Kinetic and Structural Studies on Flavin-dependent Enzymes involved in Glycine Betaine Biosynthesis and Propionate 3-nitronate Detoxification

Francesca Salvi
KINETIC AND STRUCTURAL STUDIES ON FLAVIN-DEPENDENT ENZYMES INVOLVED IN GLYCINE BETAINE BIOSYNTHESIS AND PROPIONATE 3-NITRONATE DETOXIFICATION

by

FRANCESCA SALVI

Under the Direction of Giovanni Gadda, PhD

ABSTRACT
Flavin-dependent enzymes are characterized by an amazing chemical versatility and play important roles in different cellular pathways. The FAD-containing choline oxidase from Arthrobacter globiformis oxidizes choline to glycine betaine and retains the intermediate betaine aldehyde in the active site. The reduced FAD is oxidized by oxygen. Glycine betaine is an important osmoprotectant accumulated by bacteria, plants, and animals in response to stress conditions. The FMN-containing nitronate monooxygenase detoxifies the deadly toxin propionate 3-nitronate which is produced by plants and fungi as defense mechanism against herbivores. The catalytic mechanism of fungal nitronate monooxygenase (NMO) was characterized, but little is known about bacterial NMOs.
In this dissertation the crystal structure of choline oxidase in complex with glycine betaine was solved and the roles of the residue F357 in the oxidative half reaction investigated by combination of steady state kinetics, rapid kinetics, pH, mutagenesis, substrate deuterium and solvent isotope effects, viscosity effects and molecular dynamics simulations. Expression trials of human choline dehydrogenase were carried out and a homology model based on choline oxidase was generated. A bacterial nitronate monooxygenase from *Pseudomonas aeruginosa* PAO1 Pa-NMO was kinetically and structurally characterized and four conserved motifs were described that identify Class 1 NMO. Two hypothetical NMOs from *P. aeruginosa* PAO1 and one from *Helicobacter pylori* not carrying the motifs of Class 1 NMO were cloned and tested for nitronate monooxygenase activity.

The crystal structure of choline oxidase in complex with glycine betaine highlighted two different conformations of loop 250-255 at the dimer interface that is proposed to control substrate access to the active site. The side chain of F357 was associated with a slow isomerization in the oxidation of the reduced FAD of choline oxidase. The first crystal structure of an NMO highlighted active site residues for site-directed mutagenesis studies. The gene function prediction for NMO was improved with the four conserved motifs of Pa-NMO. The two hypothetical NMOs from *P. aeruginosa* PAO1 were shown to possess a different enzymatic activity and to identify two distinct classes of enzymes. The hypothetical NMO from *H. pilory* did not exhibit NMO activity and contained iron as cofactor.

INDEX WORDS: Flavin, Choline oxidase, Nitronate monooxygenase, Propionate 3-nitronate, Choline dehydrogenase, Gene function prediction
KINETIC AND STRUCTURAL STUDIES ON FLAVIN-DEPENDENT ENZYMES INVOLVED IN GLYCINE BETAINE BIOSYNTHESIS AND PROPIONATE 3-NITRONATE DETOXIFICATION

by

FRANCESCA SALVI

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the College of Arts and Sciences Georgia State University 2015
KINETIC AND STRUCTURAL STUDIES ON FLAVIN-DEPENDENT ENZYMES
INVOLVED IN GLYCINE BETaine BIOSYNTHeSIS AND PROPIONATE 3-NITROnATE DETOXIFICATION

by

FRANCESCA SALVI

Committee Chair: Giovanni Gadda

Committee: Irene T. Weber
Donald Hamelberg

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2015
DEDICATION

To my husband Roberto Orrù
ACKNOWLEDGEMENTS

I thank my supervisor and my committee members for their support, all my lab members and collaborators for the time spent together. The work of this dissertation was supported by an MBDAF fellowship.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................................................... v

LIST OF TABLES ......................................................................................................................................................... xii

LIST OF FIGURES ........................................................................................................................................................ xiv

LIST OF SCHEMES ........................................................................................................................................................ xix

1 CHAPTER I: INTRODUCTION ........................................................................................................................................... 1

1.1 General features of flavin-dependent enzymes ................................................................................................. 1

1.2 Structural features of selected flavin-dependent enzymes .................................................................................... 10

1.3 Oxidases, monooxygenases, and dehydrogenases .............................................................................................. 17

1.4 Glycine betaine biosynthesis .................................................................................................................................. 32

1.5 Nitronate monooxygenases and the detoxification of propionate 3-nitronate. .............................................................. 41

1.6 Specific goals ........................................................................................................................................................... 52

1.7 References ............................................................................................................................................................ 56

2 CHAPTER II: CRYSTAL STRUCTURE OF CHOLINE OXIDASE IN COMPLEX WITH THE REACTION PRODUCT GLYCINE BETAIN. ............................................................................................................ 74

2.1 Abstract ................................................................................................................................................................. 74

2.2 Introduction .......................................................................................................................................................... 74

2.3 Materials and methods ........................................................................................................................................... 77

2.4 Results ................................................................................................................................................................. 79
3 CHAPTER III: INVOLVEMENT OF F357 IN AN ISOMERIZATION PARTIALLY RATE LIMITING FOR FLAVIN OXIDATION IN CHOLINE OXIDASE.........................98

3.1 Abstract .................................................................................................................98

3.2 Introduction.............................................................................................................99

3.3 Materials and methods..........................................................................................102

3.4 Results.....................................................................................................................107

3.5 Discussion...............................................................................................................121

3.6 Conclusion..............................................................................................................127

3.7 References..............................................................................................................129

4 CHAPTER IV: HUMAN CHOLINE DEHYDROGENASE: MEDICAL PROMISES AND BIOCHEMICAL CHALLENGES.................................................131

4.1 Abbreviations .........................................................................................................131

4.2 Abstract ..................................................................................................................131

4.3 Introduction .............................................................................................................132

4.4 Physiological roles of human CHD .........................................................................135

4.5 Medical relevance of CHD .....................................................................................137
4.6 Subcellular localization of mammalian CHD

4.7 Glucose-Methanol-Choline enzyme oxidoreductase superfamily

4.8 Purification of CHD

4.9 Biochemical properties of CHD

4.10 Homology model of human CHD

4.11 Comparison of CHD with CHO

4.12 Conclusions

4.13 References

5 CHAPTER V: EXPRESSION TRIALS OF RECOMBINANT HUMAN CHOLINE DEHYDROGENASE

5.1 Abbreviations

5.2 Abstract

5.3 Introduction

5.4 Materials and methods

5.5 Results and discussion

5.6 Conclusions

5.7 References

6 CHAPTER VI: THE COMBINED STRUCTURAL AND KINETIC CHARACTERIZATION OF A BACTERIAL NITRONATE MONOOXYGENASE FROM Pseudomonas aeruginosa PAO1 ESTABLISHES NMO CLASS I AND II
6.1 Abbreviations ........................................................................................................188
6.2 Abstract ...................................................................................................................188
6.3 Introduction ...........................................................................................................189
6.4 Experimental procedures ..................................................................................192
6.5 Results ..................................................................................................................197
6.6 Discussion ............................................................................................................209
6.7 Acknowledgements ............................................................................................217
6.8 References ..........................................................................................................217

7 CHAPTER VII: PA1024 FROM Pseudomonas aeruginosa PAO1 IS A
NADH:FERRICYANIDE REDUCTASE AND NOT A NITRONATE
MONOOXYGENASE....................................................................................................223

7.1 Abbreviations .......................................................................................................223
7.2 Abstract .................................................................................................................223
7.3 Introduction .........................................................................................................224
7.4 Experimental procedures ..................................................................................226
7.5 Results ..................................................................................................................232
7.6 Discussion ............................................................................................................241
7.7 Conclusions .........................................................................................................248

8 CHAPTER VIII: THE HYPOTHETICAL NITRONATE MONOOXYGENASE
PA0660 FROM Pseudomonas aeruginosa PAO1 ENCODES A DIAPHRASE ..............254
10.1 References: .........................................................................................................................320
LIST OF TABLES

Table 1.1: Family of flavoprotein oxidases. Modified from (82) .............................................. 21
Table 1.2: Families of flavoprotein monooxygenases, adapted from (128) ................................. 23
Table 1.3: Plants and fungi containing free 3-nitropropionic acid or its precursor forms. Adapted from (197) .................................................................................................................. 43
Table 2.1: X-ray diffraction data collection and model refinement statistics. .............................. 81
Table 3.1. Kinetic parameters of choline oxidase wild-type (1, 28) and variant enzymes prepared in this study at pH 10.0 and 25 °C with choline as a substrate................................................................................ 112
Table 3.2. Multiple deuterium kinetic isotope effects on the kinetic parameter $k_{cat}/K_{ox}$ at pL 7.0 and 25 °C of wild-type choline oxidase (7) and of the double variant enzyme M62A/F357A prepared in this study. .......................................................................................................................... 114
Table 3.3. Pre-steady-state kinetic parameters of variant enzymes prepared in this study and choline oxidase wild-type (1, 28) at pH 10.0 and 25 °C with choline as a substrate. ...... 116
Table 3.4: Apparent steady-state kinetic parameters for choline at atmospheric oxygen, 25 °C, in 50 mM sodium phosphate pH 7.0 ........................................................................................................... 118
Table 4.1: Medical relevance of human CHD .................................................................................... 138
Table 4.2: Purification attempts of CHD from various cellular sources ........................................ 144
Table 4.3: GMC enzyme superfamily cofactor content; $^a$PEG, polyethylene glycol; $^b$n.d., not determined ........................................................................................................................................ 146
Table 4.4: Proposed roles for putative active site residues of human CHD ................................. 152
Table 5.1: Oligonucleotide primers used for PCR amplification of hCHD ................................. 168
Table 6.1: Substrate specificity of Pa-NMO in 50 mM potassium phosphate, pH 7.5, 30 °C, and atmospheric oxygen (i.e., 0.23 mM) ........................................................................................................ 199
Table 6.2: X-ray diffraction data collection and model refinement statistics. .........................202
Table 6.3: Conserved motifs in the NMO enzyme family. .............................................207
Table 7.1: Oligonucleotide primers used for PCR amplification of \textit{pa1024}. ..................227
Table 7.2: Apparent steady-state kinetic parameters of PA1024 with potassium ferricyanide or menadione as electron acceptors. ........................................................................................................237
Table 7.3: Conserved motifs in the protein sequence of PA1024. ....................................239
Table 8.1: Conserved motifs in the protein sequence of PA0660. ....................................270
Table 9.1: Oligonucleotide primers used for PCR amplification of \textit{hp0773}. ....................284
Table 9.2: Inductively coupled plasma optical emission spectrometry (ICP-OES) of purified HP0773. ...................................................................................................................................292
LIST OF FIGURES

Figure 1.1: Some of the different biological roles of flavoproteins (modified from (6)) ..........1
Figure 1.2: UV-visible absorption spectra of different ionization states of the flavin cofactor in
glucose oxidase (Panel A) and of D-amino acid oxidase (Panel B). Adapted from \(^{(1)}\) and \(^{(27)}\)......5
Figure 1.3: Crystal structures of proteins with the most common FMN- and FAD- binding
domains. ........................................................................................................................................12
Figure 1.4: Overall fold of GMC members. ..................................................................................16
Figure 1.5: Bent conformation of the oxidized FAD in monoamine oxidase B (PDB 1GOS). ....17
Figure 1.6: Reaction catalyzed by D-arginine dehydrogenase. ..................................................26
Figure 1.7: Reaction catalyzed by dihydroorotate dehydrogenase. ...........................................28
Figure 1.8: Reaction the reaction catalyzed by acyl-CoA dehydrogenases. ................................29
Figure 1.9: Binding site of OpuC for glycine betaine (PDB 3PPP).............................................35
Figure 1.10: Structure of choline oxidase. ....................................................................................39
Figure 1.11: Biochemical mechanism of succinate dehydrogenase inhibition. .......................46
Figure 1.12: Oxidation mechanism of propionate 3-nitronate in NMOs. .................................49
Figure 2.1: The hydrophobic cluster on the solvent accessible surface of choline oxidase........77
Figure 2.2: Interactions of glycine betaine in the active site of choline oxidase and the
conformation of the FAD cofactor in the structure of choline oxidase presented in this study. ...82
Figure 2.3: Different conformations of loop 250-255 in the crystal structures of choline oxidase
in complex with glycine betaine and with DMSO (PDB 2JBV). ..................................................84
Figure 2.4: F253 positions in open and closed conformations of loop 250-255. .......................85
Figure 2.5: CAVER \(^{(37)}\) analysis of tunnels from the active site to the enzyme surface..........87
Figure 3.1: Entrance to the active site of choline oxidase.........................................................101
Figure 3.2: UV-visible absorbance spectra of the enzymes M62A/F357A (A), F357A (B), M62A (C), and wild-type choline oxidase (D). ................................................................. 108

Figure 3.3. Solvent viscosity effect on $k_{cat}/K_{ox}$ at pH 10.0 and 25 °C, with glycerol as viscosogen from 4.5% to 24% v/v. ......................................................................................................................... 110

Figure 3.4. Steady-state kinetics of enzymes M62A/F357A, M62A, and F357A performed at 25 °C in 50 mM sodium pyrophosphate pH 10.0 ......................................................................................................................... 111

Figure 3.5: pH dependence of the kinetic parameter $k_{cat}/K_{ox}$ determined at 25 °C for the double variant M62A/F357A (solid circles) and for wild-type choline oxidase$^{(i)}$ (empty circles). ....... 114

Figure 3.6. Anaerobic reduction of the variant enzyme M62A/F357A with choline or deuterated choline as substrate. ......................................................................................................................... 115

Figure 3.7 Probability distributions of the side-chain torsional angles for M62, L65, V355, F357, and M359 belonging to the hydrophobic cluster of the dimeric form of choline oxidase (PDB entry 4MJW) without substrate bound. Depicted are the chi 1 angles for wild-type choline oxidase (red), F357A variant (blue), and M62A variant (green)......................................................... 120

Figure 3.8 Newly proposed gating mechanism of the dimeric form of choline oxidase, involving F357 and E312, presumed to control an entry site of oxygen into the active site. .................. 121

Figure 4.1: Phylogenetic tree......................................................................................................................... 132

Figure 4.2: Physiological roles of human CHD in catabolism of choline and synthesis of glycine betaine. CHD: choline dehydrogenase. ......................................................................................... 135

Figure 4.3: Overall structures and active sites of members of the GMC oxidoreductase enzyme superfamily. ......................................................................................................................... 142

Figure 4.4: Alignment of the amino acid sequences of human CHD and A. globiformis CHO. 149

Figure 4.5: Homology model of human CHD and comparison of its active site with CHO. ..... 150
Figure 4.6: Timeline of biochemical and medical interest on CHD. .................................................. 154

Figure 5.1: Location of human choline dehydrogenase gene on chromosome 3. ..................... 164

Figure 5.2: Synthesized gene coding for hCHD with codon usage optimized for recombinant expression in E. coli. ................................................................................................................. 167

Figure 5.3: Multiple sequence alignment with Clustal W2 of protein sequences of choline dehydrogenase enzymes from different sources. ......................................................................................................................... 174

Figure 5.4: Design of five deletion mutants at the N terminal of human choline dehydrogenase. .................................................................................................................................................. 175

Figure 5.5: Three dimensional model of hCHD shown as gray cartoon with the FAD cofactor as yellow sticks. ........................................................................................................................................... 176

Figure 5.6: SDS-PAGE analysis of the crude extracts of small scale expression trials. .......... 179

Figure 5.7: SDS-PAGE analysis of the expression trial in Rosetta(DE3)pLysS ...................... 181

Figure 5.8: Elution samples from the affinity column. .............................................................. 183

Figure 6.1: SDS-PAGE analysis of recombinant purified Pa-NMO. ........................................ 197

Figure 6.2: Anaerobic reduction of purified Pa-NMO with 1 mM P3N. ................................. 198

Figure 6.3: Oxygen consumption during turnover of purified Pa-NMO with 1 mM 3NPA or 1 mM P3N. ........................................................................................................................................... 199

Figure 6.4: Steady-state kinetics of Pa-NMO with P3N as substrate at 30 °C and pH 7.5 ....... 200

Figure 6.5: Overall structure of Pa-NMO. .................................................................................. 203

Figure 6.6: The active site of Pa-NMO from different views. .................................................. 205

Figure 6.7: Multiple sequence alignment of protein sequences annotated as NMO enzymes from both prokaryotic and eukaryotic sources. .............................................................................. 206
Figure 6.8: Neighbor joining tree based on percentage of identity of enzymes annotated as NMO from both prokaryotic and eukaryotic sources. .................................................................207

Figure 6.9: Representation of the four conserved motifs identified in this study in the crystal structure of Pa-NMO.....................................................................................................................208

Figure 7.1: Overall three dimensional structure of the gene product PA1024 (PDB code 2GJL) shown as cyan cartoon, with the FMN cofactor in yellow sticks. ......................................................225

Figure 7.2: Panel A: SDS-PAGE analysis of purified PA1024 (lane 2). ........................................232

Figure 7.3: Oxygen consumption monitored with 1 mM P3N or 3-NPA as substrate. ............233

Figure 7.4: Anaerobic reduction with NADH of PA1024 in 20 mM potassium phosphate, 200 mM sodium chloride, pH 7.0 and 25 °C. .................................................................235

Figure 7.5: Apparent steady-state kinetics of PA1024 at saturating NADH concentration (100 µM) by varying the concentrations of potassium ferricyanide (panel A) and menadione (panel B)...........................................................................................................................236

Figure 7.6: Conserved motifs highlighted in the crystal structure of PA1024. .......................238

Figure 7.7: Comparison of the motifs on the crystal structures of PA1024 and Pa-NMO. ......240

Figure 7.8: Multiple sequence alignments of protein sequences from prokaryotic and eukaryotic sources currently annotated as hypothetical nitronate monoxygenases.........................241

Figure 8.1: Panel A shows the SDS-PAGE analysis of purified PA0660 (lane 1).......................263

Figure 8.2: Time-resolved absorption spectra of the anaerobic reduction of PA0660 with 200 µM NADPH. .........................................................................................................................265

Figure 8.3: Anaerobic reduction of PA0660 by NAD(P)H in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, 25 °C. .................................................................266
Figure 8.4: Time resolved absorption spectra of PA0660 mixed with stoichiometric NADPH at atmospheric oxygen at 25 °C. .......................................................... 269

Figure 8.5: Multiple sequence alignment of hypothetical nitronate monooxygenases carrying the conserved motifs identified in the protein sequence of PA0660. .................................................. 271

Figure 9.1: SDS-PAGE analysis of samples from the purification by affinity chromatography of recombinant HP0773. ........................................................................................................ 290

Figure 9.2: UV-visible absorption spectrum of purified HP0773 in 50 mM HEPES-Na, pH 8.0, 10% v/v glycerol, 300 mM NaCl at 25 °C. ............................................................................. 291

Figure 9.3: Mass spectrometric analysis of the extracted flavin cofactor of HP0773. .......... 291

Figure 9.4: Monitoring of oxygen consumption for turnover of HP0773 with 10 mM P3N and 10 mM 3NPA. .................................................................................................................. 293

Figure 9.5: Monitoring of the absorbance at 340 nm of NADPH in turnover with HP0773 at atmospheric oxygen. ........................................................................................................ 293

Figure 9.6: Homology model of HP0773. .............................................................................. 295

Figure 9.7: Protein BLAST search of the amino acid sequence of HP0773 against the Swissprot database. .................................................................................................................. 298
LIST OF SCHEMES

Scheme 1.1: Chemical structures of riboflavin, FMN, and FAD, as adapted from reference (1) ....2

Scheme 1.2: Redox and ionization states of the isoalloxazine ring, as modified from reference (27) .................................................................4

Scheme 1.3: Two half reactions of the catalytic cycle of flavoenzymes catalyzing redox reactions ..............................................................................................................................................................................................................................................7

Scheme 1.4: Schematic representation of recurrent structural features surrounding the flavin cofactor in flavoenzymes performing CH dehydrogenation reactions. Adapted from (3) .............14

Scheme 1.5: Reaction of flavin hydroquinones with oxygen and different fates (pathways a and b) of the radical pair. Adapted from (119) and (117). ..................................................................................................................18

Scheme 1.6: Recurrent structural features related with oxygen reactivity observed in flavin-dependent oxidases and monooxygenases. Modified from (117) and (119). ........................................................20

Scheme 1.7: Molecular structures of choline, betaine aldehyde, glycine betaine. ......................32

Scheme 1.8: Catalytic mechanism of choline oxidation to glycine betaine by choline oxidase from A. globiformis. ..................................................................................................................................................37

Scheme 1.9: Chemical structure of 3-nitropropionic acid and of two nitroalkanes and their conjugated bases. ........................................................................................................................................................................42

Scheme 1.10: Branched mechanism for reaction of nitronate monooxygenase from C. saturnus with propionate 3-nitronate. .................................................................................................................................................................51

Scheme 2.1: The oxidation of choline to glycine betaine by choline oxidase. .........................75

Scheme 3.1: Choline oxidation followed by the reaction of the anionic hydroquinone with oxygen. ........................................................................................................................................................................99

Scheme 4.1: The three biosynthetic pathways for glycine betaine from choline. ....................134
Scheme 4.2: Choline metabolism. .................................................................136

Scheme 6.1: Schematic representation of the interactions established by the FMN cofactor with Pa-NMO. ........................................................................................................211

Scheme 9.1: Detoxification of the mitochondrial toxin P3N by the FMN-dependent NMO to malonic semialdehyde (MSA), which can be converted to acetyl-CoA (8) ..........................................................282

Scheme 9.2: Potential substrates tested for the enzymatic activity of HP0773 ........................................299
CHAPTER 1: INTRODUCTION

1.1 General features of flavin-dependent enzymes

The research of this thesis is entirely performed on different enzymatic systems belonging to the category of flavoproteins. This particular category of enzymes requires a flavin cofactor for catalysis which is remarkably versatile in the types of chemistry of the reaction catalyzed, with the ability to act both as nucleophile and electrophile and to perform one- and two-electron transfers \(^1, 2\). This catalytic versatility and the presence of a cofactor that can be used as a spectroscopic probe inside the protein render flavoenzymes an attractive and unique tool for mechanistic studies \(^3, 4\). The question of the role of the protein environment on the accurate fine-tuning of the properties of the flavin cofactor is a key point of mechanistic and structural investigations of flavoproteins, and it is still not fully understood \(^3\). Since their first discovery and initial characterization in the 1930s \(^1\) flavoproteins have been related to diverse biological roles (Figure 1.1) among which energy production, biodegradation, chromatin remodeling, DNA repair, apoptosis, protein folding, xenobiotic detoxification, and neural development \(^5, 6\), and they represent 1-3 % of prokaryotic and eukaryotic genes \(^7\).

![Figure 1.1: Some of the different biological roles of flavoproteins (modified from (6))](image-url)
The building block of the flavin cofactor is riboflavin and two derivatives of this vitamin are naturally occurring in flavoproteins: flavin adenine dinucleotide (FAD), and to a lesser extent, flavin mononucleotide (FMN), \(^{(8)}\) as shown in Scheme 1.1.

**Scheme 1.1**: Chemical structures of riboflavin, FMN, and FAD, as adapted from reference \(^{(1)}\)

The precursor riboflavin (vitamin B\(_2\)) is an essential nutrient for animals, while it is synthesized by bacteria fungi, and plants \(^{(8-10)}\). Therefore in animals riboflavin must be obtained from the diet, with the main sources being milk, dairy, and meat products, followed by cereal, fish, and certain fruits and vegetables \(^{(10-12)}\). Riboflavin deficiency has been described as a consequence of a diet poor in vitamin B\(_2\) and it leads to serious medical consequences, such as anemia \(^{(13)}\), impaired iron absorption \(^{(14)}\), and an increased risk of cardiovascular diseases \(^{(10)}\). Free riboflavin, either present in the food or generated through FAD hydrolysis by the phosphatases of
enterocytes, is absorbed mainly in the small intestine \(^{(12, 15)}\) via a transporter that starts showing a saturating behavior at \(\approx 30\) mg of riboflavin per meal \(^{(16)}\). Riboflavin is transported in blood bound to albumin and to certain immunoglobulins \(^{(12)}\), and is then processed by riboflavin kinase to yield FMN, which is in turn substrate of FAD synthetase to produce FAD \(^{(10)}\). Most flavoproteins (\(\approx 90\%\)) bind the FAD or FMN cofactor tightly but noncovalently \(^{8}\). The remaining group of flavoproteins (\(\approx 10\%\)) displays a covalent attachment of the cofactor via either a link between the N1 or N3 atom of a histidine and the 8\(\alpha\) atom of the isoalloxazine ring or a link between the thiol group of a cysteine and the 8 or 6-position of the isoalloxazine ring \(^{(8)}\). The role of the covalent linkage of the flavin cofactor has been investigated \(^{(17)}\) and site-directed mutagenesis studies have shown its involvement in modulation of the properties of the cofactor \(^{(18, 19)}\). The reactive moiety of the cofactor is the 7,8-dimethyl-isoalloxazine ring (Scheme 1.1), which is an amphipathic molecule, with the xylene ring representing the hydrophobic moiety and the pyrimidine ring the hydrophilic one \(^{(1, 3)}\). The isoalloxazine ring is an electron-deficient heterocycle and therefore can undergo a nucleophilic attack, specifically at positions C4a, N5, and C6 (Scheme 1.1), with C4a and N5 being kinetically relevant, while C6 is usually involved in the formation of covalent linkage of the flavin cofactor rather than catalysis \(^{(5)}\). The electron distribution of the isoalloxazine ring can be significantly perturbed by the interactions between the cofactor and the protein environment, due to the high polarizability of this tricyclic system \(^{(5)}\). Naturally occurring modified versions of the isoalloxazine moiety have been reported in literature \(^{(15, 20, 21)}\), such as in the case of 6-OH or 8-OH-FAD in pig liver glycolate oxidase \(^{(22)}\), electron transfer flavoprotein \(^{(22, 23)}\), human apoptosis-inducing protein AMID \(^{(24)}\), and of 6-OH-FMN in trimethylamine dehydrogenase \(^{(25)}\). Modifications on the isoalloxazine ring drastically alter the reactivity of the flavin cofactor and in most cases the 6-OH- and 8-OH-flavins are
catalytically inactive (20). Synthetic flavin analogs have been developed since the mid-1960s as tool for deflavination/reconstitution studies to create spectroscopic, chemically active, or mechanistic probes such as 8-mercaptoflavin, or 8-halogen-substituted flavins (2). This approach has been applied to gain information regarding flavoprotein folding and mechanism (26). As introduced above, a flavin cofactor can have different oxidation and ionization states, as shown in Scheme 1.2 (1, 5).

Scheme 1.2: Redox and ionization states of the isoalloxazine ring, as modified from reference (27)

The one-electron reduced form can exist in neutral form, called blue semiquinone, or in anionic form with the negative charge localized on the N1 locus in the case of the red semiquinone (1, 5). Despite the fact that the semiquinoid form is known to be unstable in solution (28), the protein environment in certain enzymes is able to stabilize this particular state of the cofactor, as described in nitronate monooxygenases (29). The pK<sub>a</sub> value of 8.5 for this dissociation equilibrium can also be significantly perturbed by the protein surrounding of the enzymes that are able to stabilize the semiquinoid form of the flavin cofactor (1). As a result only the blue or red semiquinone may be observed for a specific enzyme in the pH range where the enzyme is
stable \(^{(1)}\). The two-electron reduced form of flavins, called hydroquinone, has also been described in the neutral or the anionic form with a pK\(_a\) in solution of 6.5 \(^{(1, 5)}\). A significant tool of the mechanistic study of flavoproteins is the possibility to exploit the spectroscopic properties of the flavin cofactor as a probe naturally buried in the protein scaffold \(^{(1)}\). Free flavins in solution have intrinsic and well-characterized absorbance and fluorescence properties due to the chromophore isoalloxazine, which are perturbed to a different extent in any enzyme by the protein environment \(^{(30-34)}\). The different redox and ionization states described above (Scheme 1.2) have different absorption spectra, as shown by the example of glucose oxidase \(^{(1)}\) and of D-amino acid oxidase \(^{(27)}\) in Figure 1.2, that allow monitoring the state of the cofactor during catalysis \(^{(1)}\).

![Figure 1.2](image)

**Figure 1.2:** UV-visible absorption spectra of different ionization states of the flavin cofactor in glucose oxidase (Panel A) and of D-amino acid oxidase (Panel B). Adapted from \(^{(1)}\) and \(^{(27)}\).

The oxidized flavin shows two characteristic absorption peaks at 360 and 450 nm due to a higher and lower energy electronic transition \(\pi \rightarrow \pi^*\), respectively \(^{(32, 35)}\). These electronic properties are the cause of the characteristic yellow color of oxidized flavoproteins and of the name given to this class of enzymes (from the Latin word *flavus*, yellow) \(^{(35)}\). The flavin semiquinone has remarkably different spectroscopic features in his anionic form with two absorption peaks in the 370 and 470 nm regions \(^{(36)}\) and in the neutral form with a broad shoulder at 610 nm \(^{(35)}\). The
two-electron reduced form of the cofactor shows the disappearance of the absorbance peak at 450 nm and is essentially colorless \(^{(35)}\). Additionally flavoproteins can display other electronic states called charge transfers which are due to partial charge transfer between a ligand in the active site and one redox form of the flavin cofactor \(^{(27)}\). In particular, charge transfer complexes are formed as a consequence of the appropriate HOMO/LUMO orbitals overlap of either an aromatic, electron rich compound as donor and oxidized flavin as acceptor, or of one electron deficient molecule as acceptor and reduced flavin as donor \(^{(27, 35)}\). The formation of a charge transfer complex is detected with the formation of an absorbance band at wavelengths longer than 500 nm and it often reports on intermediate states relevant for catalysis \(^{(5, 35)}\). Flavin cofactors are also intrinsically fluorescent, with the oxidized form of the cofactor emitting yellow-green fluorescence with the maximum emission at 520 nm \(^{(35)}\). The quantum yield of FAD is \(\approx 12\%\) of the one of FMN due to the quenching effect of the interactions between the isoalloxazine ring and the adenine moiety \(^{(35)}\). The perturbations of the fluorescence properties of the cofactor buried in the enzyme give precious insights on the interaction with the protein environment and changes in the fluorescence of the flavoprotein report on ligand binding, formation of intermediates, and sometimes on reactions not accompanied by absorption changes \(^{(5)}\). Another technique used to study the electronic properties of flavins is represented by vibrational spectroscopy \(^{(37, 38)}\), which uses the C=O stretching vibration as probe and it has been applied to the investigation of flavoproteins \(^{(39, 40)}\).

One of the peculiar features of flavin cofactors is the ability to perform both one-electron and two-electron transfers, which makes them the perfect candidates for a redox switch \(^{(1, 2)}\). As a consequence many flavoproteins (\(\approx 90\%\)) are oxidoreductases \(^{(7)}\) that carry on oxidations, reductions, monooxygenations, and electron transfers \(^{(41)}\). In solution at pH 7.0 the redox
potential for oxidized flavin/hydroquinone is -207 mV, while for the one-electron transfers of oxidized flavin/semiquinone and of semiquinone/hydroquinone the redox potential values are -314 mV and -124 mV, respectively \(^{(42)}\). The protein surrounding of the cofactor modulates the redox potential, and therefore the ability to accept electrons, of the flavin bound to the enzyme \(^{(5)}\).

The catalytic scheme of flavoproteins performing redox reactions proceeds through two half-cycles, the reductive half reaction (RHR) and the oxidative half reaction (OHR), as shown in Scheme 1.3 \(^{(1)}\).

\[
\text{Scheme 1.3: Two half reactions of the catalytic cycle of flavoenzymes catalyzing redox reactions. A and B represent two generic substrates and Fl is the flavin cofactor. Modified from \(^{(1)}\)}
\]

During the reductive half reaction the substrate \(A_{\text{red}}\) is oxidized to \(A_{\text{ox}}\) by reducing the flavin cofactor from its oxidized state \(\text{Fl}_{\text{ox}}\) to the one- or two-electron reduced \(\text{Fl}_{\text{red}}\) \(^{(1)}\). In the second half cycle (OHR) the reduced flavin \(\text{Fl}_{\text{red}}\) is reoxidized by reducing the substrate \(B_{\text{ox}}\) to its reduced form \(B_{\text{red}}\) \(^{(1)}\). In some cases A is represented by the organic substrate that undergoes oxidation and B is either oxygen or another electron acceptor with the function of restoring the oxidized flavin cofactor for turnover \(^{(43)}\). In other cases the flavin becomes reduced by NAD(P)H or other systems (substrate A) in order to be able to reduce the organic substrate B \(^{(43)}\). As mentioned above flavoproteins catalyze not only redox reactions but they can also perform covalent catalysis \(^{(5, 44)}\). Indeed \(\approx10\%\) of flavoproteins not catalyzing redox reactions have been
classified as transferases (4.3%), lyases (2.9%), isomerases (1.4%), and ligases (0.4%) \(^{(7)}\). The establishment of unique criteria to organize all flavoproteins in distinct classes is a difficult challenge. Furthermore, flavoproteins sometime are more complicated systems with the presence of auxiliary redox centers, such as a disulfide, heme, iron-sulfur clusters, thiamine diphosphate, and metals \(^{(4, 7)}\). Different reviews have presented classification schemes based on the number of electrons involved in the catalytic cycle \(^{(27)}\), on structural features \(^{(3)}\), or on the type of reaction catalyzed \(^{(5, 15)}\). The unusual flavoproteins that catalyze reactions with no net redox change are the topic of a review by Bornemann \(^{(45)}\), which summarizes the studies of chorismate synthase, photolyases, and other systems. One of the most common classifications found in literature lists the following four classes based on the reactivity of the reduced flavin with oxygen \(^{(46)}\):

1. **Oxidases**: the reduced cofactor is rapidly oxidized by oxygen with formation of H\(_2\)O\(_2\) and oxidized enzyme, for instance in the case of choline oxidase \(^{(47, 48)}\), glucose oxidase \(^{(49, 50)}\), glycolate oxidase \(^{(51, 52)}\), or D-amino acid oxidase \(^{(53, 54)}\).

2. **Monooxygenases**: the reduced cofactor reacts rapidly with oxygen to form a C4a hydroperoxide intermediate. The collapse of this C4a intermediate results in the insertion of a single oxygen atom in the substrate and in the production of H\(_2\)O and oxidized enzyme. Examples of flavin monooxygenases are phenylacetone monooxygenase \(^{(55)}\), \(p\)-hydroxybenzoate hydroxylase \(^{(56)}\), and tryptophan 7-halogenase \(^{(57)}\).

3. **Electron transferases**: the reduced cofactor reacts slowly with oxygen to yield an oxygen radical paired with the neutral flavoenzyme semiquinone radical in the absence of a physiological electron acceptor. Some well characterized electron
transf erases are flavodoxin\(^{58, 59}\), NADPH-cytochrome P-450 reductase\(^{60, 61}\), and ferredoxin-NADP\(^+\) reductase\(^{62-64}\).

4. **Dehydrogenases:** the reduced cofactor reacts very poorly with oxygen to yield O\(_2^−\) and in the case of true dehydrogenases the reactivity of reduced flavins with oxygen is completely abated. This class of enzymes relies on a physiological electron acceptor other than oxygen for turnover. Examples of dehydrogenases are D-arginine dehydrogenase\(^{65-68}\), and acyl-CoA dehydrogenases\(^{69-72}\).

The research of this thesis deals with enzymatic systems belonging to the classes of oxidases, monooxygenases, and dehydrogenases. Therefore the common features of these three classes are discussed more in details in section 1.3.

The important biological roles and the amazing catalytic versatility of flavoproteins are mirrored in the variety of applications of these enzymes. Many flavoenzymes are important drug targets\(^{73}\) for the design of antimicrobial drug agents, as in the case of thymidylate synthase\(^{74-76}\), UDP-galactopyranose mutase\(^{77, 78}\), or dihydroorotate dehydrogenase\(^{79-81}\). Furthermore \(\approx60\%\) of human flavoproteins are associated with diseases caused by gene alteration and many of them are involved in disfunctions of the mitochondria, peroxisomes, and the endoplasmic reticulum\(^{8}\).

The chemical versatility of the flavin cofactor, the ability to control the reaction of the cofactor with oxygen and the high region- and enantioselectivity of certain flavoenzymes make them suitable candidates for biocatalytic applications\(^{6, 82}\), including protein engineering for enzyme-assisted catalysis\(^{83-85}\) and green chemistry\(^{86}\), or the development of biosensors\(^{87, 88}\) as for instance the ones based on glucose oxidase\(^{89-91}\) and choline oxidase\(^{92-94}\). Many flavoenzymes
can process important pollutants\(^{(95-97)}\), thereby representing a precious asset in environmental chemistry, for instance in soil detoxification processes\(^{(98)}\).

### 1.2 Structural features of selected flavin-dependent enzymes

The first crystal structure of a flavoprotein is the one of flavodoxin in 1972\(^{(99, 100)}\), which represented an important breakthrough in the field. Before this year mechanistic studies on flavoenzymes were somehow hampered by the lack of structural data and had to rely on chemical modifications to identify putative active site residues. Since then, 1524 and 762 crystal structures are deposited in the PDB database with the cofactors FAD and FMN respectively (as of January 2015). The availability of three dimensional structures of flavoproteins from different families, of variant versions and of different states of the same enzyme has enabled to complement the mechanistic data with structural information to look at the catalytic cycle of enzymatic reactions\(^{(3, 7, 9, 101)}\). The comparison of different states of the same enzyme can highlight different conformations that can be addressed by mechanistic studies, as in the case of pyranose-2-oxidase\(^{(102, 103)}\). The structure of protein-ligand complexes in biochemical research plays an essential role in the identification of residues interacting with the substrate/product or in clarifying the binding mode of inhibitors for drug design\(^{(104)}\). The trapping of intermediates in the snapshot of a crystal structure of an enzyme is more challenging to obtain but gives precious insights in the catalytic mechanism of the enzymatic system\(^{(104)}\). Furthermore the availability of crystal structures of flavoproteins has opened the way to computational studies of this class of enzymes which address key issues such as movement of enzyme gates and diffusion of ligands to the active site\(^{(105, 106)}\). Gates can be defined as dynamics structural elements (single residues or loops) that can reversibly transition between an open and closed state, thereby regulating substrate access/product release, solvent access, and control and synchronization of reactions...
An extensive review on gates in 71 enzymes present in the PDB database highlights the importance of molecular dynamics simulations for this type of studies and how in many cases the analysis of the crystal structure alone is not enough to identify gates in enzymes \(^{[107]}\).

The research of this thesis relies on the combined approach of mechanistic, structural, and computational studies on the same enzymatic system, with the aim of gaining an interdisciplinary perspective on flavin-dependent enzymes. As mentioned before flavoenzymes stand out in the comparison with other cofactor dependent enzymes for the variety of reactions performed \(^{(1)}\).

This diversity is maintained in the structural features, as currently 23 structural clans of the PFAM classification are present in the crystal structures of all flavoproteins in the PDB, while for instance only 5 clans have been described for pyridoxal 5’-phosphate-dependent enzymes \(^{(7)}\).

The most common folds for FMN-dependent enzymes are the \((\beta/\alpha)_8\)-barrel (clan TIM_barrel), and the flavodoxin-like fold (clan Flavoprotein), while the most represented topologies in FAD-dependent enzymes belong to the clan NADP_Rossmann, and, to a lesser extent, to the clan FAD_PCMH \(^{(7)}\). Figure 1.3 shows one example of the most common clans, namely old yellow enzyme from \textit{Saccharomyces pastorianus} (PDB 1OYC), flavodoxin from \textit{Desulfovibrio vulgaris} (PDB 1FX1), human glutathione reductase (PDB 3GRS), and uridine diphospho-N-acetylenopyruvylglucosamine reductase from \textit{Escherichia coli} (PDB 1MBT) for the clans TIM_barrel, Flavoprotein, NADP_Rossmann, and FAD_PCMH, respectively \(^{(7)}\).
Figure 1.3: Crystal structures of proteins with the most common FMN- and FAD- binding domains. The clans TIM_barrel, Flavoprotein, NADP_Rossmann, and FAD_PCMH are represented by old yellow enzyme from Saccharomyces pastorianus, PDB 1OYC (panel A), flavodoxin from Desulfovibrio vulgaris, PDB 1FX1 (panel B), human glutathione reductase, PDB 3GRS (panel C), and uridine diphospho-N-acetylenopyruvilglucosamine reductase from Escherichia coli, PDB 1MBT (panel D). Adapted from (7).

Remarkably FMN- or FAD-dependent enzymes that catalyze unusual non redox reactions are also characterized by unique folds, not shared by the other flavoproteins in the PDB, as in the case of chorismate synthase (7).
The comparison of the folding topologies of the crystal structures of flavoproteins present to date in the PDB with the physiological function, when known, has highlighted the fact that there is little or no correlation between folding topology and the type of reaction catalyzed (3). This discrepancy leads to the hypothesis that evolution “recycled” some common folds to develop new types of catalysts, as in the case of D-amino acid oxidase and p-hydroxy-benzoate hydroxylase with same overall three-dimensional structure and different functions (3). On the other hand some flavoproteins with diverse folding topologies carry out the same chemical reaction, such as flavocytochrome b2 and D-amino acid oxidase, thereby drawing the attention to similar architectures of key active site residues to perform similar catalysis (3). The active site of enzymes is designed to perform different tasks: ligand recognition, modulation of the properties of the ligand and of the overall polarity of the active site, and desolvation (3). In the case of flavoproteins the residues of the active sites participate also in recruiting the flavin and in fine tuning its redox potential. Some common features in the environment of the isoalloxazine ring, possibly involved in modulating the reactivity, have been described in a review study by Fraaije and Mattevi (3) (Scheme 1.4), which compares the structures of 13 flavoenzymes catalyzing a dehydrogenation reaction on different substrates.
Scheme 1.4: Schematic representation of recurrent structural features surrounding the flavin cofactor in flavoenzymes performing CH dehydrogenation reactions. Adapted from (3)

In all these structures, ranging from cholesterol oxidase to medium-chain acyl-CoA dehydrogenase, the N1-C2 locus of the flavin is close (≤3.5 Å) to a positive charge represented by the side chains of arginine and lysine, or the N terminus of an α-helix, or a cluster of peptide nitrogens (3). This positive charge was found to be important for binding of the cofactor and for tuning the redox potential of the isoalloxazine ring (3). Another conserved structural feature is the interaction of the N5 atom of the isoalloxazine ring with a hydrogen bond donor on the face of the flavin that is not exposed to the active site cavity. The donor is usually represented by a nitrogen atom of the main or side chain of a neighbor residue and the distance varies from 2.8 to 3.4 Å (3). Despite the fact that such a conserved interaction with a reactive atom of the cofactor most likely plays an important role in catalysis, it has not been clarified yet the specific effect of this geometric arrangement in the active site (3). The respective orientation of the C-H moiety of the molecule involved in the dehydrogenation reaction and of the N5 atom of the flavin is also a recurrent feature in these 13 structures of flavoenzymes catalyzing a dehydrogenation reaction,
with a distance N5-CH and an angle N10-N5-CH varying from 3.0 to 3.8 Å and from 96 to 117 degree, respectively \(^{(3)}\).

One enzyme superfamily involved in dehydrogenations and structurally and kinetically well characterized is the GMC (Glucose-Methanol-Choline) oxidoreductase superfamily, which groups different flavoproteins performing the oxidation of an alcohol to the corresponding aldehyde or carboxylic acid, with 8 members with crystal structures deposited in the PDB database (Figure 1.4) \(^{(108, 109)}\).

These 8 flavoproteins, namely glucose oxidase, pyranose 2-oxidase, cholesterol oxidase, cellobiose dehydrogenase, choline oxidase, aryl-alcohol oxidase, formate oxidase, and pyridoxine 4-oxidase, share a conserved FAD binding domain. However they present significant differences in the substrate binding domain, as expected from the fact that the substrates are structurally unrelated alcohols, such as sugars, cholesterol, and choline. Furthermore some additional structural elements near the entrance of the active site, and therefore possibly involved in substrate access to the active site, are present only in choline oxidase \(^{(105)}\), aryl-alcohol oxidase \(^{(110)}\), and glucose oxidase \(^{(110)}\). According to different mechanistic studies on enzymes belonging to this enzyme superfamily, the common strategy for alcohol oxidation is a hydride transfer mechanism \(^{(108, 111)}\). Consistent with a shared catalytic strategy some common features in active site geometry have been observed, primarily the conserved Asn-His or His-His pair in front of the cofactor (Figure 1.4) \(^{(108)}\).
Figure 1.4: Overall fold of GMC members. 
*Arthrobacter globiformis* choline oxidase (S101A) PDB 3NNE (1), *Brevibacterium sterolicum* cholesterol oxidase PDB 1COY (2), *Phanerochaete chrysosporium* cellobiose dehydrogenase PDB 1KDG (3), *Plerotus eryngii* aryl-alcohol oxidase PDB 3FIM (4), *Aspergillus niger* glucose oxidase PDB 1CF3 (5), *Trametes ochracea* pyranose 2-oxidase PDB 2IGK (6), *Aspergillus oryzae* formate oxidase PDB 3Q9T (7) and *Mesorbium loti* pyridoxine 4-oxidase PDB 3T37 (8). The conserved active site residues His-Asn (A) the pair His-His (B) of choline oxidase and glucose oxidase are shown.

Structural studies on flavoproteins added important knowledge to other biochemical studies. Different conformations of the flavin cofactor could be seen only by X-ray crystallography, such as in the case of monoamine oxidase B, where the oxidized flavin is in a bent conformation.
(Figure 1.5) possibly relevant for catalysis as it exposes the atoms C4a and N5 in a constrained geometry\textsuperscript{(112)}.

**Figure 1.5**: Bent conformation of the oxidized FAD in monoamine oxidase B (PDB 1GOS). FAD is shown as yellow sticks and the $2F_c-F_o$ map is shown in blue contoured at 0.4 e/Å\textsuperscript{3}.

Furthermore, the application of single-crystal Raman spectroscopy\textsuperscript{(113)} on the FMN-dependent enzyme NrdI from *Bacillus cereus*\textsuperscript{(114)} and of microspectrophotometry\textsuperscript{(115)} on the FAD-dependent phenylacetone monooxygenase (PAMO) from *Thermobifida fusca*\textsuperscript{(116)} has paved the way for obtaining snapshots of different states of the cofactor during X-ray data collection. Recently the biochemical characterization of prokaryotic enzymes, and therefore also of bacterial flavoproteins, have been influenced by the rapid development of structural genomic consortia\textsuperscript{(7)}. These consortia have largely contributed to a rapid increase in number of crystal structures of bacterial flavoproteins, but the gap between the high throughput approach for protein crystallization and the lack of further biochemical characterization has led to the presence in the PDB database of protein structures without known function and the related problem of not reliable function prediction\textsuperscript{(7)}.

