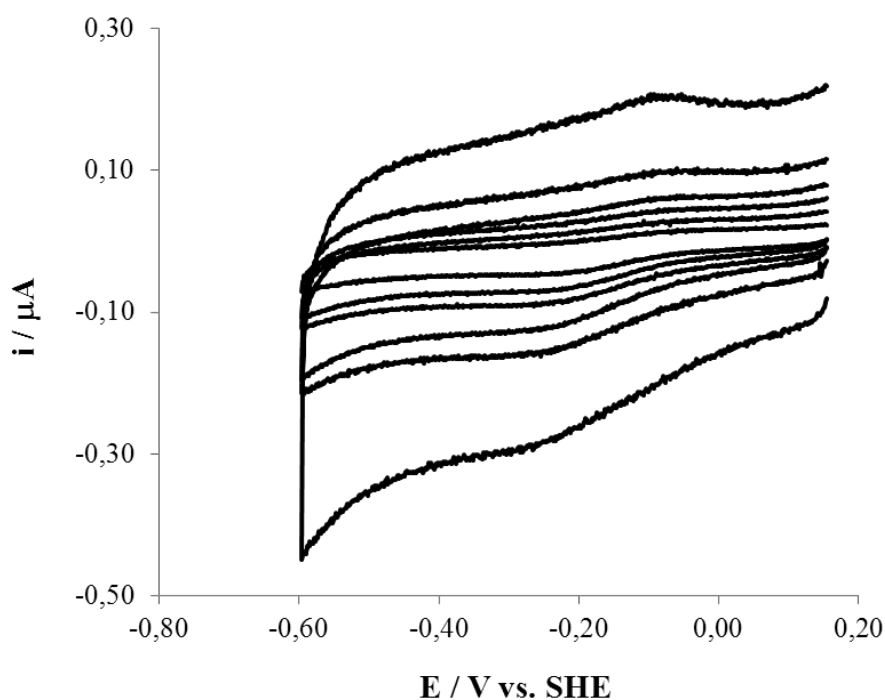
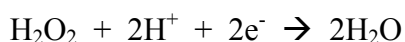
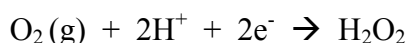


Figure III. 11 – Cyclic voltammograms of *C.necator* NapAB adsorbed to the pyrolytic graphite electrode in non-turnover conditions and 10 mM Tris-HCl pH 7.6 and 0.1M NaCl (scan rates: 5, 10, 15, 20, 30 and 50 mV s^{-1}).

Turn-over conditions

The catalytic activity of the oxidized *Cn* NapAB was investigated in the potential range 160 mV to -600 mV vs SHE in the presence of increasing concentrations of potassium nitrate in the electrolyte solution (10mM Tris-HCl pH 7.6 and 0.1M NaCl). Several protein/electrode interfaces were tested. Most experiments were performed with the enzyme immobilized either on the UTPGEh and using a rotation speed of 3000 rpm or on the PGEh electrode with the electrolyte homogenized with a magnetic stirrer. In other experiments the enzyme was entrapped in a dialysis membrane and in this configuration a gold electrode was also tested. Some experiments were done with the enzyme freely diffusing in solution. The promoters polylysine (PLL), 4,4'-dipyridil and the surfactant didodecyldimethylammonium bromide (DDAB) were tested. The use of an ionic liquid as a promoter (0.4M [C₄mim]Cl⁻, described in Chapter 4), was also investigated. Additionally, enzyme samples from different purification batches were tested.

In all situations, the addition of nitrate to the voltammetric cell, in successive proportions of variable concentrations (ranging from 10 μM to 10 mM), produced no coherent and/or measurable alteration in the cyclic voltammograms. In some experiments with the graphite electrodes, and upon nitrate addition, a catalytic wave developed at a potential close to that of the non-turnover signal. However, once the solutions were well deaerated, no changes could be detected on the CVs upon nitrate addition. It was also observed that upon air or water addition a catalytic wave developed at the same potential. Therefore, the appearance of this wave can be attributed to the presence of dissolved oxygen in the nitrate solution. Since dissolved oxygen is readily reduced in the potential range 160 to -600 vs SHE according to:



Increasing amounts of a well-deaerated solution of hydrogen peroxide were added to the voltammetric cell with the enzyme immobilized in the UTPGEh.

In the presence of H₂O₂, a catalytic wave is observed on the cyclic voltammograms and the current increases with increasing H₂O₂ concentration in the range of 50 μM to 500 μM. For the highest concentrations the wave loses its peak shape and becomes sigmoidal with a half wave potential value $E_{1/2} \approx -0.155$ mV vs SHE, a value close to that of the non-turnover signal and therefore must

correspond to the same redox process (Figure III. 12). For each H_2O_2 concentration the catalytic current was measured at the peak potential and the data fitted to the Lineweaver-Burk plot. A Michaelis-Menten constant (K_m) of $100 \mu\text{M}$ and a maximum catalytic current i_{max} of $7.5 \times 10^{-7} \text{ A}$ were determined. With the amount of enzyme adsorbed this catalytic current was converted into the turnover number, $k_{\text{cat}} = 5.5 \text{ s}^{-1}$. This catalytic activity was totally unexpected, certainly not related to the Mo catalytic site but probably connected to the hemes present in the NapB subunit. As described by Paes de Sousa and collaborators [110], altered forms of *c*-type cytochromes, with either Met-His or bis-His-coordination can be induced by interaction with pyrolytic graphite electrodes. This interaction promotes the displacement of one of the axial ligands and the altered forms display peroxidatic activity. Both $E^{0'}$ and $E_{1/2}$ values do not compare with the redox potentials of the heme cofactors displayed in Table III.3, which is in agreement with the formation of an altered form of the hemes. Moreover, the shape of the catalytic voltammograms and the turnover number estimated for the catalytic activity of NapAB towards H_2O_2 are similar to those reported by Paes de Sousa. The protein may be actually facing the electrode through the small di-heme *c*-type subunit NapB and electrons are being directly transferred between the hemes and the electrode. This transfer occurs most probably through the solvent exposed heme (Figure III. 3).

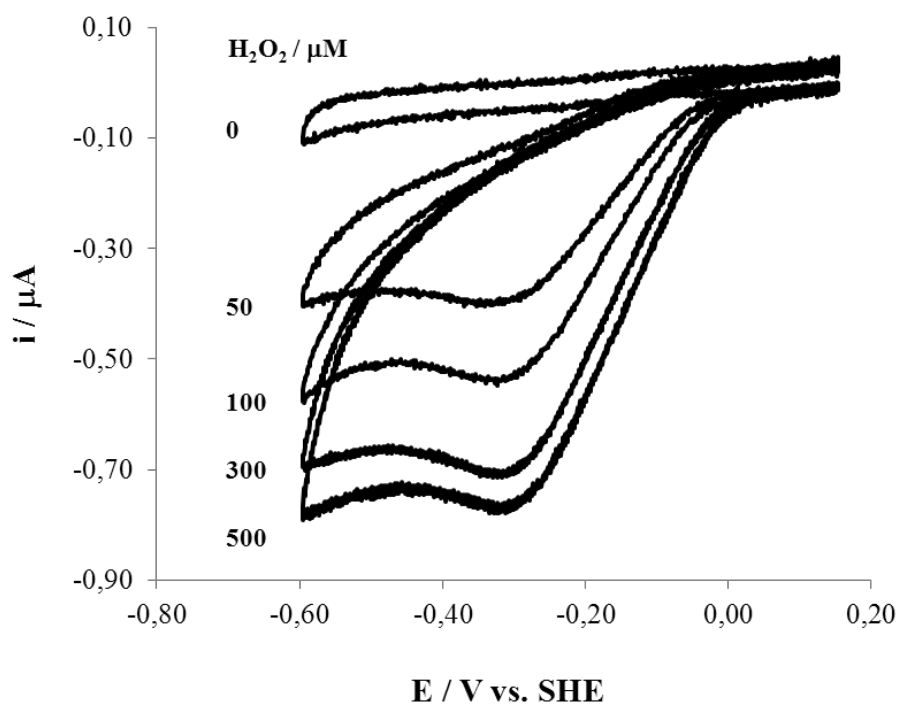


Figure III. 12 - Cyclic voltammograms of *C.necator* NapAB adsorbed at a graphite electrode in the presence of increasing concentrations of H_2O_2 , in 10 mM Tris-HCl pH 7.6 and 0.1 M NaCl ($v = 20 \text{ mV s}^{-1}$, 3000 rpm).

The results reported for *C.necator* NapAB are different from those observed for other nitrate reductases investigated by protein film voltammetry using graphite electrodes. In NapAB from *R.sphaeroides* non-turnover-signals are never observed, but the catalytic activity towards nitrate is always detected. This activity reaches a maximum at a given potential and then decreases again and depending on the experimental conditions (pH and T) the maximum occurs in the range of -100 mV to -200 mV vs SHE. In the case of *Rs* NapAB, the catalytic activity was also proven to be inhibited by high substrate concentrations, which constitutes an unprecedented fact for enzymes of the DMSO reductase family of Mo proteins. These results suggest that the large catalytic subunit NapA, where nitrate is reduced to nitrite, is involved in the voltammetric response [111-113].

The reason for the different voltammetric behavior between *C.necator* and *R.sphaeroides* NapAB when adsorbed on graphite electrodes cannot be explained on the basis of simple electrostatic electrode/protein interactions. The electrostatic surface potential around the exposed-funnel in the case of *C.necator* is markedly positively charged, as previously mentioned (Figure III. 6), but does not seem to interact with the negatively charged graphite surface. In the case of NapAB from *R.sphaeroides*, the electrostatic surface potential is clearly more anionic and neomycin has been shown to improve the voltammetric signal stability [113]. Therefore, a favorable electrostatic interaction must exist between the electrode surface and the negatively charged region through the NH^{3+} groups of neomycin, so that electrons are exchanged with the catalytic active site. In both enzymes, the NapB subunits are very similar and the electrostatic surface potential around the exposed heme(s) is positive. Our voltammetric data suggest that in the case of *C.necator* NapAB, a domain for electron transfer was built up through interactions with the C-O functionalities of the graphite surface and the region around the exposed heme. However, this affected the enzyme properties and the peroxidatic activity displayed by *C.necator* NapAB does not, or may not, have biological relevance.

C.necator NapAB is located on the periplasmic compartment of the cell, thus, exposition to different concentrations of nitrate substrate during natural biochemical processes (or in the culture medium) may explain some of its indifference towards the presence of nitrate in the electrolytic cell. The protein may have developed a protection mechanism, leading to a required protein pre-activation for nitrate consumption. New experiments are necessary to answer this question.

CHAPTER 4

Protein Crystallization using Ionic Liquids

The work described in this chapter was the subject of the following publication:

C. Coelho, J. Trincão, and M. J. Romão (2010) “The use of ionic liquids as crystallization additives allowed to overcome nanodrop scaling up problems: A success case for producing diffraction-quality crystals of a nitrate reductase” *Journal of Crystal Growth*, 312, 714-719.

4.1 INTRODUCTION

4.1.1 Limitations in Protein Crystallization

Protein crystallization is a key step in determining a macromolecule three-dimensional structure. Many proteins are difficult to crystallize. The propensity to form, an orderly packed array of molecules or not in the form of a crystal can be related with some intrinsic properties. The protein molecular size might be important, since larger proteins are more likely to have segments of residues with limited mobility, which hinder crystallization. Another important feature is protein solubility, since the protein must be able to decrease its solubility in the precipitant to start nucleation. Protein crystallization can also be determined by chemical interaction between macromolecules, usually the larger the differences between various possible contacts, the easier it is to produce a good crystal. There are also external protein factors responsible for limitations in obtaining crystals. The most common is generally the limited amount of biomaterial available. Some proteins are very hard to produce and difficult to purify and the final yield can be very low, limiting the number of screens that can be setup by the crystallographer to achieve crystallization conditions.

Recent advances in robotics and computer control, driven by some structural genomics initiatives around the world, have been developed in order to overcome some of these difficulties. The use of dispensing nanodrop crystallization robots is one of them and is becoming very popular. In the particular case of the nitrate reductase from *C.necator*, due to the limited amount of protein available initially, the High Throughput Crystallization Facility at EMBL Grenoble, which offers automated nanovolume crystallization screening on a service basis, was requested for the crystallization setups [86]. This facility has the particularity of performing the crystallization experiments, which can be followed up after by users, through automated imaging systems (through a real-time dedicated web interface). A total of 96 different crystallization conditions were tested, with the 16 μ l of pure *C.necator* nitrate reductase available at 10 mg/ml. For common manual crystallization this reduced amount of protein would only be sufficient to setup one crystallization plate of 24 different conditions. The protein sent to the robot, successfully crystallized in several different conditions and two of them in particular gave crystals that were suitable for diffraction data collection. The native structure of NapAB could be solved at 1.5Å resolution, using data from one of these robot microcrystals (Chapter 3). This case illustrates the efficient and reliable possibilities of using a crystallization robot.

Very recently an Oryx8 protein crystallization robot (Douglas Instruments) was acquired by our research group (Crystallography group at Faculdade de Ciências e Tecnologia from Universidade Nova de Lisboa) and all the nanocrystallization setups can now be performed in-house.

4.1.2 Limitations to Scale up Robot Conditions

The *C.necator* NapAB crystal structure raised new questions that lead to the need of obtaining additional crystallographic data for other enzyme forms. Corroboration of the new mechanistic evidences related with the assignment of sulfur as the 6th molybdenum ligand and the pursuit of additional mechanistic studies were the main reasons for pursuing crystallographic studies on *C.necator* NapAB [32, 33]. New experiments should include crystallographic studies for different forms of the enzyme, namely reduced forms, as well as of complexes with substrates analogues and inhibitors. As an initial approach to obtain new data, the best robot crystallization conditions were reproduced using larger drop volumes, a procedure commonly referred as scale up. Despite being an empirical procedure, in practical terms it may become difficult to accomplish and no NapAB crystals could be obtained by simply using this approach. The change in equilibration kinetics, due to the larger surface area to volume ratio of the nanodrops is possibly one of the major causes for this specific problem, together with different evaporation rates. Other external causes may be related with variations among protein batches and aging samples, commercial versus homemade solutions or a mixed combination of several factors.

Crystallographers are always searching for new and efficient crystallization protocols, which enable them to obtain better diffracting crystals. When there are no limitations, particularly concerning the amount of protein available, several different conditions are tested within the initial crystallization condition. This includes, besides screening the precipitant concentration, pH and buffers, the screening of different crystallization temperatures, and the use of different additives in the mixture. Some seeding techniques are also regularly performed, with the objective of introducing external crystallization nuclei in the protein solution to induce nucleation. For the scale up of nitrate reductase crystals, seeding experiments were also tested as well as several additives, particularly some with biological significance for the protein (substrate and inhibitors), but with no positive results. Only long thin needles were obtained, not suitable for diffraction experiments.

In the pursuit for new crystallization methodologies, some recently described compounds named ionic liquids (IL) were tested as crystallization additives. The results obtained were very promising, as explained below.

4.1.3 Ionic Liquids and Proteins

Room temperature ionic liquids (IL) are salts that do not crystallize at room temperature. They are composed of an organic cation (for example, 1-butyl-3-methylimidazolium or [C₄mim]) and any of a variety of organic/inorganic counter-anions (Cl⁻, PF₆⁻ or BF₄⁻, for example) (Figure IV. 1). Ionic liquids have been investigated as environmentally friendly replacements for organic solvents: due to their negligible vapor pressure they can be efficiently reused. They also possess high thermal and chemical stability and have been used in several areas of research such as green chemistry, enzymatic catalysis and electrochemistry, making the most of their excellent stability and unusual solvent properties. Depending on their structure, they can be immiscible or not with water. Biocatalytic reactions in ionic liquids have shown higher selectivity, faster rates and enhanced enzyme stability. Ionic liquids can be considered as “personalized solvents” because it’s easy to change their structure and thus their solvent properties. IL are more viscous than typical organic solvents, probably due to the strong intermolecular forces between solvent molecules. The charge to charge interactions is inherent in ionic liquids, but reducing van der Waals interactions by reducing the surface area of the molecule can slightly lower the viscosity. Ionic liquids with shorter alkyl chains are less viscous than those with longer ones. There are more than 1000 different ionic liquids described in the literature, from which 300 are already commercially available. They are much more expensive than organic solvents and the key to make their use more profitable will be to develop efficient recovery, isolation and reuse [114-116].

The interactions between ionic liquids and proteins have been investigated in recent years. Enzymes suspended in ionic liquids remain stable and catalytic active, even though they are not stable or active in conventional polar organic solvents. Ionic liquids have been demonstrated also to improve protein refolding following denaturation. This is particular relevant nowadays, due to the common expression of recombinant proteins in prokaryotic systems, which in the majority of the cases lack the proper refolding machinery. Due to their ionic nature, they may interact with charged groups of enzymes, either in the active site or in the surface, giving rise to protein structural changes. Unlike conventional organic solvents, ionic liquids possess no vapor pressure, and are able to dissolve many

compounds. These non-aqueous but polar solvents may also constitute a new and efficient media for enzyme catalysis. The possibility to tune ionic liquids properties, such as solubility, polarity or hydrophilicity, by selecting the appropriate combination anion/cation, makes them remarkably versatile for many applications [117-119].

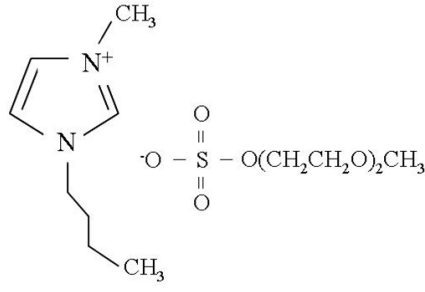
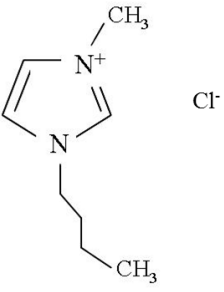
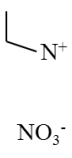
Ionic Liquids	[C ₄ mim] MDEGSO ₄	[C ₄ mim] Cl ⁻	EAN
Name	1-Butyl-3-methylimidazolium 2-(2-methoxyethoxy)ethyl sulfate	1-Butyl-3-methylimidazolium chloride	Ethylammonium nitrate
Molecular Weight (g.mol ⁻¹)	338	174	108
Formula	C ₁₃ H ₂₆ N ₂ O ₆ S	C ₈ H ₁₅ ClN ₂	C ₂ H ₈ N ₂ O ₃
Structure			

Figure IV. 1 – Some common ionic liquids used in protein crystallography.

4.1.4 Protein Crystallization and Ionic Liquids

Despite the large number of studies relating ionic liquids and protein stabilization or renaturation, there were only a few reports on the use of IL in protein crystallization, when the scale up trials for the nitrate reductase crystals were started. The first study dates from 1999 and reports the use of ethyl ammonium nitrate in the crystallization of lysozyme [120]. In a subsequent study, Pusey and collaborators described crystal improvement of four different proteins (canavalin, β -lactoglobulin B, xylanase and glucose isomerase, respectively), with the use of three different IL [121]. However, all these studies were performed in model proteins, which posed no major crystallization problems. Furthermore, only crystal morphology was assessed, and no details from diffraction quality of the improved crystals were given. Recently, a third publication was available in which the authors performed similar experiments as Pusey and co-workers, for other model proteins (lysozyme,

catalase, myoglobin, trypsin, xylanase and glucose isomerase) and a monoclonal antibody Fab in complex with a peptide [122]. In this case, improvement of crystal size was only observed for lysozyme and trypsin. For every study the authors reported that all proteins retained their catalytic activity after incubation with ionic liquids, and that crystals were obtained at a number of conditions where they were absent from IL-free control experiments. This approach seemed very promising, and although limited by the reduced amount of protein available, two water-soluble ionic liquids were tested in an attempt to scale up the robot crystallization conditions for the *C.necator* nitrate reductase.

4.2 *C.NECATOR* NAPAB: A CASE STUDY

4.2.1 Protein Expression and Purification

C.necator periplasmic nitrate reductase was expressed and purified as mentioned in chapter 3 of this Thesis (sections 3.2.1 and 3.2.2). The protein used to scale up the initial robot conditions was obtained from the 10L BioStat B-Plus reactor batch. Final protein concentration was 10 mg/ml in 10 mM Tris-HCl pH 7.6 buffer.

4.2.2 Crystallization

Crystallization drops were setup at 20°C, using the sitting-drop vapor diffusion method, as in the crystallization robot. Initial attempts were made using the same conditions for the best crystals obtained by the robot (condition 42: 0.1M Bis-Tris pH 5.5, 25% w/v PEG 3350 and condition 90: 0.2M sodium formate, 20% w/v PEG 3350), but increasing ten times the drop volume (from 0.2µl in the robot, to 2µl in manual setup). Crystallization stock solutions number 42 and 90 from Index Crystallization Screen were purchased from Hampton research. As no crystals were obtained, a PEG 3350 screen concentration was performed, which lead to the formation of micro needles and sea urchins, but no measurable crystals. Seeding experiments were also tested as well as additives (such as potassium nitrate, sodium azide, MPD, sodium malonate and sodium dithionite). Due to the continuous lack of results, we pursued a new approach, using water soluble ionic liquids (IL) that were tested as crystallization additives. Due to the small amount of protein available only two different IL were selected for the experiments: [C₄mim] Cl⁻ (1-butyl-3-methylimidazolium chloride) and [C₄mim] [MDEGSO₄] (1-butyl-3-methylimidazolium 2(2-methoxyethoxy) ethyl sulphate), both available from Fluka. The main objective of this choice was to maintain the cation ([C₄mim]) while changing the anion. The protein was incubated either with 0.2M and 0.4M of [C₄mim] Cl⁻ and [C₄mim] [MDEGSO₄], for several minutes before setting up the crystallization drops. All values for ionic liquid concentration are in respect to final concentration in protein solution. Crystal for the nitrate reductase protein with overall dimensions of 0.3 x 0.1 x 0.1 mm³ were obtained within two days, using 0.2M sodium formate, 15% PEG 3350 and 0.4M [C₄mim] Cl⁻.

To ensure that manual crystallization conditions were as similar as possible to the robot, all scale up experiments (with or without ionic liquids) were performed using the same crystallization solutions,

at the same temperature (20°C) and with the same protein batch purification. A sample from this batch (10L BioStat B-Plus reactor batch) was also sent to the crystallization robot facility in EMBL to prove that the differences from the two protein batches (robot-batch and 10L batch) were not the cause for reproducibility problems.

4.2.3 Data Collection and Analysis

The best crystals were harvested in a solution containing a higher precipitant concentration (25% PEG 3350, and 0.2M sodium formate) and allowed to stabilize for several minutes before cryo-cooling with paratone oil as cryoprotectant. Crystals were flash cooled and stored in liquid nitrogen, and several datasets were collected at the ESRF (beamlines ID 14-1 and 14-2), using ADSC Quantum-4R CCD detectors. In general crystals diffracted in a range of 2 to 3.5 Å, with some few crystals diffracting as well as the initial ones obtained at the robot (Table IV. 1). Subsequent soaking experiments, using 20 mM sodium dithionite were performed in some crystals obtained with the help of the ionic liquid, as previously mentioned in Chapter 3 (Table III. 1, dataset name DIT-IL).