### 1.3 Oxidases, monooxygenases, and dehydrogenases

As mentioned in section 1.1, three important classes of flavoenzymes are oxidases, monooxygenases, and dehydrogenases. The difference among these three classes is often
referred to as oxygen reactivity of flavins in the literature of this field \(^{(117)}\), because it is related to the ability of the reduced flavin cofactor to activate molecular oxygen. Activating molecular oxygen is a difficult task as the transfer of two electrons from the singlet hydroquinone to the triplet oxygen is spin forbidden \(^{(118)}\). Reduction of oxygen to \(\text{H}_2\text{O}_2\) or \(\text{H}_2\text{O}\) must therefore occur in a stepwise mechanism, with the first electron transfer yielding a caged radical pair flavin semiquinone/oxygen superoxide followed by a spin inversion (Scheme 1.5) \(^{(117)}\).

\[
\text{Scheme 1.5: Reaction of flavin hydroquinones with oxygen and different fates (pathways a and b) of the radical pair. Adapted from \(^{(119)}\) and \(^{(117)}\).}
\]

There are not so many catalysts that can carry on oxygen activation \(^{(117, 118)}\). Reduced flavins (hydroquinones) in solution show a slow reaction with oxygen, with second-order rate constants of 250 M\(^{-1}\) s\(^{-1}\), but the protein environment can enhance it in the case of oxidases and monooxygenases, or completely suppress it in the case of true dehydrogenases \(^{(117)}\). The pathway followed by the caged radical pair flavin semiquinone/oxygen superoxide after the first electron
transfer is different in the case of oxidases and monooxygenases (Scheme 1.5) \(^{(119)}\). In the case of oxidases the flavin semiquinone transfers another electron (associated with a proton transfer) to superoxide anion to yield oxidized flavin and H\(_2\)O\(_2\) (pathway a in Scheme 1.5). The formation of H\(_2\)O\(_2\) as a side product in flavoprotein oxidases is receiving increasing attention for its role in the generation of reactive oxygen species (ROS) \(^{(119)}\), which are involved in ageing and pathological processes \(^{(120)}\). In the case of monooxygenases the fate of the radical pair is to form a highly reactive C4a-(hydro)peroxyflavin intermediate (pathway b in Scheme 1.5), which serves as nucleophile species to insert an oxygen atom in the substrate \(^{(119)}\). The final products after the C4a-(hydro)peroxyflavin intermediate collapses are oxidized flavin and water. The case of pyranose 2-oxidase for which a C4a-(hydro)peroxyflavin intermediate was detected \(^{(121)}\) represents a remarkable exception to these two distinct pathways and raises the question of whether the reaction of hydroquinones in oxidases still proceeds via the formation of a transient C4a-(hydro)peroxyflavin intermediate that is not stabilized enough to be detected. Some common structural features have been linked to the oxygen reactivity of flavoproteins by different site-directed mutagenesis studies \(^{(117, 119)}\). In oxidases, it is important the presence of an apolar site near the flavin, such as Val464 in choline oxidase \(^{(122)}\), and of a positive charge in the active site (Scheme 1.6) \(^{(117)}\). Monooxygenases are characterized by a well-defined hydrophobic cavity, usually not described in oxidases and by a serine or threonine residue with the hydroxyl group contacting the N5 atom of the isoalloxazine ring (Scheme 1.6) \(^{(119)}\).
There is general consensus that oxygen reacts close to the C4a-N5 locus, but with a different orientation in the case of oxidases and monooxygenases (Scheme 1.6), with the edge-on geometry in oxidases being suitable for an electron transfer but not properly aligned for the orbital overlap of the adduct on the C4a position \(^{(119)}\). Accessibility of the C4a-N5 locus of the isoalloxazine ring to oxygen is another important parameter \(^{(119)}\), as shown by site-directed mutagenesis studies of gating residues close to this locus in L-galactono-1,4-lactone dehydrogenase \(^{(123)}\) and aryl-alcohol oxidase \(^{(124)}\).

Flavoprotein oxidases can be grouped in 6 families (Table 1.1) based on sequence and structural similarities \(^{(82)}\), namely Glucose-Methanol-Choline oxidase (GCM), Vanillyl Alcohol oxidase (VAO), Amine Oxidase (AO), Sulfhydryl Oxidase (SO), Acyl-CoA Oxidase (ACO), and 2-HydroxyAcid oxidase (HAO).
Table 1.1: Family of flavoprotein oxidases. Modified from (82).
GMC, Glucose-Methanol-Choline oxidase; VAO, Vanillyl Alcohol Oxidase; AO, Amine Oxidase; SO, Sulfhydryl Oxidase; ACO, Acyl-CoA Oxidase; HAO, HydroxyAcid Oxidase.

<table>
<thead>
<tr>
<th></th>
<th>GMC</th>
<th>VAO</th>
<th>AO</th>
<th>SO</th>
<th>ACO</th>
<th>HAO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cofactor</strong></td>
<td>FAD</td>
<td>FAD</td>
<td>FAD</td>
<td>FAD</td>
<td>FAD</td>
<td>FMN</td>
</tr>
<tr>
<td><strong>Typical substrates</strong></td>
<td>alcohols</td>
<td>alcohols, amines</td>
<td>amines</td>
<td>thiols</td>
<td>acyl-CoA</td>
<td>hydroxyacids</td>
</tr>
<tr>
<td><strong>Example (PDB)</strong></td>
<td>glucose oxidase (1CF3)</td>
<td>vanillyl alcohol oxidase (1VAO)</td>
<td>D-amino acid oxidase (1KIF)</td>
<td>Erv2p sulfhydryl Oxidase (1JRA)</td>
<td>acyl CoA oxidase (1IS2)</td>
<td>glycolate oxidase (1GOX)</td>
</tr>
</tbody>
</table>

The GMC-type oxidases possess a conserved N-terminal FAD-binding domain (Pfam00732), while they display a variable substrate binding domain at the C terminus (82). These enzymes catalyze the oxidation of an alcohol moiety of primary and secondary alcohols, and of sugars (111). Many members of this superfamily have been characterized with structural and mechanistic studies, mainly glucose oxidase, choline oxidase, pyranose 2-oxidase, and aryl alcohol oxidase (82). These studies highlighted that GMC-type enzymes share a common mechanism of deprotonation of the hydroxyl group by a conserved active site base followed by a hydride transfer from the Cα of the substrate to the N5 of the flavin (108). Remarkably the GMC superfamily includes also some dehydrogenases, such as cellobiose dehydrogenase, and enzymes performing non-redox reactions, namely hydroxynitrile lyase (82). This suggests that the classification in the GMC superfamily still needs to be improved and currently there are no motifs known to discriminate between GMC-type oxidases and unusual GMC members (82).

The VAO family is characterized by the presence of a conserved N-terminal FAD-binding domain (Pfam01565) with an estimated 25% of members containing a cofactor covalently linked to a histidine residue (82). VAO-type enzymes, such as alditol oxidase and reticuline oxidase, perform the oxidation of alcohols and amines, the hydroxylation of alkyl-phenols, ether bond
cleavage, and reactions of C-C and C-O bonds forming \(^{(82)}\). It is interesting to note that two versions of an enzyme performing the same reaction, i.e. cholesterol oxidase, belong one to the GMC and the other to the VAO family \(^{(82)}\).

The AO family includes enzymes with similar N-terminal FAD-binding domain (Pfam01593 and Pfam01266) that carry on amine oxidation, with some exceptions of alcohol and thioether oxidases \(^{(82)}\). The catalytic mechanism of enzymes belonging to this family, among which D-amino acid oxidase and human monoamine oxidase, is based on the proton abstraction from the amine group followed by a hydride transfer to the N5 atom of the flavin cofactor. The resulting imine product spontaneously hydrolyzes to aldehyde and amine \(^{(82)}\).

Members belonging to the SO family show an all-\(\alpha\) fold with an atypical FAD binding. The cofactor is indeed in an intermediate conformation between planar and bent and exposes the adenine moiety on the surface \(^{(125)}\). These enzymes participate in the oxidation of cysteine residues to disulfide bonds, and they are divided in the two subclasses Erv-like and Ero-like based on substrate specificity \(^{(82)}\).

The ACO family includes peroxisomal enzymes involved in fatty acid degradation, and characterized by three conserved N-terminal (Pfam02771), middle (Pfam02770), and C-terminal (Pfam00441) domains. They all contain FAD as cofactor and they start the C\(_\alpha\)-C\(_\beta\) oxidation via a C\(_\alpha\) proton abstraction by an active site glutamate, followed by a hydride transfer from the C\(_\beta\) to the N5 atom of the isoalloxazine ring \(^{(82)}\). Interestingly the fungal nitroalkane oxidase (E.C. 1.7.3.1) is grouped in the ACO family \(^{(82)}\). This FAD-dependent enzyme catalyzes the oxidation of primary and secondary nitroalkanes via the abstraction of a proton from the C\(_\alpha\) of the substrate, followed however by a nucleophilic attack of the resulting carbanion to the N5 atom of
the flavin\textsuperscript{(126)}. This covalent adduct collapses after the attack of a water molecule and leads to the formation of reduced flavin and carbonyl product\textsuperscript{(126)}.

Enzymes belonging to HAO family, such as glycolate oxidase, contain non covalent FMN as a cofactor buried in a TIM barrel fold and they perform the oxidation of aromatic and aliphatic 2-hydroxy acids to the corresponding 2-oxo acids\textsuperscript{(82, 111)}. An active site arginine has been shown to be involved in binding the carboxylate moiety of the substrate, and a conserved active site histidine is located less than 5 Å from the isoalloxazine\textsuperscript{(111)}. The reaction seems to proceed via a hydride transfer mechanism despite the fact that previous studies suggested a carbanion mechanism for this family of enzymes\textsuperscript{(82, 111, 127)}. Notable exceptions to this classification of flavoprotein oxidases are pyruvate oxidase, pyridoxal 5'-phosphate oxidase, NADPH oxidase, and xanthine oxidase\textsuperscript{(82)}.

Flavoprotein monooxygenases represent the largest family of flavoenzymes characterized so far\textsuperscript{(7)}, and they can be classified in 8 groups from A to H (Table 1.2), according to structural and functional properties\textsuperscript{(128)}.

\textbf{Table 1.2:} Families of flavoprotein monooxygenases, adapted from\textsuperscript{(128)}.

<table>
<thead>
<tr>
<th>Cofactor</th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>FAD</td>
</tr>
<tr>
<td>B</td>
<td>FAD</td>
</tr>
<tr>
<td>C</td>
<td>FMN</td>
</tr>
<tr>
<td>D</td>
<td>FAD/FMN</td>
</tr>
<tr>
<td>E</td>
<td>FAD</td>
</tr>
<tr>
<td>F</td>
<td>FAD</td>
</tr>
<tr>
<td>G</td>
<td>FAD</td>
</tr>
<tr>
<td>H</td>
<td>FMN</td>
</tr>
</tbody>
</table>
Groups A and B monooxygenases are single-component enzymes with a Rossmann fold that rely on NAD(P)H for reduction of the FAD cofactor prior to the reaction with oxygen to form the C4a-(hydro)peroxyflavin intermediate. Group A members, such as para-hydroxybenzoate hydroxylase, perform hydroxylation and sulfoxidation reactions, and are often involved in the degradation of polyaromatic compounds (128). Group B can be divided in four subgroups, namely Baeyer–Villiger monooxygenases, such as phenylacetone monooxygenase, flavoprotein monooxygenases, such as mammalian flavomonooxygenase, N-hydroxylating monooxygenases, such as ornithine monooxygenase, and YUCCA enzymes that are involved in the growth hormone auxin biosynthetic pathway in plants (128, 129).

The first two subgroups of Group B are characterized by a broad substrate specificity linked to a high regio- and stereo-selectivity and they can carry out the conversion of a keto group to an ester moiety (116), or different types of heteroatom oxygenation (130). The N-hydroxylating monooxygenases are instead characterized by higher substrate specificity and they are involved in the formation of siderophores (131).

Groups C, D, E, and F are represented by two-component systems where reduced flavin is generated by the component with reductase activity and then channeled to the component with monooxygenase activity for the formation of the C4a-(hydro)peroxyflavin intermediate (128). Group C includes FMN-dependent enzymes with a TIM-barrel fold, such as luciferase or alkanesulfonate monooxygenase, that are involved in light emission, desulfonation, sulfoxidation, epoxidation, and hydroxylation reactions (128). Group D are enzymes with an acyl-CoA dehydrogenase fold, such as 4-hydroxyphenylacetate 3-hydroxylase, that catalyze aromatic hydroxylation and N-hydroxylation reactions, while members belonging to group E show a Rossmann fold and perform epoxidation reactions, as in the case of styrene monooxygenase (128).
Group F, which is the last one representing two-component systems, is also characterized by a Rossmann fold and is involved in halogenation reactions, as for example tryptophan 7-halogenase (128). Group G and H have been recently introduced to classify unusual one-component systems that utilize the substrate rather than NAD(P)H as source of electrons to reduce the flavin cofactor (128). Group G is represented by enzymes with a monoamine oxidase fold, as in the case of tryptophan 2-monooxygenase, where the FAD cofactor is reduced by an amino acid substrate, which is converted to an imine intermediate, subsequently decarboxilated to amide (128). The stabilization of C4a-(hydro)peroxyflavin intermediate in this class of enzyme has not been experimentally observed yet. Group H is represented by TIM-barrel FMN-dependent enzymes that can be further divided in two subgroups, based on functional properties (128). The first subgroup is represented by lactate 2-monooxygenase which oxidizes L-lactate to acetate, via an enzyme-pyruvate complex that reacts with oxygen leading to the decarboxylation of pyruvate and to the insertion of oxygen atoms into acetate and water (128). No observation of a stable C4a-(hydro)peroxyflavin intermediate has been reported yet. The second subgroup is represented by the detoxifying enzymes nitronate monooxygenases, which use molecular oxygen and catalyze one-electron transfer from the toxin propionate 3-nitronate or other nitronates to the flavin cofactor (128). This catalytic strategy represents a remarkable exception in the family of flavinmonooxygenases, as it stabilizes the intermediate anionic flavin semiquinone rather than the C4a-(hydro)peroxyflavin (29).

A plethora of mechanistic, structural, and computational studies and reviews have focused on flavin-dependent oxidases and monooxygenases, due also to the different biocatalytic applications. Much less is known instead on flavin-dependent dehydrogenases and no extensive review and classification of this family is available. As mentioned above some GMC-type
enzymes are dehydrogenases, as in the case of choline dehydrogenase and cellobiose dehydrogenase but defined structural features to distinguish GMC oxidases and dehydrogenases have not been identified yet \(^{(82)}\). From site-directed mutagenesis studies of L-galactono-1,4-lactone dehydrogenase it has emerged how some structural features may be involved in preventing oxygen access to the C4a-N5 locus of the flavin in the case of dehydrogenases \(^{(123)}\). Indeed the mutation of an alanine residue close to the C4a-N5 locus of the flavin to glycine increased the ability of the reduced cofactor to react with oxygen 400 folds \(^{(123)}\).

Some flavin-dependent dehydrogenases have been kinetically and structurally characterized, as in the case of D-arginine dehydrogenase, acyl-CoA dehydrogenase, and dihydroorotate dehydrogenase. D-arginine dehydrogenase catalyzes the enantioselective oxidation of D-amino acids via a ping-pong bi-bi mechanism with phenazine methosulfate (PMS) as electron acceptor, while the physiological electron acceptor is still unknown (Figure 1.6) \(^{(66)}\). This enzyme is able to process all D-amino acids except the negatively charged D-aspartate and D-glutamate, with D-arginine being the best substrate \(^{(65)}\).

**Figure 1.6:** Reaction catalyzed by D-arginine dehydrogenase. Scheme of the oxidation of D-amino acids by D-arginine dehydrogenase in panel A, where R represents the side chain of the amino acid. In panel B the two different conformations of Y53 in D-arginine dehydrogenase (PDB
3NYC) are shown as blue sticks, FMN is in yellow and iminoarginine in magenta. The cavity of the active site is shown as grey surface.

Mechanistic studies with pH and deuterium isotope effects combined with a computational approach have established that the enzyme binds preferably the zwitterionic form of the substrate and that reduction of the flavin, which is partially rate limiting for turnover, occurs via hydride transfer after N-H bond cleavage \(^{68, 132}\). The investigation was carried out with D-leucine as substrate as the reduction of the enzyme by D-arginine is too fast to be monitored by a stopped-flow spectrophotometer \(^{68}\). Solvent viscosity studies on \(k_{\text{cat}}/K_m\) with D-leucine are consistent with the presence of a slow isomerization to yield an enzyme-substrate complex competent for catalysis, while the inverse hyperbolic pattern observed on \(k_{\text{cat}}\) shows that a slow isomerization before product release is partially rate limiting \(^{68}\). Crystal structures of D-arginine dehydrogenase ligand free and in complex with the products imino-arginine and imino-histidine have been solved to 1.06 Å, 1.30 Å, and 1.30 Å resolution, respectively \(^{65}\). From the comparison of the crystal structures of the free enzyme and of the enzyme-product complex different conformations of a “lid” loop with Y53 have been identified, with the side chain of Y53 moving from the active site cavity to the protein surface \(^{65}\).

Dihydroorotate dehydrogenases (DHODs) are FMN-dependent enzymes that catalyze the conversion of dihydroorotate to orotate in the pyrimidine biosynthetic pathway \(^{81}\). Class 1 DHODs are cytosolic enzymes that use fumarate (Class 1A) or NAD\(^+\) (Class 1B) as electron acceptor for oxidizing the reduced FMN \(^{81}\). Class 1A DHODs are homodimers, while Class 1B members are heterotetramers that contain FAD and iron-sulfur clusters in addition to FMN \(^{81}\). Class 2 DHODs are membrane-bound enzymes that rely on ubiquinone as electron acceptor for turnover \(^{133}\). Remarkably humans possess only Class 2 DHODs, and Class 1 DHODs are essential for pathogens, therefore representing an important drug target for selective
antimicrobial agents \textsuperscript{(134)}. Furthermore inhibitors of human Class 2 DHOD have been investigated for the treatment of cancer and autoimmune diseases \textsuperscript{(134)}. Reduction of FMN occurs via hydride transfer from the C6 position of DHO to the N5 atom of FMN, triggered by C5 pro-S proton abstraction by the active site base cysteine in Class 1 or serine in Class 2 DHODs \textsuperscript{(81)}.

\textbf{Figure 1.7:} Reaction catalyzed by dihydroorotate dehydrogenase. Scheme of dihydroorotate oxidation in Panel A. Panel B shows a view of the complex of dihydroorotate dehydrogenase with the product orotic acid (PDB 1F76), with FMN in yellow and orotic acid in gray.

Mechanistic studies with multiple deuterium isotope effects on the reductive half reaction of DHODs have established that the deprotonation step and the hydride transfer reaction follow a concerted mechanism in Class 1A \textsuperscript{(81)}, and a stepwise mechanism in Class 2 \textsuperscript{(133)}. An alternative interpretation of the multiple deuterium isotope effects observed in Class 2 DHODs is a concerted mechanism with significant quantum mechanical tunneling \textsuperscript{(133)}.

Crystal structures of bacterial and human DHODs have been solved at high resolution \textsuperscript{(135, 136)} with different structures of complexes with inhibitors that gave precious insights for structure-based design of selective inhibitors \textsuperscript{(136-138)}.

Acyl-CoA dehydrogenases are mitochondrial enzymes involved in fatty acid β-oxidation and their deficiency is associated with an accumulation of fatty acids in the body leading to severe
metabolic disorders with described symptoms as lethargy, encephalopathy, seizure, and apnea \(^{(139-141)}\). Acyl-CoA dehydrogenases involved in fatty acid metabolism are classified based on the substrate specificity in short, medium, long, and very long chain acyl-CoA dehydrogenases (SCAD, MCAD, LCAD, and VLCAD) \(^{(142, 143)}\). Some acyl-CoA dehydrogenases are instead involved in amino acid metabolism, namely iso(3)valeryl-CoA dehydrogenase is involved in leucine metabolism, iso(2)valeryl-CoA dehydrogenase for isoleucine, isobutyryl-CoA dehydrogenase for valine, and glutaryl-CoA dehydrogenase for lysine and tryptophan \(^{(142)}\). Acyl-CoA dehydrogenases are soluble homotetramers with the exception of VLCAD, which is a homodimer bound to the inner mitochondrial membrane \(^{(142)}\). They all perform α-β dehydrogenations of acyl-CoA thioesters via concerted proton abstractions by a conserved glutamate and hydride transfer to FAD, resulting in the insertion of a \textit{trans} double bond in the product \(^{(143)}\). The reduced FAD is oxidized by the electron-transferring flavoprotein (ETF), which is in turn connected to the electron transport chain \(^{(144)}\).

**Figure 1.8:** Reaction the reaction catalyzed by acyl-CoA dehydrogenases. Scheme of the reaction catalyzed by acyl-CoA dehydrogenases (Panel A). Panel B shows the docking site for ETF (PDB 1T9G); the structure of MCAD is shown as gray cartoon, FAD is in yellow sticks, the conserved W166 is in blue spheres, E212 and E359 are represented with green sticks.
Crystal structures of pig MCAD (145), rat SCAD (146), bacterial SCAD (147), human iso(3)valeryl-CoA and iso(2)valeryl-CoA (148) are available, with different complexes with substrates or inhibitors in the case of MCAD (145, 149). The comparison of the crystal structure of MCAD free and bound to ligands highlighted different conformations of the side chains of the catalytic base E376 and of a glutamate and tyrosine residues involved in binding, while no major changes are observed in the tertiary structure (142). The protein complex between acyl-CoA dehydrogenases and ETF has been investigated by pH, ionic strength, and chemical modification, suggesting the importance of electrostatics in the protein-protein interaction (150). A model based on the analysis of the surface of the crystal structure of MCAD identifies a putative binding site for ETF in a cleft lined with charged residues above the cavity where the FAD is buried (143), which was later confirmed in the crystal structure of human MCAD in complex with ETF (PDB 1T9G) (151). A tryptophan residue, conserved as tryptophan or phenylalanine in the other acyl-CoA dehydrogenases, lies in this cleft, aligned with the dimethylbenzene moiety of FAD and possibly plays a role in the electron transfer between the FAD of MCAD and the flavin cofactor of ETF (Figure 1.8) (143, 151). The crystal structure of human MCAD in complex with ETF also supports the importance of electrostatics in the protein-protein interaction, with E212 or E359 on the surface of MCAD interacting with R249 on ETF (Figure 1.8) (151).

Acyl-CoA dehydrogenases represent an important model for the investigation of the structural features related to oxygen reactivity, due to the available crystal structure of their peroxisomal equivalent acyl-CoA oxidases (152). Acyl-CoA oxidases catalyze the same reaction as acyl-CoA dehydrogenases, share the same substrate specificity and mechanistic studies have shown that the reductive half-reaction is very similar (142). Despite the similar active site architecture, the reduced FAD cofactor is able to react with molecular oxygen in the case of acyl-CoA oxidases.
and the main structural differences possibly related to oxygen reactivity have been summarized in a review study by Kim and Miura (142):

- **Desolvation of the active site.** The N-terminal domain and C-terminal domain of acyl-CoA oxidase are rotated 13° compared with MCAD. This movement results in a wider active site cavity and in less hydrogen bond interactions between the FAD and the protein moiety in acyl-CoA oxidase compared to acyl-CoA dehydrogenase. In the case of MCAD, the crystal structure of the enzyme in complex with ligands shows the displacement of water molecules observed instead in the active site of the free enzyme as a consequence of ligand binding. The wider active site cavity in acyl-CoA oxidase allows molecular oxygen to fit in for the formation of the ternary complex. Indeed ligand free reduced MCAD displays some oxygen reactivity, consistent with the complete desolvation of the active site participating in preventing the reaction of the reduced cofactor with oxygen. Furthermore, a partial oxidase activity in SCAD increases by decreasing the size of substrate analogues (153). The tetrameric oligomerization state of acyl-CoA dehydrogenases participates too in protecting the active site from solvent access, while the active sites in dimeric acyl-CoA oxidase are more exposed to solvent.

- **Physical accessibility of the cofactor to oxygen.** As mentioned before, it is generally accepted that in order to oxidize the reduced flavin cofactor oxygen needs to be localized close to the isoalloxazine ring. Therefore steric hindrance to access the isoalloxazine ring can prevent reaction with oxygen. In the case of acyl-CoA dehydrogenases a conserved tryptophan (W166 in MCAD) covers the xylene moiety of the FAD, possibly interfering with oxygen access to the reduced cofactor. Interestingly bacterial SCAD from *M. elsdenii* (154), which displays some oxygen reactivity, shows the substitution of the
conserved tryptophan with a less bulky phenylalanine. However the fact that this tryptophan residue is conserved also in the case of acyl-CoA oxidase (W174) suggests that this particular steric hindrance to the isoalloxazine ring does not play a major role in oxygen access to the reduced cofactor.

- **Prevention of the interaction with ETF.** Helix S of acyl-CoA oxidase, which is not present in the crystal structure of MCAD, covers the surface of the binding site for ETF. Therefore it is possible that this structural element prevents the interaction with electron acceptors similar to ETF.

### 1.4 Glycine betaine biosynthesis

Glycine betaine ($N,N,N$-trimethylglycine) is a zwitterionic amino acid derivative (Scheme 1.7) named after the plant *Beta vulgaris* (sugar beets) from which it was first isolated and discovered in the 19th century ($^{155}$) and its role as a common osmotic effector in bacterial cells has been established since 1983 ($^{156,157}$).

![Molecular structures of choline, betaine aldehyde, glycine betaine.](image)

Scheme 1.7: Molecular structures of choline, betaine aldehyde, glycine betaine.

Most bacteria possess adaptation systems based on osmoprotectants as stress response to environmental conditions such as high salinity and extreme temperature ($^{158}$). In high salinity conditions the efflux of water from the cytosol and the consequent decrease in cell volume trigger initially an uptake of potassium ions and synthesis of glutamate as counter ion in order to act as osmotic agent and promote water retention ($^{159}$). However, high intracellular concentrations of potassium glutamate are detrimental for the stability of most enzymatic systems, haltering
important metabolic pathways. Therefore, this initial increase of intracellular potassium glutamate is usually followed by the accumulation of “compatible solutes”, small organic molecules that allow water retention without significant perturbation of the metabolism of the cell. Examples of compatible solutes are polyols, free amino acids, and amino acids derivatives, among which glycine betaine. The use of the same small group of compatible solutes is shared by most organisms, regardless of the phylogenetic distance, ranging from bacteria and unicellular algae to plants and animals, suggesting strong selective pressure in the evolution of this adaptation system. A remarkable exception is represented by halobacteria which are able to adapt to very high salinity environments (up to 5 M) by relying only on increased intracellular concentrations of inorganic salts, as high as 7 molal in the case of species of Halobacterium. The enzymatic systems of halobacteria are unusually tolerant towards high concentration of inorganic salts and typically require salt concentration higher than 1 M to reach maximal activity. The amino acids composition of enzymes from halobacteria shows a high percentage of aspartate, glutamate, and weakly hydrophobic residues that enable the proper protein folding only in high salt conditions. This section will focus on the compatible solute glycine betaine, which is present in bacteria, plants, and animals. Some bacteria like Escherichia coli accumulate glycine betaine in the cytosol only in high salinity conditions as osmoprotectant and they are not able to use this compound as carbon or nitrogen source, while it has been reported that Rhizobium melitoti and Pseudomonas aeruginosa can use glycine betaine also as a nutrient. Certain plants such as spinach, sugar beets, barley, and maize naturally produce glycine betaine in response to abiotic stress and many economically relevant plants such as rice, tomato and potato, have been genetically engineered and characterized for the biosynthesis of glycine betaine, therefore leading to a wealth of studies and information on the physiological role
of glycine betaine in plants \(^{(162-164)}\). It has been shown that glycine betaine in plants works as compatible solute for water retention in high salinity, provides protection to the photosynthetic machinery and to the reproductive organs, helps stabilizing macromolecular complexes, and induces the expression of genes coding for ROS scavenging enzymes \(^{(162, 164)}\). In mammals the osmoprotectant role of glycine betaine is mainly required in the cells of the kidney inner medulla that are characterized by a continuous osmotic stress in the process of urine concentration and by an intracellular glycine betaine concentration ranging from 36 to 50 mM that protects these cells from apoptosis triggered by excessive shrinking \(^{(165, 166)}\). The osmoprotectant role of glycine betaine has been described also in liver cells during dehydration \(^{(167)}\) and in corneal epithelial cells in case of elevated tears osmolarity, which is associated with the pathology of dried eye \(^{(168)}\). In mammals glycine betaine plays an important role as methyl donor in the conversion of homocysteine to methionine catalyzed by betaine homocysteine methyltransferase \(^{(155)}\), linking glycine betaine to the folate one-carbon metabolism \(^{(169)}\). Deficiency of glycine betaine can lead to accumulation of homocysteine in blood, typical of the condition called homocystinuria and linked to cardiovascular diseases \(^{(155)}\). Glycine betaine can be imported in the cell (in the case of animals obtained from the diet) or synthesized from the precursor choline \(^{(155)}\). Choline and glycine betaine deficiencies have been related to different human pathologies, such as fatty liver, cognitive dysfunctions, cardiovascular diseases, metabolic syndrome, and cancer \(^{(169)}\). The first uptake systems for glycine betaine in bacteria have been described in *Salmonella typhimurium* and *E. coli* and initially discovered for the uptake of the compatible solute proline \(^{(159)}\). Both organisms possess three transporters for proline, namely PutP, ProP, and ProU \(^{(159)}\). While PutP is devoted only to the uptake of proline as nutrient in physiological conditions and is inhibited by high osmolarity, ProP and ProU ensure the uptake of proline, glycine betaine and other
compatible solutes in high salinity conditions \(^{(159)}\). Three transporters for glycine betaine, OpuA, OpuC, and OpuD, have been identified in *Bacillus subtilis*, with OpuA and OpuC being multicomponents and similar to the ProU type transporters identified in *E. coli* \(^{(170)}\). OpuD represents instead a different type of glycine betaine transporter, with only one protein chain and is induced by high salinity \(^{(170)}\).

Figure 1.9: Binding site of OpuC for glycine betaine (PDB 3PPP). Glycine betaine is shown as cyan sticks, the side chains of Q39, Y91, Y241, Y137, and Y217 are in green sticks.

Despite the fact that early studies on the translocation of \(^{14}\)C-labeled glycine betaine in barley plants \(^{(171)}\) showed a rapid uptake of this compound and an active and regulated translocation process from the initial application site, very little is known on specific transporters of glycine betaine in plants \(^{(164)}\). Preliminary studies have reported that the gene product LeProT1 from tomato is able to transport glycine betaine, proline, and with less affinity \(\gamma\)-amino butyric acid (GABA), while the GABA transporter ProT2 from *Arabidopsis* is inhibited by glycine betaine \(^{(164)}\). In animals, exogenous glycine betaine is obtained from the diet, in humans mainly from
wheat, spinach, shellfish, and sugar beets, and it is rapidly absorbed in the ileum with a bioavailability close to 100% \(^{(172)}\). A specific transporter for glycine betaine and GABA, BGT1, has been identified in different animal species, among which mouse, dog, and human \(^{(173)}\). Renal BGT1 is based on Na\(^+\) and Cl\(^-\) coupled transport and is induced by high osmolarity via transcription regulation and post-translational modifications \(^{(165,173)}\).

Choline, which represents the precursor for glycine betaine biosynthesis in different organisms, can be obtained from hydrolysis of phosphatidylcholine, from the diet (mainly beef, chicken, liver and eggs), or it can be synthesized \textit{de novo} \(^{(155)}\). Bacteria can produce phospholipases in order to obtain free choline and this process seems to be relevant for bacterial survival in host-pathogen interactions \(^{(161)}\). The biosynthesis of glycine betaine from choline occurs via a two-step oxidation with the cytotoxic compound betaine aldehyde as intermediate \(^{(162,174)}\). The first oxidation of choline to betaine aldehyde is carried out by a soluble iron-sulfur choline monooxygenase (CMO) in plants and by a putative flavin-dependent membrane bound choline dehydrogenase (CHD) in bacteria and animals \(^{(162)}\). The conversion of betaine aldehyde to glycine betaine is then performed by a soluble NAD\(^+\)-dependent betaine aldehyde dehydrogenase (BADH) \(^{(175)}\). In plants this biosynthetic pathway is localized in chloroplasts \(^{(164)}\), while in animals CHD and BADH have been identified in mitochondria \(^{(176,177)}\). The soluble and flavin-dependent choline oxidase (CHO) from \textit{Arthrobacter globiformis} is the only enzymatic system currently known with the ability to perform both oxidation steps by retaining betaine aldehyde in the active site \(^{(48,162)}\). An alternative pathway of biosynthesis of glycine betaine from glycine has been described only in the extreme halophytic phototrophic bacteria \textit{Actinopolyspora halophile} and \textit{Ectothiorhodospira halochloris}, where glycine undergoes three methylation steps to sarcosine, dimethylglycine, and glycine betaine by two methyl transferases \(^{(164)}\).
Despite the fact that human choline dehydrogenase has been associated to different pathologies (178-180), very little is known on this enzyme from both prokaryotic and eukaryotic sources due to issues in purification and in vitro stabilization of the protein. Chapter IV covers in details the medical relevance and the biochemical knowledge on choline dehydrogenase. In contrast, the soluble choline oxidase from *A. globiformis*, which belongs to the GMC superfamily of flavin dependent oxidases, has been extensively characterized from a mechanistic point of view (181).

Choline oxidation proceeds through two hydride transfer steps from the Cα of choline and of betaine aldehyde to the N5 of the FAD cofactor after proton abstraction from the hydroxyl group by a catalytic base (Scheme 1.8) (48).

**Scheme 1.8**: Catalytic mechanism of choline oxidation to glycine betaine by choline oxidase from *A. globiformis*.

Substrate deuterium kinetic isotope effects on the reductive half reaction show that the two hydride transfers represent the main rate limiting steps in the overall turnover of the enzyme and they occur quantum mechanically within a highly preorganized active site, as supported by the temperature profile of the substrate deuterium kinetic isotope effects on $k_{cat}$ and $k_{cat}/K_m$ (48, 182).

After the hydride transfer reaction the reduced FAD is oxidized by oxygen, with a synchronous transfer of a hydride from the N5 atom of FAD and of a proton from a solvent exchangeable site to yield hydrogen peroxide without detection of stable reaction intermediates (183). In the steps...
that involve the reoxidation of the reduced enzyme a partially rate limiting isomerization was detected by solvent viscosity effect on the kinetic parameter $k_{cat}/K_{ox}$\textsuperscript{(183)}. The intermediate betaine aldehyde, which is not released to the solution\textsuperscript{(47)}, is hydrated to gem-diol before the second hydride transfer\textsuperscript{(184)}. The mechanism of the enzyme was established to be a sequential bi-bi mechanism\textsuperscript{(47)}, with product release not rate limiting in turnover and occurring after oxidation of the reduced FAD by oxygen\textsuperscript{(185)}. The FAD cofactor is covalently attached to H99 via a 8α-N(3)-histidyl linkage which contributes to the preorganization of the active site for the hydride tunneling\textsuperscript{(19)}. CHO crystallizes as a dimer, consistent with the dimeric oligomerization state of the enzyme in solution\textsuperscript{(186)}, and three dimensional structures of wild type in complex with DMSO, V464A free enzyme and S101A in complex with acetate (with 4 dimers in the asymmetric unit) variants are available\textsuperscript{(181)}. The overall folding of the monomer of choline oxidase (PDB 2JBV) is similar to the other members of the GMC superfamily, with an FAD binding and a substrate binding domain. The dimer is asymmetric with the two active sites located at the poles of the dimer and the dimer interface stabilized by hydrophobic and ionic interactions\textsuperscript{(187)}. The active site shows a hydrophobic pocket formed by the aromatic residues W61, W331, V464, and Y465, the polar residue S101 and the ionizable residues E312, H351, and H466 (Figure 1.10). The FAD cofactor shows a fairly planar conformation only in the crystal structures of the V464A and S101A variants, while in case of wild type choline oxidase it is distorted with a C4a adduct, most likely a flavin C4a-OH or C4a-OO(H) adduct promoted by X-ray exposure during data collection of the crystal\textsuperscript{(188)}. 
A study with substrate and product analogs established the importance of the trimethylammonium moiety of choline and its positive charge for binding and suggested the presence of an hydrophobic cage for the positioning of the three methyl substituents. Site-directed mutagenesis studies of the residue E312 showed the involvement of this negatively charged side chain in binding and positioning of the substrate in the active site. The pH profile of the kinetic parameters $k_{\text{cat}}$ and $k_{\text{cat}}/K_m$ established the presence of a catalytic base in the reductive half reaction of choline oxidation involved in the formation of the alkoxide species. The more likely candidates for this role, based on the crystal structure of choline oxidase and on the conserved histidine residues in the GMC superfamily, are the residues H351 and H466. The pH profile of the variants H351A and H466A, which represent the removal of the putative base, was still consistent with the presence of the catalytic base. The mutation to alanine mainly affects the binding and positioning of choline and the hydride transfer reaction in the case of H351 and the overall polarity of the active site and the stabilization of the transition state.
for choline oxidation in H466\(^{(191)}\). However, the mutation to alanine generates an empty space that could be filled by a water molecule acting as surrogate base. A recent investigation of the variant H466Q, which is a more filling substitution of the imidazole ring, identified H466 as the catalytic base of choline oxidase\(^{(192)}\). No ionizable side chains are involved in the oxidative half reaction, with the positively charged trimethylammonium group of the intermediate betaine aldehyde and of the product glycine betaine playing an important role in oxygen activation, as shown by the 80-fold decrease in the kinetic parameter \(k_{\text{cat}}/K_{\text{ox}}\) with the neutral and isosteric substrate analog 3,3-dimethylbutanol\(^{(122, 189)}\). The hydrophobic side chain of V464 plays an important role in localization of oxygen, as shown by the 50-fold decrease in \(k_{\text{cat}}/K_{\text{ox}}\) in the variant V464A\(^{(122)}\). S101 is located less than 4 Å away from the N5 atom of FAD and different substitutions at this position highlight the importance of the hydrophilic character of S101 for the hydride transfer reaction\(^{(193, 194)}\). The removal of the hydroxyl group of S101 by site-directed mutagenesis to alanine affects the reductive half-reaction but improves the reaction of reduced FAD with oxygen, as shown by the 3-fold increase in \(k_{\text{cat}}/K_{\text{ox}}\) and 20-fold decrease in \(K_{\text{ox}}\)\(^{(193, 194)}\). A mechanistic investigation of the role of N510, which is 4.7 Å away from the N3 atom of FAD and is part of the conserved His-Asn catalytic pair in the GMC superfamily, concludes that this residue is important for both reductive and oxidative half reactions and for the timing of the OH and CH bond cleavage\(^{(195)}\).

Despite the detailed structural and mechanistic studies performed on choline oxidase some aspects of this enzymatic system still need to be addressed, such as substrate access to the active site. Molecular dynamics studies on the free monomer of choline oxidase (PDB 2JBV) suggest that substrate access to the active site is controlled by the movement of the side chains of M62, L65, V355, F357, and M359, forming a hydrophobic cluster on the surface of choline oxidase.
The electrostatic potential of the protein surface around the hydrophobic cluster, which is rich in negatively charged glutamate residues, probably plays a role in attracting and directing the positively charged choline through this hydrophobic gate \(^{(105)}\). A crystal structure of choline oxidase in complex with a physiologically relevant ligand, which could give some insights on gating mechanism for substrate binding, is not available.

Choline oxidase from \textit{A. globiformis} shares 37\% sequence identity with choline oxidase from \textit{Aspergillus fumigatus} and the active site residues that are relevant for catalysis/substrate positioning are fully conserved \(^{(196)}\). The biochemical characterization of recombinant choline oxidase from \textit{A. fumigatus} established that this enzyme contains covalently bound FAD and is able to oxidize both choline and betaine aldehyde using oxygen as electron acceptor with a ternary complex mechanism \(^{(196)}\). Interestingly, choline oxidase from \textit{A. fumigatus} releases betaine aldehyde to the solution, in contrast to the ability of the bacterial enzyme to retain the intermediate in the active site and perform the second oxidation to glycine betaine \(^{(196)}\). Indeed, the characterization of the glycine betaine biosynthesis in \textit{A. fumigatus} shows the presence of a betaine aldehyde dehydrogenase in the same operon, which catalyzes the oxidation of this toxic intermediate to glycine betaine using NAD\(^+\) \(^{(196)}\).

### 1.5 Nitronate monooxygenases and the detoxification of propionate 3-nitronate

Propionate 3-nitronate, which is well known for its neurological toxicity, is the conjugated base of the naturally occurring nitro compound 3-nitropropionic acid or \(\beta\)-nitropropionic acid (Scheme 1.9) \(^{(197)}\).
Despite the fact that the toxic effect of propionate 3-nitronate in the form of the glycoside hyptagenic acid has been known since 1920 (198) and its chemical structure was established in 1949 (199), the biosynthesis and physiological role of this compound are not fully characterized yet (197). Precursor forms of 3-nitropropionic acid found in nature are glycosides of this compound and 3-nitropropanol, which is not toxic until oxidized to 3-nitropropionic acid by aldehyde dehydrogenase (200). The glycosides of 3-nitropropionic acid and also of 3-nitropropanol are hydrolized by the gut flora and studies on rodents show that the toxicity of aglycone is higher than the glycoside (LD$_{50}$ of 77 mg/Kg bw compared to LD$_{50}$ of >2500 mg/Kg bw) (201), supporting the notion that the glycoside is an inactive precursor of the toxin until release by hydrolysis. As shown in Table 1.3, which lists some sources of glycosides, 3-nitropropanol, and free 3-nitropropionic acid, this nitrotoxin has been so far isolated from fungi and plants (197).
Table 1.3: Plants and fungi containing free 3-nitropropionic acid or its precursor forms. Adapted from [157].

<table>
<thead>
<tr>
<th>PLANTS</th>
<th>FUNGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astragalus sp.</td>
<td>Arthrinium aureus</td>
</tr>
<tr>
<td>Indigofera spicata</td>
<td>Astragalus magdalenae</td>
</tr>
<tr>
<td>Astragalus falcatus</td>
<td>Arthrinium phaeospermum</td>
</tr>
<tr>
<td>Indigofera endecaphylla</td>
<td>Astragalus mollissism</td>
</tr>
<tr>
<td>Cirsium varia</td>
<td>Arthrinium sacchari</td>
</tr>
<tr>
<td>Lotus angustissimus</td>
<td>Aspergillus flavus</td>
</tr>
<tr>
<td>Coronilla viminalis</td>
<td>Arthrinium serieuses</td>
</tr>
<tr>
<td>Lotus uliginosus</td>
<td>Aspergillus oryzae</td>
</tr>
<tr>
<td>Corynocarpus laevigatus</td>
<td>Arthrinium terminalis</td>
</tr>
<tr>
<td>Securigera orientalis</td>
<td>Aspergillus wentii</td>
</tr>
<tr>
<td>Hippocrepis balearica</td>
<td>Astragalus miser</td>
</tr>
<tr>
<td>Securigera parviflora</td>
<td>Penicillium atrovenetum</td>
</tr>
<tr>
<td>Hippocrepis comosa</td>
<td>Astragalus fastidius</td>
</tr>
<tr>
<td>Securigera securidaca</td>
<td>Pestalotia palmarum</td>
</tr>
<tr>
<td>Hippocrepis emerus</td>
<td>Astragalus pomennesis</td>
</tr>
<tr>
<td>Securigera varia</td>
<td>Phomosis sp.</td>
</tr>
<tr>
<td>Hiptage madablota</td>
<td>Astragalus amphioxys</td>
</tr>
<tr>
<td>Scorpiurus muricatus</td>
<td>Septoria cirsii</td>
</tr>
<tr>
<td>Viola odorata</td>
<td></td>
</tr>
<tr>
<td>Scorpiurus vermiculates</td>
<td></td>
</tr>
</tbody>
</table>
Despite the importance of this metabolic poison only preliminary studies with isotope-labeled precursors have been performed on the biosynthesis of 3-nitropropionic acid in the fungus *Penicillium atrovenetum*\(^{(202)}\) and in the plant *Indigofera spicata*\(^{(203)}\). The results of these studies suggest a different biosynthetic route for 3-nitropropionic acid, with aspartate as precursor in fungi and malonate in plants, but the characterization of these possible pathways has not been carried out yet\(^{(197)}\). The physiological role of 3-nitropropionic acid has been assigned mainly to a defense tool against herbivores in plants based on the known toxicity to animals and to its higher content in parts of the plants that are more easily reached by animals\(^{(197)}\). Furthermore, the presence of 3-nitropropionic acid may be involved in the nitrification process as its enzymatic oxidation, detected in many plants from where the toxin has been isolated, leads to formation of nitrite or nitrate\(^{(197)}\). Incubation of isolated rat liver mitochondria with 3-nitropropionic acid and with its conjugate base established that the toxicity of this metabolic poison is only due to its deprotonated form propionate 3-nitronate irreversibly inhibiting succinate dehydrogenase\(^{(204)}\), drawing attention to the chemical properties of this molecule.

The presence of a nitro group in nitroalkanes has an electronic withdrawing effect on the C\(\alpha\) position that confers acidity to the C\(\alpha\)-H bond\(^{(200)}\). As a result the deprotonation of this C\(\alpha\)-H bond occurs with pK\(_a\) values close to the physiological range (9.1 in the case of 3-nitropropionic acid) instead of >25 in the absence of the nitro group\(^{(205)}\). The C\(\alpha\)-H deprotonation of 3-nitropropionic acid to yield propionate 3-nitronate follows the principle of nonperfect synchronization, which is characterized by different molecular processes in a chemical reaction not being synchronized in the transition state\(^{(206)}\). In the case of the ionization of 3-nitropropionic acid these molecular processes are the C\(\alpha\)-H bond cleavage and the delocalization of the electrons between the C\(\alpha\) and nitrogen atoms\(^{(197)}\). The delocalization of electrons stabilizes
the product and renders the deprotonation thermodynamically favorable with pKₐ close to the physiological range, but it lags behind the Cα-H bond cleavage and the asynchronous transition state is associated with a high intrinsic energy barrier (206). As a consequence, the equilibration of neutral nitroalkanes and deprotonated nitronates (Scheme 1.9) is slow, with second-order rate constants ranging from 2 to 6 M⁻¹ s⁻¹ for deprotonation and from 15 to 75 M⁻¹ s⁻¹ for reprotonation, and this allows the separate investigation of the acid and base forms by preventing their rapid interconversion during initial rate measurements (197).

Many intoxication cases of 3-nitropropionic acid have been reported since its first description in 1920, both in humans and livestock (197). In the case of humans the intoxication episodes occurred through ingestion of food contaminated by fungi containing the toxin, with most documented examples due to moldy sugarcane in the northern regions of China (884 cases, of which 10% fatal) (207). Poisoning of animals and livestock with forages containing 3-nitropropionic acid have been described in chickens, meadow voles, mice, pigeons, pigs, possums, rabbits, rats, ponies, (208-210) and it represents a significant damage for cattle losses in the western regions of United States, Canada, and northern regions of Mexico (211, 212).

The toxic effects described in these poisoning cases included poor coordination, foaming of the mouth, incontinence, excessive salivation, weight loss, increased heart rate and respiratory distress (197, 213). The neurological damages linked to the administration of 3-nitropropionic acid have been extensively investigated in animal models (197, 200), and the induced lesions are very similar to the ones associated with Huntington’s disease, leading to the use of this compound for models of this pathology (214). The toxicity of propionate 3-nitronate arises from the irreversible inhibition of succinate dehydrogenase, which represents both a key enzyme of the Krebs cycle and Complex II in the oxidative phosphorylation pathway, thus dramatically hampering ATP
production \(^{(204)}\). Inhibition of fumarase, another component of the Krebs cycle, has also been reported \(^{(215)}\). The toxic effects of this mitochondrial damage, which particularly affects the striatum \(^{(216)}\) with additional impairment to the brain blood barrier \(^{(217)}\), proceeds with perturbed intracellular calcium homeostasis, glutameric activation of the NMDA receptor and increase of oxygen reactive species, leading to mitochondrial membrane depolarization and apoptosis \(^{(218,219)}\). The release of nitrite by the systemic denitrification of 3-nitropropionic acid is the cause of methemoglobinemia, which consists in the oxidation of ferrous ion in hemoglobin to ferric in methemoglobin, losing the ability to transport oxygen \(^{(200)}\). This condition may lead to cyanosis, but it has been shown to be transient and not responsible of the fatal effects of propionate 3-nitronate poisoning \(^{(200)}\). The biochemical mechanism of succinate dehydrogenase inhibition by propionate 3-nitronate has been suggested to be related to the structure of the toxin being an analogue of succinate but not fully elucidated yet, despite the availability of a crystal structure of this enzyme in complex with the inhibitor (Figure 1.11) \(^{(220)}\).

![Chemical structures of 3-nitropropionate and succinate are shown in panel A, and a view of the crystal structure of succinate dehydrogenase in complex with 3-nitropropionic acid (PDB 1YQA) is shown in panel B. FMN and 3-nitropropionic acid are in yellow and cyan, respectively.]

**Figure 1.11**: Biochemical mechanism of succinate dehydrogenase inhibition.
The spectroscopic characterization of succinate dehydrogenase inactivated by propionate 3-nitronate and the lack of chemical modifications on the flavin cofactor in the crystal structure rule out the formation of a nucleophilic attack of the nitronate on the flavin cofactor with the formation of a N5 covalent adduct \(^{(220,221)}\). A mechanism proposed by Coles et al. is based on the oxidation of propionate 3-nitronate to 3-nitroacrylate, which then reacts with a cysteine residue \(^{(221)}\). The crystal structure of the enzyme-inhibitor complex highlights instead a modified R297, but the resolution of 2.3 Å is not high enough to determine the chemical modification of the side chain of R297 \(^{(220)}\).