Dataset name	DIT-IL (2007)	Nap_IL (2008)	Nap_IL (2009)
Beamline (ESRF)	ID14-2	ID 14-1	ID14-2
Space Group	C2	C2	C2
Cell constants	$a = 119.4 \text{ \AA}, b = 71.4 \text{ \AA},$ $c = 128.4 \text{ \AA}$ and $\beta = 121.0^\circ$	$a = 135.5 \text{ \AA}, b = 71.8 \text{ \AA},$ $c = 114.9 \text{ \AA}$ and $\beta = 110.5^\circ$	$a = 121.7 \text{ \AA}, b = 71.2 \text{ \AA},$ $c = 121.7 \text{ \AA}$ and $\beta = 115.0^\circ$
Resolution (Å)	1.7	2.7	2.0
Matthew's coefficient (Å ³ /Da)	2.32	2.45	2.23
Solvent content (%)	46.90	49.75	44.95
Crystallization conditions	0.2M sodium formate 15% PEG 3350 0.4M [C ₄ mim] Cl ⁻ 20mM sodium dithionite	0.2M sodium formate 15% PEG 3350 0.4M [C ₄ mim] Cl ⁻	0.2M sodium formate 15% PEG 3350 0.4M [C ₄ mim] Cl ⁻ 20mM sodium dithionite

Table IV. 1 – Crystal parameters and X-ray diffraction data statistics for the best diffracting crystals obtained using ionic liquids.

4.3 RESULTS AND DISCUSSION

All attempts to reproduce in house the crystals obtained at the High Throughput crystallization robot, but using larger volume drops had failed. The only crystallization results obtained were precipitate, small urchins and thin needles (Figure IV. 2). To overcome this problem, two different ionic liquids ($[C_4mim] Cl^-$ and $[C_4mim] [MDEGSO_4]$) were tested as crystallization additives. In both cases decreased nucleation was observed, and with 0.4M of $[C_4mim] [MDEGSO_4]$, the needles were larger although still very thin for diffraction experiments (Figure IV. 3). With 0.4M $[C_4mim] Cl^-$ as crystallization additive in the protein solution, good results were achieved and large single crystals were obtained in a reproducible manner (Figure IV. 4). When using a lesser amount of this IL in the protein solution (0.2M), no good crystals were obtained (Figure IV. 3). The changes in the counter ion, as well as the ionic liquid concentration, were crucial for successful optimization of crystallization.

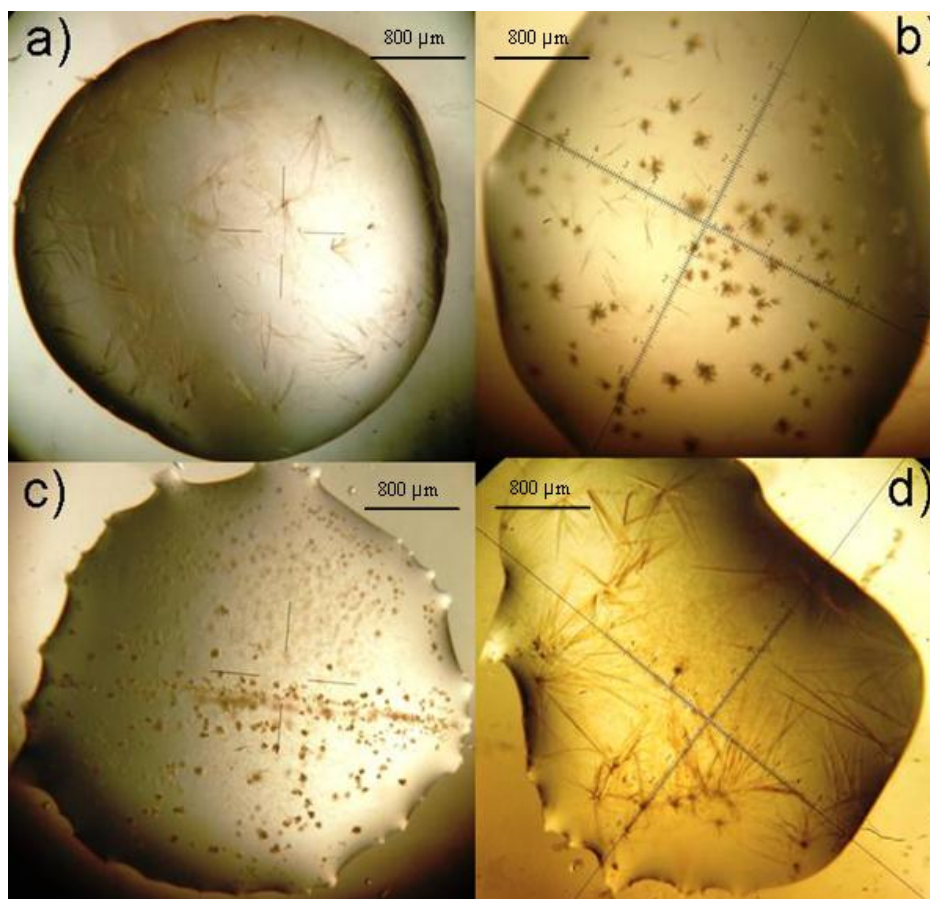


Figure IV. 2 - Microdrop crystallization in the absence of ionic liquids: (a) condition 0.2M sodium formate and 15% PEG 3350, protein concentration 5 mg/ml; (b) sea urchins obtained using 0.2M sodium formate and 12.5% PEG 3350; (c) streak seeding from drop (b); (d) long but fine needles obtained in the same condition (a) but with protein concentration 10 mg/ml.

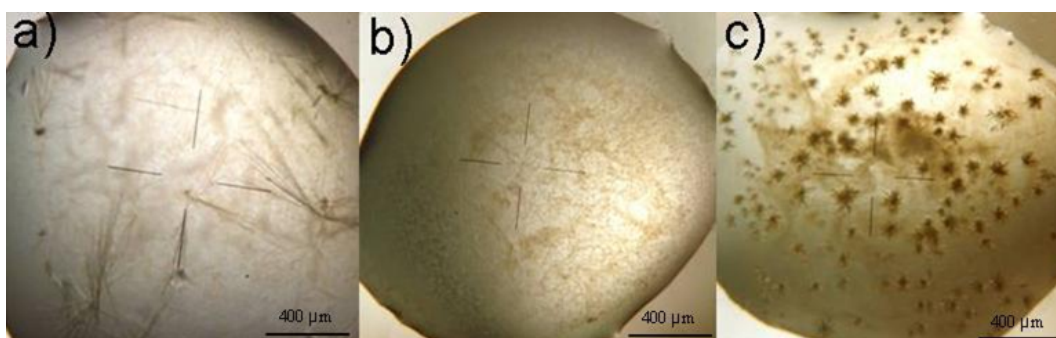


Figure IV. 3 - Microdrop crystallization in the presence of (a) 0.4M [C₄mim][MDEGSO₄] (b) 0.2M [C₄mim][MDEGSO₄] and (c) 0.2M [C₄mim]Cl. Crystallization conditions are 0.2M sodium formate, 15% PEG 3350.

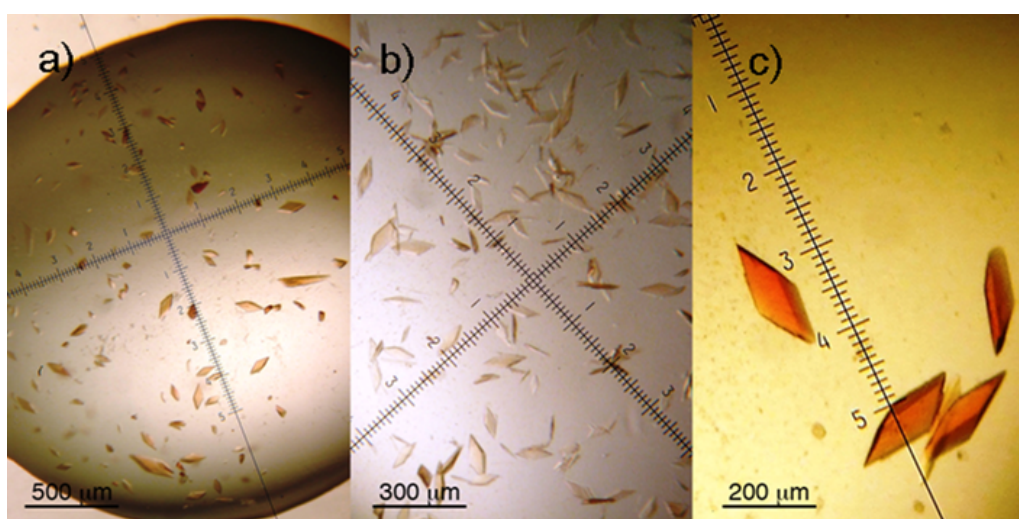


Figure IV. 4 - Microdrop crystals obtained in the presence of 0.4M [C₄mim]Cl revealing crystallization reproducibility. Crystallization conditions are 0.2M sodium formate and 15% PEG 3350. Approximate crystal dimensions 0.3 x 0.1 x 0.1 mm³.

The maximum resolution for each dataset was very dependent not only in intrinsic crystal quality but also on crystal handling and cryo-cooling. *C.necator* NapAB crystals were very sensitive, and typical diffraction observed ranged between 2-3Å, with only about 20% of the crystals diffracting isotropically to the same resolution as the ones obtained from the robot. Though the majority of the crystals seemed unique, diffraction data showed two lattices after data collection for some of the crystals, probably a consequence from crystal manipulation and cooling. To avoid overlaps during data collection, the X-ray detector was moved further away from the crystal to increase the distance and the ability to process data all the way to the detector corner (Figure IV. 5).

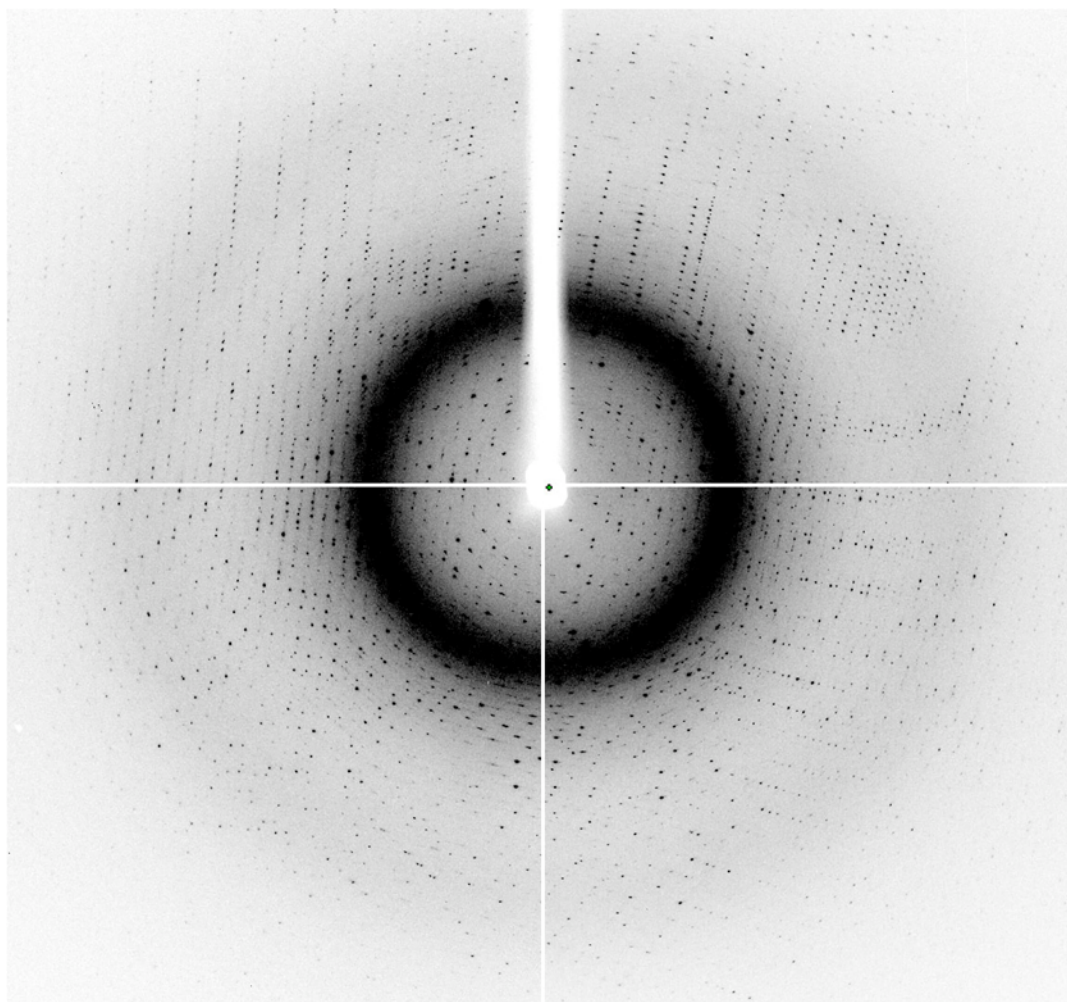


Figure IV. 5 - *C.necator* NapAB protein diffraction pattern of a crystal obtained in-house, using 0.4M [C₄mim] Cl. Resolution at edge is 2Å and dataset name Nap_IL (2009) from Table IV.1

All crystals from *C.necator* nitrate reductase grown in the presence of the ionic liquid belong to the same space group C2 as the native crystal, obtained in the crystallization robot. Despite belonging to the same space group, differences can be observed in cell constants dimensions. Particularly for the case described in the previous Chapter 3 (dataset name DIT-IL), where differences in unit cell parameters account for a decrease of 16% and 13% in the *a* and *b* axes respectively, while the *c* axis showed an increase of 33% (Table IV. 1). This leads to a small decrease in unit cell volume (9%), but large differences in crystal packing (Figure IV. 6). The differences in cell dimensions are accompanied by differences in the relative positioning of the individual molecules in the crystal lattice, induced by the presence of the ionic liquid in the protein. These differences clearly demonstrate the interaction between the ionic liquids and the protein molecules, which in this particular case were probably necessary to hold the conformational changes that occurred in the soaking experiments, with the reducing agent sodium dithionite.

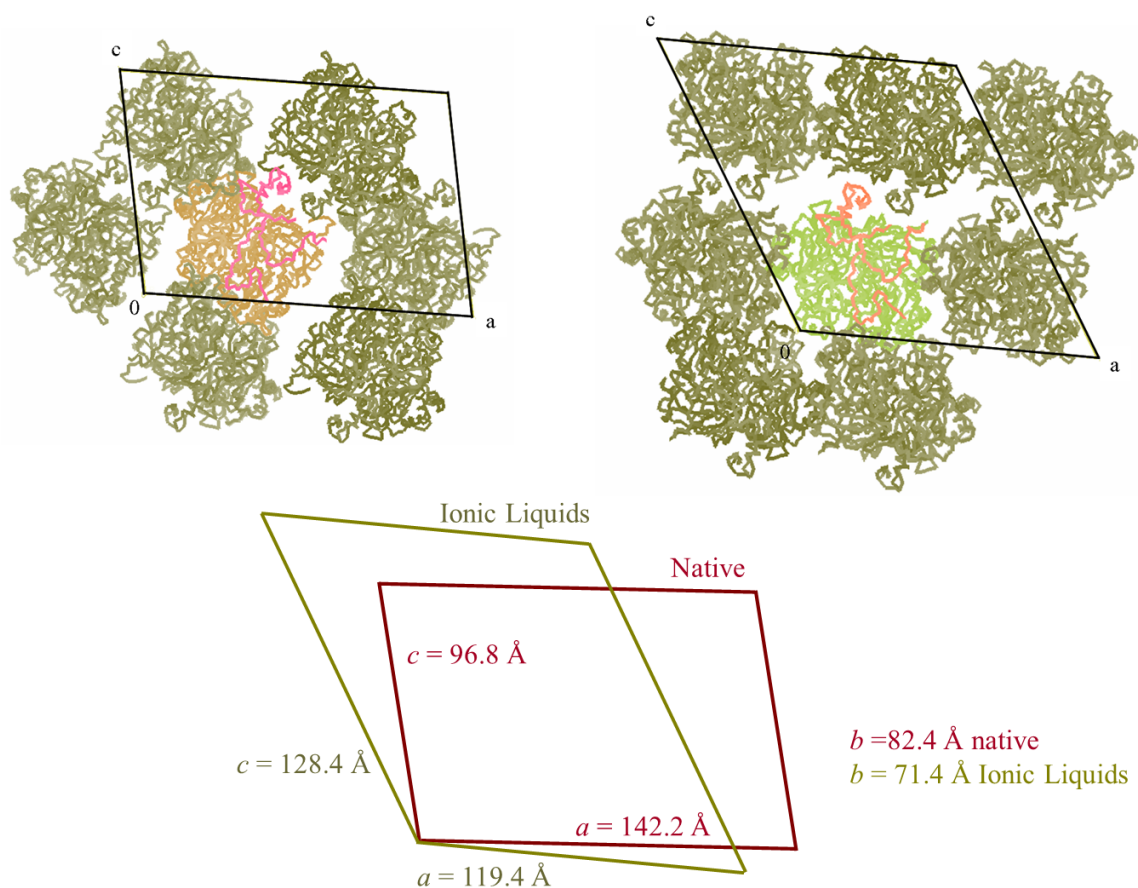


Figure IV. 6 - Changes in the crystal packing. Upper Left: crystal packing for the native protein, with unit cell constants $a = 142.2\text{\AA}$, $b = 82.4\text{\AA}$, $c = 96.8\text{\AA}$ and $\beta = 100.7^\circ$ (46.9% solvent); Upper Right: packing of crystals obtained with IL, unit cell constants $a = 119.4\text{\AA}$, $b = 71.4\text{\AA}$, $c = 128.4\text{\AA}$ and $\beta = 121.0^\circ$ (51.5% solvent content); Center: comparison of the two unit cell constants, projected along the b axis. With IL, a and b axes decrease respectively 16% and 13%, while c axis increased 33%.

After careful inspection of the electron density map, no IL molecules could be clearly identified. Also in the recent work of Judge and co-workers [122], where crystals of two model proteins and a Fab complex were obtained in the presence of ionic liquids, no IL molecules could be assigned in the crystal structures. This might be due to some disorder in the crystal, to unspecific binding or even no binding of ionic liquid molecules to the protein. Although they could not be identified in the crystal structure, IL molecules must affect specific crystallization contacts within the crystal. This could explain the large differences observed in the cell constants for crystals grown in the presence of IL, resulting from changes in the relative packing of the molecules. The way ionic liquids affect protein crystallization cannot be explained by a single mechanism, but rather a combination of several different factors. Changes in growth kinetics and in solution properties might be important, but also variations in the vapor equilibration rate. The presence of ionic liquid can conduce a slower

evaporation rate, and consequently to different molecule packing. Ionic liquids can also be responsible for changes in protein solubility as previously mentioned.

The difference encountered when changing from chloride (Cl⁻) to a larger anion such as 2-(2-methoxyethoxy) ethyl sulphate (MDEGSO₄⁻), might be explained by recent studies elucidating the differences among the structures of the [C₄mim] with different anions. Ionic liquids have localized structures around each ion, constituted by pairs of ions and not dissociated ions as previously thought [123, 124]. Comparing the results obtained for the difficult case of the nitrate reductase, as well as those from model proteins using the cation [C₄mim], it is easily observable that it has the largest success rate.

It is important to note that due to the rather limited amount of protein (before the 300L industrial bacterial growth), and the successful results obtained with 0.4M [C₄mim] Cl⁻, a more extensive screen was not performed. It is also important to mention that with the sample from the 10L batch that was sent again to the robot facility in EMBL, it was possible to reproduce the initial microcrystals obtained. This proves that the difference from the two protein batches was not the cause of reproducibility problems during scale up attempts.

4.4 CONCLUSIONS

The use of dispensing robots is becoming a very common and popular practice in the protein crystallography community. These robots possess several impressive features, such as allowing for very fast screening of a wide range of conditions, in controlled environments, and using minimal amounts of protein. Although very powerful in the initial crystallization process, they are generally not used in the following crystallization optimization steps. In addition, the increasing need to scale up from nanodrops to regular drops is becoming quite problematic. Despite the several advances in protein crystallography it is still not possible to define general crystallization protocols, since each individual protein has its intrinsic properties and requirements for successful crystallization. The crystallization of *C.necator* NapAB constitutes the first reported case on the use of ionic liquids to solve a real difficult case scenario. The use of ionic liquid as a crystallization additive allowed overcoming the volume scale up problem. The reproducibility of good quality crystals will allow pursuing the mechanistic studies on the periplasmic nitrate reductase. The results here reported may define a strategy that can be applied to many other cases. These organic salts can also be used to optimize crystal properties, such as size and morphology, to improve data diffraction or enhance protein stability. Depending on the amount of pure protein available, when screening crystallization conditions using ionic liquids, one should try different cation/anion combinations, such as using larger versus smaller ones, while maintaining the same cation. The final concentration of IL in the protein solution should be between 0.2M and 0.6M, the most studied range of concentrations for protein-IL interactions.

One clear remarkable sign of the importance and usefulness of ionic liquids in protein crystallization is the recent market launch of a screen of additives for protein crystallization, composed entirely of ionic liquids (Ionic Liquid Screen from Hampton Research).

CHAPTER 5

General Conclusions and Future Perspectives

5.1 GENERAL CONCLUSIONS

Mouse aldehyde oxidase homologue1 (mAOH1) and *C.necator* periplasmic nitrate reductase (*Cn* NapAB) are two molybdopterin dependent enzymes. Due to the characteristics and specificities of their catalytic active sites the two proteins belong to different Mo containing enzymes families. mAOH1 is a protein homologue to aldehyde oxidase and belongs to the xanthine oxidase family, while *Cn* NapAB belongs to the DMSO reductase family. The three dimensional structure for the two proteins was determined by X-ray protein crystallography and studies regarding function, chemical behavior and structure specificities were performed for the nitrate reductase.

The recombinant mAOH1 enzyme was used to optimize the crystallization conditions for the native enzyme from mouse liver. After extensive trials, a usable dataset was collected at 2.9Å resolution for the native protein crystals. The structure was solved by molecular replacement using the homologous protein xanthine oxidase (XO). This constitutes the first structure ever obtained for an aldehyde oxidase. It is known that human AO (AOX1) is involved in drug metabolizing processes, possibly through xenobiotic detoxification, and the recognition of its importance has been increasing in recent years. Several pharmaceutical companies have been studying the protein interaction with a large number of different drugs. The range of substrates is much broader in aldehyde oxidase than in xanthine oxidase, despite their remarkable overall similarity. With the mAOH1 crystal structure, we were able to identify crucial residues in the protein active site responsible for the binding of different substrates, particularly large and bulky N-heterocycles. Also, the differences in the active site funnel, which leads to the molybdopterin cofactor, could be addressed. The mAOH1 funnel is much wider and able to accommodate some substrates that are unable to be metabolized by XOR. An inter-conversion mechanism (XDH→XO) exists in mammalian xanthine oxidoreductases, which was never described for other XO family members, including mAOH1. Some important protein residues involved in this inter-conversion mechanism were assigned in mAOH1, and their comparison with bovine XOR helped to explain the lack of such mechanism in aldehyde oxidases.