The pharmacokinetics of 3-nitropropionic acid is not well characterized because classical radiolabeled ADME (Absorption, Distribution, Metabolism, Excretion) studies have not been carried out \(^{(200)}\). It is known that free nitropropionic acid and 3-nitropropanol are rapidly absorbed by the gastrointestinal tract, with the maximal blood concentration of these compounds in cattle reached in one hour \(^{(222)}\). The biotransformation processes described to date are only the oxidation of 3-nitropropanol to the corresponding carboxylic acid with a half-life of \(\leq 10\) minutes, and the systemic denitrification of these compounds leading to methemoglobinemia \(^{(200)}\). There is no antidote currently available for poisoning by propionate 3-nitronate and the fast absorption of this toxic compound coupled with an irreversible inhibition mechanism render the research of an antidote a difficult challenge \(^{(200)}\). Indeed the attempt to contrast the poisoning of 3-nitropropanol by preventing its oxidation to 3-nitropropionic acid with administration of alcohol dehydrogenase inhibitors immediately after exposure to 3-nitropropanol failed to even diminish the toxic effects, while a prior treatment with alcohol dehydrogenase inhibitors before administration of 3-nitropropanol prevented any toxic effect \(^{(222,223)}\).
The ability to detoxify propionate 3-nitronate in nature has been observed in plants, fungi, and bacteria \(^{(197)}\). Enzymes able to oxidize propionate 3-nitronate or 3-nitropropionic acid were purified for the first time in 1987 from the fungus *Penicillium atrovenetum* \(^{(224)}\) and in 1999 from the plant *Hippocrepis comosa* \(^{(225)}\). The enzyme from *P. atrovenetum* has been shown to be a dimeric FMN-dependent enzyme that catalyzes the oxidation of only the anionic form propionate 3-nitronate and that based on the detection of H\(_2\)O\(_2\) has been classified as propionate 3-nitronate oxidase \(^{(224)}\). The enzyme from *H. comosa* was partially purified by leaf extracts and shown to oxidize 3-nitropropionic acid with maximal activity at pH 4.8 and production of H\(_2\)O\(_2\) \(^{(225)}\). No protein sequences and further characterization studies of these two enzymes are available. The most well-known system for propionate 3-nitronate detoxification is represented by nitronate monooxygenases (NMO), which were previously named 2-nitropropane dioxygenases before reclassification by IUBMB in 2010 \(^{(197)}\). NMOs are FMN-dependent enzymes that catalyze the oxidation of propionate 3-nitronate by transferring one electron from the substrate to the flavin cofactor and use oxygen as second substrate (Figure 1.12) \(^{(197)}\).
Figure 1.12: Oxidation mechanism of propionate 3-nitronate in NMOs. Transfer of one electron from propionate 3-nitronate to the oxidized FMN to yield a radical intermediate and anionic flavosemiquinone. The black line and the red line spectra represent typical UV-visible absorbance spectra for oxidized flavin and red semiquinone.

Monooxygenase activity is supported by the lack of detection of H₂O₂ (29) and recently NMOs have been classified as Group H of flavin-dependent monooxygenases (128). To date only fungal NMOs have been rigorously characterized from a kinetic point of view. The first NMO characterized is the FMN-dependent enzyme from *Neurospora crassa* in 2005 (226) but the identification of propionate 3-nitronate as the physiological substrate was recognized only in 2012 (227). Therefore the enzyme was initially kinetically characterized with alkynitronates, specifically ethyl nitronate (EN) (226). The enzyme, which is a homodimer of 80 kDa, oxidizes nitronates from 2 to 6 carbon length, and to a lesser extent the neutral forms (nitroalkanes) (226). Site-directed mutagenesis studies of H196 establish that this residue acts as a base for the deprotonation of the neutral nitroethane (228). Furthermore, the fate of the anionic form ethynitronate in the active site branches between release to the solution and oxidation to
acetaldehyde (228). NMO stabilizes the anionic flavosemiquinone during catalysis, which represents a remarkable exception in the field of flavin-dependent monooxygenases that usually rely on the formation of a C4a-(hydro)peroxyflavin intermediate (29). Production and release of hydrogen peroxide or superoxide were not detected with the exception of propyl-1-nitronate for both and propyl-2-nitronate for superoxide (226). In 2012, the enzymatic activity of fungal NMOs from *N. crassa* and from *Cyberlindera saturnus* (previously named *Williopsis saturnus*) with propionate 3-nitronate was established for the first time and the physiological role of the gene coding for NMO from *N. crassa* confirmed to be propionate 3-nitronate detoxification (227). The mechanism of propionate 3-nitronate oxidation has been extensively investigated in NMO from *C. saturnus* (229). This enzyme is able to oxidize only the anionic alkyl nitronates with propionate 3-nitronate being the best substrate with second order rate constants of $10^6$ M$^{-1}$ s$^{-1}$, close to the diffusion limit (229). The enzyme follows a sequential bi-bi steady-state kinetic mechanism with oxygen reacting before release of the first product from the active site and no pH dependence of the kinetic parameters $k_{cat}$, $k_{cat}/K_m$ for propionate 3-nitronate and oxygen was observed in the pH range from 5.5 to 10.5 (229). The reductive half reaction is too fast to be monitored in a stopped-flow spectrophotometer with the anionic flavosemiquinone forming in the dead time of the instrument (2.2 ms) even at 7°C (229). The reaction with oxygen is also very fast with second order rate constants in the order of $10^7$ M$^{-1}$ s$^{-1}$, while solvent viscosity studies on the steady-state kinetic parameters establish substrate binding and product release as partially rate limiting steps (229). After the first electron transfer that yields the substrate radical and the anionic flavosemiquinone a transient peroxinitro acid was detected and monitored with an increase in absorbance at 300 nm (229). As shown in Scheme 1.10 two possible pathways lead to this
intermediate: one in which oxygen reacts first with the anionic flavosemiquinone and another one in which oxygen is activated instead by the substrate radical \(^{(229)}\).

![Scheme 1.10](image)

**Scheme 1.10**: Branched mechanism for reaction of nitronate monooxygenase from *C. saturnus* with propionate 3-nitronate.

FMN\(_{ox}\) and FMN\(_{semq}\) indicate the oxidized form of the cofactor and the anionic flavosemiquinone and E stands for enzyme. Modified from \(^{(229)}\).

The toxin propionate 3-nitronate has been isolated from plants and fungi, however 85% of the genes annotated as hypothetical nitronate monooxygenases in the GenBank belong to bacteria. The physiological role of bacterial NMOs has been shown to be protection from the toxic effect of propionate 3-nitronate and in some bacteria, such as *Pseudomonas aeruginosa* PAO1, this metabolic poison can be used as the only carbon and nitrogen source \(^{(227, 230)}\). In contrast to the well-characterized fungal NMOs, very little is known on bacterial NMOs, with the exception of a
preliminary enzymatic characterization of a NMO from *Pseudomonas* sp. strain JS189 presented at a conference \(^{(231)}\). A crystal structure of the hypothetical NMO PA1024 from *P. aeruginosa* PAO1 is available in the PDB database (PDB 2GJL) \(^{(232)}\). However, the enzyme PA1024 was not kinetically characterized, leaving the classification of the enzyme as hypothetical \(^{(232)}\). Similarly, the importance for virulence of the hypoxia-induced gene Rv1894c in *Mycobacterium tuberculosis* was investigated in guinea pigs, but the enzymatic activity of the purified protein with propionate 3-nitronate was not established and the hypothetical annotation as NMO not confirmed \(^{(233)}\). Therefore, there is a lack of biochemical data on bacterial nitronate monooxygenases. Furthermore the gene function prediction of the \(\approx\)5000 genes in the GenBank annotated as hypothetical NMO is not reliable as it is based only on modest overall protein sequence similarity with the enzymes from *N. crassa* and *C. saturnus*.

1.6 Specific goals

The research of this thesis focuses on flavin-dependent enzymes belonging to two pathways: the oxidation of choline to the osmoprotectant glycine betaine and the detoxification of the metabolic poison propionate 3-nitronate.

In particular, the specific goals related to enzymes involved in the oxidation of choline are the following:

1) **Structural determination of choline oxidase from *A. globiformis* in complex with a physiologically relevant ligand.**

The reaction catalyzed by choline oxidase from *A. globiformis* has been extensively characterized by previous mechanistic and structural studies on this enzyme \(^{(181)}\). However, some aspects related to substrate access to the active site have not been fully elucidated yet. A crystal
structure of this enzyme in complex with a physiologically relevant ligand would shed light on which structural elements play a role in controlling substrate access to the active site. This is discussed in Chapter II.

2) Site-directed mutagenesis of hydrophobic residues lining the entrance to the active site of choline oxidase from *A. globiformis* and kinetic studies of the variant enzymes.

A previous study of molecular dynamics performed on choline oxidase highlighted the movement of the side chains of an hydrophobic cluster (M62, L65, V355, F357, and M359) on the surface of choline oxidase. The movements of these side chains regulate the radius of the entrance to the active site, with the side chains of M62 and F357 facing each other. Single and double variant enzymes in which the side chains of M62 and F357 were mutated to alanine were generated and kinetically characterized. This kinetic investigation was complemented by computational studies and is presented in Chapter III.

3) Expression trials of human choline dehydrogenase.

The oxidation of choline in humans is performed in mitochondria by the membrane associated choline dehydrogenase, which has been recently associated to different human pathologies. The biochemical knowledge of choline dehydrogenase is very limited due to the scarce *in vitro* stability of this enzyme, which has hampered its characterization. Chapter IV presents a review study of the medical relevance of human choline dehydrogenase, summarizes the previous attempts to purify the enzyme from different sources, and provides an homology model of human choline dehydrogenase based on the crystal structure of the bacterial enzyme choline oxidase. Chapter V presents different expression trial of the recombinant human choline dehydrogenase in *E. coli*. 
The specific goals related to the investigation of the enzymes involved in the detoxification of propionate 3-nitronate are the following:

4) **Structural and kinetic characterization of the bacterial nitronate monooxygenase Pa-NMO from *Pseudomonas aeruginosa* PAO1.**

Previous studies on fungal nitronate monooxygenases focused on the kinetic characterization of the enzymes from *Neurospora crassa* (Nc-NMO) and *Cyberlindera saturnus* (Cs-NMO), establishing a different substrate specificity of the two fungal enzymes \(^{(197)}\). No crystal structure of nitronate monooxygenases was available, which would give precious information for site-directed mutagenesis. Furthermore, very little biochemical information was available for bacterial nitronate monooxygenases, despite the fact that 85% of genes annotated in the GenBank as hypothetical nitronate monooxygenases belong to bacteria. Chapter VI presents a combined kinetic and structural characterization of the nitronate monooxygenase Pa-NMO from *P. aeruginosa* PAO1, and the identification of four conserved motifs in Pa-NMO and CsNMO that establish Class 1 NMO.

5) **Expression, purification, and biochemical characterization of hypothetical nitronate monooxygenases not belonging to Class 1 NMO.**

The gene function prediction of hypothetical nitronate monooxygenases is based on modest overall protein sequence similarity to the two fungal enzymes Cs-NMO and Nc-NMO, which belong to two different classes (Class 1 and Class 2 NMO, respectively). This leads to a scarcely reliable gene function prediction. Furthermore, with the current annotation many organisms possess more genes predicted to code for nitronate monooxygenases. It is important to establish if this redundancy is indicative of different isoforms of nitrate monooxygenases or if it is due to the unreliable gene function prediction. *P. aeruginosa* PAO1 possess the gene *pa4202* coding
for Pa-NMO, and other two genes, *pa1024* and *pa0660*, annotated as hypothetical nitronate monooxygenases. Chapter VII and VIII present the cloning, expression, purification and biochemical characterization of the hypothetical nitronate monooxygenases PA1024 and PA0660. Chapter IX describes the cloning, expression, purification and characterization of the hypothetical nitronate monooxygenase HP0773 from the human pathogen *Helicobacter pylori*, which represents a possible drug target.

General conclusions on the results obtained in this dissertation research are presented in Chapter X.
1.7 References


212. Mathews, F. P. (1940) *Poisoning in sheep and goats by sacahuiste (Nolina texana) buds and blooms*, Texas Agricultural Experiment Station, College Station, Tex.


CHAPTER II: CRYSTAL STRUCTURE OF CHOLINE OXIDASE IN COMPLEX WITH THE REACTION PRODUCT GLYCINE BETAINÉ

(This chapter has been published verbatim in Salvi, F., Wang, Y. F., Weber, I. T., and Gadda, G., (2014), *Acta Crystallographica D* 70(Pt 2):405-13); the author contributed to design the study, crystallization trials and data processing, and wrote the manuscript.

2.1 Abstract

Choline oxidase from *Arthrobacter globiformis*, involved in the biosynthesis of glycine betaine from choline, has been extensively characterized in its mechanistic and structural properties. Despite the knowledge gained on the enzyme the details of substrate access to the active site are not fully understood. The “loop-and-lid” mechanism described for the Glucose-Methanol-Choline enzyme superfamily has not been confirmed for choline oxidase. Instead, a hydrophobic cluster on the solvent accessible surface of the enzyme has been proposed by molecular dynamics to control substrate access to the active site. Here, the crystal structure of the enzyme was solved in complex with glycine betaine at pH 6.0 at 1.95 Å resolution, allowing for a structural description of the ligand-enzyme interactions in the active site. This structure is the first one of choline oxidase in complex with a physiologically relevant ligand. The two protein structures are virtually identical with the exception of a loop at the dimer interface, which assumes two distinct conformations. The different conformations of loop (250-255) define different accessibilities of the proposed active site entrance delimited by the hydrophobic cluster on the other subunit of the dimer suggesting a role for regulating substrate access to the active site.

2.2 Introduction

Choline oxidase (E.C. 1.1.3.17) from *Arthrobacter globiformis* catalyzes the oxidation of choline to glycine betaine (Scheme 2.1), which is a ubiquitous osmoprotectant in bacteria, plants and
animals (1). The enzyme is important in biotechnological applications for the genetic engineering of economically relevant plants to potentiate osmotic stress resistance (2, 3) and the developing of sensors for the detection of choline and derivatives in biological fluids (4). The oxidation of choline to glycine betaine catalyzed by choline oxidase occurs via two hydride transfer reactions with the rate limiting steps represented by the two flavin reductions (Scheme 2.1) (5).

![Scheme 2.1: The oxidation of choline to glycine betaine by choline oxidase.](image)

Studies of the temperature dependence of the substrate kinetic isotope effect in choline oxidase demonstrated that the hydride transfer in alcohol oxidation occurs quantum mechanically within a preorganized enzyme-substrate complex (6). The reaction intermediate betaine aldehyde has been shown to predominantly exist in solution in the gem-diol form (7) and to stay bound in the active site of the enzyme in bacteria (Scheme 2.1) (8). In fungi, instead, it is released to bulk solvent when the enzyme turns over with choline (9). Choline oxidase contains FAD covalently linked to the protein through H99 (10). The enzyme is grouped in the Glucose-Methanol-Choline (GMC) enzyme oxidoreductase superfamily, (11) which includes a variety of FAD-dependent enzymes that oxidize unrelated alcohols and share similar three-dimensional structures (12). The mechanism of action of bacterial choline oxidase has been extensively characterized (13) with
structural and mechanistic studies showing the importance of residues S101, \(^{(14)}\) E312, \(^{(13)}\) H351, \(^{(15)}\) V464, \(^{(16,17)}\) H466, \(^{(18)}\) and N510 \(^{(19)}\) in the active site.

The crystallographic structure of wild-type choline oxidase was previously reported (PDB 2JBV) from single crystals obtained at pH 8.5 \(^{(13)}\). It shows a distorted flavin with an O atom covalently linked to the flavin C(4a) atom of the isoalloxazine ring and contains DMSO from the crystallization cocktail in the active site \(^{(13)}\). Two crystal structures of active site mutants are also available from crystallization conditions at pH 6.0, e.g., the S101A enzyme in complex with acetate (PDB 3NNE) \(^{(20)}\) and the V464A enzyme devoid of ligands (PDB 3LJP) \(^{(16)}\). In all cases the enzyme crystallizes as a dimer and biochemical studies established a dimeric state for the enzyme in solution \(^{(21)}\). The enzyme active site is completely secluded from bulk solvent, \(^{(13,20)}\) raising the question of how the substrate accesses the active site. A “loop-and-lid” mechanism has been proposed to control substrate access in the homotetrameric pyranose 2-oxidase \(^{(22-25)}\) and monomeric cholesterol oxidase, \(^{(26-28)}\) which are members of the GMC enzyme superfamily.

The corresponding loop that covers the active site in choline oxidase (e.g., residues 74-85) is well defined in the available crystal structures and it was shown to be static over 60 ns in molecular dynamics simulations \(^{(29)}\). In contrast, rapid dynamic motions of a hydrophobic cluster composed of M62, L65, V355, F357 and M359 on the solvent accessible surface above the FAD cofactor were observed \(^{(29)}\). The side chains of these residues delimit the entrance of a tunnel leading to the active site (Figure 2.1) and their motions regulate the radius of this entrance.
In this study, we report the three-dimensional structure of choline oxidase in complex with glycine betaine solved to 1.95 Å resolution, allowing for the description of key interactions of active site residues with the reaction product of choline oxidation. The FAD cofactor is not modified on its C(4a) atom, as in the published structure of the wild-type enzyme \(^{13}\). The dimer structure of the enzyme in its complexes with glycine betaine and DMSO showed different conformations for a solvent accessible loop covering the hydrophobic cluster of the other subunit, underlying the importance of the dimeric state of the enzyme for catalysis.

### 2.3 Materials and methods

**Materials.** Escherichia coli strain Rosetta (DE3)pLysS was obtained from Novagen (Madison, WI). Magnesium acetate and PEG 6000 were purchased from Sigma-Aldrich (St. Louis, MO),
calcium chloride and choline chloride were from ICN Biomedicals (Irvine, CA) and glycerol was from Thermo Fisher Scientific (Waltham, MA). All other reagents were of the highest purity commercially available.

**Crystallization and X-Ray Data Collection.** Recombinant choline oxidase from *A. globiformis* was expressed in *E. coli* strain Rosetta (DE3)pLysS and purified as described previously \(^{(21)}\). The protein was stored in 20 mM Tris-Cl, pH 8.0. A single crystal was grown by the hanging drop vapor diffusion method at room temperature by mixing 1 µL of choline oxidase (6.6 mg/mL) with 1 µL of reservoir solution containing 0.1 M magnesium acetate (pH 6.0), 50 mM calcium chloride, 2.5% v/v glycerol and 10% w/v PEG 6000. The enzyme drop was equilibrated against 500 µL of reservoir solution and crystal growth was observed in 1 day. The crystal was soaked in the reservoir solution with 1M choline chloride and 20% glycerol as cryoprotectant for ~1 min and frozen immediately in liquid nitrogen. X-ray data were collected at 100 K on beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory.

**Structure Determination and Model Refinement.** The X-ray data were integrated and scaled using HKL2000 \(^{(30)}\). The structure was solved by molecular replacement using PHASER \(^{(31)}\) in the CCP4i suite of programs \(^{(32)}\) with the previously published crystal structure of choline oxidase as the initial model (PDB 2JBV) \(^{(13)}\). The crystal structure was refined with Refmac \(^{(33)}\) and manual adjustment and rebuilding were performed using the program COOT \(^{(34)}\). In both subunits glycine betaine was refined with 0.5 occupancy as suggested by the weak electron density. Higher peaks in the electron density map were observed for the two O atoms of the carboxylate group, while lower density peaks were observed for the other atoms. The crystal structure was deposited with PDB ID 4MJW.
Protein structures were superimposed on C\textsubscript{\alpha} atoms by using SUPERPOSE of the CCP4 suite\textsuperscript{(35)}. Figures of the structures were generated with PYMOL (http://www.pymol.org) and CCP4mg\textsuperscript{(36)}. The detection of tunnels to the active site was performed with the software CAVER\textsuperscript{(37)}. The number of approximating balls was set at 12, the minimum probe radius was 1.0 Å, the shell depth 4 Å, the shell radius 3 Å, the clustering threshold 3.5, and the starting point was set on the N5 atom of the flavin cofactor with a maximum distance of 3 Å and a desired radius of 5 Å.

2.4 Results

\textit{Structure of Choline Oxidase in Complex with Glycine Betaine.} The structure of choline oxidase in complex with the reaction product glycine betaine at pH 6.0 was solved in the space group P\textsubscript{4}32\textsubscript{1}2, after soaking the protein crystals in a solution with 1 M choline chloride. The space group was the same previously seen for the structures at pH 8.5 of the enzyme co-crystallized with DMSO bound at the active site (PDB 2JBV)\textsuperscript{(13)} and the active site variant V464A devoid of ligands (PDB 3LJP) at pH 6.0\textsuperscript{(16)}. The structure was refined to an $R$-factor of 0.15 and resolution of 1.95 Å. The crystallographic data and refinement statistics are presented in Table 2.1.

The enzyme-product complex crystallized as a homodimer with each monomer consisting of FAD- and substrate-binding domains. Overall, the fold of the protein polypeptide was the same as those previously described for choline oxidase in complex with DMSO\textsuperscript{(13)} and two active site variants, e.g., the S101A enzyme in complex with acetate\textsuperscript{(20)} and the V464A enzyme devoid of ligands\textsuperscript{(16)}. Superimposition of the C\textsubscript{\alpha} atoms of the enzyme in its complexes with glycine betaine and DMSO yielded rmsd values of 0.50 and 0.48 Å for the A and B chains over 526 and 529 amino acid residues, respectively, indicating that the two structures were practically identical. Most of the symmetrical H-bonding and electrostatic interactions between the two
monomers in each dimeric structure, which were previously described in the DMSO structure (13) are maintained in the enzyme-glycine betaine structure. The main interactions at the dimer interface are two symmetric sets of ionic pairs, i.e. listed in parentheses as ranges for subunits A and B, D358-R396 (2.9 Å to 3.1 Å) and R363-D397 (2.8 Å to 3.0 Å), and one symmetric set of hydrogen bonds, i.e., T256-E370 (2.8 Å in subunit A/2.9 Å in subunit B). Non-polar interactions of F253 with residues on the other subunit differ in the complexes with glycine betaine and DMSO (vide infra).
<table>
<thead>
<tr>
<th>Table 2.1: X-ray diffraction data collection and model refinement statistics.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>No. of reflections (total)</td>
</tr>
<tr>
<td>No. of reflections (unique)</td>
</tr>
<tr>
<td>(R_{\text{merge}}) (%)</td>
</tr>
<tr>
<td>(&lt;I/\sigma&gt;)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Redundancy</td>
</tr>
<tr>
<td>(CC_{1/2})</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
</tr>
<tr>
<td>(R_{\text{work}}/ R_{\text{free}}) (%)</td>
</tr>
<tr>
<td>No. of atoms</td>
</tr>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>Ligand (FAD, BET)</td>
</tr>
<tr>
<td>Solvent</td>
</tr>
<tr>
<td>Isotropic B factors (Å²)</td>
</tr>
<tr>
<td>Protein (main chain)</td>
</tr>
<tr>
<td>Protein (side chain)</td>
</tr>
<tr>
<td>FAD</td>
</tr>
<tr>
<td>Glycine betaine</td>
</tr>
<tr>
<td>Solvent molecules</td>
</tr>
<tr>
<td>r.m.s. deviation</td>
</tr>
<tr>
<td>Bonds length (Å)</td>
</tr>
<tr>
<td>Angle distance</td>
</tr>
<tr>
<td>Ramachandran plot results</td>
</tr>
<tr>
<td>Number of residues in favored region</td>
</tr>
</tbody>
</table>
**FAD-Binding Site.** The electron density map of the FAD cofactor is well defined and clearly indicates the covalent linkage to H99. FAD is buried in the flavin-binding domain and occupies the same position previously established in other crystal structures of the enzyme. A notable difference between the structures of the wild-type enzyme is that the isoalloxazine in the structure presented here does not contain an O atom covalently linked to the C(4a) atom, as previously reported for the enzyme in complex with DMSO (Figure 2.2) \(^{(13, 38)}\).

![Figure 2.2: Interactions of glycine betaine in the active site of choline oxidase and the conformation of the FAD cofactor in the structure of choline oxidase presented in this study. Panel A and B: glycine betaine in subunit B is shown in sticks in gray and the omit map contoured at 0.4 e/Å³ is in blue, FAD is shown in yellow sticks and water is shown as a red sphere. The side chains of the residues close to glycine betaine are displayed as cyan sticks and labelled; hydrogen bonds are highlighted with black dashes and the distance in Å is shown. For clarity Y465 is not shown. Panel C: FAD is shown as yellow sticks and the electron density map (2Fo-Fc) in blue contoured at 0.7 e/Å³.](image-url)
Consequently, the conformation of the isoalloxazine in the glycine betaine complex is fairly planar, with a slight V-shaped bend on the N(5)-N(10) axis. This is consistent with the flavin being in the reduced state, as observed for a number of other flavoenzymes \(^\text{(39, 40)}\), in agreement with the result that although choline was added to the protein crystal the product of its oxidation is found in the active site of the enzyme. The reduced state of the FAD cofactor is further confirmed by the change in color from yellow to colorless of the crystal that was observed on exposure to X-rays.

**Glycine Betaine Binding.** After refinement of the atoms belonging to the protein, FAD cofactor and waters, electron density not consistent with water molecules for shape and distance was present in front of the *re* face of the isoalloxazine ring. This electron density was tentatively modeled with various components of the crystallization cocktail, the substrate choline, the reaction product glycine betaine or the reaction intermediate betaine aldehyde. The product glycine betaine gave the best fit in the electron density (Figure 2.2). The carboxylate oxygen atoms were a good fit to the electron density near the flavin, while choline has only a single hydroxyl group in the corresponding position. The planar carboxylate group of glycine betaine also was a good fit in the electron density that is consistent with the sp\(^2\) character of the carboxylate carbon atom. This part of the electron density could not be modeled with sp\(^3\) hybridized carbon atom as present in choline. The presence of the product glycine betaine in the active site rather than choline is also in agreement with the notion that the reaction with choline is fast \(^\text{(5)}\) and that product release would be more difficult in the crystal.

The interactions of glycine betaine in the active site of choline oxidase are shown in Figure 2.2. The closest contact between the ligand and the flavin is through an O atom of the carboxylate of glycine betaine, which is 2.7 Å in subunit A and 2.6 Å in subunit B away from the N(5) atom of
FAD. Other interactions of the ligand carboxylate are with the O(4) atom of FAD (3.4 Å in subunit A/3.3 Å in subunit B), the side chain amide of N510 (2.8/3.0 Å), and the Nε2 atom of H466 (3.1/3.3 Å). Further contacts of the ligand carboxylate are with the hydroxyl of S101 (3.3/3.6 Å). The Cα atom of glycine betaine is close to the side chains of V464 (3.4/3.6 Å) and H351 (3.3/3.6 Å). The positively charged trimethylammonium moiety of glycine betaine is proximal to the aromatic side chains of W61 (4.0/3.9 Å), W331 (3.9/3.6 Å), and Y465 (3.9/4.1 Å). The trimethylammonium group of glycine betaine is (4.0/3.8 Å) away from the carboxylate of the side chain of E312.

**Conformations of the Loop Comprising Residues 250-255.** As illustrated in Figure 2.3, the superimposition of the structures of choline oxidase in complex with glycine betaine and DMSO showed two different conformations for loop 250-255 at the dimer interface in both subunits of the dimers.

**Figure 2.3:** Different conformations of loop 250-255 in the crystal structures of choline oxidase in complex with glycine betaine and with DMSO (PDB 2JBV).

Panel A: the superimposed structures are shown as light grey cartoon for choline oxidase in complex with glycine betaine and dark grey cartoon for choline oxidase in complex with DMSO (PDB 2JBV). The FAD cofactor of choline oxidase in complex with glycine betaine is shown in yellow sticks; the conformations of the main chains of loop 250-255 are highlighted in red for choline oxidase in complex with DMSO (PDB 2JBV) and in blue for choline
oxidase in complex with glycine betaine. For clarity, only loop 250-255 of subunit B is shown. Panel B: the main chain of the residues of loop 250-255 and the side chain of F253 of choline oxidase with glycine betaine are shown as blue sticks and the electron density of the omit map in blue (contoured at 0.5 e/Å³) and in grey (contoured at 0.33 e/Å³); the same atoms in the structure with DMSO (PDB 2JBV) are shown as red sticks.

The two conformations of loop 250-255 are significantly different with a largest distance between the Cα atoms of the residue A252 of 7.5 Å in subunit A and 7.7 Å in subunit B in the two superimposed structures. In the structure of the enzyme-glycine betaine complex, the side chain of F253 is partially disordered beyond Cγ with atoms visible in the electron density only at lower contour levels. Similar disorder of the F253 side chain is also observed in the enzyme-DMSO complex. The electron densities of the remaining portions of loop 250-255 are instead defined well in both complexes. By lowering the contour limits of the 2Fo-Fc electron density maps to 0.5 e/Å³ (1.49 rmsd) and 0.33 e/Å³ (0.99 rmsd) for the enzyme-DMSO and -glycine betaine complexes in Figure 3B (the former is not shown for clarity), it is evident that the side chain of F253 points to different directions in the two structures (Figure 2.4).

**Figure 2.4:** F253 positions in open and closed conformations of loop 250-255. The side chains of the hydrophobic cluster M62, L65, V355, F357 and M359 of chain A are shown in green sticks, while F253 of chain B is shown as blue sticks. FAD is in yellow sticks and the flavin rings are accessible from the surface through the tunnel to the active site. The surface is shown in gray in subunit A and in blue in subunit B. Panel A: choline oxidase in complex with glycine betaine (open conformation) is shown as grey surface and glycine...
betaine is shown in magenta sticks. Panel B: choline oxidase in complex with DMSO (closed conformation) is shown as grey surface and DMSO is shown in magenta sticks.

In the complex with glycine betaine, the side chain of F253 is close to W330 (3.5/3.6 Å) of the other subunit. In the enzyme-DMSO complex, instead, the aromatic ring of F253 stacks on the side chains of M62 (4.0/4.2 Å), L65 (3.6/4.3 Å), F357 (4.7/3.6 Å), and M359 (3.1/3.4 Å) of the other subunit. We define the conformation observed in the structure with glycine betaine in the active site as “open” and the “closed” conformation is seen in the previously published structure of the enzyme in complex with DMSO (PDB 2JBV) (13). Analysis of the published structures of choline oxidase demonstrates that both open and closed conformations are present in the two subunits of the V464A enzyme dimer at pH 6.0 (PDB 3LJP) (16), whereas the S101A enzyme (PDB 3NNE) (20) crystallized with acetate in the active site at pH 6.0 displays only the closed conformation in all eight subunits present in the crystallographic structure (data not shown).

**Predicted Tunnels to the Active Site.** One tunnel with radius of 1.0 Å was identified using CAVER that connects the surface of choline oxidase in complex with glycine betaine with the active site of the enzyme, as shown in Figure 2.5.
Figure 2.5: CAVER (32) analysis of tunnels from the active site to the enzyme surface. Panel A: choline oxidase in complex with glycine betaine; panel B: choline oxidase in complex with DMSO. The hydrophobic cluster (M62, L65, V355, F357, M359), and the residue D358, are shown in blue sticks, FAD is shown in yellow sticks, and the side chain of F253 of the other subunit is shown as red stick. The tunnel with the entrance regulated by the side chains of the hydrophobic cluster M62, L65, V355, F357, and M359 is colored orange in panel A.

This tunnel is delimited by the side chains of the hydrophobic cluster M62, L65, V355, F357 and M359, which was previously predicted to gate the access of the substrate to the active site through molecular dynamics (41). In the structure of the wild-type enzyme in complex with DMSO no tunnel was predicted going through the hydrophobic cluster, primarily due to steric hindrance by the side chain of F253 from the other subunit of the dimer (Figure 2.5B). Thus, it appears that the conformation of the side chain of F253 from the other subunit of the dimeric structure determines if the tunnel is open or closed.

2.5 Discussion

The present study reports the first crystal structure of choline oxidase in complex with glycine betaine, the product of the oxidation of choline catalyzed by the enzyme. Prior to this study two other structures of choline oxidase in complex with components of the crystallization solution
(DMSO or acetate) were reported \(^\text{(13, 20)}\), but none with physiologically relevant molecules. Within the GMC enzyme oxidoreductase superfamily, the structure of fungal pyranose 2-oxidase from *Peniophora* sp. is the only other example of an alcohol oxidase in complex with a reaction product, e.g., 2-keto-\(\beta\)-D-glucose \(^\text{(42)}\). The structure of the choline oxidase-glycine betaine complex is therefore an important breakthrough that complements previous mechanistic investigations on the catalytic roles of several amino acid residues in the active of the enzyme \(^\text{(10, 13-16, 18-20, 43-47)}\). The comparison of the structure of the enzyme-product reported here with those previously obtained for choline oxidase allowed the identification of conformational and topological differences at the dimer interface, with implications for the mechanism of substrate access to the active site. We assume the conformational differences arise from the different crystallization conditions or presence of mutations, since the compared crystal structures were refined in the same space group with similar cell parameters, however, we cannot rule out the possible influence of crystal contacts in constraining the enzyme conformation.

The closest interaction of the carboxylate of glycine betaine with the flavin is with the N(5) atom of the isooaloxazine (Figure 2.2). The carboxylate group also interacts with the protein through the side chains of H466 and N510 (Figure 2.2). Despite its proximity, the carboxylate of glycine betaine likely does not interact with the hydroxyl of S101 due to a non-optimal orientation for H-bonding. It is noteworthy to consider that the carboxylate C atom of glycine betaine is the same C atom that in choline is oxidized in the reaction catalyzed by choline oxidase, i.e., the C\(\alpha\) atom carrying the hydroxyl O atom and from which the hydride ion that reduces the flavin originates. Thus, all the interactions between the carboxylate of glycine betaine in the active site are consistent with the mechanism for the oxidation of choline, in that: \(\alpha\) an hydride ion is transferred from the choline C\(\alpha\) atom to the flavin N(5) atom, as suggested by kinetic isotope
effects (5, 6, 48); b) the serine hydroxyl stabilizes the transition state for the proton transfer reaction that converts choline to choline alkoxide, as suggested by mutagenesis of S101 (14); c) the positively charged imidazolium of H466 stabilizes the alkoxide reaction intermediate in the oxidation of choline, as suggested by activity rescuing of the H466A mutant at low pH (18); d) N510 is important for the relative timing for the cleavages of the OH and CH bonds of choline, as suggested by multiple kinetic isotope effects studies of choline oxidase with N510 replaced with alanine or histidine (19).

The position of H466 near the carboxylate of glycine betaine suggests this residue may be the catalytic base that carries out substrate activation to the alkoxide species by catalyzing the cleavage of the alcohol OH bond. In this respect, the previous characterization of the mutant enzymes H466A (18) or H351A (15) showed that an active site base was still present in the active sites of these enzymes, as suggested by pH profiles of $k_{cat}/K_{choline}$ (15, 18). However, a water molecule acting as surrogate base could occupy the space of the imidazole side chains of the histidine residues once they are substituted with alanine in the mutant enzymes. This hypothesis is being currently investigated on a mutant choline oxidase in which H466 is replaced with glutamine. H466 is fully conserved in the active sites of the GMC enzymes, and its counterpart in pyranose 2-oxidase and aryl-alcohol oxidase has been shown to act as base (49, 50).

The trimethylammonium moiety of glycine betaine interacts with the side chains of W61, Y465 and W331 (Figure 2.2). The trimethylammonium group is preserved in the oxidation of choline to glycine betaine, therefore similar interactions with the enzyme are expected for this group in the product and substrate complexes. The trimethylammonium group is also (4.0/3.6 Å) away from the side chain of E312. The importance of E312 for substrate binding was established through mechanistic studies of the enzyme variant E312Q, which showed a 500-fold increase in
the $K_d$ value for choline in rapid kinetics $^{13}$. An independent mechanistic study with choline analogs carrying one or two methyls on the amine portion of the molecule demonstrated the importance of hydrophobic interactions between the methyls and active site residues in choline oxidase $^{51}$. This is in agreement with W61, Y465 and W331 forming hydrophobic interactions with the trimethylammonium moiety of glycine betaine, as observed in the structure of the enzyme-glycine betaine complex. It is noteworthy that the carboxylate of acetate in the previously published crystal structure of the S101A enzyme $^{20}$ has a different orientation from the carboxylate in glycine betaine, further consistent with the importance of the trimethylammonium moiety in ligand binding.

The C$_\alpha$ atom of glycine betaine, which corresponds to the C$_\beta$ atom of choline before its oxidation by the enzyme, is close to both H351 and V464 (Figure 2.2). A mechanistic investigation of H351 through site directed mutagenesis showed that this residue is important for substrate binding and positioning and contributes to the stabilization of the transition state for the hydride transfer reaction catalyzed by the enzyme $^{15}$. A mechanistic study on V464 mutants replaced with threonine or alanine showed that the size and hydrophobic character of this residue are important for the localization of O$_2$ close to the FAD C(4a) atom allowing the reoxidation of the reduced flavin in turnover $^{16}$.

The two structures of choline oxidase in complex with glycine betaine and DMSO highlight distinct conformations of the loop of residues 250-255, which is located at the poles of the dimer interface close to the proposed site of access to the active site of the other subunit delimited by the hydrophobic cluster M62, L65, V355, F357, and M359. In the crystal structure of the V464A enzyme without any bound ligand, loop 250-255 is present in both the open conformation in one subunit and in the closed one in the other subunit. This indicates the variability of this loop in the
different structures of the enzyme. The closed conformation results in a constricted access of the tunnel entrance delimited by the hydrophobic cluster to access the active site. Thus, the active site is secluded from the bulk solvent when loop is in the closed conformation, which is probably required for the hydride tunneling and for the reactivity of the reduced flavin cofactor with oxygen. Interestingly, this loop assumes an open conformation when the product is bound in the active site. In the case of the previously published crystal structure of wild-type enzyme with DMSO the loop 250-255 is present in a closed conformation, whereas in the crystal structure of choline oxidase with the natural product glycine betaine the same loop assumes an open conformation (Figure 2.4). We hypothesize that loop 250-255 swings between the open and closed conformations in the ligand free form of the enzyme and it is then stabilized in the closed conformation upon substrate binding to provide an optimal environment for the hydride tunneling and for oxygen reactivity. It is possible that the water involved in the reaction is already present in the active site when choline is bound with the closed loop. The water molecule would become reactive for hydration only upon formation of the aldehyde intermediate. Alternatively, the open conformation of the loop 250-255 observed in the crystal structure of the enzyme in complex with glycine betaine could be triggered by the formation of the product in the active site and play a role in product release.

Software tools that compute tunnels in structures of proteins can give important insights on substrate access to the active site and on residues located around bottlenecks of tunnels (52). The predicted tunnel in the structure of choline oxidase in complex with glycine betaine through the entrance delimited by the side chains of the hydrophobic cluster M62, L65, V355, F357, and M359 is likely utilized by choline and glycine betaine to enter and leave the active site cavity (Figure 2.5). The side chain of D358 is located near the entrance of the tunnel, consistent with a
role of electrostatics in guiding positively charged choline into the active site as described in a previous computational study of choline oxidase \(^{(29)}\). In the case of the enzyme complex with DMSO, no tunnel was predicted as a consequence of the closed conformation of loop 250-255 and the rotated side chain of F253 (Figure 2.5). In the free enzyme devoid of ligands, exemplified by the structure of the V464A enzyme, which is almost identical to that of the wild-type enzyme, both open and closed conformations of loop 250-255 are present in the two subunits, with the tunnel entrance open or constricted, respectively (data not shown).

### 2.6 Conclusion

The crystal structure of choline oxidase in complex with the reaction product glycine betaine was solved to a resolution of 1.95 Å. This crystal structure is the first reported for choline oxidase with the physiological product bound and the second structure of a flavoenzyme of the GMC superfamily with a natural ligand in the active site. The fact that choline oxidase has been extensively characterized kinetically allows the comparison of the new structural data for glycine betaine in the active site with the kinetic data obtained by previous studies. The residues that were concluded to be important for catalysis by kinetic studies are confirmed in their role by their interactions with glycine betaine observed in the present crystal structure. Moreover, the present study suggests the direction of further studies on choline oxidase, such as investigation of H466 as the catalytic base in the active site of the enzyme. The prediction of tunnels highlights the role of the hydrophobic cluster M62, L65, V355, F357, and M359 as the bottleneck of the tunnel leading to the active site. Further studies will be undertaken using site-directed mutagenesis of the side chains of the hydrophobic cluster M62, L65, V355, F357, and M359.

One main difference is described between the crystal structures of choline oxidase with DMSO \(^{(13)}\) and with glycine betaine. The shift of the loop 250-255 and in particular the highly flexible
side chain of F253 define an open and a closed conformation that are consistent with a gating mechanism to control the access to the active site through a pore delimited by the hydrophobic residues M62, L65, V355, F357 and M359. This study sets the stage for future studies of site-directed mutagenesis of F253 to investigate its role in controlling substrate access to the active site. The different conformations of loop 250-255, especially of F253 that covers the active site of the other subunit, suggest an important role played by the dimeric state of the enzyme in controlling substrate access.

2.7 Acknowledgements

X-ray data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. We thank Elvira Romero for purifying the enzyme, Johnson Agniswamy for mounting the protein crystals for data collection and for data collection, and Robert Harrison for insightful discussions. This work was supported in part by Grant MCB-1121695 from the NSF (G.G.).

2.8 References


CHAPTER III: INVOLVEMENT OF F357 IN AN ISOMERIZATION PARTIALLY RATE LIMITING FOR FLAVIN OXIDATION IN CHOLINE OXIDASE

(The author carried out all the experiments described in this Chapter, with the exception of the molecular dynamics simulations, which were carried out by Isela Rodriguez and Dr. Donald Hamelberg)

3.1 Abstract

Choline oxidase from *Arthrobacter globiformis*, catalyzes the oxidation of choline to glycine betaine using oxygen as electron acceptor. A previous mechanistic study of the oxidative half reaction of choline oxidase established that the reaction of the reduced FAD cofactor with oxygen occurs without stabilization of detectable intermediates and with a synchronous transfer of one hydrogen and one proton to oxygen to yield hydrogen peroxide. A partially rate limiting isomerization of the reduced enzyme during the reaction with oxygen was detected by solvent viscosity studies. In this study we hypothesized that the side chains of M62 and F357 at the entrance to the active site of choline oxidase may be related to the slow isomerization detected and we engineered a double variant enzyme M62A/F357A. The kinetic characterization of the double variant enzyme showed the lack of the isomerization detected in wild-type choline oxidase, and a lack of saturation with oxygen concentrations as high as 1 mM, while most kinetic parameters were similar to wild-type choline oxidase. The kinetic characterization of the single variant enzymes established that only the side chain of F357 plays a role in the isomerization of choline oxidase in the oxidative half reaction. Molecular dynamics studies suggested that a possible explanation for the slow isomerization related to F357 is that the phenyl ring of this residue participates in a gating mechanism for a narrow tunnel proposed as regulating oxygen access to the reduced cofactor.
3.2 Introduction

Choline oxidase (EC 1.1.3.17, choline:oxygen 1-reductase) catalyzes the oxidation of choline to glycine betaine through two subsequent hydride transfer reactions to the flavin cofactor. The hydride transfers represent the rate limiting steps in the overall turnover of this enzyme (Scheme 3.1). After each hydride transfer the reduced flavin cofactor is oxidized by oxygen (Scheme 3.1).

Scheme 3.1: Choline oxidation followed by the reaction of the anionic hydroquinone with oxygen. RHR, reductive half reaction; Sub. KIE, substrate kinetic isotope effect; Sol. KIE, solvent kinetic isotope effect.

Recent reviews have summarized the important contributions of many mechanistic, structural and computational studies on the oxygen reactivity of flavin hydroquinones. It was
concluded that important features for the reaction with oxygen are a positive charge and of a non-polar environment near the flavin cofactor, which in choline oxidase are represented by the positively charged reaction product of the oxidation of choline and V464, respectively.\(^4\) The reaction of reduced flavin with oxygen is initiated by a single electron transfer to oxygen that yields the radical pair \(\text{O}_2^\bullet^-\) and flavosemiquinone (Scheme 3.1).\(^5,6\) A mechanistic investigation of the oxygen reactivity of choline oxidase concluded that the oxidation of the hydroquinone to oxidized flavin and hydrogen peroxide occurs without stabilization of any detectable intermediate with a synchronous transfer of a hydrogen from the N5 atom of the cofactor and of a proton from a solvent exchangeable site to oxygen (Scheme 3.1).\(^7\) Moreover the inverse hyperbolic behavior of the solvent viscosity effect on \(k_{\text{cat}}/K_{\text{ox}}\) on choline oxidase established the presence of an isomerization partially rate limiting in the reaction of reduced flavin with oxygen \(^7\), which was not observed in other oxidases. Indeed solvent viscosity studies performed on bovine serum amine oxidase detected no effect of viscosity on \(k_{\text{cat}}K_{\text{ox}}\), indicative of oxygen diffusion being not rate limiting.\(^8\) In contrast, a kinetic investigation on glucose oxidase showed a linear positive effect of viscosity on \(k_{\text{cat}}/K_{\text{ox}}\), which is consistent with oxygen access to the active site being diffusion controlled \(^9\). The isomerization detected in choline oxidase was not further investigated and could be possibly related to a gating process for oxygen accessing the site of the reaction with the reduced flavin.

Specific hydrophobic cavities for diffusion of a gas have been visualized in crystals of copper amine oxidase saturated with xenon.\(^10\) Tunnels for oxygen have been investigated with molecular dynamics and site-directed mutagenesis studies in copper amine oxidase\(^10,11\), cytochrome c oxidase\(^12\), 12/15-lipoxygenase\(^13\), D-amino acid oxidase\(^14\), cholesterol oxidase type I and II\(^15,16\), and aryl-alcohol oxidase\(^17\). In the case of choline oxidase a tunnel to the
active site for choline access and glycine betaine release has been identified by molecular
dynamics studies with the entrance regulated by the side chains of a hydrophobic cluster (M62,
L65, V355, F357, and M359) at the dimer interface (Figure 3.1 panel A)\cite{18}.

**Figure 3.1:** Entrance to the active site of choline oxidase.
Panel A and B: the side chains of M62, L65, V355, F357, and M359 of subunit B are shown as green spheres, the
surface of subunit A at the dimer interface is shown as magenta surface, the side chain of F253 of subunit B is
highlighted by magenta sticks, FAD is in yellow sticks. Panel A represents the open conformation of loop 250-255
as observed in choline oxidase in complex with the product glycine betaine (PDB 4MJW) and panel B represents the
closed conformation observed in choline oxidase in complex with DMSO (PDB 2JBV). Panel C: tunnel A and B are
shown as orange and gray, respectively, glycine betaine as blue sticks, the side chains of M62, F357, E312, and
V464 are shown as green sticks, FAD is in yellow sticks.
A recent study on the crystal structure of choline oxidase in complex with the product glycine betaine has highlighted two different conformations of loop 250-255, with the side chain of F253 in the closed conformation covering the proposed entrance to the active site of the other monomer (Figure 3.1 panel A and B). The side chains of F357 and M62 face each other in the hydrophobic cluster and possibly play a role in the partially rate-limiting isomerization detected in the reaction of reduced choline oxidase with oxygen. This represents the first kinetic investigation of a partially rate-limiting isomerization in the oxygen reactivity of flavooxidases. Furthermore the relative contribution of the two side chains of M62 and F357 was analyzed using the inverse thinking approach on double variant enzymes. This approach is based on taking as a reference point the kinetic parameters of the double variant enzyme and then analyze how much of the activity of the wild-type protein is restored in the single variants.

3.3 Materials and methods

**Materials.** *Escherichia coli* strain Rosetta(DE3)pLysS was from Novagen (Madison, WI). QIAprep Spin Miniprep kit and QIAquick PCR purification kit were from Qiagen (Valencia, CA). *Pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA) and *DpnI* enzyme from New England BioLabs (Ipswich, MA). Oligonucleotides for site-directed mutagenesis of the mutant genes were from Sigma Genosys (The Woodlands, TX). Choline chloride was from ICN Pharmaceutical Inc. (Irvine, CA), 1,2-[\(^2\)H\(_4\)]choline bromide (98%) and sodium deuterioxide were bought from Isotec Inc. (Miamisburg, OH). Deuterium oxide (99.9%) and deuterium chloride (99.5%) were purchased from Cambridge Isotope Co. (Andover, MA). All other reagents used were of the highest purity commercially available.

**Site-directed mutagenesis, protein expression and purification.** The single mutant genes *codA*/M62A and *codA*/F357A were prepared by site-directed mutagenesis using as template the
pET20b(+) /codA plasmid harboring the wild-type gene. The amplification of the mutagenic primers by *Pfu* DNA polymerase was checked by agarose gel electrophoresis and the PCR products were purified by using the QIAquick PCR Purification Kit from Qiagen. The purified PCR products were treated with *Dpn*I at 37 °C for 2 h and then used to transform DH5α strain of *E. coli*. In the case of the double mutant gene codA/M62AF357A the single mutant gene codA/F357A was used as a template to insert the second mutation M62A. The presence of the desired mutations was confirmed by DNA sequencing at the Cell, Protein, and DNA core facility at Georgia State University. The recombinant single variant M62A and F357A and the double variant M62AF357A enzymes were expressed in *E. coli* strain Rosetta(DE3)pLysS and purified to homogeneity as previously described for the wild-type choline oxidase(21) in presence of 10% (v/v) glycerol to increase the stability of the enzymes.

**Enzyme assays.** UV-visible absorbance was recorded with an Agilent Technologies diode-array spectrophotometer Model HP 8453 PC equipped with a thermostated water bath. The determination of the extinction coefficient of the enzyme-bound flavin and of the covalent linkage of the FAD cofactor to the protein have been carried out as previously described. (22) The enzymatic activity was measured with the method of initial rates (23) by following oxygen consumption with a computer-interfaced oxygen electrode from Hansatech, as previously described for wild-type choline oxidase. (7) The measured initial rates were normalized for the enzyme-bound flavin with the experimentally determined $\varepsilon_{450} = 11,050 \text{ M}^{-1} \text{ cm}^{-1}$, $\varepsilon_{454} = 11,180 \text{ M}^{-1} \text{ cm}^{-1}$, and $\varepsilon_{450} = 11,500 \text{ M}^{-1} \text{ cm}^{-1}$ for the enzymes M62A/F357A, F357A, and M62A, respectively. The steady-state kinetics experiments were carried out at 25 °C by varying the concentration of both choline from 0.02 to 50 mM, and oxygen from 44 to 1000 µM. The reaction mixture containing buffer and choline was equilibrated to the desired oxygen concentration by bubbling
an O₂/N₂ gas mixture, and the reaction was started by addition of choline oxidase to a final concentration between 95 and 600 nM. The pH profiles of the steady-state kinetic parameters were obtained by using sodium pyrophosphate or sodium phosphate at a final concentration of 50 mM for the different pH values. The apparent steady-state kinetic parameters for choline were determined at atmospheric oxygen and 25°C, in 50 mM sodium phosphate, pH 7.0.

Multiple deuterium kinetic isotope effects (KIEs) on the kinetic parameter \( \text{app}(k_{\text{cat}}/K_{\text{ox}}) \) were carried out on the double variant M62A/F357A with the procedure described above for the enzymatic activity by using 1,2-[\(^2\)H₄]choline bromide (98%) and buffers containing 99.9% D₂O, as described for the wild-type enzyme.\(^{(7)}\) The pH values of the buffers containing D₂O were adjusted by adding 0.4 to the measured pH value.\(^{(24)}\) The experiment was performed at fixed saturating concentration (40 or 60 mM) of choline or 1,2-[\(^2\)H₄]choline bromide and varying oxygen from 38 to 834 µM.

Solvent viscosity studies on the kinetic parameter \( \text{app}(k_{\text{cat}}/K_{\text{ox}}) \) were performed at 25 °C in 50 mM sodium pyrophosphate, pH 10.0, with the procedure described above using glycerol as viscogen, in the same conditions used for wild-type choline oxidase.\(^{(7)}\) The viscosity values at 25 °C of the reaction mixture at the different concentrations of glycerol were calculated using relative viscosities at 20 °C available from Lide.\(^{(25)}\) The solvent viscosity study on the kinetic parameter \( \text{app}(k_{\text{cat}}/K_{\text{ox}}) \) of the double variant enzyme M62AF357A was also performed at pH 7.0, 8.0 or 9.0 in 50 mM sodium phosphate or pyrophosphate and 25 °C.

The reductive half reaction was performed with an SF-61DX2 Hi-Tech KinetAsyst high-performance stopped-flow spectrophotometer, thermostated at 25 °C, which was made anaerobic by an overnight treatment with the oxygen scavenging system glucose (5 mM)/glucose oxidase (1 µM) in 100 mM sodium pyrophosphate, pH 6.0. The rate constants for flavin reduction were
measured at 25 °C and pH 10.0 by monitoring the decrease in absorbance at 450 nm. The enzyme was passed through a PD10 column equilibrated in 50 mM sodium pyrophosphate, pH 10.0, and loaded in a tonometer which was made anaerobic by 25 cycles of degassing by alternating vacuum and flushing with argon. The syringes containing choline or 1,2-[\textsuperscript{2}H\textsubscript{4}]choline in 50 mM sodium pyrophosphate, pH 10.0, were made anaerobic by flushing them with argon for 30 min. Glucose (2 mM)/ glucose oxidase (0.5 µM) were added to enzyme, buffer, and substrates solutions to scavenge any trace of oxygen. After mixing anaerobically an equal volume of enzyme and substrate the concentration of enzyme was 11 µM and that of substrate ranged from 0.07 to 12 mM.

\textbf{Data analysis.} The kinetic parameters at a fixed concentration of organic substrate and varying oxygen concentrations were obtained from the Michaelis-Menten equation using KaleidaGraph (Synergy Software, Reading, PA). Steady-state kinetic data at varying choline and oxygen concentrations were fit to equation 1 in the case of the enzymes M62AF357A and M62A and to equation 2 for the enzyme F357A using Enzfitter software (Biosoft, Cambridge, U.K.).

\[
\frac{v_0}{e} = \frac{k_{\text{cat}} [\text{CH}][\text{O}_2]}{K_{\text{ch}} [\text{O}_2] + K_{\text{ox}} [\text{CH}]+[\text{CH}][\text{O}_2]+K_{\text{ia}} K_{\text{ox}}}
\]

Eq. 1

\[
\frac{v_0}{e} = \frac{k_{\text{cat}} [\text{CH}][\text{O}_2]}{K_{\text{ox}} [\text{CH}]+[\text{CH}][\text{O}_2]+K_{\text{ia}} K_{\text{ox}}}
\]

Eq. 2

Equation 1 represents a sequential steady-state kinetic mechanism where \( v_0 \) represents the initial velocity, \( e \) is the concentration of enzyme, \( k_{\text{cat}} \) is the first-order rate constant for enzyme turnover at saturating concentration of both substrates, \( K_{\text{ch}} \) and \( K_{\text{ox}} \) are the Michaelis constants for choline and oxygen, respectively, and \( K_{\text{ia}} \) is a kinetic constant that accounts for the intersecting line
pattern in the double reciprocal plot. Equation 2 represents a sequential steady-state kinetic mechanism of the type described by Equation 1 when $K_{ch} \ll K_{ia} K_{ox}$.

Solvent viscosity effects on $k_{cat}/K_{ox}$ were calculated by dividing the value of this kinetic parameter determined in the absence of the viscogen by the value determined in the presence of glycerol.

Stopped-flow traces were fit to Equation 3 or Equation 4 which represent a single and double exponential process, respectively. $A$ represents the absorbance at 450 nm at time $t$, $B_1$ and $B_2$ are the amplitudes of the decrease in absorbance associated with the first and second phases, $k_{obs1}$ and $k_{obs2}$ represent the observed rate constants for the change in absorbance, and $C$ is an offset value accounting for the nonzero absorbance value at infinite time.

$$A = B_1 \exp (-k_{obs1} t) + C$$ \hspace{1cm} \text{Eq. 3}

$$A = B_1 \exp (-k_{obs1} t) + B_2 \exp (-k_{obs2} t) + C$$ \hspace{1cm} \text{Eq. 4}

Concentration dependence of the observed rate constants for flavin reduction was analyzed with Equation 5, where $S$ represents the concentration of organic substrate, $k_{red}$ is the rate constant for flavin reduction at saturating substrate concentration, and $K_d$ is the dissociation constant for substrate binding.

$$k_{obs} = \frac{k_{red} S}{K_d + S}$$ \hspace{1cm} \text{Eq. 5}

**Molecular dynamics.** Three 110 ns molecular dynamics simulations, for substrate-free dimeric wild-type choline oxidase, were conducted by using Amber 10/14 molecular simulations program package: a 110 ns MD simulation on the substrate-free wild-type choline oxidase, a 110 ns MD simulation of the F357A variant, and a 110 ns MD simulation of the M62A variant. PDB
4MJW, resolution of 1.95 Å, of choline oxidase crystal structure from *A. globiformis* was used for all simulations. All of the simulations were carried out on substrate-free choline oxidase in the explicit TIP3P water model, in a periodic octahedron box, and with the force-field 14SB at a temperature of 298 K. The systems were neutralized with chloride ions. All simulations were also carried out using the NTP ensemble at a constant pressure of 1 bar. Initial minimization was performed on the systems for 1000 steps with harmonic constraints of 100 kcal/mol/Å², followed by a short 0.5 ps MD simulation with position constraints of 50 kcal/mol/Å², and a 1 ns MD simulation with position constraints of 25 kcal/mol/Å². Final equilibration was conducted for 1 ns without any harmonic constraints. The SHAKE algorithm was used to constrain bonds involving hydrogen and the Langevin thermostat, with a collision frequency of 1 ps⁻¹, was used to maintain control of the temperature at 298 K. The electrostatic interactions were evaluated via the particle mesh Ewald method and a cutoff of 9.0 Å was used for the nonbonded interactions. A time step of 2 fs was used for the all three MD simulations to integrate Newton’s equation of motion. PTRAJ, analysis program given by amber tools, and Xmgrace, plotting program, were used to represent the chi dihedral angle distribution of the hydrophobic cluster amino acids produced by the MD simulations. To obtain information about tunnel and channel formation within the enzyme, PTRAJ was used to convert the existing trajectory frames into PDB format, and CAVER 3.0, software for tunnel analysis and visualization, was used to calculate tunneling.

### 3.4 Results

**Protein expression and purification.** The variant enzymes M62A/F357A, M62A, and F357A were expressed in *E. coli* and purified as described for wild-type choline oxidase. Glycerol (10% v/v) was added throughout the purification steps to increase the stability of the enzymes. All three variant enzymes stabilized the anionic form of the flavin semiquinone during
purification, which was fully oxidized by extensive dialysis at pH 6.0, as in the case of wild-type choline oxidase.\(^{(21, 26)}\) The UV-visible absorbance spectra of the fully oxidized enzymes (Figure 3.2) had maxima in the 360 and 450 nm region, which are characteristic of flavoproteins.

The variant enzymes showed small (≤4 nm) hypsochromic shifts of the high energy flavin band compared to wild-type choline oxidase (Figure 3.2). The enzymes M62A/F357A, and M62A showed negligible (≤2 nm) hypsochromic shifts of the low energy flavin band, while F357A showed a negligible (≤2 nm) bathochromic shift of the low energy flavin band. In all variant enzymes treatment of the protein with trichloroacetic acid followed by centrifugation yielded a
yellow pellet of denatured holoenzyme and a clear supernatant, consistent with the flavin cofactor being covalently bound to the protein, as in the case of wild-type choline oxidase. These results suggest the mutations did not dramatically affect the protein environment surrounding the flavin.

**Solvent viscosity effect on** $k_{cat}/K_{ox}$. Solvent viscosity studies were performed on the variant enzymes prepared in this study in order to monitor the slow isomerization occurring during the oxidation of the cofactor by oxygen in choline oxidase (Scheme 3.1). Indeed wild-type choline oxidase showed an inverse solvent viscosity effect on the parameter $k_{cat}/K_{ox}$. The effect of glycerol as a viscogen on the parameter $k_{cat}/K_{ox}$ was investigated at 25 °C in a pH range from 7.0 to 10.0 for the double variant enzyme M62A/F357A, in the same conditions of the wild-type protein. No solvent viscosity effect was observed in this pH range (Figure 3.3). At 25 °C and pH 10.0 the single variant F357A displayed the same behavior of the enzyme M62A/F357A with no viscosity effect on $k_{cat}/K_{ox}$, while the single variant M62A shows an inverse viscosity effect comparable to the one previously detected in wild-type choline oxidase (Figure 3.3).
Figure 3.3. Solvent viscosity effect on $k_{\text{cat}}/K_{\text{ox}}$ at pH 10.0 and 25 °C, with glycerol as viscogen from 4.5% to 24% v/v.

The normalized $k_{\text{cat}}/K_{\text{ox}}$ is shown as a function of relative viscosity for the enzyme M62A/F357A (A), F357A (B), M62A (C), and wild-type choline oxidase (D). The theoretical behavior of a fully diffusion-limited reaction is represented by the dashed line with a slope of 1 and the result expected by a lack of diffusion control is shown by the dashed line with slope of 0. Error bars are shown in red.

**Steady-state kinetics.** The steady-state kinetics of the variant enzymes was investigated by monitoring oxygen consumption at varying concentrations of choline and oxygen at pH 10.0 and 25 °C. The pH value of 10.0 was chosen because it represents the independent region for the kinetic parameters in choline oxidase and all mutant variants investigated thus far, such as H99N, S101A, E312Q, H351A, H466A, V464A, N510A. (1, 27) Differently from wild-type choline oxidase in the case of the variant enzymes F357A and M62A/F357A it was not possible to achieve saturation with oxygen concentrations as high as 1 mM (Figure 3.4), consistent with a $K_{\text{ox}}$ larger than this value, while the enzyme M62A displayed a decreased $K_{\text{ox}}$ compared to the wild-type (Table 3.1).
Figure 3.4. Steady-state kinetics of enzymes M62A/F357A, M62A, and F357A performed at 25 °C in 50 mM sodium pyrophosphate pH 10.0.
Panel A: dependence of the initial rate on oxygen concentration for the enzyme M62A/F357A at choline concentrations of 20 (solid circles), 7 (solid squares), 2 (empty circles), 1 (solid triangles), 0.4 (empty squares) mM. Panel B: dependence of the initial rate on oxygen concentration for the enzyme M62A at choline concentrations of 1 (solid circles), 0.2 (empty squares), 0.1 (solid triangles), 0.02 (solid squares) mM. Panel C: Lineweaver-Burk plot of $e/v_0$ as a function of the inverse of choline concentration at oxygen concentrations of 822 (solid circles), 666 (solid triangles), 496 (solid squares), 277 (empty circles) μM.
Table 3.1. Kinetic parameters of choline oxidase wild-type (1, 28) and variant enzymes prepared in this study at pH 10.0 and 25 °C with choline as a substrate.

* a the parameter could not be determined because it was impossible to achieve saturation with 1 mM oxygen. b the parameter $K_{ch}$ could not be determined because for the enzyme F357A $K_{ch} \ll K_{ia} K_{ox}$ (Eq. 2).

<table>
<thead>
<tr>
<th>Kinetic parameter</th>
<th>M62A/F357A</th>
<th>F357A</th>
<th>M62A</th>
<th>Wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>nd$^a$</td>
<td>nd$^a$</td>
<td>19.1 ± 0.9</td>
<td>60 ± 1</td>
</tr>
<tr>
<td>$k_{cat}/K_{ox}$ (M$^{-1}$ s$^{-1}$)</td>
<td>19,300 ± 200</td>
<td>16,800 ± 200</td>
<td>89,000 ± 10,000</td>
<td>86,400 ± 3,600</td>
</tr>
<tr>
<td>$k_{cat}/K_{m}$ (M$^{-1}$ s$^{-1}$)</td>
<td>65,000 ± 23,000</td>
<td>nd$^b$</td>
<td>150,000 ± 23,000</td>
<td>237,000 ± 9,000</td>
</tr>
<tr>
<td>$K_{ox}$ (μM)</td>
<td>nd$^a$</td>
<td>nd$^a$</td>
<td>210 ± 20</td>
<td>690 ± 30</td>
</tr>
<tr>
<td>$K_{ch}$ (mM)</td>
<td>1.0 ± 0.4</td>
<td>nd$^b$</td>
<td>0.13 ± 0.02</td>
<td>0.25 ± 0.01</td>
</tr>
<tr>
<td>$K_{ia}$ (mM)</td>
<td>3.7 ± 0.7</td>
<td>6.6 ± 0.1</td>
<td>0.10 ± 0.02</td>
<td>0.14 ± 0.01</td>
</tr>
</tbody>
</table>

The best fit for the steady-state kinetic data of the enzymes M62A and M62A/F357A was obtained with Eq. 1, representing a sequential steady-state kinetic mechanism, as in the case of wild-type choline oxidase. (1) The best fit for the enzyme F357A was obtained with Eq. 2, representing a particular case of sequential steady-state kinetic mechanism when $K_{ch} \ll K_{ia} K_{ox}$, as shown by the lines intersecting at the origin in the Lineweaver-Burk plot in Figure 3.4 panel C. The steady-state kinetic parameters determined at pH 10.0 and 25 °C for the variant enzymes are shown in Table 3.1 in comparison with wild-type choline oxidase. (28) The variant enzymes F357A and M62A/F357A showed an increase in $K_{ox}$ and a 5 times decrease in the kinetic parameter $k_{cat}/K_{ox}$ compared to wild-type choline oxidase (1), while the variant enzyme M62A displayed kinetic parameters similar to the wild-type enzyme.

Multiple deuterium kinetic isotope effects on $k_{cat}/K_{ox}$. The relative timing for the transfer of the hydrogen and the proton to oxygen to yield H$_2$O$_2$ was investigated through the use of multiple kinetic isotope effects with 1,2-[$^2$H$_4$]choline and deuterated water on the parameter $k_{cat}/K_{ox}$, as
previously performed on wild-type choline oxidase.\(^7\) The kinetic parameter \(\text{app}(k_{\text{cat}}/K_{\text{ox}})\) was determined at choline concentrations of 40 mM and 60 mM and varying oxygen at 25 °C in 50 mM sodium phosphate, pH 7.0, yielding similar values of 20500 ± 800 and 20000 ± 300 M\(^{-1}\) s\(^{-1}\) for choline and 17700 ± 100 and 17600 ± 200 M\(^{-1}\) s\(^{-1}\) for 1,2-\([^{2}\text{H}_{4}]\)choline. Similar results were obtained by repeating this experiment in 50 mM sodium pyrophosphate, pH 8.0, with \(\text{app}(k_{\text{cat}}/K_{\text{ox}})\) values of 27300 ± 400 and 27800 ± 2000 M\(^{-1}\) s\(^{-1}\) for choline and 15400 ± 1800 and 13000 ± 2800 M\(^{-1}\) s\(^{-1}\) for 1,2-\([^{2}\text{H}_{4}]\)choline. The similar values of \(\text{app}(k_{\text{cat}}/K_{\text{ox}})\) obtained confirm that the enzyme M62A/F357A was fully saturated by the organic substrate and that \(\text{app}(k_{\text{cat}}/K_{\text{ox}})\) approaches well the true kinetic parameter \(k_{\text{cat}}/K_{\text{ox}}\). The use of D\(_2\)O to determine MKIE could yield not only an isotope effect on the kinetic parameter investigated, but also a viscosity effect due to the increase in viscosity of the solution by substituting H\(_2\)O with D\(_2\)O, and a pH effect due to perturbations of observed p\(K_a\) values\(^{24, 29, 30}\). The possibility of a pH effect upon substituting H\(_2\)O with D\(_2\)O have been ruled out by establishing that the kinetic parameter \(k_{\text{cat}}/K_{\text{ox}}\) is pH independent in H\(_2\)O (Figure 3.5) as well as in D\(_2\)O from pD 7.0 to 9.0 (with values of 18,400 ± 400, 17,900 ± 1,100, and 24,000 ± 1,000 M\(^{-1}\) s\(^{-1}\)).
Figure 3.5: pH dependence of the kinetic parameter $k_{\text{cat}}/K_{\text{ox}}$ determined at 25 °C for the double variant M62A/F357A (solid circles) and for wild-type choline oxidase$^{(1)}$ (empty circles).

The contribution of a viscosity effect to the isotope effect of D$_2$O was also ruled out by showing that no viscosity effect on $k_{\text{cat}}/K_{\text{ox}}$ is observed in the pH range from 7.0 to 10.0 (Figure 3.3). A substrate and solvent kinetic isotope effects $k_{\text{cat}}/K_{\text{ox}}$ were measured for the variant enzyme M62A/F357A. The substrate isotope effect on $k_{\text{cat}}/K_{\text{ox}}$ is generated by the fact that the N5 atom of the isoalloxazine ring of the FAD cofactor is deuterated upon reduction by deuterated choline (Scheme 3.1). The deuterium at the N5 position is then transferred from the reduced flavin to oxygen, yielding a substrate isotope effect on $k_{\text{cat}}/K_{\text{ox}}$. As previously discussed for wild-type choline oxidase, the determination of a substrate isotope effect on $k_{\text{cat}}/K_{\text{ox}}$ in aqueous buffer implies that the transfer of the deuterium to oxygen occurs prior to the potential wash-out of the deuterium atom from the flavin cofactor to the solvent.$^{(7)}$ As shown in Table 3.2 the measured isotope effects on the double variant M62A/F357A at pH 7.0 and 25 °C are not significantly different from the measured isotope effects of the wild-type enzyme.$^{(7)}$

Table 3.2. Multiple deuterium kinetic isotope effects on the kinetic parameter $k_{\text{cat}}/K_{\text{ox}}$ at pH 7.0 and 25 °C of wild-type choline oxidase$^{(7)}$ and of the double variant enzyme M62A/F357A prepared in this study.

<table>
<thead>
<tr>
<th>KIE</th>
<th>M62A/F357A</th>
<th>Wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^D(k_{\text{cat}}/K_{\text{ox}})_{\text{H}_2\text{O}}$</td>
<td>1.2 ± 0.1</td>
<td>1.4 ± 0.1</td>
</tr>
<tr>
<td>$^D(k_{\text{cat}}/K_{\text{ox}})_{\text{D}_2\text{O}}$</td>
<td>1.7 ± 0.2</td>
<td>1.7 ± 0.1</td>
</tr>
<tr>
<td>$^\text{D}<em>2\text{O}(k</em>{\text{cat}}/K_{\text{ox}})_{\text{H}}$</td>
<td>1.1 ± 0.1</td>
<td>1.2 ± 0.1</td>
</tr>
<tr>
<td>$^\text{D}<em>2\text{O}(k</em>{\text{cat}}/K_{\text{ox}})_{\text{D}}$</td>
<td>1.6 ± 0.2</td>
<td>1.5 ± 0.1</td>
</tr>
<tr>
<td>$^\text{D}<em>2^\text{D}<em>2\text{O}(k</em>{\text{cat}}/K</em>{\text{ox}})$</td>
<td>1.9 ± 0.3</td>
<td>2.1 ± 0.4</td>
</tr>
<tr>
<td>$^D(k_{\text{cat}}/K_{\text{ox}})<em>{\text{H}<em>2\text{O}} \times ^\text{D}<em>2\text{O}(k</em>{\text{cat}}/K</em>{\text{ox}})</em>{\text{H}}$</td>
<td>1.3 ± 0.1</td>
<td>1.7 ± 0.2</td>
</tr>
</tbody>
</table>
**Reductive half reaction.** The oxidation of choline and 1,2-[^2]H_h[ strained choline was performed anaerobically for the enzymes M62A, F357A, and M62A/F357A at pH 10.0 and 25 °C with the use of a stopped-flow spectrophotometer by monitoring the decrease in absorbance at 450 nm. In the case of the enzyme M62A/F357A the decrease in absorbance at 450 nm exhibited a biphasic behavior (Figure 3.6).

**Figure 3.6.** Anaerobic reduction of the variant enzyme M62A/F357A with choline or deuterated choline as substrate. Panel A shows the stopped-flow traces obtained with 6 mM choline (red line) or deuterated choline (black line). The traces were fit to Eq. 3 or 4. For clarity, one experimental point every 5 is shown (vertical line). Panel B shows the concentration dependence of the observed rate constants for flavin reduction with choline (solid circles) or deuterated choline (solid squares) as substrate.
Flavin reduction is represented by the phase with the highest amplitude, accounting for more than 90% of the total change, and for which a deuterium substrate kinetic isotope effect was observed. The concentration dependence of the observed rate constants of this phase was analyzed with Eq. 5 and $k_{\text{red}}$ value of $45.1 \pm 0.5$ s$^{-1}$ and $K_d$ of $1.64 \pm 0.06$ mM with choline as a substrate were determined (Table 3.3). The measured substrate kinetic isotope effect on $k_{\text{red}}$ is $6.4 \pm 0.4$.

Table 3.3. Pre-steady-state kinetic parameters of variant enzymes prepared in this study and choline oxidase wild-type$^{(1,28)}$ at pH 10.0 and 25 °C with choline as a substrate.

<table>
<thead>
<tr>
<th>Kinetic parameter</th>
<th>M62A/F357A</th>
<th>F357A</th>
<th>M62A</th>
<th>Wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{\text{red}}$ (s$^{-1}$)</td>
<td>$45.1 \pm 0.5$</td>
<td>$85.1 \pm 1.2$</td>
<td>$92.1 \pm 3.1$</td>
<td>$93 \pm 1$</td>
</tr>
<tr>
<td>$K_d$ (mM)</td>
<td>$1.64 \pm 0.06$</td>
<td>$2.6 \pm 1$</td>
<td>$&lt; 0.07$</td>
<td>$0.29 \pm 0.01$</td>
</tr>
</tbody>
</table>

A small first phase, within 2 ms and with amplitude less than 10% of the total change, was observed with 12 and 6 mM of organic substrate. This small phase might be due to substrate binding as already described for variants of Class 2 Dihydroorotate Dehydrogenase.$^{(31)}$ The spectral changes were too small for a reliable fit of rate constants. Sometimes a final small slow phase was observed which was not concentration dependent, with amplitude less than 10% of the total change, and average $k_{\text{obs}}$ of $0.1$ s$^{-1}$. Because this phase is much slower than $k_{\text{cat}}$ with an estimated value of $60$ s$^{-1}$ in the same experimental conditions, it is considered not kinetically relevant and possibly due to the presence of a fraction of damaged enzyme or consumption of a small amount of contaminating oxygen.

In the case of the enzyme F357A flavin reduction occurred with a monophasic behavior and the observed rate constants were fit to Eq. 5, yielding a $k_{\text{red}}$ value of $85.1 \pm 0.2$ s$^{-1}$ and a $K_d$ value of $2.6 \pm 0.1$ mM (Table 3.3). The reductive half reaction of the enzyme M62A showed a biphasic pattern, with the fast and slow phases accounting for approximately 70% and 30% of the total
change in absorbance at saturating choline concentration, respectively. The fast phase was assigned to flavin reduction, as it accounts for most of the absorbance change at 450 nm. The observed rate constants for the fast phase were fit to Eq. 5 yielding a $k_{\text{red}}$ value of $92.1 \pm 3.1 \text{ s}^{-1}$ and a $K_d$ value less than 0.07 mM (Table 3.3). In order to achieve pseudo-first order conditions it was not possible to lower choline concentration less than 0.07 mM and therefore to experimentally determine $K_d$. The second phase showed hyperbolic concentration dependence with a limiting value of $13 \pm 0.3 \text{ s}^{-1}$ and an equilibrium constant less than 0.07 mM. A similar slow phase accounting for 30-40% of the total change in absorbance at 450 nm was previously observed in the active site variant of choline oxidase V464A$^{(32)}$ and it was assigned to the rate limiting interconversion of an incompetent to the competent form of enzyme-substrate complex. The slow phase observed in the reductive half reaction of the single variant M62A is most likely due to the rate limiting interconversion of an incompetent to the competent form of enzyme-substrate complex. The fast and slow phases are seven times separated; for this reason it is not possible to rule out that the observed rate constants of the two phases are affecting each other to a certain degree. The fitting of the observed rate constants for this variant enzyme is therefore considered a good approximation of the $k_{\text{red}}$ value rather than an exact value and it is considered only to estimate the loss compared with wild-type choline oxidase.

**Apparent steady-state kinetics.** Apparent steady-state kinetics were performed at atmospheric oxygen, 25 °C and pH 7.0, in order to investigate the effect of the removal of the side chains M62 and F357 on the activity of choline oxidase in physiologically relevant conditions. Table 3.4 presents the apparent kinetic parameters for choline of the enzymes M62A, F357A, M62A/F357A in comparison with wild-type choline oxidase$^{(1)}$. In the case of the double variant
M62A/F357A there was a 3 times decrease in the \( \text{app} k_{\text{cat}} \) value, a 20 times increase in the \( \text{app} K_m \) value, and a 60 times decrease in the \( \text{app}(k_{\text{cat}}/K_m) \) compared to wild-type choline oxidase.

**Table 3.4:** Apparent steady-state kinetic parameters for choline at atmospheric oxygen, 25 °C, in 50 mM sodium phosphate pH 7.0.

<table>
<thead>
<tr>
<th></th>
<th>( \text{app} k_{\text{cat}} ) (s(^{-1}))</th>
<th>( \text{app} K_m ) (mM)</th>
<th>( \text{app}(k_{\text{cat}}/K_m) ) (M(^{-1}) s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>M62AF357A</td>
<td>6.8 ± 0.3</td>
<td>8.5 ± 1.0</td>
<td>800 ± 100</td>
</tr>
<tr>
<td>F357A</td>
<td>2.5 ± 0.1</td>
<td>11.9 ± 1.3</td>
<td>210 ± 20</td>
</tr>
<tr>
<td>M62A</td>
<td>12.5 ± 0.3</td>
<td>0.8 ± 0.1</td>
<td>15600 ± 1000</td>
</tr>
<tr>
<td>wild-type</td>
<td>20 ± 1</td>
<td>0.4 ± 0.1</td>
<td>50000 ± 13000</td>
</tr>
</tbody>
</table>

**Molecular dynamics.** A CAVER analysis of the crystal structure of choline oxidase (PDB 4MJW) highlighted two tunnels, A and B, leading to the isoalloxazine ring, with the side chain of F357 shaping the intersection between these two tunnels (Figure 3.1 panel C). Tunnel A coincides with the previously proposed choline access to the active site, delimited by the side chains of the hydrophobic cluster M62, L65, V355, F357, and M359, while the narrow tunnel B gated by F357 and E312 possibly represents a tunnel for oxygen access to the active site. In order to take into consideration the importance of the dimer interface to regulate substrate access to the active site in this study molecular dynamic simulation were performed on dimeric choline oxidase (PDB 4MJW) in an effort to accurately and efficiently study the dynamics of the residues in the hydrophobic cluster and a newly proposed gating mechanism, involving F357 and E312, presumed to control an entry site of oxygen into the active site. Figure 3.7 displays the probability distributions of the side chain torsional angles for M62, L65, V355, F357, and M359 belonging to the hydrophobic cluster of the dimeric form of ligand-free choline oxidase (PDB entry 4MJW).
Figure 3.7 Probability distributions of the side-chain torsional angles for M62, L65, V355, F357, and M359 belonging to the hydrophobic cluster of the dimeric form of choline oxidase (PDB entry 4MJW) without substrate bound. Depicted are the chi 1 angles for wild-type choline oxidase (red), F357A variant (blue), and M62A variant (green).

Depicted are the chi 1 angle distributions for wild-type choline oxidase (red), F357A variant (blue), and M62A variant (green). The sampled conformations for the chi 1 angles for the residues of the hydrophobic cluster in the M62A enzyme variant strongly resemble those sampled by the residues in wild-type choline oxidase (Figure 3.7). This suggests that the mutation of M62 to alanine does not significantly affect the gating mechanism of the hydrophobic cluster proposed for choline access and glycine betaine release (tunnel A in Figure 3.1 panel C). On the other hand, the chi 1 angle distributions sampled for the F357A variant hydrophobic cluster residues deviate significantly from those sampled by the wild-type enzyme, consistent with the side chain of F357 playing an important role in the gating mechanism. Furthermore, Figure 3.8 illustrates the proposed tunnel for oxygen access controlled by the residues F357 and E312. The movements of the phenyl ring of F357 determine an open and closed state of the narrow hydrophobic tunnel gated by F357 and E312. The open state is shown in panel A and D, while the closed state is highlighted in panel C and E.
Figure 3.8 Newly proposed gating mechanism of the dimeric form of choline oxidase, involving F357 and E312, presumed to control an entry site of oxygen into the active site. Panel A and C show the open and closed state, respectively. Panel B illustrates the proposed tunnel calculated by CAVER 3.0 highlighted in red (VDW). Panel D and E represent the open and closed state with residues surrounding the tunnel highlighted in transparent (VDW). Choline oxidase is represented in white cartoon, the hydrophobic cluster residues are highlighted in green (VDW), the F357/E312 gating mechanism is highlighted in yellow (Surf).

3.5 Discussion

The effect of the removal of the side chains M62 and F357 at the entrance to the active site of choline oxidase was investigated by creating a double variant enzyme M62A/F357A. The kinetic parameter $k_{\text{cat}}/K_{\text{ox}}$ reports on the kinetic steps where the two electron-reduced form of the flavin cofactor reacts with oxygen (Scheme 3.1). A mechanistic investigation on this kinetic parameter with the use of solvent viscosity, pH and multiple kinetic isotope effects is carried out for the double variant enzyme M62A/F357A of choline oxidase. A partially rate limiting isomerization in $k_{\text{cat}}/K_{\text{ox}}$ was previously detected in wild-type choline oxidase. This study investigated which
side chains are related to the slow isomerization occurring during flavin oxidation by combining for the first time site-directed mutagenesis and solvent viscosity studies to monitor the isomerization.\(^{(7)}\) Indeed other mechanistic characterizations of the oxygen reactivity of oxidases included solvent viscosity studies as a control for isotope effects and did not observe the inverse hyperbolic behavior detected in choline oxidase.\(^{(8, 9)}\) The slow isomerization during flavin oxidation could be related to a gating mechanism for oxygen. The detection of two tunnels leading to the isoalloxazine ring suggested the side chains of M62 and F357 at the entrance to the active site as possibly involved in controlling oxygen access to the reduced cofactor and in the slow isomerization detected in wild-type choline oxidase. This hypothesis was confirmed by the fact that the inverse effect on solvent viscosity on \(k_{\text{cat}}/K_{\text{ox}}\) was not detected in the double variant M62A/F357A (Figure 3.3).

The multiple kinetic isotope effects on \(k_{\text{cat}}/K_{\text{ox}}\) at pH 7.0 were not significantly different from wild-type choline oxidase\(^{(7)}\) (Table 3.2), showing that the mechanism of flavin oxidation has remained unchanged (Scheme 3.1), with the two transfers of a proton and a hydrogen to oxygen to yield \(\text{H}_2\text{O}_2\) and oxidized FAD occurring in the same step.\(^{(7)}\) The \(k_{\text{cat}}/K_{\text{ox}}\) value showed no pH dependence (Figure 3.5) as previously observed for the wild-type enzyme.\(^{(1)}\) This is due to the fact that in choline oxidase the positive charge for oxygen activation is held by the non-ionizable positively charged trimethylammonium group of the product of choline oxidation.\(^{(2, 33)}\) The catalytic efficiency \(k_{\text{cat}}/K_{\text{ox}}\) for the reaction of the reduced cofactor with oxygen is slightly impaired, as shown by the decrease of almost five times of this kinetic parameter compared to wild-type choline oxidase\(^{(1)}\) (Table 3.1). A higher decrease in \(k_{\text{cat}}/K_{\text{ox}}\) was not expected as the active site environment involved in oxygen reactivity, such as V464\(^{(4)}\), is not altered in the variant enzymes generated in this study. It was not possible to achieve saturation with oxygen in the case
of the double variant M62A/F357A due to an increased Michaelis constant for oxygen $K_{ox}$ larger than 1 mM. A similar behavior of increase of the Michaelis constant for oxygen was previously described in aryl alcohol oxidase$^{(17)}$, cholesterol oxidase$^{(15, 16)}$ and 12/15 lipooxygenase$^{(13)}$ in the case of mutagenesis of residues at the bottleneck of tunnels for oxygen. Indeed there was a 8 times increase in $K_{ox}$ in the variant L367F of 12/15 lipooxygenase$^{(13)}$, and in the variant F501A of aryl-alcohol oxidase the $K_{ox}$ was estimated to be 3.6 mM$^{(17)}$. In the case of cholesterol oxidase type I the $K_{ox}$ value of the variant F359W increased of 2 times compared to the wild-type enzyme, and lack of saturation with oxygen concentrations as high as 1.2 mM was observed for the variants N485D and G347N$^{(16)}$. A similar case was observed with cholesterol oxidase type II, with lack of saturation with oxygen reported for the variants E311D and E311Q$^{(15)}$. Solvent viscosity studies on the kinetic parameter $k_{cat}/K_{ox}$ were not performed on these investigations of oxygen reactivity.

The binding and oxidation of the organic substrate choline is marginally affected by the removal of M62 and F357, as shown by the slight increase of six times of the dissociation constant for choline binding $K_d$ and the negligible decrease of two times of the rate for flavin reduction $k_{red}$ compared to the values determined for the wild-type enzyme$^{(1)}$ (Table 3.3). Similarly to wild-type choline oxidase$^{(1)}$, the reductive half reaction in the double variant enzyme M62A/F357A represents the main rate limiting step as supported by the large substrate isotope effect of 6.4 measured for the reductive half reaction at 25 °C and pH 10.0, which represents the pH independent region for kinetic steps in choline oxidase.$^{(1)}$

In order to establish the relative contribution of the side chains M62 and F357 the Mildvan’s approach$^{(20, 34)}$ was applied to the double variant enzyme M63A/F357A. The inverse thinking on double variant enzymes is based on the approach of comparing the kinetic parameters of the
double variant enzyme to the wild-type one, followed by the analysis of how much of the wild-type activity is restored by putting back a side chain per time in two single variant enzymes.\(^{(20)}\) For this approach the single variant enzymes M62A and F357A were created and the steady-state kinetic parameters were determined at pH 10.0 and 25 °C (Table 3.1). The slight decrease in the kinetic parameter \(k_{\text{cat}}/K_{\text{ox}}\) observed in the double variant M62A/F357A was completely restored by putting back the side chain of F357 in the single variant enzyme M62A as shown by the measured value of \(k_{\text{cat}}/K_{\text{ox}}\) of 89000 M\(^{-1}\) s\(^{-1}\) which is very similar to the value of 86400 M\(^{-1}\) s\(^{-1}\) of wild-type choline oxidase (Table 3.1). Furthermore the measured \(k_{\text{cat}}/K_{\text{ox}}\) of the single variant F357A showed the same decrease observed for the double variant M62A/F357A (Table 3.1), confirming the role of F357 as slightly affecting the catalytic efficiency of the reaction of the reduced cofactor with oxygen. The effect on the Michaelis constant \(K_{\text{ox}}\), which in the enzyme M62A/F357A increased above 1 mM, showed an opposite direction in the two single variant enzymes created in this study. Indeed the single variant M62A showed a three times smaller \(K_{\text{ox}}\) (Table 3.1) compared to wild-type choline oxidase, while the single variant F357A exhibited a lack of saturation up to 1 mM oxygen, consistent with a much larger \(K_{\text{ox}}\). These observations are consistent with the side chain M62 providing basically no additional effect as a limiting case of partial antagonism, as inferred from the opposing effects on \(K_{\text{ox}}\). The side chain of F357 is the only one involved in the partially rate limiting isomerization in \(k_{\text{cat}}/K_{\text{ox}}\). This conclusion is supported by the fact that by putting back the side chain of F357 in the single variant M62A the inverse solvent viscosity effect on \(k_{\text{cat}}/K_{\text{ox}}\) observed in wild-type choline oxidase is completely restored (Figure 3.4). Remarkably the single variant enzyme F357A exhibited the same behavior of the double variant enzyme M62A/F357A with no solvent viscosity effect on \(k_{\text{cat}}/K_{\text{ox}}\) (Figure 3.4), further confirming the role of F357 in the slow isomerization in \(k_{\text{cat}}/K_{\text{ox}}\).
The rate of choline oxidation was unaffected in the single variant enzymes M62A and F357A, with an antagonistic effect on the binding for choline shown by a ten times increase in $K_d$ for F357A and a decrease of at least three times in $K_d$ for M62A compared to the wild-type enzyme (Table 3.3). A ten times change in $K_d$ corresponds to an energy loss of 1.5 kcal mol$^{-1}$, which does not affect significantly substrate binding. The two times decrease in $k_{\text{red}}$ observed in the double variant M62A/F357A compared to the single variants M62A and F357A is possibly due to the removal of both side chains causing an increase in solvent accessibility to the active site that may perturb the hydride transfer reaction.

The kinetic characterization of variants of choline oxidase presented in this study established the side chain of F357 to be involved in the slow isomerization of the reduced enzyme-intermediate complex in the oxidation of FAD by oxygen. The mechanistic investigation of the enzymes M62A/F357A and F357A showed that the disappearance of this isomerization does not significantly affect catalysis, as the reductive half reaction in these variants enzymes is essentially unaffected and the catalytic efficiency of the oxidative half reaction is only slightly impaired. This observation is consistent with the slow isomerization in the oxidative half reaction being relevant to oxygen access to the active site. A CAVER analysis and molecular dynamics simulations on choline oxidase were carried out to investigate possible gating mechanism for oxygen.

Molecular dynamics simulations on dimeric wild-type choline oxidase and on the in silico variants M62A and F357A showed that only the removal of F357 side chain leads to perturbations of the dynamic motions of the side chains of the hydrophobic cluster of choline oxidase, which is proposed to be the gating mechanism for tunnel A for choline access and glycine betaine release (Figure 3.7). This is consistent with the experimental data of the kinetic
characterization of the variant enzymes M62A, F357A, and M62A/F357A, that concluded that the mutation of F357 was playing a more preponderant role compared to M62. However, the perturbations of the dynamic motions of the proposed gating mechanism for tunnel A only slightly affected the binding of choline (increase in $K_d$ of ten times in the F357A variant). The characterization of the recombinant enzyme F357A and the double variant M62A/F357A highlighted instead a perturbation in the kinetic steps related to the reaction of the reduced cofactor with oxygen, with the disappearance of a partially rate-limiting isomerization detected in wild-type choline oxidase and a linear behavior up to 1 mM oxygen. As discussed above the increased in $K_{ox}$ was previously associated in aryl-alcohol oxidase, cholesterol oxidase type I and II, and 12/15 lipoxygenase with site directed mutagenesis of residues at the bottleneck of tunnels proposed to regulate oxygen access to the active site $^{(13, 15-17)}$. The effects on oxygen reactivity detected in this study can be explained by the fact that F357, in pair with E312, is also participating to a second gating mechanism for tunnel B (Figure 3.1 panel C), a narrow tunnel we propose can be related to oxygen access to the active site. As shown in Figure 3.8, the side chain of F357 moves between two states during the simulation on wild-type choline oxidase, and these movements appear to regulate the bottleneck of tunnel B to the active site pocket. It is also possible that the movements of the phenyl ring of F357 play a role in directing the diffusion of oxygen in the direction of the C4a atom of the flavin cofactor.

The removal of the side chain of F357, but not of M62, resulted in a slight decrease in $k_{cat}/K_{ox}$ but caused a lack of saturation with oxygen at concentrations up to 1 mM. This feature is important for the efficiency of choline oxidase at atmospheric oxygen, which represents a physiologically relevant condition. Indeed Table 3.4 shows that the apparent catalytic efficiency for choline in
physiological conditions decreased only three times with the removal of the side chain of M62 but it decreased more than 200 times with the removal of the side chain of F357.

3.6 Conclusion

A previous mechanistic study of the oxidative half reaction of wild-type choline oxidase established the presence of a partially rate limiting isomerization in $k_{\text{cat}}/K_{\text{ox}}$. This study addressed the question of which residues are involved in this isomerization by a mechanistic investigation of variant versions of choline oxidase and by molecular dynamics studies. The side chains of F357 and M62 are located at the entrance to the active site of choline oxidase facing each other and the position of F357 at the intersection between two tunnels computed by CAVER suggests their possible involvement in controlling this isomerization. Following the inverse thinking on double variant enzymes by Mildvan(20, 34) the enzyme M62A/F357A was generated and kinetically characterized, then the individual side chains were restored in the single variant enzymes M62A and F357A. The side chain of F357 was shown to be the one related to the partially rate limiting isomerization in the reaction of the reduced flavin cofactor with oxygen by solvent viscosity studies on $k_{\text{cat}}/K_{\text{ox}}$. Computational studies showed the detection of a narrow tunnel (tunnel B) gated by F357 and E312, and suggested that one possible explanation of the partially rate limiting isomerization in $k_{\text{cat}}/K_{\text{ox}}$ is the movement of the aromatic ring of F357 to control access of oxygen to the reduced cofactor. The characterization of the variant enzyme M62A highlighted no major effect on the kinetic parameters, while the removal of the side chain of F357 resulted in the disappearance of the slow isomerization in the oxidation of the reduced FAD and in the lack of saturation with oxygen. Furthermore the physiological impact of the removal of side chain F357 associated with a mechanistically small decrease in efficiency in the reaction with oxygen was investigated by apparent steady-state kinetics at
atmospheric oxygen and pH 7.0 and highlighted a decrease of more than 200 times in the apparent catalytic efficiency for choline. This study complements the mechanistic study of oxygen reactivity performed on wild-type choline oxidase and it represents the first instance of using solvent viscosity effects to monitor an isomerization in the mechanistic characterization of oxygen reactivity in flavin dependent oxidases.
3.7 References


CHAPTER IV: HUMAN CHOLINE DEHYDROGENASE: MEDICAL PROMISES
AND BIOCHEMICAL CHALLENGES

(This chapter has been published verbatim in Salvi, F., and Gadda, G., (2013), Arch. Biochem. Biophys. 537(2):243-52)

4.1 Abbreviations
CHD, choline dehydrogenase; CHO, choline oxidase; CMO, choline monooxygenase; BADH, betaine aldehyde dehydrogenase; PMS, phenazine methosulfate; GMC, Glucose-Methanol-Choline; PQQ, pyrroloquinoline quinone.

4.2 Abstract
Human choline dehydrogenase (CHD) is located in the inner membrane of mitochondria primarily in liver and kidney and catalyzes the oxidation of choline to glycine betaine. Its physiological role is to regulate the concentrations of choline and glycine betaine in the blood and cells. Choline is important for regulation of gene expression, the biosynthesis of lipoproteins and membrane phospholipids and for the biosynthesis of the neurotransmitter acetylcholine; glycine betaine plays important roles as a primary intracellular osmoprotectant and as methyl donor for the biosynthesis of methionine from homocysteine, a required step for the synthesis of the ubiquitous methyl donor S-adenosyl methionine. Recently, CHD has generated considerable medical attention due to its association with various human pathologies, including male infertility, homocysteinuria, breast cancer and metabolic syndrome. Despite the renewed interest, the biochemical characterization of the enzyme has lagged behind due to difficulties in the obtainment of purified, active and stable enzyme. This review article summarizes the medical relevance and the physiological roles of human CHD, highlights the biochemical knowledge on the enzyme, and provides an analysis based on the comparison of the protein sequence with that of bacterial choline oxidase, for which structural and biochemical information is available.
4.3 Introduction

Human choline dehydrogenase (CHD; E.C. 1.1.99.1) is a nuclear-encoded, mitochondrial enzyme involved in choline metabolism. According to the NCBI gene database, the gene coding for CHD has been identified on chromosome 3 (gene location 3p21.1). From a study of Haubrich et al. on human tissues, the enzymatic activity of CHD has been detected mainly in kidney, with 6-times lower levels in liver. Other tissues like blood, spleen and heart displayed very low CHD activity, whereas no detectable activity was reported for muscle and fat tissue. CHD is also present in bacteria, fungi and other mammals (Figure 4.1), but it is notably absent in plants where an iron-sulfur choline monooxygenase (CMO; E.C. 1.14.15.7) has been identified instead.