Several crystal structures from periplasmic nitrate reductases proteins have been solved in the last twelve years, both for monomeric and heterodimeric proteins. These structures were solved at a resolution that couldn't unequivocally assign all the Mo ligands, and the 6th molybdenum position was believed to be occupied by an O/OH group. The heterodimeric nitrate reductase from *C.necator*

is homologous to the previous protein structures reported, but its structure could be solved at 1.5Å resolution, constituting the highest resolution structure obtained for a nitrate reductase so far. With high resolution, clear electron density was observed for an atom heavier than oxygen at the Mo 6th position. The *B*-factor analysis provided corroboration and the ligand could be unambiguously assigned as a sulfur atom. The same procedure was used for the monomeric nitrate reductase from *D.desulfuricans* in recent crystallographic studies, with the same result. It is likely that this constitutes a common feature for all periplasmic nitrate reductases. In addition, the partially reduced form of the enzyme could be obtained at 1.7Å, and has particular importance in explaining the protein reaction mechanism. This NapAB partially reduced form was obtained using ionic liquids as described in detail in the case study reported. These novel compounds have proven to be useful as crystallization additives. They were also decisive to achieve the partially reduced form of the enzyme in this specific case. The *C.necator* NapAB spectroscopic and electrochemical studies also revealed some unexpected features of the protein, particularly the different hemes potentials observed when comparing with the homologous heterodimeric NapAB from *R.sphaeroides*. Moreover, the lack of a catalytic voltammetric response towards the natural substrate nitrate was unexpected, since it was already described and reported for the homologous *Rs* NapAB.

5.2 FUTURE PERSPECTIVES

When solving a new protein three-dimensional structure, some answers can be given to the universe of pre-existing questions regarding that particular structure. Nevertheless, many more questions are raised afterwards. This becomes very clear, when a crystallographer is able to solve its first protein crystal structure.

There are still many open and important questions directly related to what was explored in this Thesis, particularly regarding the mAOH1 protein. To answer some of these burning questions, one of the first steps is to try to improve the protein crystals, in order to extend diffraction. The main goal is to perform crystal soaking experiments, particularly with the important compounds that have been described to be metabolized by aldehyde oxidases, in particular those of pharmacological relevance. The ability to analyze where these compounds bind in the protein active site would be of great interest. With the now available coordinates, some substrate docking studies and molecular simulations can also be performed simultaneously, and the results compared afterwards. With the

increasing importance of aldehyde oxidase for the pharmaceutical companies, drug design studies should not be forgotten as an aim to our studies.

Biochemical studies have shown different specificities among the different aldehyde oxidase isoforms, particularly between mAOH1 and mAOH2. It would also be of great value to solve the crystal structure for the mAOH2 isoform and perform structural comparisons. The amount of protein is very limited, but with the recently available and improved expression system it should be feasible. Also, the crystallization robot can be helpful in these particular cases. Ionic liquids should again be helpful in this task, not only as crystallization additives, but also as protein stabilizing agents, as has been described for several enzymes.

Also for the nitrate reductase protein, there are still interesting experiments waiting to be done and several aspects needing to be solved. The complete reduced form of the enzyme would be very good to achieve and further experiments using the anaerobic chamber can be of special help in this case. The ionic liquids should also be useful. With the now available large *Cn* NapAB protein quantity, a broader ionic liquids screen can be performed, in the pursuit to better understand their particular properties and protein interactions in protein crystallization. Assigning ionic liquids in the electron density maps would also be very interesting and such experiments must be performed.

Nevertheless, the main and more important accomplishment to be done in the near future is to be able to solve the crystal structure of the human Aldehyde oxidase, from which diffracting crystals were very recently obtained. This constitutes a very challenging project, and hopefully all the necessary conditions are met, so that this project can be pursued.

REFERENCES 6

1. Hille, R., *Molybdenum and tungsten in biology*. Trends Biochem Sci, (2002). 27(7): p. 360-7.
2. Mendel, R.R. and Bittner, F., *Cell biology of molybdenum*. Biochim Biophys Acta, (2006). 1763(7): p. 621-35.
3. Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R., and Meyer, O., *Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster*. Science, (2001). 293(5533): p. 1281-5.
4. Hille, R., *Molybdenum enzymes containing the pyranopterin cofactor: an overview*. Met Ions Biol Syst, (2002). 39: p. 187-226.
5. Romão, M.J., *Molybdenum and tungsten enzymes: a crystallographic and mechanistic overview*. Dalton Trans, (2009)(21): p. 4053-68.
6. Schwarz, G., Mendel, R.R. and Ribbe, M.W., *Molybdenum cofactors, enzymes and pathways*. Nature, (2009). 460(7257): p. 839-47.
7. Leimkuhler, S., Wuebbens, M.M. and Rajagopalan, K.V., *Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor*. J Biol Chem, (2001). 276(37): p. 34695-701.
8. Magalon, A., Fedorb, J.G., Walburgera, A., and Weiner, J.H., *Molybdenum enzymes in bacteria and their maturation*. Coordination Chemistry Reviews, (2011). 255(9-10): p. 1159-78.
9. Hille, R., Nishino, T. and Bittner, F., *Molybdenum enzymes in higher organisms*. Coord Chem Rev, (2011). 255(9-10): p. 1179-205.
10. Wahl, B., Reichmann, D., Niks, D., Krompholz, N., Havemeyer, A., Clement, B., Messerschmidt, T., Rothkegel, M., Biester, H., Hille, R., Mendel, R.R., and Bittner, F., *Biochemical and spectroscopic characterization of the human mitochondrial amidoxime reducing components hmARC-1 and hmARC-2 suggests the existence of a new molybdenum enzyme family in eukaryotes*. J Biol Chem, (2010). 285(48): p. 37847-59.
11. Brondino, C.D., Romao, M.J., Moura, I., and Moura, J.J., *Molybdenum and tungsten enzymes: the xanthine oxidase family*. Curr Opin Chem Biol, (2006). 10(2): p. 109-14.
12. Hille, R., *Molybdenum-containing hydroxylases*. Arch Biochem Biophys, (2005). 433(1): p. 107-16.
13. Enroth, C., Eger, B.T., Okamoto, K., Nishino, T., and Pai, E.F., *Crystal structures of bovine milk xanthine dehydrogenase and xanthine oxidase: structure-based mechanism of conversion*. Proc Natl Acad Sci U S A, (2000). 97(20): p. 10723-8.
14. Romão, M.J., Archer, M., Moura, I., Moura, J.J., LeGall, J., Engh, R., Schneider, M., Hof, P., and Huber, R., *Crystal structure of the xanthine oxidase-related aldehyde oxido-reductase from D. gigas*. Science, (1995). 270(5239): p. 1170-6.
15. Garattini, E. and Terao, M., *Increasing recognition of the importance of aldehyde oxidase in drug development and discovery*. Drug Metab Rev, (2011).
16. Garattini, E., Fratelli, M. and Terao, M., *Mammalian aldehyde oxidases: genetics, evolution and biochemistry*. Cell Mol Life Sci, (2008). 65(7-8): p. 1019-48.
17. Terao, M., Kurosaki, M., Marini, M., Vanoni, M.A., Saltini, G., Bonetto, V., Bastone, A., Federico, C., Saccone, S., Fanelli, R., Salmona, M., and Garattini, E., *Purification of the aldehyde oxidase homolog 1 (AOH1) protein and cloning of the AOH1 and aldehyde oxidase homolog 2 (AOH2) genes. Identification of a novel molybdo-flavoprotein gene cluster on mouse chromosome 1*. J Biol Chem, (2001). 276(49): p. 46347-63.
18. Garattini, E., Mendel, R., Romao, M.J., Wright, R., and Terao, M., *Mammalian molybdo-flavoenzymes, an expanding family of proteins: structure, genetics, regulation, function and pathophysiology*. Biochem J, (2003). 372(Pt 1): p. 15-32.
19. Feng, C., Tollin, G. and Enemark, J.H., *Sulfite oxidizing enzymes*. Biochim Biophys Acta, (2007). 1774(5): p. 527-39.

20. Kisker, C., Schindelin, H., Pacheco, A., Wehbi, W.A., Garrett, R.M., Rajagopalan, K.V., Enemark, J.H., and Rees, D.C., *Molecular basis of sulfite oxidase deficiency from the structure of sulfite oxidase*. *Cell*, (1997). 91(7): p. 973-83.
21. Kappler, U. and Bailey, S., *Molecular basis of intramolecular electron transfer in sulfite-oxidizing enzymes is revealed by high resolution structure of a heterodimeric complex of the catalytic molybdopterin subunit and a c-type cytochrome subunit*. *J Biol Chem*, (2005). 280(26): p. 24999-5007.
22. Schrader, N., Fischer, K., Theis, K., Mendel, R.R., Schwarz, G., and Kisker, C., *The crystal structure of plant sulfite oxidase provides insights into sulfite oxidation in plants and animals*. *Structure*, (2003). 11(10): p. 1251-63.
23. Karakas, E. and Kisker, C., *Structural analysis of missense mutations causing isolated sulfite oxidase deficiency*. *Dalton Trans*, (2005)(21): p. 3459-63.
24. Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K.V., and Rees, D.C., *Crystal structure of DMSO reductase: redox-linked changes in molybdopterin coordination*. *Science*, (1996). 272(5268): p. 1615-21.
25. Moura, J.J., Brondino, C.D., Trincao, J., and Romao, M.J., *Mo and W bis-MGD enzymes: nitrate reductases and formate dehydrogenases*. *J Biol Inorg Chem*, (2004). 9(7): p. 791-9.
26. Gonzalez, P.J., Rivas, M.G., Brondino, C.D., Bursakov, S.A., Moura, I., and Moura, J.J., *EPR and redox properties of periplasmic nitrate reductase from *Desulfovibrio desulfuricans* ATCC 27774*. *J Biol Inorg Chem*, (2006). 11(5): p. 609-16.
27. Richardson, D.J., Berks, B.C., Russell, D.A., Spiro, S., and Taylor, C.J., *Functional, biochemical and genetic diversity of prokaryotic nitrate reductases*. *Cell Mol Life Sci*, (2001). 58(2): p. 165-78.
28. Bertero, M.G., Rothery, R.A., Palak, M., Hou, C., Lim, D., Blasco, F., Weiner, J.H., and Strynadka, N.C., *Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A*. *Nat Struct Biol*, (2003). 10(9): p. 681-7.
29. Jormakka, M., Richardson, D., Byrne, B., and Iwata, S., *Architecture of NarGH reveals a structural classification of Mo-bisMGD enzymes*. *Structure*, (2004). 12(1): p. 95-104.
30. Dias, J.M., Than, M.E., Humm, A., Huber, R., Bourenkov, G.P., Bartunik, H.D., Bursakov, S., Calvete, J., Caldeira, J., Carneiro, C., Moura, J.J., Moura, I., and Romão, M.J., *Crystal structure of the first dissimilatory nitrate reductase at 1.9 Å solved by MAD methods*. *Structure*, (1999). 7(1): p. 65-79.
31. Arnoux, P., Sabaty, M., Alric, J., Frangioni, B., Guigliarelli, B., Adriano, J.M., and Pignol, D., *Structural and redox plasticity in the heterodimeric periplasmic nitrate reductase*. *Nat Struct Biol*, (2003). 10(11): p. 928-34.
32. Coelho, C., Gonzalez, P.J., Moura, J.G., Moura, I., Trincao, J., and Joao Romao, M., *The Crystal Structure of *Cupriavidus necator* Nitrate Reductase in Oxidized and Partially Reduced States*. *J Mol Biol*, (2011). 408(5): p. 932-48.
33. Najmudin, S., Gonzalez, P.J., Trincao, J., Coelho, C., Mukhopadhyay, A., Cerqueira, N.M., Romao, C.C., Moura, I., Moura, J.J., Brondino, C.D., and Romao, M.J., *Periplasmic nitrate reductase revisited: a sulfur atom completes the sixth coordination of the catalytic molybdenum*. *J Biol Inorg Chem*, (2008). 13(5): p. 737-53.
34. Gonzalez, P.J., Correia, C., Moura, I., Brondino, C.D., and Moura, J.J., *Bacterial nitrate reductases: Molecular and biological aspects of nitrate reduction*. *J Inorg Biochem*, (2006). 100(5-6): p. 1015-23.
35. Stolz, J.F. and Basu, P., *Evolution of nitrate reductase: molecular and structural variations on a common function*. *ChemBiochem*, (2002). 3(2-3): p. 198-206.
36. Ferry, G., *Dorothy Hodgkin, A Life*. 1998, London: Granta Books.
37. Rhodes, G., *Crystallography made crystal clear: a guide for users of macromolecular models*. 1993, California: Academic Press.