Figure 4.1: Phylogenetic tree.
CHD from Homo sapiens (NP_060867.2), Brucella suis (YP_005154241.1), Vibrio vulnificus (ADV86038.1), Ochrobactrum anthropic (ABS15421.1), Bos Taurus (NP_001192493.1), Rattus norvegicus ( NP_942026.1), Mus musculus (NP_001129712.1), Gallus gallus (XP_414335.3), Escherichia coli (BAE76094.1), Pseudomonas aeruginosa (NP_254059.1), Pan troglodytes (XP_001173164.1), Macaca mulatta (NP_001244738.1), Bacillus selenitireducens (YP_003698280.1), Serratia plymuthica (YP_004504883.1), Halomonas elongata (YP_003896928.18.1), Aspergillus oryzae (EIT82252.1), Felix catus (XP_003982381.1), Cavia porcellus (XP_003480015.1), Sinorhizobium meliloti (YP_004548030.1), Mesorhizobium opportunistum (YP_004613943.1),
*Burkholderia graminis* (ZP_02886806.1), *Granulicella mallensis* (YP_005059092.1), *Danio rerio* (XP_002663301.1).

CHD from eukaryotic organisms is associated with the inner mitochondrial membrane on the matrix side \(^{(4-6)}\). The homologous protein from prokaryotic sources is associated with the cytosolic side of the cell membrane \(^{(7)}\).

From a medical point of view, human CHD is of interest due to its association with various pathologies, including male infertility \(^{(8)}\), homocysteinuria \(^{(9)}\), and cancer \(^{(10,11)}\). The enzyme also readily metabolizes choline when the latter is administered as a pharmacological agent, thereby limiting its potential therapeutic use \(^{(2)}\). Great attention has also been directed to bacterial CHD for medical, biotechnological and fundamental reasons. The medical interest is primarily due to the fact that the prokaryotic enzyme plays an important role for the ability of bacteria to grow in environments with high salinity, such as human infection sites, and thus represents a potential pharmaceutical target for combination therapy \(^{(12)}\). Biotechnological applications primarily focus on the genetic engineering of glycine betaine biosynthesis from choline to provide osmotic stress resistance in economically relevant plants and the detection of choline and its derivatives in biological fluids \(^{(13,14)}\). Biochemical interest stems mainly from the comparison of the catalytic strategies utilized in the oxidation of choline by the three types of enzymes that carry out the reaction: CHD, CMO, and choline oxidase (CHO; E.C. 1.1.3.17) (Scheme 4.1) \(^{(15)}\). In organisms that oxidize choline using CHD or CMO, a second enzyme, i.e., betaine aldehyde dehydrogenase (BADH; E.C. 1.2.1.8), is responsible for further oxidation of betaine aldehyde to glycine betaine \(^{(16)}\). In contrast, CHO catalyzes both steps of oxidation of choline to betaine aldehyde and betaine aldehyde to glycine betaine \(^{(17)}\).
Scheme 4.1: The three biosynthetic pathways for glycine betaine from choline.

Despite considerable interest in both human and prokaryotic CHD, the biochemical characterization of the enzyme has significantly lagged behind its medical and biotechnological applications due to high instability of the enzyme once it is removed from the inner mitochondrial membrane. This has prevented the obtainment of active and stable purified enzyme for structural-functional studies.

In this review article, we present the physiological roles, medical relevance, and cellular localization of human CHD; we provide a summary of the attempts and difficulties of purifying CHD, summarize the available biochemical knowledge on the enzyme, and introduce a comparison of the CHD protein sequence with that of bacterial choline oxidase, for which structural and biochemical information is available. A model of the 3D structure of CHD built on the known 3D structure of CHO previously obtained by X-ray crystallography is also presented.
4.4 Physiological roles of human CHD

Human CHD catalyzes the oxidation of choline to betaine aldehyde, which is further oxidized to glycine betaine by a second enzyme, BADH (16). This enzymatic reaction achieves the dual physiological role of regulating the concentration of free choline in cells and bodily fluids and synthesizing a metabolite that is relevant to both osmoprotective and methylating processes (Figure 2.2).

![Diagram](image)

**Figure 4.2:** Physiological roles of human CHD in catabolism of choline and synthesis of glycine betaine. CHD: choline dehydrogenase. BA: betaine aldehyde, BADH: betaine aldehyde dehydrogenase.

Since 1983 (18), the osmoprotectant role of glycine betaine has been studied in regard to the regulation of cell volume and the increase in stability of intracellular macromolecules (19). In humans and mammals glycine betaine plays an important osmoprotectant role in the kidney, where it was shown to increase in concentration during dehydration (20), and in the liver, where it regulates hepatocellular hydration (21). Fatty liver is often observed as a consequence of a lower intake of choline with the diet due most likely to a decrease in the concentration of phosphatydilcholine, which is necessary for the synthesis of the very-low-density lipoprotein.
(VLDL)\(^{22}\). In rodents, it has been established that fatty liver is developed when they are fed on a low choline diet \(^{23}\). It is interesting to note that humans, animals, plants and many microorganisms share the use of glycine betaine as the primary tool for protection from various forms of stress besides osmotic pressure due to swift variations in the concentrations of osmolites in cellular environments, including high and low temperatures, reactive oxygen species and other forms of cellular insults \(^{24}\). In humans and mammals, glycine betaine plays also an important role as a methyl group donor involved, for example, in the methylation of homocysteine to methionine (Scheme 4.2) \(^{19}\). The concentration of glycine betaine is thus considered important for metabolic syndrome, lipid disorders, diabetes, vascular diseases and the development of the embryo \(^{25}\).

**Scheme 4.2:** Choline metabolism.

EC 1.1.99.1, CHD; EC 1.2.1.8, BADH; EC 1.5.8.4, dimethylglycine dehydrogenase; EC 2.1.1.5, betaine-homocysteine methyltransferase; EC 2.1.1.13, methy1tetrahydrofolate-homocysteine methyltransferase; EC 2.1.1.17, phosphatidylethanolamine N-methyltransferase; EC 2.1.1.37, DNA (cytosine-5-)-methyltransferase; EC 2.3.1.6, choline acetyltransferase; EC 2.5.1.6, S-adenosylmethionine synthetase; EC 2.7.1.32, choline kinase; EC 2.7.7.15,
choline-phosphate cytidylyltransferase; EC 2.7.8.2, diacylglycerol choline phosphotransferase; EC 3.1.1.4, phospholipase A2; EC 3.3.1.1, adenosylhomocysteinase; DMG, dimethylglycine, Me-THF, methyl tetrahydrofolate, SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine, THF, tetrahydrofolate.

The regulation of the concentration of choline in tissues and blood is very important as choline plays key roles in different pathways. Choline is involved in the epigenetic regulation of gene expression through DNA methylation (26), as shown for example in the global hypomethylation of hepatic DNA of rats fed with a low choline diet (27), in the biosynthesis of lipoproteins and membrane phospholipids and in the biosynthesis of the neurotransmitter acetylcholine (Scheme 4.2) (28). It is therefore important for the integrity of cell membranes, lipid metabolism and nerve function. Choline is considered an important nutrient for fetal and brain development (29-31) as shown for example by the different rate of development of the hippocampus in the fetal brain of rodent models in the case of low and high maternal choline intake (32). Choline is a constituent of phospholipids involved in signal transduction, such as phosphatidylcholine and plasmalogen, and of the phospholipid platelet activating factor (33). The metabolism of choline is also interrelated with the metabolism of folate and it has been shown that the folate content in the liver of choline deficient rats decreased by 31% compared to control rats (Scheme 4.2) (34). To the best of our knowledge the utilization of choline by other xenobiotic oxygenases present in the liver, such as for example cytochrome P450 or flavin monooxygenase, has not been described in the literature. It would be interesting to evaluate whether these enzymes can also oxidize choline.

4.5 Medical relevance of CHD

In the past 5 years, attention on human CHD and its involvement in various pathologies has grown considerably (Table 4.1).
Table 4.1: Medical relevance of human CHD.

<table>
<thead>
<tr>
<th>Medical relevance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male infertility</td>
<td>(35) (8) (36)</td>
</tr>
<tr>
<td>Homocysteinuria</td>
<td>(9) (19)</td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td>(28)</td>
</tr>
<tr>
<td>Choline as pharmacological agent</td>
<td>(2) (39)</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>(10) (11)</td>
</tr>
</tbody>
</table>

The enzyme has been associated with male infertility in multiple independent studies (8, 35, 36). In 2010, Johnson et al. established a correlation between the activity of CHD and male fertility by deleting the corresponding gene in mice (35). The absence of CHD activity resulted in diminished sperm motility that greatly affected the reproductive ability of mice, with only one out of eleven CHD(-/-) mice being able to reproduce. Mitochondrial alterations were described in testis as well as liver, kidney and heart (35). Polymorphisms in the human gene have recently been identified and associated with decreased activity of human CHD and alterations in human sperm (e.g., rs12676, Leu-78 →Arg) (8, 36).

Impairments in human CHD activity have been associated with homocysteinuria, an accumulation of homocysteine that represents an independent risk factor for cardiovascular diseases (9, 19). This observation is consistent with glycine betaine being the main methyl donor in the conversion of homocysteine to methionine catalyzed by the enzyme betaine-homocysteine methyltransferase (Scheme 4.2) (37). The biosynthesis of methionine is important for the metabolism of methyl groups, because it is necessary for the synthesis of the widely used methyl donor S-adenosylmethionine (SAM) (Scheme 4.2) (38).
A population-based study in 2008 monitored the blood level of choline and its metabolite glycine betaine in relation to components of metabolic syndrome, such as percent body fat, blood pressure, serum lipids, etc. \(^{(28)}\). A correlation was found between high concentrations of choline, low concentrations of glycine betaine in blood and a high cardiovascular disease risk profile. It was shown that the amount of choline and glycine betaine supplied in the diet did not have a significant effect on the blood concentrations of these compounds. Therefore, the authors proposed that altered concentrations of choline and glycine betaine in blood should arise from a malfunction of the mitochondrial biosynthesis of glycine betaine rather than dietary patterns.

Human CHD is important for the catabolic utilization of choline when the latter is administered as a pharmacological agent, because choline is involved in the stimulation of cholinergic neuronal activity (Scheme 4.2) \(^{(2)}\) and in restoring phosphatidylcholine levels in the neuron membrane, thus displaying a neuroprotective action relevant for diseases such as memory and cognitive deficits \(^{(39)}\). Some studies have suggested a mechanism of neuroprotection associated with the supply of exogenous choline to be based on the fact that in case of choline deficiency the brain may degrade the membrane phospholipids of the neurons in order to recycle choline for the production of acetylcholine (Scheme 4.2) \(^{(39)}\). CHD, predominantly active in the two main detoxifying organs liver and kidney, determines the half-life of choline in blood. In a study in which \(^{[3]}\)H-methyl-choline was administered intravenously to guinea pig the main metabolite of choline detected in blood was shown to be glycine betaine, with kidney and liver removing about 50% of the administered dose within 3 minutes following injection \(^{(2, 40)}\). The rapid turnover of choline when administered as a drug is clearly not desirable, since it limits the therapeutic action and requires the administration of higher doses. More recently, the administration of choline as pharmacological agent has evolved into the use of cytidine 5′-diphosphocholine as choline donor.
This compound is readily hydrolyzed in the intestine to yield choline and cytidine, which are rapidly absorbed yielding increased plasmatic levels of the compounds. Both compounds can cross the blood-brain barrier and be utilized for the re-synthesis of cytidine 5'-diphosphocholine (Scheme 4.2) in the brain. Cytidine 5'-diphosphocholine showed promising results in clinical trials for stroke therapy, memory impairment in the elderly, recovery from brain injury, Alzheimer’s disease, glaucoma and gave improvements in bradykinesia and muscular rigidity \(^{39}\). Cytidine 5'-diphosphocholine is sold as dietary supplement in the United States and as a prescription drug both in Japan and Europe \(^{39}\).

A recent population-based study showed that the metabolic oxidation of choline is related to the risk of developing breast cancer \(^{10}\). It was concluded that the dietary intake of higher doses of choline in women is related to a lower risk of developing breast cancer. In that study, increased risk factor for breast cancer was also associated with the polymorphism +432G>T (rs12676) in the human gene coding for CHD \(^{10}\). Curiously, this is the same polymorphism that has been linked to altered sperm mobility patterns and altered mitochondrial morphology in human sperm associated with sterility \(^{8,36}\). A study of prognostic biomarkers for breast cancer identified the expression of CHD among three human genes controlled by estrogens, and showed that this is a strong predictor of the outcome of treatment with tamoxifen in early-stage (ER)-positive breast cancer patients (CHD originally reported as Genbank accession number AI240933) \(^{11}\).

### 4.6 Subcellular localization of mammalian CHD

The subcellular localization of mammalian CHD has been studied in rat. Experiments of gradient centrifugation have shown the association of the enzyme with the inner membrane of mitochondria \(^{4,41}\). The N-terminal sequencing of CHD extracted from rat liver mitochondria demonstrated that mature CHD begins with amino acid 35. In that study it was suggested that
CHD contains an N-terminal cleavable mitochondrial targeting presequence of 34 amino acids and it was hypothesized that two cleavage sites may be present for recognition and processing by Mitochondrial Processing Protease and Inner Membrane Protease \(^{(41)}\). These data are in agreement with the notion that mitochondrial proteins encoded by nuclear genes are synthesized on cytosolic ribosomes and include in their primary structure specific leader sequences for import to the mitochondria and localization in either the outer or inner membranes, intermembrane space or matrix \(^{(42)}\). Based on these studies, it is presumed that human CHD is similarly localized on the inner mitochondrial membrane.

### 4.7 Glucose-Methanol-Choline enzyme oxidoreductase superfamily

CHD is one of the enzymes that were originally grouped in the Glucose-Methanol-Choline (GMC) enzyme oxidoreductase superfamily based on primary structure alignment \(^{(43)}\). This superfamily of enzymes was established in 1992 when the analysis of the protein sequences of *Drosophila melanogaster* glucose dehydrogenase, *Escherichia coli* CHD, *Aspergillus niger* glucose oxidase and *Hansenula polymorpha* methanol oxidase highlighted the fact that these proteins are homologs \(^{(44)}\). Members of this family are flavoenzymes and catalyze the oxidation of a variety of alcohols with different chemical structures. Over the years several enzymes have been added to the superfamily, and the crystal structures of eight members are currently available, including *A. globiformis* CHO \(^{(45)}\), *A. niger* glucose oxidase \(^{(46)}\), *Brevibacterium sterolicum* cholesterol oxidase \(^{(47)}\), *Phanerochaete chrysosporium* cellobiose dehydrogenase \(^{(48)}\), *Trametes ochracea* pyranose-2-oxidase \(^{(49)}\), *Pleurotus eryngii* aryl-alcohol oxidase \(^{(50)}\), *Aspergillus oryzae* formate oxidase \(^{(51)}\), and *Mesorhizobium loti* pyridoxine 4-oxidase \(^{(52)}\).

Despite sharing limited sequence similarity, these enzymes all share similar overall 3D structures and highly conserved catalytic sites (for a recent review see \(^{(53)}\)) (Figure 4.3).
Figure 4.3: Overall structures and active sites of members of the GMC oxidoreductase enzyme superfamily. *Arthrobacter globiformis* CHO (S101A) PDB 3NNE (A), *Brevibacterium sterolicum* cholesterol oxidase PDB 1COY (B), *Phanerochaete chrysosporium* cellobiose dehydrogenase PDB 1KDG (C), *Plerotus eryngii* aryl-alcohol oxidase PDB 3FIM (D), *Aspergillus niger* glucose oxidase PDB 1CF3 (E), *Trametes ochracea* pyranose 2-oxidase PDB 2IGK (F), *Aspergillus oryzae* formate oxidase PDB 3Q9T (G) and *Mesorbium loti* pyridoxine 4-oxidase PDB 3T37 (H).

Not surprisingly, major differences are present in the substrate domains of these enzymes, primarily because the alcohol substrates are structurally unrelated \(^{(54)}\). Given the difficulty in the obtainment of purified, active and stable enzyme, structural information derived from X-ray
crystallography is not available for CHD. However, based on the similar active site architectures and the available biochemical data on other members of the superfamily that have been characterized in depth, including CHO, it is reasonable to expect that CHD utilizes a catalytic mechanism for the oxidation of choline similar to that of CHO, in which alcohol oxidation occurs by hydride transfer \(^{(17)}\). As illustrated in a later section, several amino acids whose role has been elucidated in CHO by site-directed mutagenesis are conserved in human CHD, allowing proposing similar roles for the residues in the active site of the membrane-bound dehydrogenase.

### 4.8 Purification of CHD

The characterization of CHD from a variety of cellular sources has so far been limited by the fact that it is difficult to obtain stable, active and highly purified enzyme once it is removed from its cellular location. The association of the enzyme with the membrane represents a further challenge for the purification and in vitro stability of the enzyme. We are aware of only one report on the partial purification of recombinant human CHD expressed in E. coli, described in a Ph.D. Thesis \(^{(55)}\). In that study, the final enzyme yield was low, with 2 mg of purified CHD out of 24 liters of bacterial cell culture, and the limited stability of the purified enzyme prevented further characterization of human CHD other than partial characterization of the cofactor and generation of antibodies \(^{(55)}\).

Most of the biochemical and kinetic properties of CHD have been determined in experiments on crude mitochondrial fractions and not purified enzyme, primarily from rat liver. Other studies have been carried out on oysters’ mitochondrial fractions without further purification of the enzyme \(^{(56)}\). The first attempt at obtaining purified enzyme was conducted by Williams et al. in 1953, who succeeded to isolate CHD from rat liver upon extraction with choleate \(^{(57)}\). It was later recognized, however, that this protocol did not yield a soluble form of the enzyme, but rather a
dispersion of insoluble particles\textsuperscript{(58)}. In 1959, Rendina and Singer developed a new extraction method for CHD from rat liver that was based on the use of \textit{Naja naja} venom as source of phospholipase A to solubilize CHD from the mitochondrial membrane, but without further purification of the protein\textsuperscript{(59)}. In 1980, a partial purification from rat liver mitochondria was reported by Tsuge \textit{et al.} upon extraction with the detergent Triton X-100 followed by DEAE-Sepharose and Choline-Sepharose affinity column chromatography\textsuperscript{(60)}. However, the preparation of the enzyme analyzed by SDS-PAGE showed three bands and demonstrated poor \textit{in vitro} stability\textsuperscript{(60)}. In 1985, Ameyama \textit{et al.} partially purified the enzyme from dog liver with less than 1\% yield\textsuperscript{(61)}. Despite multiple efforts to purify CHD from various eukaryotic sources the problems associated with the scarce solubility and \textit{in vitro} stability of the mammalian enzyme are still a major challenge (Table 4.2).

\begin{table}
\centering
\caption{Purification attempts of CHD from various cellular sources.}
\begin{tabular}{lcc}
\hline
\textbf{Source} & \textbf{Year} & \textbf{Reference} \\
\hline
Rat liver & 1980 & \textsuperscript{(60)} \\
Dog liver & 1985 & \textsuperscript{(61)} \\
\textit{P. aeruginosa} & 1994 & \textsuperscript{(62)} \\
Recombinant \textit{H. elongata} & 2003 & \textsuperscript{(63)} \\
Recombinant \textit{E. coli} & 2010 & \textsuperscript{(64)} \\
\hline
\end{tabular}
\end{table}

The purification of CHD from prokaryotic sources has revealed challenges similar to those encountered with the mammalian enzyme. A protocol for the partial purification of CHD from \textit{P. aeruginosa} has been reported, but the authors stated limited stability of the enzyme and the co-purification of the second enzyme of the biosynthetic pathway, BADH, as byproduct of this protocol\textsuperscript{(62)}. An attempt to obtain a pure preparation of CHD was performed by our group in
In that study, recombinant protein from *Halomonas elongata* was expressed in *E. coli* and partially purified to ~70% homogeneity by treatment with ammonium sulfate followed by DEAE-Sepharose column chromatography. A subsequent partial purification of recombinant CHD from *E. coli* was carried out in 2010 by Rajan *et al.* by applying the purification protocol developed for CHD from *H. elongata*.

### 4.9 Biochemical properties of CHD

In 1980, Tsuge *et al.* noted in their study on the purification of rat liver CHD that “all of the characteristics reported thus far have been obtained using a relatively crude preparation and a highly purified preparation is an urgent priority in the field” (60). As illustrated in the previous section, this remains true in 2013 and little is known on the biochemical properties of CHD, with no reports on the human enzyme other than genetic studies with only an isolated determination of the enzymatic activity from tissue homogenates (2). Moreover, the limited biochemical knowledge on CHD has been acquired on either mitochondrial fractions enriched with the enzyme or partially purified and unstable preparations of enzyme.

The cofactor content of CHD has been investigated, but unequivocal evidence on its identity is not available. Based on the presence of a glycine box GXGXXG at the N-terminus of the protein, it has been proposed that FAD is the cofactor of CHD. This would agree well with what is known on other GMC members, each containing FAD as cofactor (Table 4.3).
Table 4.3: GMC enzyme superfamily cofactor content; aPEG, polyethylene glycol; b n.d., not determined.

<table>
<thead>
<tr>
<th>GMC member</th>
<th>Genbank code</th>
<th>Glycine box</th>
<th>Cofactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choline oxidase</td>
<td>AAP68832.1</td>
<td>GGGSAG</td>
<td>FAD</td>
</tr>
<tr>
<td>Glucose oxidase</td>
<td>1CF3_A</td>
<td>GGGLTG</td>
<td>FAD</td>
</tr>
<tr>
<td>Cholesterol oxidase</td>
<td>1COY_A</td>
<td>GSGYGG</td>
<td>FAD</td>
</tr>
<tr>
<td>Cellobiose dehydrogenase</td>
<td>1KDG_A</td>
<td>GAGPGG</td>
<td>FAD</td>
</tr>
<tr>
<td>Aryl-alcohol oxidase</td>
<td>3FIM_B</td>
<td>GGGNAG</td>
<td>FAD</td>
</tr>
<tr>
<td>Pyranose 2-oxidase</td>
<td>2IGK</td>
<td>GSGPIG</td>
<td>FAD</td>
</tr>
<tr>
<td>Formate oxidase</td>
<td>3Q9T_A</td>
<td>GGGTAG</td>
<td>FAD</td>
</tr>
<tr>
<td>Pyridoxine 4-oxidase</td>
<td>3T37_A</td>
<td>GGGSAG</td>
<td>FAD</td>
</tr>
<tr>
<td>PEG&lt;sup&gt;a&lt;/sup&gt; dehydrogenase</td>
<td>BAE96591.1</td>
<td>GAGSAG</td>
<td>FAD</td>
</tr>
<tr>
<td>Choline dehydrogenase</td>
<td>AAH34502.1</td>
<td>GAGSAG</td>
<td>n.d.&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

However, direct evidence for the presence of the flavin has not been reported in the partially purified preparations of CHD, with the enzymes from *P. aeruginosa*<sup>(62)</sup> and dog<sup>(61)</sup>, which contain the glycine box, proposed to use pyrroloquinoline quinone (PQQ) instead of a flavin. We are aware of two Ph.D. Theses showing UV-visible absorbance spectra of partially purified CHD from *E. coli* with a peak in the 450 nm region consistent with the presence of a flavin cofactor but further characterization of the cofactor has not been reported<sup>(55, 65)</sup>. Based on the high number of cysteine residues, on the location of the eukaryotic enzyme in the mitochondrial inner membrane next to the respiratory chain and on early studies of the presence of labile sulfur groups<sup>(60)</sup> it is often assumed in the introduction of many studies that the enzyme contains iron-sulfur clusters<sup>(41)</sup>. However, no evidence is available to conclude unequivocally that iron-sulfur clusters are present in CHD, with the notion that the enzyme is only partially purified.
contributing further to raise questions of whether contaminant proteins may instead harbor some of the reported cofactors.

The electron acceptor in the reaction of choline oxidation catalyzed by CHD is not known. It has been reported that oxygen is not the preferred electron acceptor even though the enzyme is able to utilize it \(^{(63)}\). The most commonly used electron acceptor in the enzymatic assays on CHD from various cellular sources \(^{(56, 59, 60, 62, 63)}\) is phenazine methosulphate (PMS). Other electron acceptors, such as cytochrome c and ferricyanide, have been tested by Rendina and Singer \(^{(59)}\) and by Barrett \textit{et al} \(^{(66)}\) on rat CHD extracted from mitochondria, but the highest enzymatic activity was measured in the presence of PMS. It has been proposed that the enzyme can utilize coenzyme Q as electron acceptor based on the cellular localization of the enzyme close to Complex II on the inner mitochondrial membrane, but further investigation is necessary to validate this hypothesis \(^{(60, 67)}\).

The substrate specificity of partially purified rat liver CHD was investigated by Tsuge \textit{et al.} and enzymatic activity was determined only with choline or betaine aldehyde as substrate \(^{(60)}\). In agreement with these data CHD from \textit{H. elongata} can also use betaine aldehyde as substrate besides choline \(^{(63)}\). \textit{In vitro} inhibition of partially purified rat liver CHD was observed with 2-dimethylaminoethanol, monoethanolamine, semicarbazide and to lesser extents L-malate and glycine betaine \(^{(60)}\).

Preliminary kinetic studies on partially purified CHD from rat liver demonstrated an increase in the specific activity of the enzyme with choline at alkaline pH \(^{(60)}\). A steady state kinetic characterization with choline and PMS as electron acceptor yielded a \(K_m\) value of 7 mM for choline and allowed to propose a Ping-Pong Bi-Bi steady state kinetic mechanism for partially purified rat CHD \(^{(60)}\). For the enzyme from \textit{H. elongata} the kinetic parameters were determined
at pH 7.0 at fixed concentrations of either PMS or oxygen as electron acceptor, thus no information on the steady state kinetic mechanism for bacterial CHD is available \(^{63}\).

### 4.10 Homology model of human CHD

We have generated an homology model of human CHD with the SWISS-MODEL server \(^{68}\) using CHO from \(A.\ globiformis\) as template. The two enzymes share 30\% sequence identity (Figure 4.4) and catalyze the same oxidation reaction of choline.

The active site mutant of CHO with Ser-101 replaced with alanine (PDB 3NNE) was used instead of the wild type enzyme, since the flavin in the latter crystallographic structure is present in an unusual C(4a)-adduct. The N-terminal and internal extra peptides present in the sequence of CHD were deleted \textit{in silico} for the construction of the homology model, because they are absent in CHO. Two other homology models of human CHD were generated using pyridoxine 4-oxidase (PDB 3T37) or aryl-alcohol oxidase (PDB 3FIM) as templates, yielding structures that were practically superimposable with that generated using CHO (data not shown). Figure 4.5 illustrates the overall structure and the active site of the homology model of human CHD, superimposed with the tridimensional structure of another active site mutant of CHO, i.e., Val-464-Ala.

The overall structure of the model of human CHD appears to be similar to the conserved overall folding of the GMC superfamily members (Figure 4.3) and it does not highlight any transmembrane helix domain or other well-defined membrane-binding domain. Therefore, we hypothesize that human CHD is not an integral membrane protein.
Figure 4.4: Alignment of the amino acid sequences of human CHD and A. globiformis CHO. Main gaps are boxed in blue, the glycine box is contoured in red, active site residues of CHO conserved in CHD are highlighted by a red star and active site residues of CHO not conserved in CHD are highlighted by a blue star.
Figure 4.5: Homology model of human CHD and comparison of its active site with CHO. Panel A: superimposition of human CHD in dark grey and CHO variant Val-464-Ala (PDB 3LJP) in light grey. The CHD residue Leu-78, associated with the polymorphism rs12676, Leu-78 → Arg, is shown in magenta. Panel B: superimposition of the active site of human CHD (magenta) with the active site of CHO variant Val-464-Ala (green); labels are for CHD (top line) and CHO (bottom line). The FAD cofactor of CHO is shown in yellow.

Other possible ways of association of CHD with the inner mitochondrial membrane can be considered. For example, an insertion of a short hydrophobic anchor of the N terminal region of the protein in the membrane, similar to the C terminal anchor described for human monoamine oxidase B \(^{(69)}\). Alternatively, hydrophobic interactions between exposed hydrophobic regions of the protein and the membrane or ionic interactions between positively charged residues and the polar heads of phospholipids. Given the difficulties encountered to extract CHD from the membrane it is also possible that there is a covalent attachment of the protein to a phospholipid, as in the case of rodent neural cell adhesion molecule (NCAM) \(^{(70)}\).

The homology model of human CHD allows us to propose the localization of Leu-78, which is relevant to the polymorphism rs12676 associated with male infertility, on the surface of the enzyme (Figure 4.4). Such a replacement of a hydrophobic residue with a positively charge one
would locally alter the polarity of the enzyme surface, perhaps decreasing the stability of the enzyme.

### 4.11 Comparison of CHD with CHO

In contrast to CHD, CHO from *A. globiformis* has been extensively investigated in its structural, biochemical, kinetic and mechanistic properties \(^{(71-82)}\). This primarily stems from the fact that the oxidase is a soluble protein, which can be purified to high yields in stable and active form \(^{(12)}\). This offers the opportunity to compare the amino acid sequences of human CHD and bacterial CHO and, because the two enzymes catalyze the oxidation of the same substrate, hypothesize that the residues conserved in the active site of the two enzymes have similar roles in catalysis. Alignment of the amino acid sequences of human CHD and *A. globiformis* CHO returns 30% identical and 28% similar residues, as shown in Figure 4.4. Several of the active site residues previously investigated in CHO by using site-directed mutagenesis and biochemical, structural and mechanistic tools are conserved in CHD, including Glu-312 \(^{(45)}\), His-351 \(^{(83)}\), His-466 \(^{(84)}\) and Asn-510 \(^{(77)}\) (Table 4.4 and Figure 4.5).

Thus, we propose that Glu-339 and His-401 of CHD participate in substrate binding and positioning for the subsequent hydride transfer, His-511 modulates the polarity of the active site and stabilizes the transition state for the oxidation of choline to betaine aldehyde, and Asn-555 is important for choline and FAD oxidation.
<table>
<thead>
<tr>
<th>Residue in CHD</th>
<th>Corresponding residue in CHO</th>
<th>Proposed role in CHO</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu-134</td>
<td>His-99</td>
<td>Covalent link of the flavin cofactor</td>
<td>(74)</td>
</tr>
<tr>
<td>Ala-136</td>
<td>Ser-101</td>
<td>Activation of the alcohol substrate for the hydride transfer reaction</td>
<td>(80, 81)</td>
</tr>
<tr>
<td>Glu-339</td>
<td>Glu-312</td>
<td>Substrate binding</td>
<td>(45)</td>
</tr>
<tr>
<td>His-401</td>
<td>His-351</td>
<td>Substrate binding and positioning</td>
<td>(83)</td>
</tr>
<tr>
<td>Ala-509</td>
<td>Val-464</td>
<td>Oxidation of the flavin cofactor</td>
<td>(78)</td>
</tr>
<tr>
<td>His-511</td>
<td>His-466</td>
<td>Modulation of active site polarity and stabilization of the transition state for choline oxidation</td>
<td>(84)</td>
</tr>
<tr>
<td>Asn-555</td>
<td>Asn-510</td>
<td>Oxidation of choline and FAD</td>
<td>(77)</td>
</tr>
</tbody>
</table>

Three residues investigated in the active site of CHO are notably different in CHD, i.e. His-99, Ser-101, and Val-464, which are replaced in CHD by Leu-134, Ala-136 and Ala-509, respectively. His-99 is the site of covalent attachment of the flavin to CHO (74), allowing proposing that a flavin, if present, in CHD would not be covalently linked to the protein. Ser-101 contributes to the optimization of the overall turnover of CHO, which requires the fine-tuning of four consecutive half-reactions for the oxidations of choline to betaine aldehyde and betaine aldehyde to glycine betaine, each followed by the oxidation of the flavin by oxygen (81). Since a second enzyme, i.e., BADH, catalyzes the oxidation of betaine aldehyde to glycine betaine, such
a fine-tuning of multiple half-reactions is not required in CHD. Val464 provides a nonpolar side chain in CHO site that is required to guide oxygen in proximity of the C(4a) atom of the flavin, where it will subsequently react with the reduced flavin \(^{(78)}\). This conclusion derived from site-directed mutagenesis studies in which Val-464 was replaced with either threonine or alanine \(^{(78)}\). Interestingly, an alanine is present in lieu of Val-464 in CHD (Table 4.4), allowing proposing that lack of a valine may be partly responsible for the low reactivity of CHD toward oxygen.

Another difference that emerges from the comparison of the amino acid sequences of the oxidase and the dehydrogenase is the presence in CHD of extra peptides at the N-terminus and at position 357-381 that are not present in CHO (Figure 4.4). The extra peptide at the N-terminus of CHD may be relevant to the import and the targeting of the enzyme to the matrix of mitochondria. This would agree well with the observation that mature rat liver CHD in mitochondria lacks the first 34 amino acids at the N-terminal end \(^{(41)}\). The second extra peptide, with a length of 25 residues, may be responsible for the membrane association of CHD, although no evidence is available to back up this hypothesis at this stage.

4.12 Conclusions

Figure 4.6 provides a timeline of the major developments on CHD since the first report in 1937 on the oxidation of choline by rat liver \(^{(85)}\).

The mammalian enzyme responsible for the oxidation of choline and its connection to the mitochondrial respiratory chain were recognized in the 50s, and the first partial purifications of CHD were carried out in the 80s. The difficulties encountered in the purification of the mammalian and subsequently of the prokaryotic enzyme led only to preliminary biochemical characterizations of CHD, with a progressive fading in interest by biochemists toward the enzyme.
Since 2007, however, human CHD has received considerable interest by the medical community due to the involvement of the enzyme in a number of human pathologies, including male infertility, homocysteinuria, breast cancer and metabolic syndrome. A biochemical characterization of the enzyme is thereby required to gain knowledge on the molecular bases linking CHD to these pathologies. For example, in vitro structural and kinetic characterizations of the CHD variants associated with polymorphisms would be fundamental to provide a biochemical rationale for the enzyme malfunction, which may be due to several factors such as protein aggregation, increased protein instability, decreased enzymatic activity, more limited access of the substrates to the enzyme active site, etc.

To date, a biochemical characterization of CHD has been severely hampered by the lack of obtainment of stable, active and highly purified enzyme from either mammalian or prokaryotic sources. Several strategies have been developed with other enzymes to express “difficult” proteins associated with membranes. A choice for CHD could be to use the yeast *Pichia pastoris,*...
which can allow import and localization of the recombinant enzyme in the mitochondria. This protocol has been successfully employed for human liver monoamine oxidase A (86). Other approaches to overexpress recombinant human CHD in *Escherichia coli* could exploit the use of the C43 strain, which was selected for its high tolerance towards membrane proteins (87). Alternatively, a fusion protein of CHD with maltose-binding protein or other types, such as SUMO (small ubiquitin-related modifier) protein could be engineered to increase the solubility of human CHD. This latter approach was successful for the bacterial expression of the 5-lipoxygenase-activating protein (FLAP) or the severe acute respiratory syndrome coronavirus (SARS-CoV) membrane protein (88). Thus, developing a successful protocol for the purification of stable and active enzyme is an absolute requirement to advance our knowledge on this old and important enzyme.

### 4.13 References


Gratson, A. ((2005)) Cloning, heterologous expression, and characterization of human and Escherichia coli choline dehydrogenases, In *Nutritional Sciences*, p 114, University of Illinois, Urbana-Champaign


Powell, N. L. N. ((2003)) Bacterial choline-oxidizing systems: characterization of enzymes involved in stress tolerance and the roles of cytochrome P450s 1A1, 1A2, and 2D6, and the GSTM1 genes in treatment resistant depression, In *Dept of Chemistry*, p 252, Georgia State University, Atlanta.


CHAPTER V: EXPRESSION TRIALS OF RECOMBINANT HUMAN CHOLINE DEHYDROGENASE

5.1 Abbreviations

PCR, polymerase chain reaction; DMSO, dimethyl sulfoxide; CHD, choline dehydrogenase; CHO, choline oxidase; IPTG, isopropyl β-D-1-thiogalactopyranoside; PMSF, phenylmethanesulfonylfluoride.

5.2 Abstract

Human choline dehydrogenase (hCHD) is a mitochondrial membrane protein related to severe diseases such as male infertility, homocysteinuria, metabolic syndrome, and cancer. The challenges encountered in the previous purification attempts of choline dehydrogenases from different sources have so far hampered the biochemical characterization of this important enzyme. In this study a bioinformatics analysis of the protein sequence of mitochondrial CHD from different sources suggests the presence of a cleavable mitochondrial targeting presequence at the N terminus of the protein and of a highly conserved pattern in the N terminal domain that is possibly related to a post-translational modification or in recruiting of a cellular factor. This study focuses on different expression conditions for hCHD and five deletion variants at the N terminal end of the protein. A partial purification of the truncated Δ1-30hCHD is achieved, but the scarce purity and incorporation of the cofactor prevented further characterization of the protein.
5.3 Introduction

The gene coding for human choline dehydrogenase (hCHD) is on chromosome 3, location 3p21.1 on the complementary strand (Figure 5.1) \(^{(1)}\). CHD is conserved in different species with the bacterial enzyme known to be associated to the cell membrane and the eukaryotic one reported as located on the inner mitochondrial membrane \(^{(2)}\). A recent study has shown that hCHD is also present on the outer membrane of the mitochondria \(^{(3)}\).

![Figure 5.1: Location of human choline dehydrogenase gene on chromosome 3. Adapted from NCBI gene database.](image)

The protein sequence of hCHD carries the signatures of the Glucose-Methanol-Choline oxidoreductase enzyme superfamily \(^{(4)}\), which includes FAD-dependent enzymes catalyzing the oxidation of structurally unrelated alcohols. hCHD catalyzes the oxidation of choline to betaine aldehyde in the glycine betaine biosynthetic pathway \(^{(3)}\). As reviewed in Chapter IV, hCHD plays an important role in controlling the concentration of choline and glycine betaine in the body, which are related to lipid metabolism, synthesis of the neurotransmitter acetylcholine, epigenetic regulation, osmoregulation, and methylation processes \(^{(5)}\). hCHD has been related to severe diseases such as male infertility \(^{(6-8)}\), homocysteinuria \(^{(9, 10)}\), metabolic syndrome \(^{(11)}\), and cancer \(^{(12, 13)}\). Recently hCHD has been shown to play a role in triggering one known pathway for mytophagy in response to the administration of the mitochondrial toxin CCCP (carbonyl cyanide \(m\)-chlorophenylhydrazone) \(^{(3)}\). However very little is known on the biochemical characterization of the enzyme due to the lack of purified and stable preparations of hCHD. As described in
Chapter IV, previous purification attempts of CHD from different sources have encountered many difficulties in the obtainment of highly purified and stable preparations of enzyme probably due to its association with the membrane. Recombinant overexpression of eukaryotic membrane proteins in *Escherichia coli* often fails to achieve a good yield of stable purified protein \(^{(14)}\). Indeed expression of membrane eukaryotic proteins requires a variety of machinery such as translocon components and cellular chaperones, which may not be present in *E. coli* \(^{(14)}\). Some strains of *E. coli* have been engineered to express chaperones, such as SHuffle T7 Express lysY, and to improve membrane protein expression such as C43(DE3) \(^{(15)}\). This study focused on the expression trials of recombinant hCHD in *E. coli*, engineering deletion mutants at the N terminus of the protein and testing different *E. coli* strains for protein expression. A construct of hCHD carrying a His-tag was designed and affinity chromatography was employed for purification of the recombinant enzyme. Furthermore, bioinformatics analysis of the protein sequence of mitochondrial CHD revealed a putative cleavable mitochondrial targeting presequence at the N terminal end of the protein and a highly conserved pattern in the N terminal domain that are not present in bacterial CHD.

### 5.4 Materials and methods

**Materials.** The plasmid pMK-RQ containing the synthetic gene coding for human choline dehydrogenase and *E. coli* strain DH5α were purchased from Invitrogen Life Technologies (Grand Island, NY). The enzymes *BamHI*, *NdeI*, *DpnI*, T4 DNA ligase and *E. coli* strain SHuffle T7 Express lysY were from New England BioLabs (Ipswich, MA), *Pfu* DNA polymerase from Stratagene (La Jolla, CA), and oligonucleotides from Sigma Genosys (The Woodlands, TX). Tween 20 and calf intestinal alkaline phosphatase (CIAP) were from Promega (Madison, WI), and HiTrapTM Chelating HP 5 mL affinity column was from GE Healthcare (Piscataway, NJ).
QIAspin Spin Miniprep kit and QIAquick PCR purification kit were from Qiagen (Valencia, CA), and choline chloride from ICN Pharmaceutical Inc. (Irvine, CA). BugBuster Master Mix and *E. coli* strains BL21(DE3)pLysS and Rosetta(DE3)pLysS were from Novagen (Madison, WI). *E. coli* strain C43(DE3) was a kind gift from Dr. Zehava Eichenbaum, Georgia State University, Atlanta. All other reagents used were of the highest purity commercially available.

**Subcloning of wild-type and N-truncated mutants of hCHD in pET20b(+)**. The synthesized gene coding for hCHD with codon usage optimized for recombinant expression in *E. coli* (Figure 5.2) was purchased by Invitrogen Life Technologies in the kanamycin-resistant delivery vector pMK-RQ designed with restriction sites for *NdeI* and *BamHI* at the 5’ and 3’ ends of the open reading frame, respectively.
Figure 5.2: Synthesized gene coding for hCHD with codon usage optimized for recombinant expression in *E. coli*.

The construct pMK-RQ/hCHD was transformed into *E. coli* strain DH5α for amplification. A single colony was used to inoculate 5 mL of Luria-Bertani medium with kanamycin (30 µg/mL) which was then incubated at 37 °C for 15 h. The plasmid was extracted from the wet cell paste obtained by centrifugation at 6,000 g for 10 min with the use of the QIAquick Spin miniprep kit.
All purification steps in the subcloning procedures were performed with the QIAquick PCR purification kit, unless specified otherwise. The full length gene was cut with double digestion by \textit{NdeI} and \textit{BamHI} at 37 °C for 2 h, and purified. The deletion mutants Δ1-17, Δ1-20, Δ1-29, Δ-30, and Δ1-38 were amplified by PCR (the primers are listed in Table 5.1) in presence of 2% DMSO with an initial denaturation step at 95 °C, followed by 18 cycles as: denaturation for 45 s at 95 °C, annealing for 1 min at 57 °C (the annealing temperature decreases of 0.2 °C each cycle), and extension for 5 min at 68 °C, followed by a final step at 68 °C for 10 min.

\textbf{Table 5.1:} Oligonucleotide primers used for PCR amplification of hCHD. Restriction sites for \textit{NdeI} and \textit{BamHI} enzymes are underlined.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>hCHDreverse</td>
<td>CGCGGATCCCTTAACGCTGGGTGGCCAGGGGTACG</td>
</tr>
<tr>
<td>Δ1-17forward</td>
<td>GGATTCATATGGGCTCTGGGTACGAG</td>
</tr>
<tr>
<td>Δ1-20forward</td>
<td>GGATTCATATGCAGCAGCAGGCAGCGGTCAGCTGGGTACG</td>
</tr>
<tr>
<td>Δ1-29forward</td>
<td>GGAATTCATATGCTGGCAAGCGCAGGTAGCG</td>
</tr>
<tr>
<td>Δ1-30forward</td>
<td>GGAATTCATATGGCAAGCGCAGCTGGGTACG</td>
</tr>
<tr>
<td>Δ1-38forward</td>
<td>GGAATTCATATGGATGAAATATAGCTATGTTTGGT</td>
</tr>
</tbody>
</table>

Successful amplification of the primers by \textit{Pfu} DNA polymerase was checked by agarose gel electrophoresis and the PCR products were purified and digested with \textit{DpnI} for 2 h to ensure complete removal of the template plasmid (i.e., wild-type) containing the full length gene. After another purification step the amplicons were subjected to a double digestion by \textit{NdeI} and \textit{BamHI} at 37 °C for 2 h and further purified. The ampicillin-resistant vector pET20b(+) was digested with \textit{NdeI} and \textit{BamHI} at 37 °C for 2 h and purified prior to be processed with 1 unit of calf intestinal alkaline phosphatase (CIAP) at 37 °C for 30 min. After a purification step with the QIAquick PCR purification kit the cut and dephosphorylated vector pET20b(+) was ligated to
the insert with incubation for 15 h at 16 °C with T4 DNA ligase. A volume of 5 µL of the ligation mixtures was used to transform *E. coli* strain DH5α and the resulting colonies grown at 37 °C on Luria-Bertani agar plates containing 50 µg/mL ampicillin were screened for the presence of the desired insert by DNA sequencing at the GSU core facility. The DNA sequencing confirmed the correct insertion of the gene in the vector pET20b(+) and the absence of undesired mutations for the constructs pET20b(+)/hCHD full length, pET20b(+)/Δ1-17hCHD, pET20b(+)/Δ1-20hCHD, pET20b(+)/Δ1-29hCHD, pET20b(+)/Δ1-30hCHD, and pET20b(+)/Δ1-38hCHD.

**Expression trials of hCHD and deletion variants.** The constructs pET20b(+)/hCHD full length, pET20b(+)/Δ1-17hCHD, pET20b(+)/Δ1-20hCHD, pET20b(+)/Δ1-29hCHD, pET20b(+)/Δ1-30hCHD, and pET20b(+)/Δ1-38hCHD were used to transform *E. coli* strains Rosetta(DE3)pLysS, SHuffle T7 Express lysY, or BL21(DE3)pLysS. Single colonies were used to inoculate 5 mL of Luria-Bertani or Terrific broth containing 50 µg/mL ampicillin and 34 µg/mL chloramphenicol and the resulting cultures were incubated at 37 °C (30 °C in the case of *E. coli* strain SHuffle T7 Express lysY) until they reached an optical density at 600 nm of 0.6. IPTG was added to final concentrations ranging from 50 to 1000 µM and the temperature was lowered to 16 or 18 °C for 18 to 21 h. Alternatively, the incubation with IPTG was performed for 2 hours at 37 °C (30 °C in the case of *E. coli* strain SHuffle T7 Express lysY). The wet cell paste obtained by centrifugation was lysed by incubation with 900 µL BugBuster Master Mix for 20 min at room temperature. Cell free extracts were obtained by centrifugation at 10,000 g for 20 min at 4 °C. Larger scale expression trials were performed by inoculating 50 mL of Luria-Bertani broth in a flask and the resulting cultures were incubated at 37 °C until they reached an optical density at 600 nm of 0.6. IPTG was added to final concentrations of 500 and 1000 µM
and the temperature was lowered to 18 °C for 20 h. In all cases the analysis of recombinant protein overexpression was carried out by SDS-PAGE analysis of the crude and cell free extracts. Samples of the crude extracts for SDS-PAGE were prepared by resuspending the wet cell paste recovered by centrifugation of an aliquote of the culture in sample buffer (60 mM TRIS-Cl, pH 6.6, 10% v/v glycerol, 5% w/v SDS, 5% v/v β-mercaptoethanol, 0.01% w/v bromophenol blue), and by boiling the solution for 10 min at 100 °C. Samples of the cell free extract were obtained by diluting the solution obtained by BugBuster Master Mix in sample buffer and boiling at 100 °C for 10 min. In all cases a negative control of *E. coli* not transformed with pET20b(+) containing the hCHD gene and a positive control of pET20b(+) containing the gene coding for choline oxidase (CHO) from *Arthrobacter globiformis* were treated in parallel with the CHD samples. Recombinant overexpression of CHO in Rosetta(DE3)pLysS was obtained by addition of IPTG to a final concentration of 50 μM as previously described \(^{(16)}\).