38. Rupp, B., *Biomolecular Crystallography*. 2010, New York: Garland Science.
39. Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M., *Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures*. FEBS J, (2008). 275(1): p. 1-21.
40. Chernov, A.A., *Protein crystals and their growth*. J Struct Biol, (2003). 142(1): p. 3-21.
41. Perutz, M., *Protein Structure "New approaches to disease and therapy"* 1992, New York: W.H. Freeman and Company.
42. Matthews, B.W., *Solvent content of protein crystals*. J. Mol. Biol., (1968). 33: p. 491-97.
43. Lesk, A.M., *Introduction to Protein Science*. 2004, Great Britain: Oxford University Press.
44. Vila, R., Kurosaki, M., Barzago, M.M., Kolek, M., Bastone, A., Colombo, L., Salmona, M., Terao, M., and Garattini, E., *Regulation and biochemistry of mouse molybdo-flavoenzymes. The DBA/2 mouse is selectively deficient in the expression of aldehyde oxidase homologues 1 and 2 and represents a unique source for the purification and characterization of aldehyde oxidase*. J Biol Chem, (2004). 279(10): p. 8668-83.
45. Pryde, D.C., Dalvie, D., Hu, Q., Jones, P., Obach, R.S., and Tran, T.D., *Aldehyde oxidase: an enzyme of emerging importance in drug discovery*. J Med Chem, (2010). 53(24): p. 8441-60.
46. Obach, R.S., Huynh, P., Allen, M.C., and Beedham, C., *Human liver aldehyde oxidase: inhibition by 239 drugs*. J Clin Pharmacol, (2004). 44(1): p. 7-19.
47. Garattini, E., Fratelli, M. and Terao, M., *The mammalian aldehyde oxidase gene family*. Hum Genomics, (2009). 4(2): p. 119-30.
48. Terao, M., Kurosaki, M., Saltini, G., Demontis, S., Marini, M., Salmona, M., and Garattini, E., *Cloning of the cDNAs coding for two novel molybdo-flavoproteins showing high similarity with aldehyde oxidase and xanthine oxidoreductase*. J Biol Chem, (2000). 275(39): p. 30690-700.
49. Kurosaki, M., Terao, M., Barzago, M.M., Bastone, A., Bernardinello, D., Salmona, M., and Garattini, E., *The aldehyde oxidase gene cluster in mice and rats. Aldehyde oxidase homologue 3, a novel member of the molybdo-flavoenzyme family with selective expression in the olfactory mucosa*. J Biol Chem, (2004). 279(48): p. 50482-98.
50. Terao, M., Kurosaki, M., Barzago, M.M., Varasano, E., Boldetti, A., Bastone, A., Fratelli, M., and Garattini, E., *Avian and canine aldehyde oxidases. Novel insights into the biology and evolution of molybdo-flavoenzymes*. J Biol Chem, (2006). 281(28): p. 19748-61.
51. Kurosaki, M., Demontis, S., Barzago, M.M., Garattini, E., and Terao, M., *Molecular cloning of the cDNA coding for mouse aldehyde oxidase: tissue distribution and regulation in vivo by testosterone*. Biochem J, (1999). 341 (Pt 1): p. 71-80.
52. Mahro, M., Coelho, C., Trincao, J., Rodrigues, D., Terao, M., Garattini, E., Saggiu, M., Lenzian, F., Hildebrandt, P., Romão, M.J., and Leimkühler, S., *Characterization and crystallization of mouse Aldehyde Oxidase 3 (mAOX3): from mouse liver to E.coli heterologous protein expression*. Drug Metabolism and Disposition, (2011).
53. Schumann, S., Terao, M., Garattini, E., Saggiu, M., Lenzian, F., Hildebrandt, P., and Leimkühler, S., *Site directed mutagenesis of amino acid residues at the active site of mouse aldehyde oxidase AOX1*. PLoS One, (2009). 4(4): p. e5348.
54. Weigert, J., Neumeier, M., Bauer, S., Mages, W., Schnitzbauer, A.A., Obed, A., Groschl, B., Hartmann, A., Schaffler, A., Aslanidis, C., Scholmerich, J., and Buechler, C., *Small-interference RNA-mediated knock-down of aldehyde oxidase 1 in 3T3-L1 cells impairs adipogenesis and adiponectin release*. FEBS Lett, (2008). 582(19): p. 2965-72.
55. Massey, V. and Edmondson, D., *On the mechanism of inactivation of xanthine oxidase by cyanide*. J Biol Chem, (1970). 245(24): p. 6595-8.
56. Wahl, R.C., Warner, C.K., Finnerty, V., and Rajagopalan, K.V., *Drosophila melanogaster ma-1 mutants are defective in the sulfuration of desulfo Mo hydroxylases*. J Biol Chem, (1982). 257(7): p. 3958-62.

57. Temple, C.A., Graf, T.N. and Rajagopalan, K.V., *Optimization of expression of human sulfite oxidase and its molybdenum domain*. Arch Biochem Biophys, (2000). 383(2): p. 281-7.
58. Carson, M., Johnson, D.H., McDonald, H., Brouillette, C., and Delucas, L.J., *His-tag impact on structure*. Acta Crystallogr D Biol Crystallogr, (2007). 63(Pt 3): p. 295-301.
59. Leslie, A.G.W., *Recent changes to the MOSFLM package for processing film and image plate data* Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, (1992). No. 26.
60. Long, F., Vagin, A.A., Young, P., and Murshudov, G.N., *BALBES: a molecular-replacement pipeline*. Acta Crystallogr D Biol Crystallogr, (2008). 64(Pt 1): p. 125-32.
61. Okamoto, K., Eger, B.T., Nishino, T., and Pai, E.F., *Mechanism of inhibition of xanthine oxidoreductase by allopurinol: crystal structure of reduced bovine milk xanthine oxidoreductase bound with oxipurinol*. Nucleosides Nucleotides Nucleic Acids, (2008). 27(6): p. 888-93.
62. Emsley, P. and Cowtan, K., *Coot: model-building tools for molecular graphics*. Acta Crystallogr D Biol Crystallogr, (2004). 60(Pt 12 Pt 1): p. 2126-32.
63. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J., *Phaser crystallographic software*. J Appl Crystallogr, (2007). 40(Pt 4): p. 658-74.
64. Rebelo, J.M., Dias, J.M., Huber, R., Moura, J.J., and Romao, M.J., *Structure refinement of the aldehyde oxidoreductase from Desulfovibrio gigas (MOP) at 1.28 Å*. J Biol Inorg Chem, (2001). 6(8): p. 791-800.
65. Truglio, J.J., Theis, K., Leimkuhler, S., Rappa, R., Rajagopalan, K.V., and Kisker, C., *Crystal structures of the active and alloxanthine-inhibited forms of xanthine dehydrogenase from Rhodobacter capsulatus*. Structure, (2002). 10(1): p. 115-25.
66. Kuwabara, Y., Nishino, T., Okamoto, K., Matsumura, T., Eger, B.T., and Pai, E.F., *Unique amino acids cluster for switching from the dehydrogenase to oxidase form of xanthine oxidoreductase*. Proc Natl Acad Sci U S A, (2003). 100(14): p. 8170-5.
67. McManaman, J.L., Neville, M.C. and Wright, R.M., *Mouse mammary gland xanthine oxidoreductase: purification, characterization, and regulation*. Arch Biochem Biophys, (1999). 371(2): p. 308-16.
68. Vorbach, C., Scriven, A. and Capecchi, M.R., *The housekeeping gene xanthine oxidoreductase is necessary for milk fat droplet enveloping and secretion: gene sharing in the lactating mammary gland*. Genes Dev, (2002). 16(24): p. 3223-35.
69. Huber, R., Hof, P., Duarte, R.O., Moura, J.J., Moura, I., Liu, M.Y., LeGall, J., Hille, R., Archer, M., and Romao, M.J., *A structure-based catalytic mechanism for the xanthine oxidase family of molybdenum enzymes*. Proc Natl Acad Sci U S A, (1996). 93(17): p. 8846-51.
70. Yamaguchi, Y., Matsumura, T., Ichida, K., Okamoto, K., and Nishino, T., *Human xanthine oxidase changes its substrate specificity to aldehyde oxidase type upon mutation of amino acid residues in the active site: roles of active site residues in binding and activation of purine substrate*. J Biochem, (2007). 141(4): p. 513-24.
71. Rocchia, W., Alexov, E. and Honig, B., *Extending the Applicability of the Nonlinear Poisson–Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions*. J. Phys. Chem. B, (2001). 105 (28): p. 6507–14.
72. Delano, W.L., *The PyMOL Molecular Graphics System*. DeLano Scientific LLC, Palo Alto, CA, (2008).
73. Pauff, J.M., Cao, H. and Hille, R., *Substrate Orientation and Catalysis at the Molybdenum Site in Xanthine Oxidase: CRYSTAL STRUCTURES IN COMPLEX WITH XANTHINE AND LUMAZINE*. J Biol Chem, (2009). 284(13): p. 8760-7.

74. Cao, H., Pauuff, J.M. and Hille, R., *Substrate orientation and catalytic specificity in the action of xanthine oxidase: the sequential hydroxylation of hypoxanthine to uric acid*. J Biol Chem, (2010). 285(36): p. 28044-53.
75. Rainer, C., *Genomic view of energy metabolism in Ralstonia eutropha H16*. Journal of molecular microbiology and biotechnology, (2009). 16(1-2): p. 38-52.
76. Vandamme, P. and Coenye, T., *Taxonomy of the genus Cupriavidus: a tale of lost and found*. Int J Syst Evol Microbiol, (2004). 54(Pt 6): p. 2285-9.
77. Warnecke-Eberz, U. and Friedrich, B., *Three nitrate reductase activities in Alcaligenes eutrophus*. Arch. Microbiol., (1993). 159: p. 405-09.
78. Hettmann, T., Siddiqui, R.A., von Langen, J., Frey, C., Romao, M.J., and Diekmann, S., *Mutagenesis study on the role of a lysine residue highly conserved in formate dehydrogenases and periplasmic nitrate reductases*. Biochem Biophys Res Commun, (2003). 310(1): p. 40-7.
79. Berks, B.C., Palmer, T. and Sargent, F., *The Tat protein translocation pathway and its role in microbial physiology*. Adv Microb Physiol, (2003). 47: p. 187-254.
80. Maillard, J., Spronk, C.A., Buchanan, G., Lyall, V., Richardson, D.J., Palmer, T., Vuister, G.W., and Sargent, F., *Structural diversity in twin-arginine signal peptide-binding proteins*. Proc Natl Acad Sci U S A, (2007). 104(40): p. 15641-6.
81. Brige, A., Leys, D., Meyer, T.E., Cusanovich, M.A., and Van Beeumen, J.J., *The 1.25 Å resolution structure of the diheme NapB subunit of soluble nitrate reductase reveals a novel cytochrome c fold with a stacked heme arrangement*. Biochemistry, (2002). 41(15): p. 4827-36.
82. Jepson, B.J., Mohan, S., Clarke, T.A., Gates, A.J., Cole, J.A., Butler, C.S., Butt, J.N., Hemmings, A.M., and Richardson, D.J., *Spectropotentiometric and structural analysis of the periplasmic nitrate reductase from Escherichia coli*. J Biol Chem, (2007). 282(9): p. 6425-37.
83. Cerqueira, N.M., Gonzalez, P.J., Brondino, C.D., Romao, M.J., Romao, C.C., Moura, I., and Moura, J.J., *The effect of the sixth sulfur ligand in the catalytic mechanism of periplasmic nitrate reductase*. J Comput Chem, (2009). 30(15): p. 2466-84.
84. Hofmann, M., *Density functional theory study of model complexes for the revised nitrate reductase active site in Desulfovibrio desulfuricans NapA*. J Biol Inorg Chem, (2009). 14(7): p. 1023-35.
85. Siddiqui, R.A., Warnecke-Eberz, U., Hengsberger, A., Schneider, B., Kostka, S., and Friedrich, B., *Structure and function of a periplasmic nitrate reductase in Alcaligenes eutrophus H16*. J. Bacteriol., (1993). 175: p. 5867-76.
86. Coelho, C., Gonzalez, P.J., Trincao, J., Carvalho, A.L., Najmudin, S., Hettman, T., Dieckman, S., Moura, J.J., Moura, I., and Romao, M.J., *Heterodimeric nitrate reductase (NapAB) from Cupriavidus necator H16: purification, crystallization and preliminary X-ray analysis*. Acta Crystallogr Sect F, (2007). 63(Pt 6): p. 516-9.
87. Kabsch, W., *Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector*. J Appl. Crystallogr., (1988). 21: p. 916-24.
88. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography*. Acta Crystallogr Sect. D., (1994). 50: p. 760-63.
89. Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A., *Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7*. Nat Protoc, (2008). 3(7): p. 1171-9.
90. Murshudov, G.N., Vagin, A.A. and Dodson, E.J., *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallogr D Biol Crystallogr, (1997). 53(Pt 3): p. 240-55.

91. Berks, B.C., Richardson, D.J., Reilly, A., Willis, A.C., and Ferguson, S.J., *The napEDABC gene cluster encoding the periplasmic nitrate reductase system of Thiosphaera pantotropha*. *Biochem J*, (1995). 309 (Pt 3): p. 983-92.
92. Simpson, P.J., McKinzie, A.A. and Codd, R., *Resolution of two native monomeric 90kDa nitrate reductase active proteins from Shewanella gelidimarina and the sequence of two napA genes*. *Biochem Biophys Res Commun*, (2010). 398(1): p. 13-8.
93. Hettmann, T., Siddiqui, R.A., Frey, C., Santos-Silva, T., Romao, M.J., and Diekmann, S., *Mutagenesis study on amino acids around the molybdenum centre of the periplasmic nitrate reductase from Ralstonia eutropha*. *Biochem Biophys Res Commun*, (2004). 320(4): p. 1211-9.
94. Vergnes, A., Gouffi-Belhabich, K., Blasco, F., Giordano, G., and Magalon, A., *Involvement of the molybdenum cofactor biosynthetic machinery in the maturation of the Escherichia coli nitrate reductase A*. *J Biol Chem*, (2004). 279(40): p. 41398-403.
95. Cunha, C.A., Macieira, S., Dias, J.M., Almeida, G., Goncalves, L.L., Costa, C., Lampreia, J., Huber, R., Moura, J.J., Moura, I., and Romao, M.J., *Cytochrome c nitrite reductase from Desulfovibrio desulfuricans ATCC 27774. The relevance of the two calcium sites in the structure of the catalytic subunit (NrfA)*. *J Biol Chem*, (2003). 278(19): p. 17455-65.
96. Leys, D., Tsapin, A.S., Neelson, K.H., Meyer, T.E., Cusanovich, M.A., and Van Beeumen, J.J., *Structure and mechanism of the flavocytochrome c fumarate reductase of Shewanella putrefaciens MR-1*. *Nat Struct Biol*, (1999). 6(12): p. 1113-7.
97. Matias, P.M., Saraiva, L.M., Soares, C.M., Coelho, A.V., LeGall, J., and Carrondo, M.A., *Nine-haem cytochrome c from Desulfovibrio desulfuricans ATCC 27774: primary sequence determination, crystallographic refinement at 1.8 and modelling studies of its interaction with the tetrahaem cytochrome c3*. *J Biol Inorg Chem*, (1999). 4(4): p. 478-94.
98. Carballal, S., Radi, R., Kirk, M.C., Barnes, S., Freeman, B.A., and Alvarez, B., *Sulfenic acid formation in human serum albumin by hydrogen peroxide and peroxyxynitrite*. *Biochemistry*, (2003). 42(33): p. 9906-14.
99. Fujiwara, N., Nakano, M., Kato, S., Yoshihara, D., Ookawara, T., Eguchi, H., Taniguchi, N., and Suzuki, K., *Oxidative modification to cysteine sulfonic acid of Cys111 in human copper-zinc superoxide dismutase*. *J Biol Chem*, (2007). 282(49): p. 35933-44.
100. Lim, J.C., Choi, H.I., Park, Y.S., Nam, H.W., Woo, H.A., Kwon, K.S., Kim, Y.S., Rhee, S.G., Kim, K., and Chae, H.Z., *Irreversible oxidation of the active-site cysteine of peroxiredoxin to cysteine sulfonic acid for enhanced molecular chaperone activity*. *J Biol Chem*, (2008). 283(43): p. 28873-80.
101. Dixon, H., Kellerman, G.M., Mitchell, C.H., Towers, N.H., and Linnane, A.W., *Mikamycin, an inhibitor of both mitochondrial protein synthesis and respiration*. *Biochem Biophys Res Commun*, (1971). 43(4): p. 780-6.
102. McAlpine, A.S., McEwan, A.G., Shaw, A.L., and Bailey, S., *Molybdenum active centre of DMSO reductase from Rhodobacter capsulatus: crystal structure of the oxidised enzyme at 1.82 Å resolution and the dithionite-reduced enzyme at 2.8 Å resolution*. *J. Biol. Inorg. Chem.*, (1997). 2: p. 690-701.
103. Blumberg, W.E. and Peisach, J., *Low-spin compounds of heme proteins*. In *Advances in Chemistry Bioorganic Chemistry Series 100*. 1971: American Chemical Society Publication. 271-84.
104. Zheng, Z. and Gunner, M.R., *Analysis of the electrochemistry of hemes with E(m)s spanning 800 mV*. *Proteins*, (2009). 75(3): p. 719-34.
105. Xie, H. and Cao, Z., *Enzymatic reduction of nitrate to nitrite: insight from density functional calculations*. *Organometallics*, (2010). 29: p. 436-41.
106. Gonzalez, P.J., Barrera, G.I., Rizzi, A.C., Moura, J.J., Passeggi, M.C., and Brondino, C.D., *EPR studies of the Mo-enzyme aldehyde oxidoreductase from Desulfovibrio gigas: an*

- application of the Bloch-Wangsness-Redfield theory to a system containing weakly-coupled paramagnetic redox centers with different relaxation rates.* J Inorg Biochem, (2009). 103(10): p. 1342-6.
107. More, C., Asso, M., Roger, G., Guigliarelli, B., Caldeira, J., Moura, J., and Bertrand, P., *Study of the spin-spin interactions between the metal centers of Desulfovibrio gigas aldehyde oxidoreductase: identification of the reducible sites of the [2Fe-2S]^{1+,2+} clusters.* Biochemistry, (2005). 44(34): p. 11628-35.
 108. Armstrong, F.A., Heering, H.A. and Hirst, J., *Reactions of complex metalloproteins studied by protein film voltammetry.* Chem. Soc. Rev., (1997). 26: p. 169-79.
 109. Leger, C. and Bertrand, P., *Direct electrochemistry of redox enzymes as a tool for mechanistic studies.* Chem Rev, (2008). 108(7): p. 2379-438.
 110. Paes de Sousa, P.M., Pauleta, S.R., Simoes Goncalves, M.L., Pettigrew, G.W., Moura, I., Moura, J.J., and Correia dos Santos, M.M., *Artefacts induced on c-type haem proteins by electrode surfaces.* J Biol Inorg Chem, (2011). 16(2): p. 209-15.
 111. Bertrand, P., Frangioni, B., Dementin, S., Sabaty, M., Arnoux, P., Guigliarelli, B., Pignol, D., and Leger, C., *Effects of slow substrate binding and release in redox enzymes: theory and application to periplasmic nitrate reductase.* J Phys Chem B, (2007). 111(34): p. 10300-11.
 112. Fourmond, V., Burlat, B., Dementin, S., Arnoux, P., Sabaty, M., Boiry, S., Guigliarelli, B., Bertrand, P., Pignol, D., and Leger, C., *Major Mo(V) EPR signature of Rhodobacter sphaeroides periplasmic nitrate reductase arising from a dead-end species that activates upon reduction. Relation to other molybdoenzymes from the DMSO reductase family.* J Phys Chem B, (2008). 112(48): p. 15478-86.
 113. Fourmond, V., Sabaty, M., Arnoux, P., Bertrand, P., Pignol, D., and Leger, C., *Reassessing the strategies for trapping catalytic intermediates during nitrate reductase turnover.* J Phys Chem B, (2010). 114(9): p. 3341-7.
 114. Fujita, K. and Ohno, H., *Enzymatic activity and thermal stability of metallo proteins in hydrated ionic liquids.* Biopolymers, (2010). 93(12): p. 1093-9.
 115. Park, S. and Kazlauskas, R.J., *Biocatalysis in ionic liquids - advantages beyond green technology.* Curr Opin Biotechnol, (2003). 14(4): p. 432-7.
 116. Sheldon, R.A., Lau, R.M., Sorgedraeger, M.J., Rantwijk, F., and Seddon, K.R., *Biocatalysis in ionic liquids.* Green Chemistry, (2002). 4: p. 147-51.
 117. Kragl, U., Eckstein, M. and Kaftzik, N., *Enzyme catalysis in ionic liquids.* Curr Opin Biotechnol, (2002). 13(6): p. 565-71.
 118. Lange, C., Patil, G. and Rudolph, R., *Ionic liquids as refolding additives: N'-alkyl and N'-(omega-hydroxyalkyl) N-methylimidazolium chlorides.* Protein Sci, (2005). 14(10): p. 2693-701.
 119. Summers, C.A. and Flowers, R.A., 2nd, *Protein renaturation by the liquid organic salt ethylammonium nitrate.* Protein Sci, (2000). 9(10): p. 2001-8.
 120. Garlitz, J.A., Summers, C.A., Flowers, R.A., 2nd, and Borgstahl, G.E., *Ethylammonium nitrate: a protein crystallization reagent.* Acta Crystallogr D Biol Crystallogr, (1999). 55(Pt 12): p. 2037-8.
 121. Pusey, M.L., Paley, M.S., Turner, M.B., and Rogers, R.D., *Protein Crystallization Using Room Temperature Ionic Liquids.* Crystal Growth & Design, (2007). 7(4): p. 787-93.
 122. Judge, R.A., Takahashi, S., Longenecker, K.L., Fry, E.H., Abad-Zapatero, C., and Chiu, M.L., *The Effect of Ionic Liquids on Protein Crystallization and X-ray Diffraction Resolution.* Crystal Growth & Design, (2009). 9(8): p. 3463-69.
 123. Cadena, C., Zhao, Q., Snurr, R.Q., and Maginn, E.J., *Molecular modeling and experimental studies of the thermodynamic and transport properties of pyridinium-based ionic liquids.* J Phys Chem B, (2006). 110(6): p. 2821-32.

124. Hardacre, C., Holbrey, J.D., McMath, S.E.J., Bowron, D.T., and Soper, A.K., *Structure of molten 1,3-dimethylimidazolium chloride using neutron diffraction*. J. Chem. Phys., (2003). 118: p. 273-78.