**Subcloning of Δ1-30hCHD in pET15b with N-terminal His-tag.** The deletion mutant Δ1-30hCHD was amplified by PCR as previously described and the amplicon was purified and digested with *DpnI* for 2 h to ensure complete removal of the plasmid containing the full length gene. After another purification step with the QIAquick PCR purification kit the amplicon was subjected to a double digestion by *NdeI* and *BamHI* at 37 °C for 2 h and further purified. The ampicillin-resistant vector pET15b was digested with *NdeI* and *BamHI* at 37 °C for 2 h and purified prior to be processed with 1 unit of calf intestinal alkaline phosphatase (CIAP) at 37 °C for 30 min. After a purification step with the QIAquick PCR purification kit the cut and dephosphorylated vector pET15b was ligated to the insert with incubation for 15 h at 16 °C with T4 DNA ligase. This cloning strategy introduces an N-terminal his-tag in the gene. A volume of 5 μL of the ligation mixtures was used to transform *E. coli* strain DH5α and the resulting
colonies on Luria-Bertani agar plates containing 50 μg/mL ampicillin were screened for the presence of the desired insert by DNA sequencing at the GSU core facility.

**Expression trials of Δ1-30hCHD.** The construct pET15b/Δ1-30hCHD carrying an N-terminal His-tag was used to transform *E. coli* strains Rosetta(DE3)pLysS, SHuffle T7 Express lysY, and BL21(DE3)pLysS or C43(DE3). Single colonies were used to inoculate 5 mL of Luria-Bertani or Terrific broth containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol and in some samples 1% w/v glucose or 5 mM MgSO4. For small scale expression trials of *E. coli* strains Rosetta(DE3)pLysS and BL21(DE3)pLysS the resulting cultures were incubated at 37 °C until they reached an optical density at 600 nm of 0.6. IPTG was then added to final concentrations of 500 and 1000 μM and the cultures were incubated at 37 °C for 4 h. Small scale expression trials of *E. coli* strains SHuffle T7 Express lysY and C43(DE3) were performed by growing the resulting cultures at 25 °C until they reached an optical density at 600 nm of 0.6. IPTG was then added to a final concentration of 100 μM and the cultures were incubated at 25 °C for 20 h. A negative and positive control of *E.coli* not transformed with the plasmid carrying the hCHD gene and of CHO from *A. globiformis* were performed as previously described. The SDS-PAGE analysis for the crude extracts of the small scale expression trials was performed as previously described in this section.

A larger scale expression trial in *E. coli* strain Rosetta(DE3)pLysS was performed by growing a culture in 100 mL of Luria-Bertani broth in a flask until an optical density at 600 nm of 0.6. IPTG was added to final concentrations of 50 or 800 μM and the cultures were incubated at 18 °C for 20 h. The wet cell paste of 1.2 g obtained by centrifugation was resuspended in the lysis buffer (50 mM TRIS-Cl, pH 8.0, 100 mM NaCl, 25 mM imidazole, 10% v/v glycerol, 5 mM β-mercaptoethanol, 1 mM PMSF, 5 mM MgCl2, 2 mg/mL lysozyme, 5 μg/mL RNase and 5 μg/mL
DNase) and subjected to several cycles of sonication. The cell free extract obtained by centrifugation at 10,000 g for 20 min at 4 °C was loaded onto a HiTrapTM Chelating HP 5 mL affinity column equilibrated in buffer A (50 mM TRIS-Cl pH 8.0, 100 mM NaCl, 25 mM imidazole, 10% v/v glycerol, 5mM β-mercaptoethanol). After washing the column with 10 mL of buffer A, the elution step was performed by applying 10 mL of 100% buffer B (buffer A + 500 mM imidazole). All purification steps were carried out at 4 °C. Samples of the different steps (crude extract, cell free extract, flow through, wash, and elution) were analyzed by SDS-PAGE analysis as previously described in this section.

A larger scale expression trial in *E. coli* strain SHuffle T7 Express lysY was performed by growing a culture in 100 mL of Terrific Broth in a flask at 25 °C until an optical density at 600 nm of 0.6. IPTG was added to a final concentration of 100 μM and the culture was incubated at 25 °C for 20 h. The wet cell paste of 2.5 g obtained by centrifugation was resuspended in the detergent lysis buffer (40 mM TRIS-Cl, pH 8.0, 80 mM NaCl, 20 mM imidazole, 16% v/v glycerol, 5mM β-mercaptoethanol, 0.4% v/v Tween 20, 1 mM PMSF, 5 mM MgCl2, 2 mg/mL lysozyme, 5 μg/mL RNase and 5 μg/mL DNase) and subjected to several cycles of sonication. The cell free extract obtained by centrifugation at 10,000 g for 20 min at 4 °C was loaded onto a HiTrapTM Chelating HP 5 mL affinity column equilibrated in buffer A (40 mM TRIS-Cl, pH 8.0, 80 mM NaCl, 20 mM imidazole, 16% v/v glycerol, 5mM β-mercaptoethanol, 0.4% v/v Tween 20). After washing the column with 10 mL of buffer A, the elution step was performed by applying 10 mL of 100% buffer B (buffer A + 500 mM imidazole). A negative control of SHuffle T7 Express lysY was subjected to the same treatment and passed through the affinity column. All purification steps were carried out at 4 °C. Samples of the different steps (crude
extract, cell free extract, flow through, wash, and elution) were analyzed by SDS-PAGE analysis as previously described in this section.

**Activity assay.** Choline dehydrogenase activity was measured with 10 mM choline and by using 1 mM phenazine methosulfate (PMS) as primary electron acceptor in a coupled assay with oxygen previously described for choline dehydrogenase from *Halomonas elongata* (17). Briefly, the organic substrate choline reduces the enzyme choline dehydrogenase; the reduced enzyme is oxidized by PMS and the enzymatically reduced PMS is spontaneously oxidized by molecular oxygen. The oxygen consumption is monitored by using a computer-interfaced Oxy 32 oxygen-monitoring system (Hansatech Instrument Ltd.) thermostated at 25 °C.

### 5.5 Results and discussion

**Conserved SKRL sequence in mitochondrial CHD.** Human choline dehydrogenase is a membrane mitochondrial protein involved in choline metabolism. Chapter IV provides a detailed review of the physiological role and importance of this enzyme and it highlights the scarcity of biochemical data on CHD due to the difficulties encountered so far in the purification process. The medical interest on CHD has grown in the last years mainly due to its association with male infertility (6-8), homocysteinuria (9, 10), metabolic syndrome (11), and cancer (12, 13). The protein sequence of mitochondrial CHD from different sources has two extra regions compared to bacterial CHD associated to the cell membrane, as highlighted by a multiple sequence alignment performed using ClustalW2 (18). The first extra region (boxed in blue in Figure 5.3) is represented by a longer N terminal region compared to prokaryotic CHD, and contains mainly hydrophobic and arginine residues. The second region (boxed in red in Figure 5.3) is located at position 80 in hCHD and shows a highly conserved sequence G-S-K-R-L-x-W (where x stands for any amino acid).
Figure 5.3: Multiple sequence alignment with Clustal W2 of protein sequences of choline dehydrogenase enzymes from different sources. 

*Homo sapiens* (human; NP_060867.2), *Brucella suis* (B.suis; YP_005154241.1), *Vibrio parahaemolyticus* (V.para; ZP_12109249.1), *Ochrobactrum anthropic* (O.anth; ABS15421.1), *Bos Taurus* (B.taur; NP_001192493.1), *Rattus norvegicus* (R.norv; NP_942026.1), *Mus musculus* (M.musc; NP_001129712.1), *Gallus gallus* (G.gallu; XP_414335.3), *Escherichia coli* (E.coli; BAE76094.1), *Pseudomonas aeruginosa* (P.aeru; NP_254059.1).

The full length protein (595 amino acids) is encoded by a nuclear gene and has therefore to be imported in the mitochondria \(^{(19)}\). It is hypothesized here that the extra region at the N terminal represents a cleavable mitochondrial targeting presequence which is removed by the mitochondrial proteases. This hypothesis was confirmed previously in the case of rat CHD, where it was shown by N terminal protein sequencing that the mature protein extracted from mitochondria lacks the first 34 amino acids, and two cleavage sites for the Mitochondrial Processing Protease (MPP) and the Inner Membrane Protease (IMP) were proposed \(^{(2)}\). The presence of a cleavable mitochondrial targeting presequence also in hCHD was hypothesized in a previous study \(^{(3)}\) but not experimentally confirmed, and the position of the cleavage site by mitochondrial proteases and the length of the mature hCHD were not determined. Due to the broad substrate specificity of mitochondrial proteases \(^{(19)}\) it is difficult to predict the cleavage site in hCHD, which could differ significantly from the one identified for rat CHD \(^{(2)}\). As a
consequence 5 different truncated mutants of the N terminal of hCHD have been designed (Figure 5.4).

<table>
<thead>
<tr>
<th>Variant</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>hCHD</td>
<td>MWCLLRGLRGALARGALGQQQSLGARALASAGSESRDEYSVYVGAGSAGCVLAGRLT</td>
</tr>
<tr>
<td>Δ1-17</td>
<td>ALGQQQSLGARALASAGSESRDEYSVYVGAGSAGCVLAGRLT</td>
</tr>
<tr>
<td>Δ1-20</td>
<td>QQQSLGARALASAGSESRDEYSVYVGAGSAGCVLAGRLT</td>
</tr>
<tr>
<td>Δ1-29</td>
<td>LASAGSESRDEYSVYVGAGSAGCVLAGRLT</td>
</tr>
<tr>
<td>Δ1-30</td>
<td>ASAGSESRDEYSVYVGAGSAGCVLAGRLT</td>
</tr>
<tr>
<td>Δ1-38</td>
<td>DEYSVYVGAGSAGCVLAGRLT</td>
</tr>
</tbody>
</table>

**Figure 5.4:** Design of five deletion mutants at the N terminal of human choline dehydrogenase. The figure shows the alignment of the N terminal sequence of full length human choline dehydrogenase (hCHD) and of the deletion mutants represented with the symbol Δ followed by the sequence deleted.

The Δ1-17 deletion mutant of hCHD was designed based on the fact that mitochondrial targeting presequences are usually between 15 and 50 amino acids long \(^{(19)}\) and that one possible recognition site for cleavage by MPP is an R-2 motif, one amino acid after an arginine residue \(^{(20)}\). The Δ1-20, the Δ1-29, and the Δ1-38 truncated versions of hCHD were chosen based on the bioinformatics analysis with the prediction servers for protein targeting sequences Predotar \(^{(21)}\), TargetP 1.1 \(^{(22, 23)}\), and MitoProt \(^{(24)}\). Based on the presence of another mitochondrial protease, the intermediate cleaving peptidase Icp55 \(^{(25)}\), which removes one destabilizing amino acid (Leu, Lys, Tyr, Phe, Arg, and Trp) exposed at the N terminus after the first cleavage by MPP the construct Δ1-30hCHD was engineered. The expected outcome was that the overexpression in *E. coli* of the full length CHD and some of these deletion variants would yield the recombinant precursor not soluble, in the inclusion bodies or highly unstable due to the absence of mitochondrial proteases in *E. coli*, while the truncated version of hCHD closest to the mature form of the protein would result in the expression of more soluble hCHD. However the expression trials of the truncated versions of hCHD performed in this study and described in the next paragraph did not show significant overexpression of the recombinant protein.
The second extra region in the protein sequences of mitochondrial choline dehydrogenases compared to bacterial enzymes possesses a highly conserved sequence $^{80}\text{G-S-K-R-L-x-W-K}^{87}$ as shown in Figure 5.3 (the numbering corresponds to the protein sequence of hCHD and x stands for any amino acid at that position). A bioinformatics analysis of ScanProsite (26), which allows searching for this motif in all protein sequences in the Swiss-Prot database, did not retrieve any hit other than eukaryotic choline dehydrogenases. According to the model of the crystal structure of hCHD presented in Chapter IV, the sequence $^{80}\text{G-S-K-R-L-x-W-K}^{87}$ is predicted to be on the surface of the protein (Figure 5.5).

**Figure 5.5:** Three dimensional model of hCHD shown as gray cartoon with the FAD cofactor as yellow sticks. The main chain of the conserved region $^{80}\text{G-S-K-R-L-x-W-K}^{87}$ is highlighted in red, and the side chain of the residue L78 is represented as blue sticks.

It was therefore hypothesized here that this conserved sequence present only in mitochondrial CHD may represent a consensus sequence for a post-translational modification or for the recruiting of some factors in the cells, rather than being involved in substrate recognition and
catalysis. It is also possible that this conserved motif is related to the association with the mitochondrial membrane, and that it is absent in bacterial CHD due to a different association mode with the plasma membrane compared with the mitochondrial one. It has been reported recently that human CHD plays a role in triggering the mitophagy after exposure to the toxin CCCP by changing its localization from the inner to the outer mitochondrial membrane and exposing the N-terminal domain of the mature hCHD (until amino acid 326) to the cytosol to recruit the factor SQSTM1 \(^3\). This factor is involved in one pathway known to target damaged mitochondria for degradation \(^{21}\). It is hypothesized here that the conserved \(^{80}\)G-S-K-R-L-x-W-K\(^{87}\) may be involved in this recruitment but there are no studies that allow backing up this hypothesis at this stage. It is interesting to note that the residue L78, which has been associated with male infertility \(^8\) when mutated to arginine, is located next to this conserved sequence (Figure 5.5), suggesting a possible medical relevance of mutations affecting this region of the protein sequence.

**Expression trials of hCHD and of the N-terminal deletion variants.** The full length gene coding for hCHD and the truncated mutants were subcloned in the expression vector pET20b(+). As described before it is hypothesized that the protein sequence of hCHD contains a cleavable targeting presequence at the N terminus and different truncated versions of hCHD were designed with the aim of improving the stability and solubility of the recombinant protein.

Small scale expression trials of 5 mL cultures of wild-type and truncated mutants of hCHD were performed in different conditions, using as positive control CHO from *Arthrobacter globiformis*, for which a successful protocol of recombinant overexpression in *E. coli* was previously optimized \(^{16}\). The background of native proteins in *E. coli* was estimated by SDS-PAGE
analysis and choline dehydrogenase activity of the crude extracts obtained from *E. coli* strains not transformed with the plasmid carrying the gene coding for CHD.

*E. coli* strains Rosetta(DE3)pLysS and BL21(DE3)pLysS were used here for the initial expression trials of full length and truncated versions of recombinant hCHD. Crude extracts from the small scale expression trials of full length and truncated version of human CHD using different IPTG concentrations and a short induction time at 37 °C showed no difference in the SDS-PAGE analysis from the background of native proteins in *E. coli* (data not shown).

A larger scale expression trial of full length and Δ1-30-human CHD was performed in Rosetta(DE3)pLysS with different IPTG concentrations and 20 h of induction at 18 °C. In parallel, a culture of CHO from *A. globiformis* was processed with the optimized protocol for overexpression. Samples of crude extracts were taken after 10, 30, 120 min and 20 h from addition of IPTG and analyzed by SDS-PAGE. The samples of the crude extracts from human CHD cultures collected at different times showed no difference from the background of the native proteins of *E. coli*, while samples collected at different times of the crude extract from CHO display a gradual increase in the overexpression of the recombinant protein monitored by SDS-PAGE analysis (data not shown). These results show that no overexpression of wild-type and truncated hCHD was observed in the crude extracts of the conditions tested. Based on the literature available on mitochondrial targeting presequences (15 to 50 amino acids length) it is possible that a longer deletion is necessary to generate the mature hCHD. However a deletion of more than 38 amino acids was not pursued due to the concern of exposing too much at the new N terminus the glycine box $^{47}$GAGSAG$^{52}$ proposed to be for FAD binding. It was considered more likely that the lack of overexpression in *E. coli* of a human mitochondrial protein known to be strongly associated with the membrane is linked to several issues, with the possible instability of
the precursor form being only one. In agreement with this notion not many eukaryotic membrane
protein are easily overexpressed in *E. coli*\(^{(14)}\). Based on the high cysteine contents of hCHD (13
cysteine residues in 595 amino acids) and of the use of the *E. coli* strain OrigamiB(DE3)pLysS
for the expression of recombinant CHD from *Halomonas elongata*\(^{(17)}\) the *E. coli* strain SHuffle
T7 Express *lysY* was tested. This strain is engineered with the deletions of the genes for
glutaredoxin reductase and thioredox reductase, which, similarly to Origami, enable the
formation of disulfide bonds in the cytosol\(^{(27)}\). Furthermore Shuffle strains express in the cytosol
the disulfide bond isomerase DsbC that has been shown to promote the correction of mis-
oxidized proteins into the correct form and also to act as a chaperone for the folding of proteins
non containing disulfide bonds\(^{(28,29)}\). The small scale expression trial of full length and truncated
versions of human CHD in the *E. coli* strain SHuffle T7 Express *lysY* showed no choline
dehydrogenase activity detected in the cell free extracts obtained by lysing the cells with the
commercial solution BugBuster Master Mix. SDS-PAGE analysis of these cell free extracts
(Figure 5.6) showed no significant overexpression of human CHD compared to the background
of native proteins of *E. coli*.

![Figure 5.6: SDS-PAGE analysis of the crude extracts of small scale expression trials.](attachment:figure56.png)
The constructs are pET20b(+)/hCHD (F, lane 1), pET20b(+)/Δ1-17hCHD (17, lane 3), pET20b(+)/Δ1-20hCHD (20, lane 4), pET20b(+)/Δ1-29hCHD (29, lane 5), and pET20b(+)/Δ1-30hCHD (30, lane 6). Lane 2 is the crude extract of *E. coli* SHuffle T7 Express lysY not transformed with the plasmid carrying the gene coding for hCHD.

**Expression trials Δ1-30hCHD with His-tag.** To address the issue of the low expression yield of hCHD the option of the recombinant protein carrying an affinity tag was chosen. This strategy was tested on the truncated version Δ1-30hCHD as it is the closest one to the mature form of rat CHD. Before performing a purification trial by affinity chromatography other expression conditions in small scale were tested for significant overexpression in the crude extracts. Small scale expression trials in *E. coli* strains Rosetta(DE3)pLysS and BL21(DE3)pLysS with different IPTG concentrations and 4 h induction at 37 °C yielded crude extracts that did not show overexpression of protein by SDS-PAGE analysis in comparison with the background of native proteins of *E. coli* (data not shown).

Addition of glucose and MgSO₄ to the medium was previously described to limit toxic effects of recombinant proteins to *E. coli* (30). Furthermore the use of *E. coli* strain C43(DE3), which is a natural double mutant strain of BL21(DE3) selected for its particular tolerance for overexpression of membrane proteins (15), showed good results for the overexpression of the membrane recombinant proteins subunits b and c of *E. coli* F-ATPase and the alanine-H⁺ carrier that failed in other *E. coli* strains (15). Small scale expression trials in *E. coli* strains SHuffle T7 Express lysY and C43(DE3) were performed, with addition of glucose and MgSO₄ to the liquid medium and growth and induction temperature of 25 °C. However, SDS-PAGE analysis of the crude extracts did not show overexpression of the truncated hCHD compared to the background of native proteins in *E. coli* (data not shown).

The next step was to perform a purification trial by affinity chromatography with the aim of concentrating the recombinant protein in case of low expression yield. For this purpose a larger
scale expression trial in the *E. coli* strain Rosetta(DE3)pLysS was performed by using IPTG concentration of 50 and 800 μM and incubation at 18 °C for 20 h. The cell free extract was loaded and eluted from a HiTrapTM Chelating HP 5 mL affinity column to detect the presence of a recombinant His-tagged protein and samples of the flow through, wash, and elution steps were analyzed by SDS-PAGE (Figure 5.7). The same purification experiment was performed on *E. coli* not carrying the construct pET15/Δ1-30hCHD in order to compare the presence of native proteins of *E. coli* in the elution step. The SDS-PAGE analysis of the elution sample (Figure 5.7) from the first expression trials in Rosetta(DE3)pLysS did not show any band at the expected molecular weight of 62 kDa and there was no difference from the background of native proteins of *E. coli*.

![Figure 5.7: SDS-PAGE analysis of the expression trial in Rosetta(DE3)pLysS.](image)
The order of the samples is the following: lane 1 (M) choline oxidase from *A. globiformis* (60 kDa), lane 2 (CE800) crude extract from the culture induced with 800 μM IPTG, lane 3 (CE50) crude extract from native proteins of *E. coli*, lane 4 (FT50) flow through from the culture induced with 50 μM IPTG, lane 5 (E800) elution from the culture induced with 800 μM IPTG, lane 6 (E50) elution from the culture induced with 50 μM IPTG, lane 7 (FT800) flow through from the culture induced with 800 μM IPTG, lane 8 (W800) wash from the culture induced with 800 μM IPTG, and lane 9 (W50) wash from the culture induced with 50 μM IPTG.

The absence of Δ1-30hCHD in the elution sample could be due to the fact that the recombinant protein did not bind to the column and was eluted in the flow through or wash sample. However
no difference in the SDS-PAGE analysis of the flow through and wash samples is evident compared to the background of native proteins of *E. coli*.

Another expression trial was performed in SHuffle T7 Express *lysY* by using Terrific Broth, IPTG concentration of 100 µM and with a protocol of growth at lower temperature (25 °C instead of 37 °C) that was successful for the expression of the membrane bound GMC member pyridoxine 4-oxidase (31). The detergent Tween 20 was added to the lysis buffer and to all the purification buffers. The cell free extract obtained with sonication in the presence of 0.4% v/v Tween 20 was loaded and eluted from a HiTrapTM Chelating HP 5 mL affinity column to detect the presence of a recombinant His-tagged protein and samples of the flow through, wash, and elution steps were analyzed by SDS-PAGE (Figure 5.8). In parallel a culture of *E. coli* not carrying the plasmid pET15b/Δ1-30hCHD was subjected to the same steps. The elution from the affinity column showed a band consistent with the expected molecular weight of 62 kDa for Δ1-30hCHD and not present in the background of native proteins in *E. coli* (Figure 5.8), suggesting the presence of only partially purified recombinant His-tagged Δ1-30hCHD. The SDS-PAGE analysis of the elution fraction shows the presence of contaminants in the sample, with an estimated purity of the 62 kDa protein of around 40% (Figure 5.8).
Figure 5.8: Elution samples from the affinity column. Lane 1 (E) shows the elution from the cell free extract of SHuffle T7 Express lysY transformed with the plasmid pET15/Δ1-30hCHD and lane 2 (-) the elution from the cell free extract of E. coli untransformed.

UV-visible spectroscopic analysis of this fraction did not detect the characteristic peaks of oxidized flavins in the 300-500 nm region, which suggests the recombinant Δ1-30hCHD is unable to retain the proposed FAD cofactor despite the presence of 16% v/v of glycerol in the lysis buffer and in all the purification process. In agreement with most of the enzyme being in the apo form and to the scarce purity of the sample, the elution was characterized by only 0.4 U/mL of choline dehydrogenase activity with 30 mM choline as substrate at pH 10.0 and 25 °C. The scarce purity of the sample and the issues with incorporation of the flavin cofactor led to the decision of not proceeding with further purification trials of hCHD.

5.6 Conclusions

Despite the medical interest on human choline dehydrogenase has grown in the last years, very little is known on this important enzyme from a biochemical point of view. The bottleneck to the biochemical characterization of choline dehydrogenase from different sources has constantly
been represented by issues encountered in the purification of this enzyme, as reviewed in Chapter IV. To date even the presence of an FAD cofactor in eukaryotic CHD is only suggested by the presence of a glycine box in the protein sequence but not experimentally confirmed and still debated. This study presents a bioinformatics analysis of the protein sequences of mitochondrial choline dehydrogenases compared with the bacterial enzyme, which highlights two extra regions in the protein from eukaryotic sources. The N terminal extra region is suggested to be a cleavable mitochondrial targeting presequence and different deletion mutants have been engineered to improve the expression of recombinant hCHD. Attempts to achieve significant overexpression of the deletion variants in different conditions failed suggesting that the hypothesized cleavable targeting presequence at the N terminus is not the only issue hampering the expression and purification of hCHD. A second highly conserved extra region \( G-S-K-R-L-x-W-K \) is described and analyzed for mitochondrial CHD, and suggested to be involved in post-translational modification or interaction with cellular factors. The hypothesis of one possible role of this sequence in recruiting a factor involved in mytophagy is presented but no experimental evidence is found at this stage. Different expression conditions have been screened in this study but no significant overexpression of recombinant hCHD was obtained. The use of a His-tag, of the \( E. coli \) strain SHuffle T7 Express \( lysY \), and of an expression protocol in Terrific Broth with growth at 25 °C finally allowed for the partial purification of the truncated \( \Delta 1-30hCHD \). However the lack of the detection of the suggested FAD cofactor and the consequent low enzymatic detected prevented any further characterization of \( \Delta 1-30hCHD \).

5.7 References


CHAPTER VI: THE COMBINED STRUCTURAL AND KINETIC CHARACTERIZATION OF A BACTERIAL NITRONATE MONOOXYGENASE FROM Pseudomonas aeruginosa PAO1 ESTABLISHES NMO CLASS I AND II

(This chapter has been published verbatim in Salvi F., Agniswamy J., Yuan H., Vercammen K., Pelicaen R., Cornelis P., Spain J., Weber I.T., and Gadda G., (2014), The Journal of Biological Chemistry 289(34):23764-75; the author contributed to the design of the study, to the purification, kinetic and structural characterization of PA4202, data analysis, and wrote the manuscript)

6.1 Abbreviations

NMO, nitronate monooxygenase; Pa-NMO, nitronate monooxygenase from Pseudomonas aeruginosa PAO1; Cs-NMO, nitronate monooxygenase from Cyberlindnera saturnus; Nc-NMO, nitronate monooxygenase from Neurospora crassa; P3N, propionate 3-nitronate; 3-NPA, 3-nitropropionate; ESI-Q-TOF, electrospray ionization-quadruple-time of flight; PEG, polyethylene glycol.

6.2 Abstract

Nitronate monooxygenase (NMO) oxidizes the mitochondrial toxin propionate 3-nitronate (P3N) to malonate semialdehyde. The enzyme has been previously characterized biochemically in fungi, but no structural information is available. Based on amino acid similarity 4,985 genes are annotated in the Gen-Bank as NMO. Of these, 4,424 (i.e., 89%) are bacterial genes, including several Pseudomonads that have been shown to use P3N as growth substrate. Here, we have cloned and expressed the gene pa4202 of Pseudomonas aeruginosa PAO1, purified the resulting protein and characterized it. The enzyme is active on P3N and other alkyl nitronates, but cannot oxidize nitroalkanes. P3N is the best substrate at pH 7.5 and atmospheric oxygen with $k_{cat}/K_m$ of $12 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$, $k_{cat}$ of 1300 s$^{-1}$ and $K_m$ of 110 μM. Anaerobic reduction of the enzyme with P3N yields a flavosemiquinone, which is formed within 7.5 ms, consistent with this species
being a catalytic intermediate. Absorption spectroscopy, mass spectrometry and X-ray crystallography demonstrate a tightly, non-covalently bound FMN in the active site of the enzyme. Thus, PA4202 is the first NMO identified and characterized in bacteria. The X-ray crystal structure of the enzyme was solved at 1.44 Å, showing a TIM barrel fold. Four motifs in common with the biochemically characterized NMO from *Cyberlindnera saturnus* are identified in the structure of bacterial NMO, defining Class I NMO, which includes bacterial, fungal and two animal NMOs. Notably, the only other NMO from *Neurospora crassa* for which biochemical evidence is available lacks the four motifs, defining Class II NMO.

### 6.3 Introduction

Functional annotation of prokaryotic genes based on experimental evidence represents only 0.33% of the microbial genes in COMBREX (COmputational BRidges to EXperiments) \(^{(1, 2)}\). This is primarily due to the rapid progress in the sequencing of entire genomes that supersedes by far the biochemical characterization of the gene products \(^{(1, 2)}\). This problem is exacerbated by the limited accuracy of the computational prediction of function based on amino acid sequences \(^{(1, 2)}\), mainly due to the paucity of biochemical information available for the reference sequence. Thus, the quality of functional predictions can be significantly improved by establishing rigorous reference standards of proteins with experimentally determined functions and activities.

An emerging biochemical interest related to the functional annotation challenge is the metabolism of nitro toxins, such as propionate 3-nitronate (P3N) and 3-nitropropionate (3-NPA), which at physiological pH exist in equilibrium with a pK\(_a\) of 9.1 \(^{(3)}\). P3N is a potent inhibitor of succinate dehydrogenase and fumarase in the Krebs cycle \(^{(4-6)}\) and is found in some plants and fungi as a defense against herbivores \(^{(3, 7)}\). The inhibition of essential metabolic enzymes by P3N immediately halts energy production in poisoned cells, causing neurological disorders \(^{(3)}\) and, at
sufficiently high doses, death\(^{(3)}\). Cases of P3N poisoning have been documented in humans, with mortality rates \(\sim 10\%\)\(^{(8)}\). P3N neurotoxicity has been conclusively demonstrated in animal models\(^{(9,10)}\), primarily due to the impact of P3N poisoning in the agricultural industry\(^{(8,11-13)}\). 3-NPA, with its conjugate base P3N, is indeed routinely used at low doses for \textit{in vivo} studies of animal models to investigate the molecular basis of Huntington's disease\(^{(10,14)}\). Plants and fungi associated with P3N possess detoxifying nitronate monoxygenases (NMO; E.C. 1.13.12.16) as defense from the toxin\(^{(3,15)}\).

NMOs are members of the Group H flavin-dependent monoxygenases\(^{(16)}\), and catalyze the oxidation of P3N or other nitronate analogues through a radical mechanism involving a catalytic flavosemiquinone without formation of a canonical C4a-(hydro)peroxy-flavin\(^{(17)}\). Oxidation of P3N by NMO yields malonic semialdehyde, which is an important metabolite that can be converted to acetyl-CoA, acetate or 3-hydroxypropionate and enter various catabolic or anabolic pathways\(^{(3)}\). To date the only NMOs that have been characterized biochemically are from \textit{Neurospora crassa} (Nc-NMO) and \textit{Cyberlindnera saturnus} (Cs-NMO), which was previously classified as \textit{Williopsis saturnus}\(^{(3)}\). These fungal enzymes contain FMN as cofactor, have fast catalytic turnovers in atmospheric oxygen with \(\text{app}k_{\text{cat}} \geq 430\ \text{s}^{-1}\) and high specificity constants for P3N with \(\text{app}(k_{\text{cat}}/K_{P3N}) > 10^6\ \text{M}^{-1}\text{s}^{-1}\) at pH 7.4\(^{(3)}\). Based on amino acid similarity with the fungal NMOs 4,985 genes are currently annotated in the Gen-Bank as NMO or 2-nitropropane dioxygenase, as the enzyme was officially known before reclassification in 2010 by the IUBMB. Of these, 4,424 (i.e., 89\%) are bacterial genes, including several Pseudomonads that have been recently shown to be able to use 3-NPA/P3N as a growth substrate\(^{(18)}\). It is possible that in some bacteria NMO may represent a detoxification mechanism similar to that proposed for fungi and plants or, alternatively it may be a vehicle to scavenge nitrogen from the environment.
*Pseudomonas aeruginosa* PAO1 possesses three genes annotated as hypothetical NMOs, namely *pa0660, pa1024* and *pa4202* \(^{(19)}\). Crystal structures of the gene product of *pa1024* in the presence and absence of 2-nitropropane are available at resolution of 2.3 Å \(^{(20)}\). However, biochemical evidence to unequivocally conclude that the protein is an NMO is lacking, with only a qualitative description of the enzyme being active on 2-nitropropane but no quantitative descriptions of specific activity or kinetic parameters being reported with P3N or other nitronates \(^{(20)}\). No experimental data are available on the other two hypothetical NMOs coded by *pa0660* and *pa4202*. The physiological role of a closely related homolog of *pa4202* in *Pseudomonas* sp. strain JS189 was rigorously established to be the biodegradation of 3-NPA, but the kinetics and the structure of the enzyme were not determined \(^{(18)}\).

In this study, we expressed and purified the recombinant protein coded by gene *pa4202* from *P. aeruginosa* PAO1. The biochemical, kinetic and structural characterization revealed that the product of gene *pa4202* is an NMO that is highly specific for the metabolic poison P3N. We named the enzyme Pa-NMO. The crystal structure of Pa-NMO solved to 1.44 Å resolution allowed us to identify four consensus motifs in the primary structure of the protein that are conserved in Cs-NMO for which biochemical data are available. The consensus motifs are found in 475 sequences annotated as hypothetical NMO belonging to bacteria, fungi and two animals, establishing a new class of enzymes based on functional annotation of a reference standard protein for which structural-functional information is now available. Notably, Nc-NMO does not belong to this group and, being the only other NMO for which biochemical evidence for enzymatic activity is available, it defines a separate class of NMOs.
6.4 Experimental procedures

**Materials.** Plasmid pET21a(+) containing the gene *pa4202* from *Pseudomonas aeruginosa* PAO1 with a C terminal histidine tag was prepared in a previous study \(^\text{(18)}\). *Escherichia coli* strain Rosetta(DE3)pLysS was from Novagen (La Jolla, CA). Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was from Promega, (Madison, WI), nitroalkanes were from Sigma-Aldrich (St. Louis, MO). HiTrapTM Chelating HP 5 mL affinity column and PD-10 desalting columns were from GE Healthcare (Piscataway, NJ). All other reagents were of the highest purity commercially available.

**Purification of Recombinant PA4202 Enzyme (Pa-NMO).** *E. coli* expression strain Rosetta(DE3)pLysS harboring the plasmid pET21a(+)/*pa4202* was used to inoculate 1 L of Terrific Broth medium containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol. When the cell culture incubated at 37 °C with shaking at 160 rpm reached an OD\( _{600} \) of 1, IPTG was added to a final concentration of 100 µM and the temperature was lowered to 18 °C. After 19 h, the cells were harvested by centrifugation yielding 29 g of wet cell paste, which was resuspended in 120 mL of lysis buffer (10 mM imidazole, 300 mM NaCl, 10% v/v glycerol, 1 mM phenylmethylsulfonyl fluoride, 5 mM MgCl\(_2\), 2 mg/mL lysozyme, 5 μg/mL DNase, 5 μg/mL RNase, and 20 mM sodium phosphate, pH 7.4) and subjected to several cycles of sonication. The cell free extract obtained after centrifugation at 12000 g for 20 min was loaded directly onto a HiTrapTM Chelating HP 5 mL affinity column equilibrated with buffer A (10 mM imidazole, 300 mM NaCl, 10% v/v glycerol, and 20 mM potassium phosphate, pH 7.4). After washing with 10 column volumes of buffer A at 5 mL/min, Pa-NMO was eluted with 30% buffer B (buffer A + 0.5 M imidazole). The purest fractions based on SDS-PAGE analysis were pooled, dialyzed against 10 mM Tris-Cl, pH 8.0, 10% glycerol and stored at -20 °C.
Enzyme Assays. Protein concentration was determined using the Bradford method with bovine serum albumin as standard \(^{(21)}\). UV-visible absorbance was recorded with an Agilent Technologies diode-array spectrophotometer Model HP 8453 PC equipped with a thermostated water bath; SDS-PAGE analysis of protein samples was performed using wide range molecular weight markers (6500-200000 Da) from Sigma-Aldrich. The extinction coefficient of the enzyme-bound flavin was calculated in duplicate by passing a fraction of purified Pa-NMO trough a PD-10 desalting column equilibrated in 50 mM potassium phosphate, pH 7.0, before heat denaturation at 100 °C for 30 or 35 minutes. The protein was removed by centrifugation and the concentration of the extracted FMN was determined from the UV-visible absorption spectrum of the yellow soluble fraction using an \( \varepsilon_{450nm} \) of 12,200 M\(^{-1}\) cm\(^{-1}\) for free FMN \(^{(22)}\). Cofactor identification was performed by mass spectrometry of the flavin extracted from a desalted sample of Pa-NMO in water treated at 100 °C for 30 min using a Waters Micromass Q-TOF micro (ESI-Q-TOF) in negative ion mode at the Mass Spectrometry Facility of Georgia State University.

The enzymatic activity of the purified Pa-NMO with nitronates was determined with the method of the initial rates \(^{(23)}\) following oxygen consumption with a Hansatech Instruments computer-interfaced Oxy-32 oxygen-monitoring system. The method of the initial rates ensures to prevent conversion between the nitronate form and the nitroalkane form of the substrate. Pa-NMO was prepared by desalting chromatography through a PD-10 column equilibrated with 50 mM potassium phosphate, pH 7.5, 10% v/v glycerol just prior to the kinetic analyses. Stock solutions of nitronates and nitroalkanes were prepared as previously described \(^{(24, 25)}\). Enzymatic assays were carried out at atmospheric oxygen and 30 °C (i.e., 0.23 mM oxygen) in 50 mM potassium phosphate, pH 7.5, with the initial rates normalized for the enzyme-bound flavin using the
experimentally determined ε_{443}=12,500 \text{ M}^{-1} \text{ cm}^{-1}. Enzyme concentrations ranged from 1.4 to 8.8 nM; substrates were in the range from 0.02 mM to 20 mM. Since the second-order rate constants for protonation of the nitronates are in the range 15-75 \text{ M}^{-1} \text{ s}^{-1} (^{26, 27})$, enzymatic activity assays were started with the addition of the nitronate to the reaction mixture to ensure that a negligible amount of neutral species, i.e., nitroalkane, was present during the time required to acquire initial rates of reactions (typically ~30 s).

In the case of the steady-state kinetic experiment performed at 30 °C and pH 7.5 the methods of the initial rates following oxygen consumption was applied as described for the enzymatic assays above, with the enzyme concentration ranging from 1 to 3.5 nM, the P3N concentration from 0.03 to 0.8 mM and oxygen concentration from 15 to 156 µM. No differences in the initial rates were observed when different concentrations of enzyme were used with the same concentrations of substrate and oxygen.

Time-resolved absorbance spectroscopy of the anaerobic reduction of Pa-NMO with P3N was performed with an SF-61DX2 Hi-Tech KinetAsyst high-performance stopped-flow spectrophotometer, thermostated at 30 °C, equipped with a photo-diode array detector. The stopped-flow instrument was made anaerobic by overnight incubation of glucose (5 mM)/glucose oxidase (1 µM) in sodium pyrophosphate, pH 6.0. The enzyme, freshly prepared in 50 mM potassium phosphate, pH 7.5, 10% v/v glycerol, was loaded into a tonometer, which was subjected to 25 cycles of degassing by applying vacuum and flushing with argon. The syringes containing the buffer (50 mM potassium phosphate, 10% v/v glycerol, pH 7.5) or the substrate P3N prepared in water/KOH and diluted in water were flushed for 30 min with argon before mounting onto the stopped-flow spectrophotometer. To ensure complete removal of traces of oxygen, glucose (2 mM) and glucose oxidase (0.5 µM) were present in the buffer, enzyme and
substrate solutions. The concentrations of enzyme and P3N after mixing were 15 µM and 1 mM, respectively. Spectra acquisition was set every 1.5 ms.

Data analysis. The kinetic parameters for the enzymatic assays carried out at atmospheric oxygen with P3N or different nitronates were obtained from the fitting of the experimental points to the Michaelis-Menten equation for one substrate using Kaleidagraph software (Synergy Software, Reading, PA). Steady-state kinetic data were fit with Enzfitter software (Biosoft, Cambridge, U.K.) to equation 1, which represents a ternary complex mechanism in which the second substrate, i.e., oxygen, reacts with the reduced enzyme before release of the product of P3N oxidation. $k_{\text{cat}}$ is the first-order rate constant for enzyme turnover at saturating concentration of both substrates, $K_a$ and $K_b$ are the Michaelis constants for P3N and oxygen, respectively, $K_{ia}$ is a kinetic constant that accounts for the intersecting line pattern in the double reciprocal plot, $e$ is the concentration of enzyme and $v_0$ is the initial velocity.

$$\frac{v_0}{e} = \frac{k_{\text{cat}} [\text{P3N}][O_2]}{K_a [O_2] + K_b [\text{P3N}] + [\text{P3N}][O_2] + K_{ia} K_b}$$

(1)

Protein crystallization and structure determination. Recombinant, purified Pa-NMO was crystallized by the vapor diffusion hanging drop method at room temperature. The composition of the reservoir solution was 14% w/v PEG 5000 monomethylether and 0.1 M HEPES-Na at pH 7.0. A 2 µL drop of this solution mixed with 2 µL of protein solution (14 mg mL\(^{-1}\) in 50 mM TRIS-Cl pH 8.0, 100 mM NaCl, 20 mM 2-nitropropane) was equilibrated against 1000 µL reservoir solution and crystals grew in 7-10 days. Single crystals were cryo-cooled with 22% v/v glycerol as cryoprotectant. X-ray data were collected at 100 K on beamline BM-22 of the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory. The X-ray data were integrated and scaled using HKL-2000\(^{28}\).
Molecular replacement was performed with Phaser\textsuperscript{29} in the CCP4 suite of programs\textsuperscript{30}, using as initial model the main chain atoms of nitroalkane oxidase from \textit{Streptomyces ansochromogenes} (PDB 3BW2) as initial model\textsuperscript{31}. Refinement of the crystal structure was carried out with \textit{REFMAC5}\textsuperscript{32} and manual adjustment and rebuilding was performed with \textit{Coot}\textsuperscript{33}. No electron density that could be modeled with the ligand 2-nitropropane was observed. The crystal structure was deposited as PDB entry 4Q4K. Protein structures were superimposed on C\textalpha{} atoms by using SUPERPOSE of the CCP4 suite\textsuperscript{34}. Figures of the structures were generated with PYMOL (http://www.pymol.org) and CCP4mg\textsuperscript{35}. The detection of tunnels to the active site was performed with the software CAVER\textsuperscript{36}. The number of approximating balls was set at 12, the minimum probe radius was 1.0 Å, the shell depth 4 Å, the shell radius 3 Å, the clustering threshold 3.5, and the starting point was set on the N5 atom of the flavin cofactor with a maximum distance of 3 Å and a desired radius of 5 Å.

A BlastP\textsuperscript{37} analysis of the protein sequence of Pa-NMO was carried out against non-redundant protein sequences database. The conserved motifs were designed manually based on the BlastP multiple sequence alignment of the results of the query. The multiple alignment in Figure 7 was created with Clustal omega\textsuperscript{38} and Jalview 2.8\textsuperscript{39} and the neighbor joining tree in Figure 8 was generated with Jalview 2.8\textsuperscript{39}.

\textit{Expression of Recombinant NMO from Burkholderia phytofirmans}. The genes encoding hypothetical NMOs \textit{bphyt}\_6745 and \textit{bphyt}\_4144 from \textit{B. phytofirmans} PsJN were cloned in the vector pET21a(+) in a previous study\textsuperscript{18}. \textit{E. coli} expression strain Rosetta(DE3)pLysS harboring this plasmid was used to inoculate 100 mL of Luria-Bertani medium containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol. When the cell culture incubated with shaking at 37 °C reached an OD\textsubscript{600} of 0.6, IPTG was added to a final concentration of 50 μM and the
temperature was lowered to 18 °C. One culture without the addition of IPTG was grown in parallel with the same conditions for both *bphyt_6745* and *bphyt_4144* as a control. After 19 h, the cells were harvested by centrifugation recovering a wet cell paste of 1 g, which was resuspended and lysed with the commercial solution BugBuster Master Mix from Novagen (La Jolla, CA). After incubation for 25 min at 25 °C the cell free extract was obtained by centrifugation at 12,000 g for 20 min and the enzymatic activity was tested at 30 °C and pH 7.5 with P3N, nitroethane, 1-nitropropane, or 3NPA as substrate as described above.

### 6.5 Results

**Purification of Recombinant Pa-NMO.** Pa-NMO was expressed in the *E. coli* strain Rosetta(DE3)pLysS and purified to high level as judged from SDS-PAGE analysis (Figure 6.1). The purified enzyme showed a specific activity of 542 µmol O₂ min⁻¹ mg⁻¹ with 1 mM P3N as substrate at pH 7.5 and 30 °C in atmospheric oxygen (i.e., 0.23 mM).

![Figure 6.1: SDS-PAGE analysis of recombinant purified Pa-NMO.](image)

The first lane shows the wide-range molecular weight markers and the second lane the purified Pa-NMO (calculated molecular weight from protein sequence 37 kDa).

**Cofactor Content.** The UV-visible absorption spectrum of purified Pa-NMO showed maxima at 370 nm and 443 nm, which are characteristic of flavin-containing enzymes (Figure 6.2).
Figure 6.2: Anaerobic reduction of purified Pa-NMO with 1 mM P3N.
The black curve represents the UV-visible absorption spectrum of the enzyme with peaks at 370 and 443 nm; the blue curve is the same sample immediately after anaerobic mixing with 1 mM P3N. Inset: mass spectrometric analysis of the extracted cofactor of purified Pa-NMO.

The identification of the cofactor after extraction from the enzyme was carried out by ESI mass spectrometric analysis in negative ion mode (Figure 6.2) indicating that FMN is present in the enzyme. The FMN cofactor is non-covalently bound to the protein as indicated by its complete release to bulk solvent after heat denaturation of the purified enzyme. An extinction coefficient of 12,500 M$^{-1}$cm$^{-1}$ was calculated for the enzyme-bound FMN, consistent with an FMN/enzyme stoichiometry of ~0.4.$^{(21)}$

Substrate Specificity. The enzymatic activity of purified Pa-NMO was measured with various nitronates or nitroalkanes by monitoring oxygen consumption at pH 7.5 and 30 °C. Oxygen consumption was seen with nitronates, but not with any of the nitroalkanes tested (i.e., 3NPA, nitroethane, 1-nitropropane, 1-nitrobutane, 1-nitropentane or 2-nitropropane), as illustrated by the example of 3NPA (Figure 6.3).
Figure 6.3: Oxygen consumption during turnover of purified Pa-NMO with 1 mM 3NPA or 1 mM P3N.

The best substrate, as indicated by the highest \( \text{app} k_{\text{cat}} \) and \( \text{app}(k_{\text{cat}}/K_m) \) values and the lowest \( \text{app} K_m \) value, was P3N (Table 6.1). Pa-NMO was able to oxidize primary nitronates with 2 to 5 carbons and the secondary propyl-2-nitronate, although with \( \text{app}(k_{\text{cat}}/K_m) \) values that were 60- to 400-times lower than with P3N (Table 6.1).

Table 6.1: Substrate specificity of Pa-NMO in 50 mM potassium phosphate, pH 7.5, 30 °C, and atmospheric oxygen (i.e., 0.23 mM).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>( \text{app}(k_{\text{cat}}/K_m) ), M(^{-1}) s(^{-1})</th>
<th>( \text{app} k_{\text{cat}} ), s(^{-1})</th>
<th>( \text{app} K_m ), μM</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3N</td>
<td>((1.2 \pm 0.2) \times 10^7)</td>
<td>(1300 \pm 70)</td>
<td>(110 \pm 20)</td>
</tr>
<tr>
<td>ethynitronate</td>
<td>((7 \pm 1) \times 10^4)</td>
<td>(350 \pm 30)</td>
<td>(5,000 \pm 1,000)</td>
</tr>
<tr>
<td>propyl-1-nitronate</td>
<td>((2.0 \pm 0.1) \times 10^5)</td>
<td>(1120 \pm 20)</td>
<td>(6,000 \pm 1,000)</td>
</tr>
<tr>
<td>propyl-2-nitronate</td>
<td>((2.9 \pm 0.1) \times 10^4)</td>
<td>nd (^a)</td>
<td>nd</td>
</tr>
<tr>
<td>butyl-1-nitronate</td>
<td>((3.3 \pm 0.1) \times 10^4)</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>pentyl-1-nitronate</td>
<td>((9.1 \pm 0.1) \times 10^4)</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>

\(^a\) nd, not determined as it was not possible to saturate the enzyme with substrate as high as 20 mM.

Ethynitronate and propyl-1-nitronate had \( \text{app} K_m \) values of ~5 mM, which were 50-fold larger than with P3N, whereas propyl-2-nitronate, butyl-1-nitronate and pentyl-1-nitronate could not saturate the enzyme up to 20 mM, consistent with even larger \( \text{app} K_m \) values (Table 6.1).
Steady-state kinetic mechanism. The steady-state mechanism of Pa-NMO was determined at 30 °C and pH 7.5 by varying the concentrations of both P3N and oxygen. Figure 6.4 panel A reports the plot of the initial rates of reaction versus P3N concentration at different oxygen concentration while Figure 6.4 panel B shows the double reciprocal plot of the initial rates of reaction and P3N concentration.

Figure 6.4: Steady-state kinetics of Pa-NMO with P3N as substrate at 30 °C and pH 7.5. Panel A: plot of the initial rates of reaction versus [P3N] at oxygen concentration of 15 µM (empty circle), 34 µM (solid square), 87 µM (solid triangle), and 156 µM (solid circle). Panel B: double reciprocal plot of the initial rates
of reaction versus [P3N] at oxygen concentration of 15 µM (empty circle), 34 µM (solid square), 87 µM (solid triangle), and 156 µM (solid circle). The lines represent the linear fit of the data to eq. 1.

The different slopes and y-intercept in the double reciprocal plot are consistent with a ternary complex mechanism, in which the second substrate oxygen reacts before the release of the first reaction product from the enzyme. The equation describing a ternary complex mechanism is also the best fit to the data ($R^2$ of 0.99) compared to other equations describing ping-pong or rapid equilibrium mechanisms (with $R^2 \leq 0.98$). The steady-state kinetic parameters derived from the fit of the data to equation 1 are: a turnover number $k_{cat}$ of 1450 (±10) s$^{-1}$, $K_{ma}$ of 60 (±3) µM, and Michaelis constants $K_a$ for P3N of 134 (±3) µM and $K_b$ for oxygen of 75 (±1) µM. The second-order rate constants $k_{cat}/K_m$ were calculated as $10.8 \times 10^6$ (±0.2 x10$^6$) M$^{-1}$s$^{-1}$ for P3N and $19.0 \times 10^6$ (±0.3 x10$^6$) M$^{-1}$s$^{-1}$ for oxygen.

**Anaerobic Flavin Reduction with P3N.** Anaerobic reduction of Pa-NMO with P3N was carried out in a stopped-flow spectrophotometer equipped with photo-diode array detection in 50 mM potassium phosphate, 10% v/v glycerol, pH 7.5 and 30 °C. Within 7.5 ms after mixing, the enzyme-bound FMN was fully converted to the anionic flavosemiquinone, with well-resolved maxima at 364 nm, 401 nm and 486 nm (Figure 6.2). An extinction coefficient of 15,080 M$^{-1}$cm$^{-1}$ was determined at 364 nm for the enzyme-bound anionic flavosemiquinone, in agreement with values typically observed for other anionic flavosemiquinones. Despite the molar excess of P3N the flavosemiquinone persisted anaerobically for at least 20 min, indicating that the 2-electron reduced hydroquinone or an N5-flavin adduct of the type seen with nitroalkane oxidase are not reaction intermediates in the normal catalytic pathway of Pa-NMO.

**Structure of Pa-NMO.** The crystal structure of Pa-NMO was solved by molecular replacement in the space group $P2_1$ and refined to 1.44 Å resolution with an $R$ factor of 20.3. The crystallographic data and refinement statistics are presented in Table 6.2.
Table 6.2: X-ray diffraction data collection and model refinement statistics.

\(^a\) Values in parenthesis are given for the highest resolution shell; \(^b\) \(R_{\text{merge}} = \sum_{hkl} \left| I_{hkl} - \langle I_{hkl} \rangle \right| / \sum_{hkl} I_{hkl} ; \)
\(^c\) \(R_{\text{work}} = \sum \left| F_{\text{obs}} - F_{\text{cal}} \right| / \sum_{\text{obs}} F_{\text{obs}} ; \)
\(^d\) \(R_{\text{free}} = \sum_{\text{test}} \left( \left| F_{\text{obs}} \right| - \left| F_{\text{cal}} \right| \right)^2 / \sum_{\text{test}} \left| F_{\text{obs}} \right|^2 .\)

<table>
<thead>
<tr>
<th><strong>Data Collection</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystal class and space group</td>
<td>monoclinic ( P 2_1 )</td>
</tr>
<tr>
<td>Number of molecules per asymmetric unit</td>
<td>2</td>
</tr>
<tr>
<td>Unit-cell parameters (Å, deg)</td>
<td>( a=70.12, b=54.52, c=88.51, \beta=96.01 )</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>37.6-1.44 (1.48-1.44) (^a)</td>
</tr>
<tr>
<td>Total observations</td>
<td>1193755</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>123114</td>
</tr>
<tr>
<td>( R_{\text{merge}} ) (%)</td>
<td>5.1 (41.5)</td>
</tr>
<tr>
<td>(&lt;I/\sigma(I)&gt;)</td>
<td>26.6 (3.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.9 (75.5)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.5 (3.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Refinement</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of reflections ( R_{\text{work}}/R_{\text{free}} )</td>
<td>108415/5780</td>
</tr>
<tr>
<td>( R_{\text{work}}/R_{\text{free}} ) (%)</td>
<td>20.3/24.0</td>
</tr>
<tr>
<td>Number of atoms</td>
<td>5911</td>
</tr>
<tr>
<td>Number of solvent molecules</td>
<td>593</td>
</tr>
<tr>
<td>Isotropic B factors (Å(^2))</td>
<td>25.85</td>
</tr>
<tr>
<td>Protein (main chain)</td>
<td>23.11</td>
</tr>
<tr>
<td>Protein (side chain)</td>
<td>27.44</td>
</tr>
<tr>
<td>FAD</td>
<td>23.33</td>
</tr>
<tr>
<td>Solvents</td>
<td>32.77</td>
</tr>
<tr>
<td>r.m.s. deviation</td>
<td></td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.017</td>
</tr>
<tr>
<td>Angle (deg)</td>
<td>1.99</td>
</tr>
</tbody>
</table>

Ramachandran plot results

| Residues in most favoured regions (%) | 540 (93.3%) |
| Residues in additional allowed regions (%) | 37 (6.4%) |
| Residues in generously allowed regions (%) | 2 (0.3%) |
| Residues in disallowed regions (%) | 0 (0.0%) |
The protein crystallized with a homodimer in the asymmetric unit (Figure 6.5A) with each monomer consisting of 351 well-defined residues. No electron density was observed for the histidine tag at the C terminus. Each monomer consists of an FMN-binding domain (residues 1-71, 112-249 and 333-351) and a substrate-binding domain (residues 72-111, and 250-332) as highlighted in Figure 6.5.

Figure 6.5: Overall structure of Pa-NMO. The dimeric form is shown as red and dark gray cartoon for chain A and B, respectively, with the FMN cofactor in yellow sticks (panel A). Panel B: the monomer of Pa-NMO with the substrate-binding domain highlighted as blue cartoon and the FMN-binding domain highlighted as cyan cartoon with the FMN cofactor in yellow sticks.

The N terminal and C terminal residues are positioned on the same side of the FMN binding domain facing each other at 16.3 Å and 18.7 Å distance in chain A and B respectively. The dimer interface includes eight contacts mainly in the FMN-binding domain and it does not seem to be directly relevant for catalysis as it is far from the active site pocket. Six contacts, with the distance in Å between subunit A and B shown in parentheses, are: Gln328-Gly197 (3.0-3.0), Asp221-Arg203 (2.8-2.8), and Glu230-Arg223 (2.9-2.9). Two additional contacts are established between the side chains of Arg226 of chain A and Glu230 of chain B (2.9 Å) and between
Glu160 of chain A and Arg265 chain B (2.9 Å). The FMN-binding domain shows a TIM barrel fold, which is common in flavin-dependent enzymes containing FMN, while the substrate-binding domain is composed of two parallel α-helices connected by loops (Figure 6.5). A DALI (42) search of the Protein Data Bank database highlights how the overall folding resembles the structure of nitroalkane oxidase from *Streptomyces ansochromogenes* (PDB 3BW4) (43) with a Z score of 46.6. The crystal structure of the gene product PA1024 from *P. aeruginosa* PAO1 (PDB 2GJL) (20) is listed at tenth position of the DALI (42) search, with a Z score of 34.1.

**Active site of Pa-NMO.** The FMN cofactor, which presents a slight V-shaped bend on the N5-N10 axis, is positioned in the FMN-binding domain with its *si* face oriented towards the active site pocket (Figure 6.6 panel A).

The active site is located at the cleft between the FMN-binding domain and the substrate-binding domain and shows a constellation of mainly aromatic residues, including Phe71, Tyr109, Phe134, Tyr299, Tyr303, Trp325 (Figure 6.6 panel B). The closest contacts of the protein with the isoalloxazine ring of FMN are Asn69 (3.0 Å from N3 atom of the isoalloxazine ring) and His133 (4.1 Å from N1 atom of FMN), and on the *re* face the Cα of Met20 (3.3 Å from N5 atom of the cofactor) (Figure 6.6 panel B). The side chain of His183, which is conserved among all the amino acid sequences currently annotated as NMO, is located 5.7 Å from the N5 atom of the isoalloxazine ring. The active site is open to solvent access as shown from the view of the electrostatic surface in Figure 6.6 panel C and a tunnel to the active site 5.4 Å long and of 1.4 Å radius at the bottleneck has been computed with CAVER (36). The entrance of the tunnel is delimited by the side chains of Leu21, Gly22, Trp335, Tyr303, Lys307 and Tyr109 (Figure 6.6 panel D).
Figure 6.6: The active site of Pa-NMO from different views. Panel A: the FMN cofactor is shown as yellow sticks in two different orientations. The omit map contoured at 0.5 e/A³ is shown in blue. Panel B: FMN cofactor is represented as yellow sticks and the side chains of active site residues as green sticks. Panel C: one different view of the active site is shown as electrostatic surface (T/e) with blue indicating areas of positive electrostatic potential, red areas of negative electrostatic potential and white areas of neutral electrostatic potential. FMN is represented as yellow sticks and the side chain of Lys307 as orange sticks. Panel D: the same view of the active site of panel C is shown with the residues lining the entrance to the active site highlighted as green sticks and the tunnel to the active site computed by CAVER is shown as gray surface.

Bioinformatics analysis of Pa-NMO. Most of the residues highlighted by the analysis of the crystal structure of Pa-NMO, such as Met20, Asn69, Phe71, Tyr109, His133, His183, Tyr299,
Tyr303 and Lys307, are conserved in the protein sequence of NMO (Figure 6.7 and 6.8) from C. saturnus (25) and in pnoA from Pseudomonas sp. JS189 (18), for which the physiological role was established to be the metabolism of P3N (18).

Figure 6.7: Multiple sequence alignment of protein sequences annotated as NMO enzymes from both prokaryotic and eukaryotic sources. NMO Class I is represented by sequences 1 to 7 including the biochemically characterized Pa-NMO (i.e., PA4202) and Cs-NMO (C. saturnus) and the P3N-NMO from Pseudomonas sp. JS189 (Ps JS189). The numbering of the residues is according to Pa-NMO protein sequence. Motifs I to IV identified are boxed in yellow and the conserved active site residues identified in the crystal structure of Pa-NMO are marked with a red star. Sequences 8, 9, and...
belong to PA0660, PA1024 and Nc-NMO from *N. crassa*, respectively. The sequence identifier used is: PA4202, *P. aeruginosa* PAO1 (NP_252891.1); Ps.JS189, *Pseudomonas* sp. JS189 (ACX83564.1); C.sarturnus, *Cyberlindnera sarturnus* (AAA64484.1); K.pneumoniae, *Klebsiella pneumoniae* (WP_004179795.1); C.capitata, *Ceratitis capitata* (XP_004532609.1); 6Z.rouxii, *Zygosaccharomyces rouxii* (XP_002498653.1); P.hodgsonii, *Pantholops hodgsonii* (XP_005969806.1); PA0660, *P. aeruginosa* PAO1 (NP_249351.1); PA1024, *P. aeruginosa* PAO1 (NP_249715.1); N.crassa, *Neurospora crassa* OR74A (XP_957588.1).

Figure 6.8: Neighbor joining tree based on percentage of identity of enzymes annotated as NMO from both prokaryotic and eukaryotic sources. The tree was generated with Jalview 2.8. Sequence identifier, organism and accession codes are the same as in Figure 6.7.

This observation led to the identification of four conserved motifs (Table 6.3 and Figure 6.9). A protein BLAST search was performed and 475 sequences annotated as hypothetical NMO belonging to bacteria, fungi and two animals, i.e, *Ceratitis capitata* and *Pantholops hodgsonii*, possess all of the four conserved motifs identified in this study.

Table 6.3: Conserved motifs in the NMO enzyme family.
The numbering is for Pa-NMO; the brackets identify residues that can alternatively be present in that position while the X represents a position where any amino acid is accepted.

<table>
<thead>
<tr>
<th>Motif</th>
<th>Consensus sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>\textsuperscript{14}P-I-X-Q-A-P-M-X-G-X-S-T-X-X-L-A-A\textsuperscript{30}</td>
</tr>
<tr>
<td>II</td>
<td>\textsuperscript{130}[V/I]-S-F-H-F-[G/N]-X-P\textsuperscript{137}</td>
</tr>
<tr>
<td>III</td>
<td>\textsuperscript{174}[V/I]-X-Q-G-X-E-A-G-G-H-R-G-X-F\textsuperscript{187}</td>
</tr>
<tr>
<td>IV</td>
<td>\textsuperscript{297}[P/A]-[D/E/P]-Y-P-X-Y-D-X-X-K-X-L\textsuperscript{309}</td>
</tr>
</tbody>
</table>
Figure 6.9: Representation of the four conserved motifs identified in this study in the crystal structure of Pa-NMO. The overall structure of one monomer of Pa-NMO is shown as gray cartoon and FMN cofactor is represented as yellow sticks. Motifs I, II, III, and IV are highlighted in magenta, green, blue, and orange respectively.

Motif I, which contains the conserved Met20 that makes contact with the re face of FMN, provides protein main chain atoms that extend from the re face of the FMN to the entrance to the active site. The conserved active site residue His133 is present in Motif II, which is located near the N1 atom of the FMN cofactor. Motif III represents a β strand followed by a loop in the active site of the enzyme and contains the fully conserved His183. Motif IV identifies an α helix delimiting the entrance to the active site and carrying the side chains of the conserved residues Tyr299, Tyr303, and Lys307. The protein sequences of NMO from C. saturnus and pnoA from Pseudomonas sp. JS189, which were previously established to be involved in P3N detoxification, carry the motifs identified in this study. We selected two additional hypothetical
NMOs carrying motifs I to IV, namely the gene products *bphyt_4144* and *bphyt_6745* from *B. phytofirmans* PsJN to experimentally demonstrate that the genes identified based on the bioinformatics analysis were able to oxidize P3N. The cell free extracts of the recombinant proteins expressed in *E. coli* have a specific activity with P3N of 12 U/mg and 5.5 U/mg, respectively. In contrast no oxygen consumption was detected with 1 mM nitroethane, 1-nitropropane, or 3NPA. The cell free extract of the control samples without addition of IPTG showed no activity with P3N. For comparison, the cell free extract of Pa-NMO obtained using the same procedure had a specific activity of 27 U/mg with P3N. The motifs I to IV identified in the sequence of Pa-NMO are not present in the sequence of PA0660 and PA1024, the other two putative NMOs in *P. aeruginosa* PAO1.

### 6.6 Discussion

The product of gene *pa4202* of *P. aeruginosa* PAO1 has been recombinantly expressed in *E. coli*, purified to high level, and characterized in its biochemical, kinetic and structural properties. This approach has established the first biochemical and kinetic characterization of a bacterial NMO and the first determination of the three dimensional structure of NMO using X-ray crystallography. The concomitant availability of biochemical and structural information on Pa-NMO has allowed us to identify conserved motifs that define a new class of NMO enzymes. These NMOs are primarily present in bacteria and fungi, but are also found in two animals.

**The *pa4202* gene encodes for NMO.** Pa-NMO is a flavoprotein and contains FMN non-covalently bound, as indicated by the mass spectrometric analysis of the extracted cofactor showing an m/z ratio of 455.1 in negative ion mode. No peaks were seen in the 600-1000 m/z region of the mass spectrum, consistent with absence of FAD in the enzyme. The non-covalent binding of the flavin to the protein was further confirmed by X-ray crystallography. These
findings are in keeping with all the biochemical data available on various NMOs, showing that they all utilize FMN, but not FAD, as cofactor \(^{(3, 17, 24, 25)}\).

The best substrate of Pa-NMO is P3N with steady-state kinetic parameters determined in atmospheric oxygen that compare well with values previously determined for the two fungal NMOs from \textit{N. crassa} and \textit{C. saturnus} previously characterized, with \( \text{app}(k_{\text{cat}}/K_m) \) and \( \text{app}k_{\text{cat}} \) values \( \geq 10^6 \text{ M}^{-1}\text{s}^{-1} \) and \( \geq 450 \text{ s}^{-1} \), respectively, at pH 7.5 and 30 °C \(^{(15)}\). The steady-state kinetic mechanism of Pa-NMO was established to be a ternary complex mechanism, similarly to \textit{C. saturnus} NMO \(^{(25)}\). As for the case of the \textit{C. saturnus} enzyme, bacterial Pa-NMO can also effectively oxidize other primary and secondary nitronates, although with \( \text{app}(k_{\text{cat}}/K_m) \) values 60-400 times lower than P3N, and cannot oxidize nitroalkanes \(^{(24)}\).

The oxidation of P3N by Pa-NMO is extremely fast and results in the formation of an enzyme-bound flavosemiquinone. Evidence for this conclusion comes from the anaerobic mixing of the enzyme with P3N in a stopped-flow spectrophotometer, immediately yielding within 7.5 ms a flavin species with maxima at 364 nm, 401 nm and 480 nm that are typical of anionic flavosemiquinones. This was previously observed with Cs-NMO, for which formation of an anionic flavosemiquinone was also too fast to be monitored in a stopped-flow spectrophotometer \(^{(25)}\).

**FMN-binding site.** The crystal structure of Pa-NMO was solved to 1.44 Å resolution and is the first structural analysis of an NMO enzyme, as no crystal structure is available for the well-characterized fungal enzymes from \textit{N. crassa} and \textit{C. saturnus} \(^{(25, 44)}\). The FMN-binding domain displays a TIM barrel fold, which is found in other 15 FMN-dependent proteins \(^{(45)}\). The phosphate group of the FMN is deeply buried in this domain and interacts via hydrogen bonds
with the N atoms of the main chains of Gly181, Gly218, Gly239, and with the hydroxyl group of Thr240 (Scheme 6.1).

![Scheme 6.1: Schematic representation of the interactions established by the FMN cofactor with Pa-NMO. Dashed lines represent hydrogen bonds. For clarity, H atoms and protein double bonds are not shown.](image)

In a similar way the ribityl moiety is held in place only by non-ionic interactions with the main and side chain atoms of Gln176, Gly181, and Thr240. The positioning of the ribityl and phosphate moieties of the cofactor via non-ionic interactions is conserved in 5 FMN-containing TIM barrel proteins, namely dihydroorotate dehydrogenase (PDB code 1DOR), glutamate synthase (1LM1), PA1024 (2GJL), isopenthenyl-diphosphate Δ-isomerase (1P0N), and dihydropyrimidine dehydrogenase (1H7W). Remarkably, in the other 10 FMN-containing TIM barrel proteins the binding of the phosphate group of the cofactor involves one or two salt
bridges with arginine residues. In the case of mandelate dehydrogenase (1HUV), glycolate oxidase (1GOX), lactate oxidase (2DU2), and flavocytochrome b2 (1FCB) the phosphate group establishes ionic bonds with the side chains of two arginine residues, while in the case of Old Yellow Enzyme (1OYA), tRNA dihydouridine synthase 2 (3B0U), histamine dehydrogenase (3K30), trymethylamine dehydrogenase (2TMD), 12-oxophytodienoate reductase (1ICS), and 2,4-dienoyl-CoA-reductase (1PS9) only one arginine side chain contacts the phosphate group.

The isoalloxazine ring of the cofactor exposes its si face to the active site. On the re face, the isoalloxazine ring makes contact with the main chain Cα atom of Met20. Interestingly, the N5 atom of FMN is not in contact with any residue of the protein, and is involved only in a hydrogen bond with the solvent, as already observed in lactate oxidase (PDB 2DU2). This is different from the other 14 FMN-binding TIM barrel proteins, where the N5 atom is located around 3 Å from either amide main chain atoms of the protein or from the side chain atoms of a cysteine, threonine, or lysine residue. In all the 15 FMN-binding TIM barrel proteins the N1-C2 atoms of the cofactor are located 3.0 Å from an arginine or lysine side chain, which is likely positively charged. This interaction is not observed in the crystal structure of Pa-NMO, where the closest contact of the N1 atom of FMN is the side chain of His133 at a distance of 4.0 Å, which may provide a positive charge as well. A similar interaction is observed in 12-oxophytodienoate reductase, where the flavin N1 atom is located 3.0 Å from the side chain of Arg239 and 3.5 Å from the side chain of His190. The N3 atom of FMN is 3 Å from the side chain of Asn69. A similar interaction is present in the other 15 FMN binding TIM barrel proteins where the polar side chain is either from glutamine, glutamate, asparagine, threonine, or serine.

**Active site residue His133.** In proximity of the N1 atom of FMN there is His133, which may be important for the stabilization of the catalytic anionic flavosemiquinone of Pa-NMO through an
ionic interaction of its side chain with the negatively charged N1 atom of the flavin. This histidine is conserved in the fungal NMO from *C. saturnus* (25), which was previously characterized mechanistically with P3N or ethylnitronate as substrate (25). In that enzyme, pH profiles of the steady-state kinetic parameters showed that $k_{\text{cat}}/K_m$ and $k_{\text{cat}}$ were 2 orders of magnitude larger at acidic pH values than alkaline values with ethylnitronate (25), consistent with catalysis being enhanced in the presence of a positive charge in the active site. No relevant ionizations in fungal NMO were seen in the $k_{\text{cat}}/K_m$ and $k_{\text{cat}}$ pH profiles with P3N, however similar solvent viscosity effects on the steady-state kinetic parameters suggested that the same enzyme species were present in turnover with ethylnitronate and P3N (25).

**Active site residue His183.** His183 is fully conserved among all the genes currently annotated as NMO in Gen-Bank (i.e., >4,900). In the structure of Pa-NMO, His183 points to the N5 atom of the FMN, but it is not sufficiently close to establish an interaction being at almost 6 Å. The corresponding residue His196 in Nc-NMO was previously mutated to asparagine, establishing it as the base that deprotonates nitroethane to ethylnitronate for the oxidation reaction catalyzed by the enzyme (46). Such a role, however, cannot be envisioned for Pa-NMO, or for Cs-NMO, since the latter catalyze the oxidative denitrification of P3N and other nitronates, but are not active with neutral nitroalkanes (this study and (24)). As illustrated below, Motif III is the only one present in Nc-NMO among the four identified in this study in Pa-NMO, suggesting that the two enzymes may have different active site topologies and (slightly) different catalytic strategies. Nonetheless, the reaction pathway for oxidative denitrification of ethylnitronate in Nc-NMO was unaltered upon replacing His196 with asparagine (46), suggesting that the fully conserved histidine is not required for the formation of the catalytically obligatory flavosemiquinone.
possible role for His183 in Pa-NMO may be in facilitating binding the P3N substrate, which has a double negative charge, through either electrostatic or hydrogen bond interactions.

**Other conserved active site residues.** Tyr109, Phe134, Tyr299 and Tyr303 are also conserved in the fungal Cs-NMO and other 475 putative NMOs. These residues constitute the walls of the enzyme active site, suggesting important roles in either substrate binding or catalysis. Lys307 is conserved among the residues that delimitate the access to the active site, probably because it contributes electrostatic attraction for the doubly negatively charged substrate.

**NMO Class I.** The present study identified four motifs in the sequence of Pa-NMO that are conserved in 475 sequences of putative NMOs, including the biochemically characterized Cs-NMO, as illustrated in Figure 6.7 showing multiple sequence alignment with Clustal omega of select protein sequences with the four conserved motifs highlighted in yellow (sequences 1 to 7). We define as NMO Class I the protein sequences containing the four motifs I to IV. Most of the class I NMOs functionally annotated belong to bacteria and fungi (i.e., 473), and only 2 are from animals. The fact that class I NMOs are so widespread in bacteria and fungi suggests their importance in conserved pathways, such as the detoxification of the toxin P3N and possibly physiological roles still not determined. *C. capitata*, which is commonly known as the Mediterranean fruit fly or medfly, is considered one of the most economically damaging agricultural pests due to its herbivorous diet, genetic variability and ability to invade different territories. The Tibetan antelope, *P. hodgsonii*, has instead adapted to highly inhospitable environments that are characterized by low partial pressure of oxygen and high ultraviolet radiation. In the Tibetan region, plants of the genus *Astragalus*, which is extremely rich in P3N, are known to become widespread during the frequent droughts. It is attractive to speculate that both these organisms may have acquired the NMO function as an adaptation to the
presence of the P3N toxin in the diet. A similar case was previously described for the
detoxification of pyrrolizidine alkaloids in the host plants by the flavin-dependent
monooxygenase (FMO) of the insect *Tyria jacobaea* \(^{50}\).

Interestingly, only the equivalent of Met20, Asn69, and His183 in Pa-NMO are conserved in the
other hypothetical NMOs of *P. aeruginosa* PAO1 coded by genes *pa0660* and *pa1024*, while the
motifs I to IV identified in this study are remarkably absent. A BLAST protein search
highlighted more than 500 sequences of hypothetical NMOs from bacteria and fungi
classified by the PA1024 motifs identified in the structural study by Ha *et al.* \(^{20}\). We suggest
that the protein sequences displaying the motifs identified in PA1024, which at this stage is not
confirmed to possess NMO function, should be grouped in a class different from that of Pa-
NMO. Indeed, no kinetic parameters or specific activity with 2-nitropropane or other
nitroalkanes or nitronates are available for PA1024 \(^{20}\). Thus, it is a priority to kinetically
characterize PA1024 and assess if this enzyme possesses characteristic NMO features, namely a

\[ k_{\text{cat}}/K_m \]

value for the physiological substrate P3N of \(~10^6\ \text{M}^{-1}\ \text{s}^{-1}\), the presence of FMN as
cofactor, and a flavosemiquinone catalytic intermediate \(^3\).

**NMO Class II.** Nc-NMO from *Neurospora crassa* contains only a few conserved residues
identified in NMO Class I, such as Met20, Asn69 and His183, and only parts of motifs I and III
(Figure 6.7). For this reason, we classify this enzyme as a member of NMO Class II, which
defines a much smaller group than Class I with only 10 hypothetical NMOs from fungi other
than Nc-NMO, such as from *Neurospora Tetrasperma* FGSC 2508 (Gen-Bank: EGO54545.1) or
from *Togninia minima* UCRPA7 (GenBank: EOO00344.1). Biochemical characterization of Nc-
NMO has indeed established the enzyme to be highly specific for nitronates, with a marked
preference for P3N in regard to \(k_{\text{cat}}/K_m\) and \(k_{\text{cat}}\) values \(^{44}\). However, this enzyme has the unique
ability to be able to utilize nitroalkanes as substrates, setting it aside from both bacterial Pa-NMO and fungal Cs-NMO for which only nitronates can be used as substrates (this study and (24)).

In summary, we have characterized biochemically, kinetically and structurally the protein encoded by gene *pa4202* of *P. aeruginosa* PAO1 and demonstrated that it is an NMO. The enzyme contains non-covalently associated FMN, has a TIM barrel fold, marked preference for the toxin P3N as substrate compared to other aliphatic nitronates, is not active on nitroalkanes, and is reduced to the flavosemiquinone upon anaerobic incubation with the substrate. This is the first instance in which a bacterial NMO has been characterized in its biochemical properties and in which the crystallographic structure of an NMO has been reported. The structural-functional approach has allowed us to establish rigorously a functional annotation of NMO genes thereby overcoming the inaccurate prediction of function based solely on amino acid sequence similarities. More than 450 putative NMOs from bacteria, fungi and two animals are grouped in NMO Class I, which contains four consensus motifs identified in Pa-NMO. Lacking the consensus motifs of Pa-NMO, the only other NMO for which biochemical evidence of function is available, Nc-NMO, is grouped with few other proteins in NMO Class II. A considerable number of annotated proteins in various microorganisms, i.e., >500, shares the consensus sequences identified in PA1024 by Ha et al. (20). However, there is no biochemical or kinetic evidence yet that the enzyme is an NMO. Furthermore no experimental data are available for PA0660, the third putative NMO of *P. aeruginosa* PAO1. Thus, it is imperative to kinetically characterize these proteins to establish whether they are NMO or have different functions. Future studies should also address the physiology of NMO in *P. aeruginosa*, to establish whether the enzyme is used to detoxify the toxin P3N or in other processes such as for example virulence or biofilm formation, besides as a growth substrate. Interestingly, Pa-NMO is conserved in all 12 *P.
aeruginosa genomes available in the Pseudomonas genome database (www.pseudomonas.com (51)) and even in all fluorescent pseudomonads (46 orthologues in total) (51), suggesting an important biological function. The availability of the structure of NMO sets the stage for future investigations by using site-directed mutagenesis aimed at structural-functional analyses of this unusual Class H flavin-dependent monooxygenase that can address fundamental questions on how oxygen is incorporated in the product of P3N oxidation, or what structural and biochemical determinants allow for the stabilization of a catalytic flavosemiquinone in catalysis.

6.7 Acknowledgements

X-ray data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline BM-22 at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. We thank Shirley Nishino for the construct of PA4202 and Chung-Dar Lu, Crystal Smitherman and Jacob Ball for critically reading the manuscript. This work was supported in part by Grant MCB-1121695 from the NSF (G.G.), a Molecular Basis of Disease Fellowship from Georgia State University (H.Y.), and COMBREX via the NIGMS 1RC2GM092602-01 (G.G. and J.C.S.).

6.8 References


12. Mathews, F. P. (1940) *Poisoning in sheep and goats by sacahuiste (Nolina texana) buds and blooms*, Texas Agricultural Experiment Station.


CHAPTER VII: PA1024 FROM Pseudomonas aeruginosa PAO1 IS A NADH:FERRICYANIDE REDUCTASE AND NOT A NITRONATE MONOOXYGENASE

(The author carried out all the experiments described in this Chapter with the exceptions of the determination of the extinction coefficient, and the steady-state kinetics, which were carried out by Jacob Ball)

7.1 Abbreviations

NMO, nitronate monooxygenase; Pa-NMO, nitronate monooxygenase from Pseudomonas aeruginosa PAO1; P3N, propionate 3-nitronate; 3-NPA, 3-nitropropionate; IPTG, isopropyl-1-thio-β-D-galactopyranoside.

7.2 Abstract

The protein PA1024 from Pseudomonas aeruginosa PAO1 is currently classified as 2-nitropropane dioxygenase, previous name for nitronate monooxygenase in the GenBank and PDB databases, but the enzyme was not kinetically characterized. In this study the enzymatic activity of PA1024 was investigated. Purified PA1024 did not exhibit nitronate monooxygenase activity. It displayed instead NADH:ferricyanide reductase activity, while it was not able to use NADPH as electron donor. The reductive half reaction with NADH showed a $k_{\text{red}}$ value of 18 s$^{-1}$ and a $K_d$ value estimated in the low µM range. The best artificial electron acceptor was potassium ferricyanide, with $^{\text{app}}k_{\text{cat}}$ of 20 s$^{-1}$ and $^{\text{app}}K_m$ lower than 75 µM. PA1024 was not able to reduce the azo dye methyl red, routinely used in the kinetic characterization of azoreductases. Finally, seven conserved motifs identified in the protein sequence of PA1024 are present in more than 1000 hypothetical proteins in the GenBank, currently annotated as nitronate monooxygenases.
7.3 Introduction

The discrepancy between the rapid increase in the number of sequenced genomes of prokaryotes and the slower experimental determination of protein function has resulted in the presence of a vast number of hypothetical proteins in the databases, with gene function prediction not always reliable \(^1,^2\). One case of enzyme family consisting mainly of hypothetical proteins is represented by nitronate monooxygenases (NMOs, E.C. 1.13.12.16), which include >5000 genes in the GenBank. NMOs are FMN-dependent enzymes that catalyze the detoxification of the metabolic poison propionate 3-nitronate (P3N), produced by plants and fungi as a defense mechanism against herbivores \(^3\). The kinetic mechanism of NMO has been investigated by previous studies on the fungal enzymes from *Neurospora crassa* \(^4\) and *Cyberlindnera saturnus* (previously known as *Williopsis saturnus*) \(^5\). The recent structural and kinetic characterization of the gene product PA4202 from *Pseudomonas aeruginosa* PAO1 as PaNMO identified four motifs that establish Class I NMO, with 500 sequences from bacteria, fungi, one insect and one animal \(^6\). *P. aeruginosa* PAO1 possesses two other genes coding for hypothetical NMOs, namely *pa0660* and *pa1024*, which do not carry the four motifs characteristic of Class I NMO. While there is no experimental evidence at transcript or protein level for the gene product PA0660, a crystal structure for the hypothetical protein PA1024 is available for the free enzyme (Figure 7.1) and in complex with 2-nitropropane (PDB codes 2GJL and 2GJN) at 2.0 and 2.3 Å resolution \(^7\).
The protein PA1024 was classified as 2-nitropropane dioxygenase, previous name for NMOs, based on the gene function prediction and on a qualitative enzymatic assay performed with 20 mM 2-nitropropane at pH 6.5 (7). However, the kinetic parameters with 2-nitropropane were not determined and the physiological substrate P3N not tested, because unknown at the time of the study. The structural characterization of PA1024 highlighted six motifs conserved in other hypothetical nitronate monooxygenases (7), which are different from the four motifs described for Class 1 NMO (6).

In this study we cloned from genomic DNA the gene *pa1024*, purified the His-tagged recombinant PA1024 and kinetically characterized it. We demonstrated that purified PA1024 does not exhibit NMO activity and it is instead an NADH:ferricyanide reductase (E.C. 1.6.99.-). Furthermore we reanalyzed the conserved motifs identified by Ha et al. in the protein sequence...
of PA1024\(^{(7)}\), identified a seventh conserved motif and showed that they are present in more than 1000 sequences in the non-redundant protein database.

### 7.4 Experimental procedures

**Materials.** The enzymes *Xho*I, *Nde*I, *Dpn*I, calf intestinal alkaline phosphatase (CIP), and T4 DNA ligase were from New England BioLabs (Ipswich, MA), *Pfu* DNA polymerase from Stratagene (La Jolla, CA), and oligonucleotides from Sigma Genosys (The Woodlands, TX). *Escherichia coli* strain DH5α was purchased from Invitrogen Life Technologies (Grand Island, NY). *E. coli* strain Rosetta(DE3)pLysS and the expression vector pET20b(+) were from Novagen (Madison, WI), QIAprep Spin Miniprep kit, QIAquick PCR purification kit and QIAquick gel extraction kit were from Qiagen (Valencia, CA). The genomic DNA of *P. aeruginosa* PAO1 was a kind gift from Dr. Jim Spain, Georgia Institute of Technology, Atlanta. HiTrapTM Chelating HP 5 mL affinity column was from GE Healthcare (Piscataway, NJ), isopropyl-1-thio-β-D-galactopyranoside (IPTG) was from Promega, (Madison, WI), and nitroalkanes were from Sigma-Aldrich (St. Louis, MO). All other reagents used were of the highest purity commercially available.

**Cloning.** The gene *pa1024* was amplified from the genomic DNA of *P. aeruginosa* PAO1 by PCR (the primers are listed in Table 7.1) in presence of 5% DMSO with an initial denaturation step at 95 °C, followed by 20 cycles of denaturation for 45 s at 95 °C, annealing for 30 s at 56 °C (the annealing temperature decreases of 0.2 °C each cycle), and extension for 3 min at 72 °C, followed by a final step at 72 °C for 10 min.
Table 7.1: Oligonucleotide primers used for PCR amplification of pa1024.
The restriction sites for NdeI and XhoI enzymes are underlined.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pa1024for</td>
<td>GGCCTTCCATATGATGGCGTGTTCAGG</td>
</tr>
<tr>
<td>pa1024rev</td>
<td>CCCTCGAGGACCCCGCGAGCA</td>
</tr>
</tbody>
</table>

The PCR product was purified by agarose gel extraction with the QIAquick gel extraction kit. The amplicon and the expression vector pET20b(+) were subjected to double digestion with NdeI and XhoI and purified with the QIAquick PCR purification kit. After a dephosphorylation step with 0.5 units of calf intestinal alkaline phosphatase for 30 min at 37 °C the dephosphorylated vector was purified with the QIAquick PCR purification kit and ligated to the insert with incubation for 15 h at 16 °C with T4 DNA ligase. A volume of 5 µL of the ligation mixtures was used to transform E. coli strain DH5α and the resulting colonies grown at 37 °C on Luria-Bertani agar plates containing 50 μg/mL ampicillin were screened for the presence of the desired insert by DNA sequencing at the Cell, Protein and DNA core facility at Georgia State University. The DNA sequencing confirmed the correct insertion of the gene in the vector pET20b(+) and the absence of undesired mutations.

Recombinant expression and purification. E. coli expression strain Rosetta(DE3)pLysS transformed with the construct pET20b(+)/pa1024 was used to inoculate 90 mL of Terrific Broth containing 50 μg/mL of ampicillin and 34 μg/mL of chloramphenicol, which was incubated at 37 °C for 15 h. One aliquote of 8 mL of this culture was used to inoculate 1 liter of Terrific Broth containing 50 μg/mL of ampicillin and 34 μg/mL of chloramphenicol, which was incubated at 37 °C until it reached an optical density at 600 nm of 0.8. IPTG was then added to a final concentration of 200 μM and the culture incubated at 18 °C for 20 h. The wet cell paste of 8.5 g,
recovered by centrifugation, was resuspended in 40 mL of lysis buffer (10 mM imidazole, 300 mM NaCl, 10% v/v glycerol, 1 mM phenylmethylsulfonyl fluoride, 5 mM MgCl₂, 2 mg/mL lysozyme, 5 μg/mL DNase, 5 μg/mL RNase, and 20 mM sodium phosphate, pH 7.4) and subjected to several cycles of sonication. The cell free extract obtained after centrifugation at 12000 g for 20 min was loaded onto a HiTrapTM Chelating HP 5 mL affinity column equilibrated with buffer A (10 mM imidazole, 300 mM NaCl, 10% v/v glycerol, and 20 mM sodium phosphate, pH 7.4). After washing with 10 column volumes of buffer A, four intermediate steps at 10 %, 20 %, 30 %, and 50 % of buffer B (buffer A + 0.5 M imidazole) were performed to remove possible contaminants. The recombinant protein PA1024 was eluted with 100 % buffer B. The purest fractions based on SDS-PAGE analysis were pooled, dialyzed against 10 mM Tris-Cl, pH 8.0, 100 mM NaCl, 10 % glycerol and stored at -20 °C.

**Enzyme assays.** UV-visible absorbance was recorded with an Agilent Technologies diode-array spectrophotometer Model HP 8453 PC equipped with a thermostated water bath. The extinction coefficient of purified PA1024 was determined by extracting the FMN cofactor by heat denaturation of the enzyme. After removing the denatured protein by centrifugation, the concentration of free FMN was determined spectroscopically by using $\varepsilon_{450} = 12,500 \text{ M}^{-1} \text{ cm}^{-1}$ \(^8\). The concentration of flavin-bound enzyme was determined by using the experimentally determined extinction coefficient $\varepsilon_{461\text{nm}}$ of 12,400 M\(^{-1}\) cm\(^{-1}\). The total protein concentration was determined using the Bradford method with bovine serum albumin as standard \(^9\). Initial rates were normalized for the flavin-bound enzyme concentration.

The NMO activity assay was performed as previously described \(^5, 6, 10, 11\), following the initial rate of oxygen consumption with a Hansatech Instruments computer-interfaced Oxy-32 oxygen-monitoring system at atmospheric oxygen and 30 °C (i.e., 0.23 mM oxygen) in 50 mM
potassium phosphate, pH 7.5. The use of the methods of the initial rate prevents the interconversion between the anionic form (nitronate) and the neutral form (nitroalkane) of the substrate. The reaction was started with the addition of the anionic form of the substrate (nitronate) to ensure that in the initial 30 s the amount of the neutral form of the substrate (nitroalkane) is negligible, since the second-order rate constants for protonation of the nitronates are in the range 15-75 M$^{-1}$s$^{-1}$ \cite{12,13}. Stock solutions of nitronates and nitroalkanes were prepared as previously described \cite{5,10}. Enzyme concentration was 180 nM and substrate concentration was 1 mM for P3N or 3-NPA, and 20 mM for 2-nitropropane, propyl-2-nitronate, nitroethane, or ethynitronate. A positive control for nitronate monooxygenase activity was performed in parallel with purified Pa-NMO to a final concentration of 1.4 nM and P3N 1 mM as previously described \cite{6}.

Reduction of the FMN cofactor by NAD(P)H was carried out anaerobically with an SF-61DX2 Hi-Tech KinetAsyst high-performance stopped-flow spectrophotometer, thermostated at 25 °C. Anaerobiosis of the instrument was obtained by overnight incubation of glucose (5 mM)/glucose oxidase (1 µM) in sodium pyrophosphate, pH 6.0. The enzyme was loaded on a desalting PD-10 column equilibrated with 20 mM potassium phosphate, pH 7.0, 200 mM NaCl, transferred in a tonometer and subjected to 20 cycles of degassing by applying vacuum and flushing with argon. The syringes containing the buffer (20 mM potassium phosphate, pH 7.0, 200 mM NaCl) or the substrate NAD(P)H diluted in buffer were flushed for 30 min with argon before mounting onto the stopped-flow spectrophotometer. To ensure complete removal of traces of oxygen, glucose (2 mM) and glucose oxidase (0.5 µM) were present in the buffer, enzyme and substrate solutions. Concentration of the substrate NAD(P)H was determined spectrophotometrically at 340 nm with
the coefficient $6,220 \text{ M}^{-1} \text{ cm}^{-1} \ (14)$. The concentration of the enzyme after mixing was 15 µM and of NAD(P)H ranged from 90 to 500 µM.

NADH oxidase activity was monitored at atmospheric oxygen in 200 mM sodium chloride, and 20 mM potassium phosphate, pH 7.0 by following the decrease in absorbance at 340 nm ($\varepsilon_{340} = 6,220 \text{ M}^{-1} \text{ s}^{-1}$) with the use of an Agilent Technologies diode-array spectrophotometer Model HP 8453 PC thermostated at 25 °C. The final concentrations of NADH and the enzyme were 100 µM and 0.14 µM, respectively. Turnover of PA1024 with NADH was monitored at atmospheric oxygen with the electron acceptors potassium ferricyanide or menadione in 200 mM sodium chloride, and 20 mM potassium phosphate, pH 7.0, at 25 °C. The concentration range of potassium ferricyanide was 75-500 µM, and that of menadione was 18-416 µM. NADH was kept at a constant, saturating concentration of 100 µM. In the case of menadione, which is dissolved in ethanol, the final ethanol concentration in all reaction mixtures was kept fixed at 1% to minimize possible effects on enzymatic activity. Turnover with menadione was monitored by following the decrease in absorbance at 343 nm, where menadione and menadiol are isosbestic, that corresponds to NADH consumption $(\varepsilon_{343} = 6,220 \text{ M}^{-1} \text{ s}^{-1}) \ (15)$. Turnover with potassium ferricyanide was monitored by following the decrease in absorbance at 420 nm due to potassium ferricyanide reduction $(\varepsilon_{420} = 1.04 \text{ M}^{-1} \text{ s}^{-1}) \ (16)$. Stock solutions of potassium ferricyanide in water were prepared fresh prior to the experiment.

The azoreductase activity of PA1024 was tested as previously described $(17, 18)$, by monitoring the reduction of methyl red at 430 nm ($\varepsilon_{430} = 23,360 \text{ M}^{-1} \text{ s}^{-1}$) in 200 mM sodium chloride, and 20 mM potassium phosphate, pH 7.0 at 25 °C. The final concentrations of NADH, methyl red and enzyme were 100 µM, 25 µM, and 0.20 µM. The control reaction run in the absence of the
enzyme was negligible, consistent with the non-enzymatic reaction between NADH and methyl red occurring only at low pH \(^{(19)}\).

**Data analysis.** The kinetic parameters for the enzymatic assays carried out at atmospheric oxygen with NADH were obtained from the fitting of the experimental points to the Michaelis-Menten equation for one substrate using KaleidaGraph software (Synergy Software, Reading, PA). Stopped-flow traces were fit with the software KinetAsyst 3 (TgK-Scientific, Bradford-on-Avon, U.K.) to Equation 1 which represents a double-exponential process in which \(A\) represents the absorbance at 461 nm at time \(t\), \(B_1\) and \(B_2\) are the amplitudes of the decrease in absorbance, \(k_{\text{obs}1}\) and \(k_{\text{obs}2}\) represent the observed rate constants for the change in absorbance, and \(C\) is an offset value accounting for the nonzero absorbance value at infinite time.

\[
A = B_1 \exp(-k_{\text{obs}1} t) + B_2 \exp(-k_{\text{obs}2} t) + C
\]

Eq. 1

Concentration dependence of the observed rate constants for flavin reduction was analyzed with Equation 2, where \(S\) represents the concentration of organic substrate, \(k_{\text{red}}\) is the rate constant for flavin reduction at saturating substrate concentration, and \(K_d\) is the apparent dissociation constant for substrate binding.

\[
k_{\text{obs}} = \frac{k_{\text{red}} S}{K_d + S}
\]

Eq. 2

**Bioinformatic analysis.** The analysis of the protein sequence of PA1024 was performed with BlastP \(^{(20)}\), selecting the non-redundant protein sequence database. The conserved motifs were designed manually based on the multiple sequence alignment generated by BlastP. The multiple sequence alignment reported in Figure 8 was created with Clustal Omega \(^{(21)}\) and Jalview 2.8 \(^{(22)}\).
7.5 Results

**Protein purification.** The gene *pa1024* was cloned from the genomic DNA of *P. aeruginosa* PAO1 in the expression vector pET20b(+), with the addition of an His-tag at the C terminus of the recombinant protein. The recombinant protein PA1024 was expressed in *E. coli* and purified to high yield by affinity chromatography. The presence of 100 mM sodium chloride in the storage buffer (10 mM TRIS-Cl, pH 8.0, 100 mM sodium chloride, and 10 % v/v glycerol) was necessary for the *in vitro* stability of purified PA1024. Figure 7.2 panel A shows the SDS-PAGE analysis of the purified protein, which estimates a purity of more than 90%. The UV-visible absorption spectrum of purified PA1024 shows maxima at 370 and 461 nm (Figure 7.2 panel B) which are consistent with the presence of FMN as cofactor, as already identified in the available crystal structure of PA1024 (PDB code 2GJL) (7).

![Figure 7.2](image)

Figure 7.2: Panel A: SDS-PAGE analysis of purified PA1024 (lane 2). Lane 1 represents the Bio-Rad broad range marker. Panel B: UV-visible absorption spectrum of the gene product PA1024 in 20 mM TRIS-Cl, pH 8.0, 100 mM sodium chloride, 10 % v/v glycerol, 25 °C.

The low energy band of the cofactor at 461 nm is significantly red shifted compared to free FMN (450 nm) and to Pa-NMO (443 nm), which is usually indicative of a more polar protein environment of the cofactor (23, 24). Indeed, in the crystal structure of PA1024 (PDB 2GJL) (7) there is a polar T75 located 3.8 Å from the N3 atom of FMN, while in the same position (3.6 Å
from N3 atom of FMN) there is the hydrophobic F71 in Pa-NMO (PDB 4Q4K) \(^{(6)}\). The flavin cofactor extracted by heat denaturation was released to the bulk solvent, confirming it is not covalently bound to the protein \(^{(7)}\). The ratio FMN/enzyme was ~0.9, consistent with a 1:1 stoichiometry per monomer of protein.

**Nitronate monooxygenase assay.** Nitronate monooxygenase activity was tested at pH 7.5, 30 °C and atmospheric oxygen \(i.e.\) 230 μM as previously described \(^{(5,6,10)}\). No enzymatic activity was detected with 1 mM P3N or 3-NPA, as shown in Figure 7.3 by the lack of oxygen consumption with P3N, while the dashed line in Figure 7.3 shows the control performed with the bacterial PaNMO (PA4202) \(^{(6)}\).

![Figure 7.3: Oxygen consumption monitored with 1 mM P3N or 3-NPA as substrate.](image)

Experiment performed in 50 mM potassium phosphate, pH 7.5, 30 °C, and atmospheric oxygen. The solid line represents the assay performed with 180 nM purified PA1024 and the dashed line the assay performed with 1.4 nM of purified Pa-NMO.