APPENDIX

Table 1 - Mineral medium composition for *C.necator* growth under aerobic conditions

Medium ^a	Final concentration	Amounts 10L Growth	Amounts 300L Growth
MgSO ₄ .7H ₂ O	0.02%	2 g	60 g
CaCl ₂ .2H ₂ O	0.001%	0.1 g	3 g
NH ₄ Cl	0.2%	20 g	600 g
FeCl ₃	0.0005%	0.05 g	1.5 g
Fructose	0.4%	40 g	1200 g
SL6 solution ^b	0.1%	50 ml	300 ml
H16 buffer (Na ₂ HPO ₄ / KH ₂ PO ₄)	-	90 / 15 g	2700 / 450 g
Tetracycline	10 mg/ml	0.1 g	3 g
Antifoam (Simethicone)	-	1 g	30 g

^a All solutions were autoclaved 20 min at 120°C, except the FeCl₃, the Fructose and Tetracycline solutions, that were sterilized using ultrafiltration with 0.2 µm pore diameter.

^b SL6 solution corresponds to an oligoelement solution composed by: ZnSO₄.7H₂O; MnCl₂.4H₂O; H₃BO₃; CoCl₂.6H₂O; CuCl₂.6H₂O; NiCl₂.6H₂O; Na₂MoO₄.

Figure 1 - Comparison between *C.necator* 300L and 10L growths

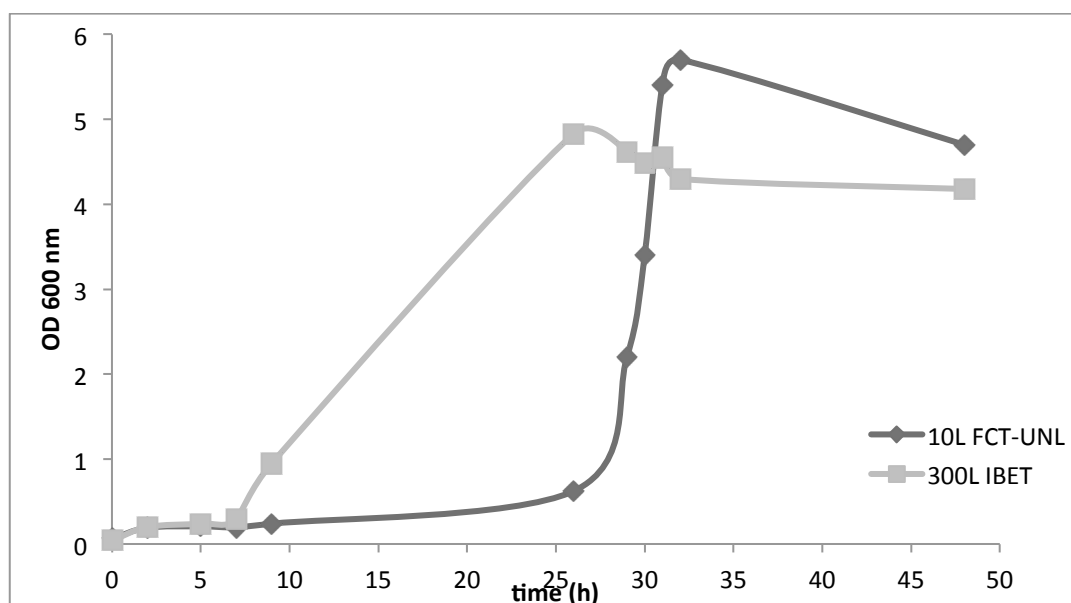


Table 2 - Index Formulation from Hampton Research (2008)

1.	0.1 M Citric acid pH 3.5, 2.0 M Ammonium sulfate
2.	0.1 M Sodium acetate trihydrate pH 4.5, 2.0 M Ammonium sulfate
3.	0.1 M BIS-TRIS pH 5.5, 2.0 M Ammonium sulfate
4.	0.1 M BIS-TRIS pH 6.5, 2.0 M Ammonium sulfate
5.	0.1 M HEPES pH 7.5, 2.0 M Ammonium sulfate
6.	0.1 M Tris pH 8.5, 2.0 M Ammonium sulfate
7.	0.1 M Citric acid pH 3.5, 3.0 M Sodium chloride
8.	0.1 M Sodium acetate trihydrate pH 4.5, 3.0 M Sodium chloride
9.	0.1 M BIS-TRIS pH 5.5, 3.0 M Sodium chloride
10.	0.1 M BIS-TRIS pH 6.5, 3.0 M Sodium chloride
11.	0.1 M HEPES pH 7.5, 3.0 M Sodium chloride
12.	0.1 M Tris pH 8.5, 3.0 M Sodium chloride
13.	0.1 M BIS-TRIS pH 5.5, 0.3 M Magnesium formate dihydrate
14.	0.1 M BIS-TRIS pH 6.5, 0.5 M Magnesium formate dihydrate
15.	0.1 M HEPES pH 7.5, 0.5 M Magnesium formate dihydrate
16.	0.1 M Tris pH 8.5, 0.3 M Magnesium formate dihydrate
17.	1.4 M Sodium phosphate monobasic monohydrate/Potassium phosphate dibasic pH 5.6
18.	1.4 M Sodium phosphate monobasic monohydrate/Potassium phosphate dibasic pH 6.9
19.	1.4 M Sodium phosphate monobasic monohydrate/Potassium phosphate dibasic pH 8.2
20.	0.1 M HEPES pH 7.5, 1.4 M Sodium citrate tribasic dihydrate
21.	1.8 M Ammonium citrate tribasic pH 7.0
22.	0.8 M Succinic acid pH 7.0
23.	2.1 M DL-Malic acid pH 7.0
24.	2.8 M Sodium acetate trihydrate pH 7.0
25.	3.5 M Sodium formate pH 7.0
26.	1.1 M Ammonium tartrate dibasic pH 7.0
27.	2.4 M Sodium malonate pH 7.0
28.	35% v/v Tacsimate pH 7.0
29.	60% v/v Tacsimate pH 7.0
30.	0.1 M Sodium chloride, 0.1 M BIS-TRIS pH 6.5, 1.5 M Ammonium sulfate
31.	0.8 M Potassium sodium tartrate, 0.1 M Tris pH 8.5, 0.5% w/v Polyethylene glycol monomethyl ether 5,000
32.	1.0 M Ammonium sulfate, 0.1 M BIS-TRIS pH 5.5, 1% w/v Polyethylene glycol 3,350
33.	1.1 M Sodium malonate pH 7.0, 0.1 M HEPES pH 7.0, 0.5% v/v Jeffamine ED-2001 pH 7.0
34.	1.0 M Succinic acid pH 7.0, 0.1 M HEPES pH 7.0, 1% w/v Polyethylene glycol monomethyl ether 2,000
35.	1.0 M Ammonium sulfate, 0.1 M HEPES pH 7.0, 0.5% w/v Polyethylene glycol 8,000
36.	15% v/v Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 2% w/v Polyethylene glycol 3,350
37.	25% w/v Polyethylene glycol 1,500
38.	0.1 M HEPES pH 7.0, 30% v/v Jeffamine M-600 pH 7.0
39.	0.1 M HEPES pH 7.0, 30% v/v Jeffamine ED-2001 pH 7.0
40.	0.1 M Citric acid pH 3.5, 25% w/v Polyethylene glycol 3,350
41.	0.1 M Sodium acetate trihydrate pH 4.5, 25% w/v Polyethylene glycol 3,350
42.	0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
43.	0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
44.	0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
45.	0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
46.	0.1 M BIS-TRIS pH 6.5, 20% w/v Polyethylene glycol monomethyl ether 5,000
47.	0.1 M BIS-TRIS pH 6.5, 28% w/v Polyethylene glycol monomethyl ether 2,000
48.	0.2 M Calcium chloride dihydrate, 0.1 M BIS-TRIS pH 5.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
49.	0.2 M Calcium chloride dihydrate, 0.1 M BIS-TRIS pH 6.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
50.	0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 5.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
51.	0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 6.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
52.	0.2 M Ammonium acetate, 0.1 M HEPES pH 7.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol

53.	0.2 M Ammonium acetate, 0.1 M Tris pH 8.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
54.	0.05 M Calcium chloride dihydrate, 0.1 M BIS-TRIS pH 6.5, 30% v/v Polyethylene glycol monomethyl ether 550
55.	0.05 M Magnesium chloride, 0.1 M HEPES pH 7.5, 30% v/v Polyethylene glycol monomethyl ether 550
56.	0.2 M Potassium chloride, 0.05 M HEPES pH 7.5, 35% v/v Pentaerythritol propoxylate (5/4 PO/OH)
57.	0.05 M Ammonium sulfate, 0.05 M BIS-TRIS pH 6.5, 30% v/v Pentaerythritol ethoxylate (15/4 EO/OH)
58.	0.1 M BIS-TRIS pH 6.5, 45% v/v Polypropylene glycol P 400
59.	0.02 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 22% w/v Polyacrylic acid sodium salt 5,100
60.	0.01 M Cobalt(II) chloride hexahydrate, 0.1 M Tris pH 8.5, 20% w/v Polyvinylpyrrolidone K 15
61.	0.2 M L-Proline, 0.1 M HEPES pH 7.5, 10% w/v Polyethylene glycol 3,350
62.	0.2 M Trimethylamine N-oxide, 0.1 M Tris pH 8.5, 20% w/v Polyethylene glycol monomethyl ether 2,000
63.	5% v/v Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 10% w/v Polyethylene glycol monomethyl ether 5,000
64.	0.005 M Cobalt(II) chloride hexahydrate, 0.005 M Nickel(II) chloride hexahydrate, 0.005 M Cadmium chloride hydrate, 0.005 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 12% w/v Polyethylene glycol 3,350
65.	0.1 M Ammonium acetate, 0.1 M BIS-TRIS pH 5.5, 17% w/v Polyethylene glycol 10,000
66.	0.2 M Ammonium sulfate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
67.	0.2 M Ammonium sulfate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
68.	0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
69.	0.2 M Ammonium sulfate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
70.	0.2 M Sodium chloride, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
71.	0.2 M Sodium chloride, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
72.	0.2 M Sodium chloride, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
73.	0.2 M Sodium chloride, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
74.	0.2 M Lithium sulfate monohydrate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
75.	0.2 M Lithium sulfate monohydrate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
76.	0.2 M Lithium sulfate monohydrate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
77.	0.2 M Lithium sulfate monohydrate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
78.	0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
79.	0.2 M Ammonium acetate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
80.	0.2 M Ammonium acetate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
81.	0.2 M Ammonium acetate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
82.	0.2 M Magnesium chloride hexahydrate, 0.1 M BIS-TRIS pH 5.5, 25% w/v Polyethylene glycol 3,350
83.	0.2 M Magnesium chloride hexahydrate, 0.1 M BIS-TRIS pH 6.5, 25% w/v Polyethylene glycol 3,350
84.	0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
85.	0.2 M Magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, 25% w/v Polyethylene glycol 3,350
86.	0.2 M Potassium sodium tartrate tetrahydrate, 20% w/v Polyethylene glycol 3,350
87.	0.2 M Sodium malonate pH 7.0, 20% w/v Polyethylene glycol 3,350
88.	0.2 M Ammonium citrate tribasic pH 7.0, 20% w/v Polyethylene glycol 3,350
89.	0.1 M Succinic acid pH 7.0, 15% w/v Polyethylene glycol 3,350
90.	0.2 M Sodium formate, 20% w/v Polyethylene glycol 3,350
91.	0.15 M DL-Malic acid pH 7.0, 20% w/v Polyethylene glycol 3,350
92.	0.1 M Magnesium formate dihydrate, 15% w/v Polyethylene glycol 3,350
93.	0.05 M Zinc acetate dihydrate, 20% w/v Polyethylene glycol 3,350
94.	0.2 M Sodium citrate tribasic dihydrate, 20% w/v Polyethylene glycol 3,350
95.	0.1 M Potassium thiocyanate, 30% w/v Polyethylene glycol monomethyl ether 2,000
96.	0.15 M Potassium bromide, 30% w/v Polyethylene glycol monomethyl ether 2,000