No enzymatic activity was detected with 20 mM nitroethane, 1-nitropropane, 2-nitropropane, or the anionic forms ethynitronate, propyl-1-nitronate, and propyl-2-nitronate. In the case of propyl-2-nitronate and ethynitronate velocities of 16 and 5 μM oxygen consumed per minute
were detected, which would correspond to enzymatic rates of 1 and 0.5 s$^{-1}$. However the same velocities were detected by adding propyl-2-nitronate and ethyl nitronate in the absence of PA1024, and they represent therefore non-enzymatic reactions.

**Reductive half-reaction with NADH.** The reduction of the FMN cofactor by NADP(H) was tested by following the decrease in absorbance at 461 nm in anaerobic conditions, at pH 7.0 and 25 °C, with the use of a stopped-flow spectrophotometer. The enzyme was fully reduced by NADH with a biphasic pattern (Figure 7.4). The first phase, which accounts for more than 90% of the total change in absorbance at 461 nm, was assigned to flavin reduction. The second phase accounts for less than 10% of the total change in absorbance at 461 nm and is not concentration dependent, with average $k_{obs}$ of 1 s$^{-1}$. This slow phase could be due to the presence of a fraction of damaged enzyme. The concentration dependence for the first phase was analyzed with Eq. 2. The enzyme is fully saturated with NADH from 90 to 500 µM, with a $k_{red}$ value of 18.7 ± 0.1 s$^{-1}$. Due to the necessity of maintaining pseudo-first order conditions it was not possible to use NADH concentrations lower than 90 µM, and therefore to determine a $K_d$ value. However, the fact that at 90 µM the enzyme is fully saturated by NADH suggests a $K_d$ value in the 5 µM range or lower.

When the enzyme was anaerobically mixed with 100 µM NADPH there were no significant changes in absorbance at 461 nm in 60 seconds. Therefore PA1024 is not able to efficiently use NADPH as substrate.
Figure 7.4: Anaerobic reduction with NADH of PA1024 in 20 mM potassium phosphate, 200 mM sodium chloride, pH 7.0 and 25 °C.
Panel A shows the reduction trace with 500 µM NADH. Panel B represents the UV-visible absorption spectra of the enzyme before and after anaerobic mixing with 500 µM NADH (solid and dashed lines, respectively). Panel C shows the observed rate constants of the first phase as a function of NADH concentration.
**Apparent steady-state kinetics with NADH.** PA1024 displayed an NADH oxidase activity of 1.0 s$^{-1}$ in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, 25 °C, and atmospheric oxygen. Turnover of PA1024 with NADH was monitored at saturating NADH concentration (100 µM) with the artificial electron acceptors potassium ferricyanide or with menadione in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, 25 °C, and atmospheric oxygen (Figure 7.5 panel A and B).

**Figure 7.5:** Apparent steady-state kinetics of PA1024 at saturating NADH concentration (100 µM) by varying the concentrations of potassium ferricyanide (panel A) and menadione (panel B). Conditions: 20 mM potassium phosphate, pH 7.0, and 200 mM sodium chloride, at atmospheric oxygen and 25 °C.
The apparent steady-state kinetics of PA1024 with potassium ferricyanide or menadione (Table 7.2) showed a saturating behavior. It was not possible to use menadione concentrations lower than 18 μM, due to the fact that the absorbance value at 343 nm is too low and the steady-state is too short for a reliable fit of the initial rate. The latter reason also applied to potassium ferricyanide with a concentration limit of 75 μM.

Table 7.2: Apparent steady-state kinetic parameters of PA1024 with potassium ferricyanide or menadione as electron acceptors. The kinetic parameters were determined with 100 μM NADH in 20 mM potassium phosphate, pH 7.0, 25 °C, and atmospheric oxygen. a $K_m$ not determined because lower than the range of substrate concentration used; b enzymatic activity not detected.

<table>
<thead>
<tr>
<th>Electron acceptor</th>
<th>$^{\text{app}}K_{\text{cat}}$ (s$^{-1}$)</th>
<th>$^{\text{app}}K_m$ (μM)</th>
<th>$^{\text{app}}K_{\text{cat}}/K_m$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potassium ferricyanide</td>
<td>19.9 ± 0.2</td>
<td>nd$^a$</td>
<td>&gt; 460,000</td>
</tr>
<tr>
<td>Menadione</td>
<td>13.6 ± 0.6</td>
<td>104 ±14</td>
<td>130,000 ± 12,000</td>
</tr>
<tr>
<td>Methyl red</td>
<td>nd$^b$</td>
<td>nd$^b$</td>
<td>nd$^b$</td>
</tr>
</tbody>
</table>

Steady-state kinetic experiments were performed by varying NADH with potassium ferricyanide or menadione as electron acceptors. The apparent kinetic parameters determined in the concentration range 25-100 μM NADH and fixed concentration of 100 μM menadione were not significantly different due to a low $K_m$ value for NADH, with an average value of initial rate of 6.73 ± 0.47 s$^{-1}$. The concentration of menadione was kept fixed at 100 μM in order to keep the absorbance value at 343 nm lower than 1.5. Similar results were obtained by varying NADH from 20 to 100 μM at fixed concentration of 100 μM ferricyanide. NADH was saturating in the range 20-100 μM with an average value of initial rate of 12.7 ± 0.7 s$^{-1}$. Due to the low $K_m$ value for NADH it was not possible to establish the steady-state kinetic mechanism.

Azoreductase activity assay. Turnover of PA1024 was monitored also with the azo dye methyl red at saturating NADH concentration (100 μM) in 20 mM potassium phosphate, pH 7.0, 200
238 mM sodium chloride, 25 °C, and atmospheric oxygen. No significant enzymatic reduction of methyl red was observed over 10 min.

**Bioinformatic analysis of PA1024.** A previous study by Ha et al. \(^{(7)}\) identified six motifs in the protein sequence of PA1024 that are conserved in many hypothetical nitronate monooxygenase. In this study, we identified motif VII located near the O2 atom of the flavin cofactor and we slightly modified the other six motifs. Table 7.3 lists the consensus sequence for these conserved motifs, which are different from the motifs identified for Class 1 NMO \(^{(6)}\) and Figure 7.6 shows the position of the seven conserved motifs in the crystal structure of PA1024 (PDB 2GJL) \(^{(7)}\).

**Figure 7.6:** Conserved motifs highlighted in the crystal structure of PA1024. PA1024 (PDB 2GJL) is shown as white cartoon in panel A and white surface in panel B. Motifs I, II, III, IV, V, VI, and VII, listed in Table 3, are shown in pink, orange, cyan, blue, black, red, and green.
Table 7.3: Conserved motifs in the protein sequence of PA1024. The numbering of the residues refers to the protein sequence of PA1024; the brackets identify residues that can alternatively be present in that position, $h$ represents position occupied by a hydrophobic residue, while $X$ represents a position where any amino acid is accepted. \(^a\) This motif was identified in a previous study by Ha et al. \(^7\). \(^b\) This motif was identified in a previous study by Ha et al. \(^7\) and modified in this study. The modifications to the consensus sequence are underlined.

<table>
<thead>
<tr>
<th>Motif</th>
<th>Consensus sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>I(^b)</td>
<td>17P-[I/V]-h-Q-G-G-M-Q-W-[V/I](^26)</td>
</tr>
<tr>
<td>II(^a)</td>
<td>66T-D-[K/R]-P-F-G-V-N-L-T-h-[L/F]-P(^78)</td>
</tr>
<tr>
<td>III(^b)</td>
<td>121[V/I]-h-H-K-C-T-X-[V/I]-R-H-A(^131)</td>
</tr>
<tr>
<td>IV(^b)</td>
<td>144I-D-G-F-E-C-[A/G]-G-H-P-G-E-X-D-[P/G/T](^159)</td>
</tr>
<tr>
<td>VI(^b)</td>
<td>225E-X-X-[T/S]-X-L-h-X-R-X-[L/M/W]-[R/H/K]-N-T-[S/A/V]-R-V(^241)</td>
</tr>
</tbody>
</table>

Figure 7.7 shows the superposition of the crystal structure of PA1024 (PDB 2GJL) \(^7\) and Pa-NMO (PDB 4Q4K) \(^6\). A BLAST search of the protein sequence of PA1024 and multiple sequence alignments (Figure 7.8) showed that more than 1000 protein sequences of hypothetical monooxygenases from bacteria, fungi, one plant (Ricinus communis; castorbean) and one animal (Pantholops hodgsonii; Tibetan antelope) carry the seven motifs listed in Table 7.3. Therefore, the signature motifs of PA1024 establish a new class of enzymes with NADH:ferricyanide reductase activity, despite their current annotation as hypothetical nitronate monooxygenases.
Figure 7.7: Comparison of the motifs on the crystal structures of PA1024 and Pa-NMO. Panel A shows the superposition the crystal structure of PA1024 (PDB 2GJL) shown as white cartoon with the conserved motifs I, II, III, IV, V, VI, and VII highlighted in pink, orange, cyan, blue, black, red, and green, and of Pa-NMO (PDB 4Q4K), shown as light orange cartoon. Panel B and C highlight motif I, IV, and VII of PA1024. In the corresponding position in the crystal structure of Pa-NMO three different motifs are present, conserved in Class I NMO, specifically motif I, III, and II for Class I NMO.
Figure 7.8: Multiple sequence alignments of protein sequences from prokaryotic and eukaryotic sources currently annotated as hypothetical nitronate monooxygenases.

The numbering of the residues is according to PA1024 sequence. Sequences 1 to 8 carry the seven conserved motifs of PA1024 listed in Table 3 and boxed in orange. Sequence 9 and 10 are Pa-NMO and PA0660 from P. aeruginosa PAO1. The following sequence identifiers are used: PA1024, P. aeruginosa PAO1 (NP_249715.1); E. coli, Escherichia coli (WP_024174515.1); N. brasiliensis, Nocardia brasiliensis NBRC 14402 (GAJ83692.1); D. desulfuricans, Deferribacter desulfuricans (WP_013008529.1); A. oryzae, Aspergillus oryzae RIB40 (XP_001823354.1); P. expansum, Penicillium expansum (KGO43292.1); R. communis, Ricinus communis (XP_002538981.1); P. hodgsonii, Pantholops hodgsonii (XP_005974730.1); Pa-NMO, i.e. PA4202 P. aeruginosa PAO1 (NP_252891.1); PA0660, P. aeruginosa PAO1 (NP_249715.1).

7.6 Discussion

A previous study on a hypothetical nitronate monooxygenase of P. aeruginosa PAO1 encoded by the gene pa4202 established Class 1 NMO with 500 sequences that share four conserved motifs in the protein sequence (6). Two other hypothetical NMOs from P. aeruginosa PAO1 are encoded by the genes pa0660 and pa1024, but they do not carry the four motifs characteristic of Class 1 NMO. The crystal structure of the gene product PA1024 as free enzyme and in complex
with 2-nitropropane was determined by Ha et al. \cite{7}, and the classification of PA1024 as 2-nitropropane dioxygenase (previous name of NMO) was based on activity detected with 10 mM propyl-2-nitronate (the enzymatic rate was not reported). In this study we purified recombinant His-tagged PA1024 and tested for nitronate monooxygenase activity, using as positive control Pa-NMO (PA4202) \cite{6} for the enzymatic assay performed. Despite its annotation PA1024 did not show nitronate monooxygenase activity (Figure 7.3). In the case of 20 mM ethynitronate and propyl-2-nitronate non-enzymatic rates of 1 and 0.5 s\(^{-1}\) were detected in the absence of the enzyme. This suggests that the activity reported by Ha et al. with propyl-2-nitronate \cite{7} may have been non-enzymatic. The presence of 2-nitropropane in complex with PA1024 (PDB code 2GJN) can be due to cocrystallization of a mother liquor component.

The lack of activity with the nitroalkanes nitroethane, 1-nitropropane, and 2-nitropropane establishes that PA1024 is not a nitroalkane oxidase \cite{25}.

PA1024 is able to efficiently use NADH as substrate. This conclusion is supported by the fact that the enzyme was fully reduced by NADH in 0.1 seconds with a \(k_{\text{red}}\) value of 18.7 s\(^{-1}\) and a \(K_d\) value in the low µM range. PA1024 displayed a clear preference of NADH over NADPH, as shown by the fact that the enzyme was not reduced by 100 µM NADPH in 60 seconds. The ability to use only NADH as substrate was previously reported in the case of azoreductases \cite{26} and for some flavin-dependent monooxygenases \cite{27}.

At atmospheric oxygen the NADH oxidase activity of PA1024 is slow, with a turnover rate of 1 s\(^{-1}\) at saturating NADH concentration, which suggests that the NADH oxidase activity is not the primary function of PA1024. Slow NADH oxidase activity in the absence of the second substrate has been reported in the case of flavin-dependent monooxygenases (uncoupling) \cite{28}, or in reductases such as old yellow enzyme \cite{29, 30}. PA1024 was able to efficiently turnover with
NADH and the artificial electron acceptors potassium ferricyanide and menadione, with \( K_m \) values estimated in the low \( \mu \)M range (Table 7.2). PA1024 can be therefore classified as NADH:ferricyanide reductase (E.C. 1.6.99.-). Other enzymes classified as NADH:ferricyanide reductase belong to the class of NADH dehydrogenases (Complex I) but they are multisubunit, membrane-bound enzymes\(^{(31)}\). On the contrary, PA1024 is a small and soluble protein. Recently enzymes with NADH:ferricyanide reductase activity were reported from bacteria with the ability to detoxify azo dyes from the environment\(^{(32-34)}\). The lack of enzymatic activity with the azo dye methyl red, routinely used in the kinetic characterization of azoreductases\(^{(17, 18)}\), supports the conclusion that PA1024 most likely is not an azoreductase.

The term diaphorase was initially used only for flavin-dependent enzymes able to use both NADH and NADPH as electron donor\(^{(35)}\). In this context PA1024 does not represent an example of classical diaphorase. Other flavin dependent enzymes with reported NAD(P)H:ferricyanide reductase activity are NADPH-ferrodoxin reductases (FNR)\(^{(36, 37)}\), ene reductases classified as old yellow enzyme\(^{(30)}\), monooxygenases\(^{(27)}\), methionine synthase reductase\(^{(38)}\). All these enzymes usually display a preference for NADPH over NADH, with few exceptions reported for monooxygenases, such as 3-hydroxybenzoate 6-hydroxylase\(^{(39)}\) or 6-hydroxy-3-succinoylpyridine 3-monooxygenase\(^{(40)}\). This feature is remarkably different from PA1024, which is able to efficiently use only NADH as substrate.

The protein sequence of PA1024 carries seven conserved motifs (Table 7.3) that are present in more than 1000 sequences currently annotated as hypothetical nitronate monooxygenases in the non-redundant protein database. These motifs are remarkably different from the ones identified in the protein sequence of Pa-NMO establishing Class 1 NMO\(^{(6)}\) and they are not present in PA0660, the other hypothetical NMO from \textit{P. aeruginosa} PAO1 (Figure 7.8). The locations of
the seven conserved motifs on the crystal structure of PA1024 \(^{(7)}\) (PDB 2GJL) are shown in Figure 7.6 and the comparison with PaNMO is shown in the sequence alignment in Figure 7.8 and in the overlay of the two crystal structures (PDB 2GJL and 4Q4K) in Figure 7.7.

Motif I in PA1024 (colored in magenta in Figure 7.6) contacts the re face of the cofactor and includes the fully conserved M23. A methionine residue contacting the face of the flavin cofactor not exposed to the active site cavity is present also in Pa-NMO (PDB 4Q4K; M20) \(^{(6)}\) and other FMN-containing enzymes with a TIM barrel fold, such as the type II isopentenyl diphosphate:dimethylallyl diphosphate isomerase (PDB 1P0N; M63) \(^{(41)}\), glutamate synthase (PDB 1LM1; M875) \(^{(42)}\), 2,4-dienoyl-CoA reductase (PDB 1PS9; M25) \(^{(43)}\), and tRNA-dihydouridine synthase (PDB 3B0U; M11) \(^{(44)}\). Therefore, the presence of this fully conserved methionine by itself in the multiple sequence alignment is not helpful in predicting protein function. However, the consensus sequence surrounding the fully conserved methionine is different in PA1024 and Pa-NMO, and possibly modulates the electronic distribution of the isoalloxazine ring, affecting the reactivity of the cofactor. For instance M23 in PA1024 is preceded by the fully conserved G22 and G21 and followed by the fully conserved bulkier W25 (Figure 7.8), while in Class 1 NMO M20 is preceded by the fully conserved P19 and followed by the fully conserved G22 (Figure 7.8).

Motif II, highlighted in orange in Figure 7.6, is located in the TIM barrel domain, expanding from the surface of the protein to the region near the N3 atom of the cofactor. The first part of the consensus sequence of motif II \(^{(66}T-D-[K/R]-P-F-G-V-N-L^{74})\) of PA1024 is conserved also in Pa-NMO (Figure 7.8), and it is therefore not specific to predict protein function, but most likely related to folding of the TIM barrel domain. The second part of motif II \(^{(75}T-h-[L/F]-P^{78})\) is remarkably different in Pa-NMO \(^{(71}F-C-H-R^{74})\) as shown in Figure 7.8. In particular the fully
conserved polar T75 is located 3.8 Å from the N3 atom of FMN in PA1024, while in the same position (3.6 Å from N3 atom of FMN) there is the hydrophobic F71 in Pa-NMO, fully conserved in Class 1 NMO. The presence of threonine or phenylalanine in the surrounding of the N3 atom of FMN seems therefore important in discriminating between the two classes of enzymes identified by PA1024 and Pa-NMO, and it could affect the polarity of the cofactor environment, modulating its reactivity \(^{(24)}\).

Motif III in PA1024, rendered in cyan in Figure 7.6, includes the C terminal part of helix \(\alpha6\) and the \(\beta\) strand \(\beta5\) of the TIM barrel domain. The side chains of the residues included in motif III are not located in the active site or near the cofactor, with the exception of the fully conserved K124 which provides a positive charge located 3.9 Å from the N1 atom of FMN. The presence of a positive charge near the N1-C2 locus of flavin cofactors in other enzymatic systems was shown to play a role in modulating the redox potential of the isoalloxazine moiety \(^{(45)}\) and also to be important for oxygen reactivity \(^{(46)}\). This motif is not conserved in Pa-NMO, where a threonine or cysteine residue is present instead of K124, but at a distance of 6.8 Å from the N1 atom of the cofactor.

Motif IV in PA1024, highlighted in blue in Figure 7.6, is located on the surface of the active site cavity, with the side chain of the fully conserved H152 in front of the flavin cofactor at 7.9 Å from the N5 atom of FMN. H152 is fully conserved also in Class 1 NMO (H183 in Pa-NMO) and it is therefore not helpful in predicting protein function. However, the consensus sequence containing this fully conserved histidine is different in Class 1 NMO and in PA1024, with the presence of a fully conserved arginine or of a fully conserved proline adjacent to the conserved histidine in Pa-NMO and PA1024, respectively.
Motif V in PA1024, colored in black in Figure 7.6, includes helix \( \alpha_8 \) and \( \beta \) strands \( \beta_7 \) and \( \beta_8 \) in the TIM barrel domain, which contact mostly the phosphate group, the ribityl moiety of FMN and the xylene moiety of the isoalloxazine ring. This motif is mostly conserved in Pa-NMO, and this is consistent with a role of the side chains involved in positioning mainly the phosphate and ribityl moiety of the cofactor, and not the reactive moiety.

Motif VI in PA1024, highlighted in red in Figure 7.6, is located in the substrate binding domain, on the surface of the protein near the possible entrance to the active site. This motif is not conserved in Pa-NMO (Figure 7.8), consistent with a different substrate specificity of the two classes of enzymes.

Motif VII of PA1024, represented in green in the TIM barrel domain in Figure 7.6, consists of the \( \alpha \) helix \( \alpha_4 \) on the surface and the strand \( \beta_4 \) in the active site cleft, with the fully conserved side chains \( 101E-T-A^{103} \) on the side of the pyrimidine moiety of the isoalloxazine. In particular the side chain of A103 is located 3.2 Å from the O2 atom of FMN, while in the same position in the crystal structure of Pa-NMO there is the ionizable side chain of H133 at 4.1 Å from the N1 atom of the cofactor, fully conserved in Class 1 NMO. As expected from the differences in the substrate binding domain between Pa-NMO and PA1024, motif IV of Class 1 NMO (boxed in yellow in Figure 7.8) located at the entrance to the active site of Pa-NMO is not conserved in PA1024.

In the multiple sequence alignment shown in Figure 7.8 a region of 29 amino acids (numbering 94 to 123) of the protein sequence of Pa-NMO corresponds to a gap in the protein sequences of hypothetical NADH:ferricyanide reductases carrying the motifs of PA1024. These 29 amino acids in the crystal structure of Pa-NMO are located on a loop on the surface of the protein that connects the TIM barrel domain with the substrate binding domain. The absence of this
structural element in the crystal structure of PA1024 contributes to the fact that the active site of PA1024 is more exposed to the solvent than the one of Pa-NMO.

In summary the last part \((^{75}\text{T}-\text{h-}[\text{L/F}]-^78\text{P})\) of motif II, motif III, motif VI, and motif VII of PA1024 are the most helpful in identifying significant differences in the two classes of enzymes represented by PA1024 and Pa-NMO, which are all currently annotated as hypothetical monooxygenases. Motif I and motif IV of PA1024 contain a fully conserved methionine and histidine residues that are not specific to the class of NADH:ferricyanide reductases identified in this study, but the different consensus sequences of these motifs can still be used to discriminate from members of Class 1 NMO.

The sequences that carry the seven motifs of PA1024 listed in Table 7.3 belong mainly to actinobacteria, proteobacteria, and fungi, and to lesser extent to deferribacteres and cyanobacteria. The bacteria that possess these hypothetical proteins identified by the BLAST search in this study are not pathogens and live mainly in soil, plants, or water. These habitats possibly suggest that this class of enzymes may be involved in some detoxification processes, which may be of relevance for biotechnological applications, other than metabolic pathways common to all bacteria. In eukaryotes hypothetical proteins carrying the motifs identified in PA1024 are present only in the plant *Ricinus communis*, and in the animal *Pantholops hodgsonii* (the Tibetan antelope). It is interesting to note that *E. coli* possesses a hypothetical nitronate monooxygenase reclassified as hypothetical NADH:ferricyanide reductase by this study. Indeed, it was previously shown the lack of ability of *E. coli* to grow in presence of the metabolic poison P3N (47) and it would be unlikely for this organism to possess an enzymatic system, NMO, able to detoxify P3N.
7.7 Conclusions

The gene function prediction of hypothetical nitronate monooxygenases is not reliable, as it is based only on modest protein sequence similarity with the fungal enzymes, and most of the hypothetical nitronate monooxygenases in the GenBank database do not carry the four conserved motifs identified for Class 1 NMO. A structural study was performed on the putative 2-nitropropane dioxygenase (previous name for nitronate monooxygenase) PA1024, but the enzyme was not kinetically characterized. The kinetic characterization in the present study demonstrated that PA1024 is not a nitronate monooxygenase, and highlighted the necessity of reclassification of the current annotation as nitronate monooxygenase in the GenBank and PDB databases. A kinetic characterization of PA1024 revealed a NADH:ferricyanide reductase activity, while it did not show azoreductase activity. Furthermore, conserved motifs previously identified in the protein sequence of PA1024 were reanalyzed and shown to be conserved in more than 1000 sequences of hypothetical proteins currently annotated as nitronate monooxygenases. The function prediction for these 1000 sequences is therefore updated to NADH:ferricyanide reductase (E.C. 1.6.99.-).
7.8 References


CHAPTER VIII: THE HYPOTHETICAL NITRONATE MONOOXYGENASE

PA0660 FROM Pseudomonas aeruginosa PAO1 ENCODES A DIAPHORASE

(The author carried out all the experiments described in this Chapter with the exceptions of the identification of the cofactor and the steady-state kinetics, which were carried out by Jacob Ball.)

8.1 Abbreviations

NMO, nitronate monooxygenase; Pa-NMO, nitronate monooxygenase from Pseudomonas aeruginosa PAO1; P3N, propionate 3-nitronate; 3-NPA, 3-nitropropionate; DCIP, 2,6-dichlorophenol indophenol; IPTG, isopropyl-1-thio-β-D-galactopyranoside.

8.2 Abstract

The hypothetical NMO PA0660 from P. aeruginosa PAO1 was cloned, expressed, and purified to high level. The purified enzyme was not able to detoxify propionate 3-nitronate, despite its annotation in the GenBank database. The kinetic characterization of PA0660 showed that it is able to use either NADPH or NADH as electron donor, with a preference for NADPH, as shown by the fact that the second order rate constant $k_{\text{red}}/K_d$ is 35 times higher with NADPH compared to NADH. The reductive half reaction with NADPH exhibited a double-exponential behavior, with a fast $k_{\text{red}}$ of 470 s$^{-1}$ and a binding constant of 120 µM for flavin reduction and a second slow phase with average $k_{\text{obs}}$ of 11.8 s$^{-1}$. The enzyme was able to turnover at 8 s$^{-1}$ with the electron acceptor DCIP at fixed concentration of 55 µM and NADH concentration ranging from 44 to 533 µM, consistent with a $K_m$ value for NADH lower than 44 µM. PA0660 was not able to reduce methyl red, the typical substrate of azoreductases. Six conserved motifs were identified in the protein sequence of PA0660 which are present in more than 1000 bacterial hypothetical NMOs which need to be reclassified to hypothetical diaphorases.
8.3 Introduction

The functional annotation of more than 5,000 hypothetical nitronate monooxygenases (NMOs, E.C. 1.13.12.16) in the GenBank is often not reliable, also due to fact that the current annotation of gene function prediction is based on only modest overall protein sequence similarity with the fungal enzymes \(^1\). Furthermore, many organisms possess more than one gene coding for hypothetical nitronate monooxygenases, according to the current annotation in the GenBank. One example is *Pseudomonas aeruginosa* PAO1, which possess three genes annotated as hypothetical nitronate monooxygenase: *pa4202*, *pa1024*, and *pa0660*. A previous structural and kinetic study on the gene product PA4202 classified this enzyme as Pa-NMO and identified four conserved motifs that established Class 1 NMO, including 500 hypothetical nitronate monooxygenases in the GenBank database \(^1\). The other two hypothetical nitronate monooxygenases PA1024 and PA0660 do not possess these four motifs.

The crystal structure of PA1024 was solved at 2.0 Å resolution in 2006 and the enzyme was initially classified as 2-nitropropane dioxygenase, previous name for nitronate monooxygenase \(^2\). However the kinetic characterization of purified PA1024 performed in our lab (Salvi, F., Ball, J., and Gadda, G., unpublished data) demonstrated that PA1024 is not a nitronate monooxygenase and the enzyme was reclassified as NADH:ferricyanide reductase (E.C. 1.6.99.-), specific only for NADH. The six conserved motifs identified by Ha et al. \(^2\) were reanalyzed and a seventh motif identified. The seven conserved motifs of PA1024 are shared by more than 1000 hypothetical nitronate monooxygenases in the GenBank database, which need to be reclassified as hypothetical NADH:ferricyanide reductases.

No evidence at transcript or protein level is present for PA0660, the third hypothetical NMO from *P. aeruginosa* PAO1.
In this study we cloned *pa0660* from genomic DNA, we expressed the recombinant His-tagged PA0660 in *Escherichia coli*, and kinetically characterized it. We showed that PA0660 is not a nitronate monooxygenase and that it displayed diaphorase activity (E.C. 1.6.99.-) with preference for NADPH. We identified six conserved motifs in the protein sequence of PA0660, which are different from the motifs identified for PA4202 and PA1024, and are conserved in more than 1000 hypothetical nitronate monooxygenases in the GenBank database.

### 8.4 Experimental procedures

**Materials.** The enzymes *BamHI*, *NdeI*, *DpnI*, calf intestinal alkaline phosphatase (CIP), and T4 DNA ligase were purchased from New England BioLabs (Ipswich, MA), and *E. coli* strain DH5α was from Invitrogen Life Technologies (Grand Island, NY). *Escherichia coli* strain Rosetta(DE3)pLysS and the expression vector pET15b were from Novagen (Madison, WI), QIAprep Spin Miniprep kit, and QIAquick PCR purification kit were from Qiagen (Valencia, CA). *Pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA), and oligonucleotides from Sigma Genosys (The Woodlands, TX). The genomic DNA of *P. aeruginosa* PAO1 was a kind gift from Dr. Jim Spain, Georgia Institute of Technology, Atlanta. HiTrapTM Chelating HP 5 mL affinity column was from GE Healthcare (Piscataway, NJ), isopropyl-1-thio-β-D-galactopyranoside (IPTG) was from Promega, (Madison, WI), and nitroalkanes were from Sigma-Aldrich (St. Louis, MO). NADPH was from Enzo Life Sciences (Farmingdale, NY), and NADH from AMRESCO (Solon, OH). All other reagents used were of the highest purity commercially available.

**Cloning.** The gene *pa0660* was amplified from the genomic DNA of *P. aeruginosa* PAO1 by PCR with the forward primer 5’GGCCTTCCCATATGATGTCTTGCCCAGCC 3’ and the reverse primer 5’GCGGATCCTCAGCGCGCCAGCGG 3’ (the restriction sites for *NdeI* and
BamHI are underlined). The DNA amplification was performed in presence of 3% DMSO with an initial denaturation step at 95 °C, followed by 20 cycles of denaturation for 45 s at 95 °C, annealing for 45 s at 56 °C (the annealing temperature decreases of 0.2 °C each cycle), and extension for 3 min at 72 °C, followed by a final step at 72 °C for 10 min.

The expression vector pET15b and the PCR product, after purification with the QIAquick PCR purification kit, were subjected to double digestion with NdeI and BamHI at 37 °C for 2 h. The digested insert and vector were purified with the QIAquick PCR purification kit. The vector was dephosphorylated with 0.5 units of calf intestinal alkaline phosphatase for 30 min at 37 °C, purified with the QIAquick PCR purification kit and ligated to the insert with incubation for 15 h at 16 °C with T4 DNA ligase. E. coli strain DH5α was transformed with the ligation mixture and the resulting colonies grown at 37 °C on Luria-Bertani agar plates containing 50 μg/mL ampicillin were screened for the presence of the desired insert by DNA sequencing at the Cell, Protein and DNA core facility at Georgia State University. The DNA sequencing confirmed the correct insertion of the gene in the vector pET15b and the absence of undesired mutations.

**Recombinant expression and purification.** A volume of 100 mL of Luria-Bertani Broth containing 50 μg/mL of ampicillin and 34 μg/mL of chloramphenicol was inoculated with E. coli strain Rosetta(DE3)pLysS carrying the construct pET15b/pa0660, and incubated at 37 °C for 15 h. A volume of 10 mL of this culture was used to inoculate 3 liters of Luria-Bertani Broth containing 50 μg/mL of ampicillin and 34 μg/mL of chloramphenicol, which was incubated at 37 °C until it reached an optical density at 600 nm of 0.7. Recombinant protein expression was induced by addition of IPTG to a final concentration of 400 μM and the culture incubated at 18 °C for 20 h. The wet cell paste of 16.5 g, recovered by centrifugation, was resuspended in 70 mL of lysis buffer (10 mM imidazole, 300 mM NaCl, 10% v/v glycerol, 1 mM phenylmethylsulfonyl
fluoride, 5 mM MgCl2, 2 mg/mL lysozyme, 5 μg/mL DNase, 5 μg/mL RNase, and 20 mM sodium phosphate, pH 7.0) and subjected to several cycles of sonication. The cell free extract obtained after centrifugation at 12000 g for 20 min was loaded onto a HiTrapTM Chelating HP 5 mL affinity column equilibrated with buffer A (10 mM imidazole, 300 mM NaCl, 10% v/v glycerol, and 20 mM sodium phosphate, pH 7.0). After washing with 10 column volumes of buffer A, the enzyme was eluted with 30% of buffer B (buffer A containing 0.5 M imidazole) and dialyzed against 10% glycerol, 10 mM sodium pyrophosphate, 10 mM sodium phosphate, pH 6.0.

**Enzyme assays.** The total protein concentration was determined using the Bradford method with bovine serum albumin as standard. UV-visible absorbance was recorded with an Agilent Technologies diode-array spectrophotometer Model HP 8453 PC equipped with a thermostatted water bath. The flavin cofactor was extracted by heat denaturation and purified by HPLC using a Shimadzu SPD-M10A system with a C18 reversed-phase column (5μ, 15 cm x 4.6 mm) equilibrated in 5% acetonitrile in water. A gradient from 5% to 100% acetonitrile developed over 1 h at 0.5 mL/min was applied, and the flavin cofactor eluted with a retention time of 5.4 minutes. The purified flavin cofactor was analyzed by MALDI-TOF mass spectrometry in negative ion mode at the Mass Spectrometry Facility of Georgia State University with a Bruker Doltonics ultrafleXtreme instrument. The matrix used was α-cyano-4-hydroxycinnamic acid (CHCA).

Initial rates were normalized for the concentration of flavin-bound enzyme calculated with the extinction coefficient of free FMN ε450=12,500 M⁻¹ cm⁻¹ α(4).

**Nitrate monooxygenase activity assay.** Nitrate monooxygenase activity was tested as previously described, by monitoring the initial rate of oxygen consumption with a
Hansatech Instruments computer-interfaced Oxy-32 oxygen-monitoring system at atmospheric oxygen and 30 °C (i.e., 0.23 mM oxygen) in 50 mM potassium phosphate, pH 7.5. The use of the methods of the initial rate prevents the interconversion between the anionic form (nitronate) and the neutral form (nitroalkane) of the substrate. Stock solutions of nitronates and nitroalkanes were prepared as previously described (1, 5-7), and the final concentration was 10 mM for propionate 3-nitronate (P3N) or 3-nitropropionic acid (3-NPA), and 20 mM for 2-nitropropane, propyl-2-nitronate, nitroethane, or ethynitronate. Enzyme concentration was 200 nM, and the reaction was started with the addition of the anionic form of the substrate (nitronate) to ensure that in the initial 30 seconds the amount of the neutral form of the substrate (nitroalkane) is negligible, since the second-order rate constants for protonation of the nitronates are in the range 15-75 M⁻¹s⁻¹ (8,9). The enzymatic assay was performed in parallel with the enzyme PaNMO (1.4 nM) and P3N (1 mM) as previously described (1).

**Reductive half reaction with NAD(P)H.** The reductive half reaction of PA0660 with NAD(P)H was performed anaerobically with an SF-61DX2 Hi-Tech KinetAsyst high-performance stopped-flow spectrophotometer, thermostated at 25 °C. After an overnight treatment with an oxygen scavenging system composed of glucose (5 mM) and glucose oxidase (1 µM) in 100 mM sodium pyrophosphate, pH 6.0, the instrument was washed with buffer (20 mM potassium phosphate, pH 7.0, 200 mM NaCl), previously flushed with argon for 30 min. Purified PA0660 was loaded on a desalting PD-10 column equilibrated with 20 mM potassium phosphate, pH 7.0, 200 mM NaCl, and subjected to 20 cycles of degassing in a tonometer by applying vacuum and flushing with argon. The syringes containing the buffer (20 mM potassium phosphate, pH 7.0, 200 mM NaCl) or the substrate NAD(P)H diluted in buffer were flushed for 30 min with argon before mounting onto the stopped-flow spectrophotometer. Glucose (2 mM) and glucose oxidase (0.5 µM) were
added to the buffer, enzyme and substrate solutions for complete removal of traces of oxygen. The concentration of the PA0660 after mixing was 11 µM and of substrate ranged from 95 to 1000 µM for NADPH (determined with the extinction coefficient $\varepsilon_{340} = 6,220 \text{ M}^{-1} \text{ cm}^{-1}$) (10). In the case of the reductive half reaction using NADH as substrate, the final concentrations of PA0660 and NADH were 16 µM and 200-1200 µM, respectively. Time-resolved absorption spectra of the anaerobic reduction of PA0660 with NADPH or NADH as substrate were recorded with the use of an SF-61DX2 Hi-Tech KinetAsyst high-performance stopped-flow spectrophotometer, thermostated at 25 °C, and equipped with a photo-diode-array detector. The same procedure described above for the preparation of enzyme and substrate was used. The concentrations after mixing of PA0660 and NAD(P)H were 16 µM and 200 µM. Time-resolved absorption spectra were recorded at 1.5 ms intervals after mixing (2.2 ms) for 0.15 s in the range from 340 to 700 nm.

**NADPH oxidase activity.** NADPH oxidase activity was monitored by following NADPH oxidation at 340 nm ($\varepsilon_{340} = 6,220 \text{ M}^{-1} \text{ cm}^{-1}$) (10) in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, at 25 °C and atmospheric oxygen. The concentrations of enzyme and NADPH were 600 nM and 100 µM, respectively.

**Diaphorase activity assay.** Turnover of PA0660 with NADPH and the artificial electron acceptor 2,6-dichloroindophenol (DCIP) was investigated by monitoring the reduction of DCIP at 600 nm ($\varepsilon_{600} = 20,600 \text{ M}^{-1} \text{ s}^{-1}$ at pH 7.0) (11) in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, at 25 °C and atmospheric oxygen. The concentrations of enzyme and DCIP were 60 nM and 55 µM. The concentration of NADPH was varied from 44 to 533 µM.

Turnover with menadione was monitored by following the decrease in absorbance at 343 nm, where menadione and menadiol are isosbestic, that corresponds to NADH consumption ($\varepsilon_{343} =$
6,220 M^{-1} s^{-1}) \textsuperscript{12} in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, at 25 °C and atmospheric oxygen. Stock solutions of menadione were prepared in ethanol and the final ethanol concentration in the reaction mixture was kept fixed at 0.6% to minimize possible effects on enzymatic activity. The concentration of enzyme and menadione used were 58-116 nM and 100 µM. The concentration of NADPH was varied from 29 to 116 µM.

Azoreductase activity assay. The azoreductase activity assay was performed by following the reduction of the azo dye methyl red at 430 nm (ε\textsubscript{430} = 23,360 M^{-1} s^{-1}) \textsuperscript{13} in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, at 25 °C and atmospheric oxygen. The concentrations of enzyme, NAD(P)H, and methyl red were 600 nM, 100 µM, and 100 µM. The non-enzymatic reaction of NAD(P)H and methyl red was negligible, in agreement with the notion that the non-enzymatic reaction is observed only at low pH \textsuperscript{14}.

Enzyme monitored single turnover. The enzyme PA0660 at a final concentration of 8 µM was mixed with a stoichiometric amount of NADPH at atmospheric oxygen. Time-resolved absorption spectra for the detection of flavin intermediates were recorded at 3 s intervals after mixing (2.2 ms) for 600 s in the range from 340 to 600 nm with an SF-61DX2 Hi-Tech KinetAsyst high-performance stopped-flow spectrophotometer equipped with a photo-diode array detector, thermostated at 25 °C. The experiment was carried out in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride.

Data analysis. The kinetic parameters for the enzymatic assays carried out at atmospheric oxygen with NADPH were obtained from the fitting of the experimental points to the Michaelis-Menten equation for one substrate using KaleidaGraph software (Synergy Software, Reading, PA). Stopped-flow traces were fit with the software KinetAsyst 3 (TgK-Scientific, Bradford-on-Avon, U.K.) to Equation 1 which describes a second exponential process in which A is the
absorbance at 447 nm at time $t$, $B_1$ and $B_2$ represent the amplitudes of the absorbance change, $k_{\text{obs}1}$ and $k_{\text{obs}2}$ define the observed rate constants, and $C$ is an offset value accounting for the nonzero absorbance value at infinite time.

$$A = B_1 \exp(-k_{\text{obs}1} t) + B_2 \exp(-k_{\text{obs}2} t) + C$$  \hspace{1cm} \text{Eq. 1}

The observed rate constants for the first phase (flavin reduction) were fit to Equation 2, where $S$ is the concentration of organic substrate, $k_{\text{red}}$ is the rate constant for flavin reduction at saturating substrate concentration, and $K_d$ is the dissociation constant for substrate binding.

$$k_{\text{obs}} = \frac{k_{\text{red}} S}{K_d + S}$$  \hspace{1cm} \text{Eq. 2}

The observed rate constants for the second phase did not show concentration dependence.

The software SPECFIT/32 (Spectrum Software Associates, USA) was used to perform a global fitting analysis of the time-resolved absorption spectra.

**Bioinformatic analysis.** The amino acid sequence of PA0660 was used to search the non-redundant protein database with BlastP\(^{(15)}\), and conserved motifs were designed manually based on the multiple sequence alignment generated by BlastP. Selected sequences were aligned with Clustal Omega\(^{(16)}\) for Figure 5, which was created with the software Jalview 2.8\(^{(17)}\).

### 8.5 Results

**Protein purification.** The recombinant His-tagged PA0660, currently annotated as hypothetical nitronate monooxygenase, was expressed in *Escherichia coli* and purified to high levels, as judged from SDS-PAGE analysis (Figure 8.1 panel A). Figure 8.1 panel B shows the UV-visible
absorption spectrum of purified PA0660, with maxima at 355 and 447 nm, which are consistent with the presence of a flavin cofactor in the oxidized state.

Figure 8.1: Panel A shows the SDS-PAGE analysis of purified PA0660 (lane 1). Lane 2 represents the Bio-Rad broad range marker. Panel B shows the UV-visible absorption spectrum of purified PA0660 in 10 mM sodium pyrophosphate, 10 mM sodium phosphate, pH 6.0, and 10 % v/v glycerol, at 25 °C.

The high-energy band is significantly blue shifted from the maximum of 375 nm of free flavin in solution \(^{(18)}\) to 355 nm in the case of FMN incorporated in PA0660. This spectral perturbation of the enzyme-bound cofactor suggests a hydrophobic environment surrounding the FMN in PA0660 \(^{(19)}\). The low-energy band also presents a small (<4 nm) hypsochromic shift from 450 to 447 nm, and a well-defined shoulder at 470 nm which are consistent with a hydrophobic environment of the pyrimidine ring of the FMN cofactor \(^{(18)}\).

Heat denaturation of PA0660 resulted in the release of the flavin cofactor to the bulk solvent, which indicates that the cofactor is not covalently bound to the protein. The identification of the flavin cofactor extracted from PA0660 was carried out by MALDI-TOF analysis with the matrix \(\alpha\)-cyano-4-hydroxycinnamic acid at the Mass Spectrometry Core Facility of Georgia State University. The detection of a peak in negative ion mode with \(m/z\) of 455.4 identifies FMN as cofactor of PA0660. No peak with \(m/z\) value of 784.5, consistent with the presence of FAD, was
detected. The ratio FMN/enzyme based on the total protein determination with the Bradford method was ~0.7, which suggests a stoichiometry of 1:1 per monomer of protein.

_Nitronate monooxygenase activity assay._ The nitronate monooxygenase activity was tested as previously described \(^{(1,5,20)}\) by monitoring initial rates of oxygen consumption with propionate 3-nitronate (10 mM) in 50 mM potassium phosphate, pH 7.5, 30 °C, and atmospheric oxygen (i.e. 230 µM). In parallel the same enzymatic assay was performed as positive control with purified Pa-NMO, for which nitronate monooxygenase activity was previously demonstrated \(^{(1)}\). While oxygen consumption was detected with the positive control with Pa-NMO, purified PA0660 exhibited no enzymatic activity with propionate 3-nitronate, or with the neutral form 3-nitropropionic acid (10 mM). PA0660 was also not able to use as substrates nitroethane, 2-nitropropane, or the anionic forms ethylnitronate, and propyl-2-nitronate with final concentration 20 mM.

_Reductive half reaction with NAD(P)H._ The anaerobic reduction of the flavin cofactor of PA0660 by NAD(P)H was monitored with a stopped-flow spectrophotometer, at pH 7.0 and 25 °C. Time-resolved absorption spectra were recorded every 1.5 ms over 0.15 s from 340 to 700 nm. When PA0660 (16 µM) was mixed anaerobically with 200 µM NADPH the trace at 7.5 ms corresponded to the fully reduced enzyme and no changes in absorbance from 340 to 700 nm were observed from 7.5 ms to 0.15 s (Figure 8.2).
Figure 8.2: Time-resolved absorption spectra of the anaerobic reduction of PA0660 with 200 µM NADPH. The experiment was performed in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, at 25 °C, and the concentration of PA0660 after mixing was 16 µM. Time-resolved absorption spectra were recorded at 1.5 ms intervals after mixing (2.2 ms) for 0.15 seconds in the range from 340 to 700 nm.

The maximal absorbance change occurred at 447 nm, which corresponds to the bleaching of the low-energy peak in the UV-visible absorption spectrum of the oxidized enzyme to the reduced enzyme, and this wavelength was used to monitor the rate of flavin reduction at different NAD(P)H concentrations. There were no changes in absorbance at 500-600 nm, indicating that no charge-transfer complex was detected under these conditions. Similar results were obtained when NADH at a final concentration of 200 µM was used as substrate. The decrease in absorbance at 447 nm exhibited a biphasic behavior with NADPH (Figure 8.3 panel A) with a fast phase accounting for ~90 % of the total absorbance change, and a slow phase representing ~10 % of the total absorbance change.
Figure 8.3: Anaerobic reduction of PA0660 by NAD(P)H in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, 25 °C.
Panel A shows the stopped-flow traces with 500 µM (blue line) and 95 µM (black line) NADPH. The inset represents the UV-visible absorption spectra of oxidized (solid line) and reduced (dashed line) enzyme. Panel B shows the stopped-flow traces with 1200 µM (blue line) and 200 µM (black line) NADH. Panel C reports the
concentration dependence of the observed rate constants of flavin reduction with NADPH (solid circles) and NADH
(empty circles) as substrate. The solid line represents the fitting to Equation 2.

This biphasic behavior was also observed in the anaerobic reduction of PA0660 by NADH (Figure 8.3 panel B), but with this substrate the fast and slow phases are not well resolved, as the observed rate constant for the fast phase is only 7 times higher than the observed rate constant for the slow phase with 1.2 mM NADH and 2.7 times with 400 µM NADH. The fast phase was assigned to flavin reduction as it accounts for ~90% of the total change in absorbance at 447 nm and displayed a hyperbolic dependence of the observed rate constants versus NADPH concentration. The $k_{\text{red}}$ and $K_d$ values obtained by the fitting of the experimental data to Eq. 2 (Figure 8.3 panel C) were $470 \pm 4 \, \text{s}^{-1}$ and $121 \pm 4 \, \mu\text{M}$, respectively, with a second-order rate constant for substrate capture $k_{\text{red}} / K_d$ of $3,880,000 \pm 92,000 \, \text{M}^{-1} \, \text{s}^{-1}$. In the case of the anaerobic reduction of PA0660 by NADH, full saturation was not achieved with concentrations of substrate as high as 1.2 mM, consistent with a $K_d$ value larger than 1.2 mM (Figure 8.3 panel C). The second-order rate constant for substrate capture $k_{\text{red}} / K_d$ for NADH was determined to be $112,000 \pm 7,000 \, \text{M}^{-1} \, \text{s}^{-1}$.

The slow phase observed in the anaerobic reduction of PA0660 by NADPH was not concentration dependent, with an average $k_{\text{obs}}$ value of $11.8 \, \text{s}^{-1}$. Similar results were obtained when NADH was used as substrate, with an average $k_{\text{obs}}$ value of $11.5 \, \text{s}^{-1}$. This second phase could represent product release or an isomerization of the enzyme-product complex.

**Diaphorase activity assay.** The diaphorase activity assay with NADPH was tested with the artificial electron acceptor DCIP at pH 7.0, 25 °C, and atmospheric oxygen. The NADPH oxidase activity of PA0660 in the same conditions was slow, with a measured rate of $0.2 \, \text{s}^{-1}$ with 100 µM NADPH. The turnover of PA0660 with NADPH was assayed at a fixed concentration of 55 µM of the electron acceptor DCIP and by varying NADPH from 44 to 533 µM. NADPH was
saturating from 44 to 533 µM with an average value of initial rate of 8.3 s\(^{-1}\). When NADPH concentrations of 107 and 44 µM were used the decrease in absorbance at 600 nm, corresponding to the reduction of DCIP, was linear only for 15 s after mixing, thereby the accuracy of the determination of the initial rate could be affected and the data are not reported.

PA0660 was able to use menadione as substrate. Turnover was monitored at fixed concentration of 100 µM menadione and NADPH was varied from 29 to 116 µM. The enzyme PA0660 was saturated by NADPH in the concentration range 29-116 µM, with an average \( \text{app} k_{\text{cat}} \) of 10.3 s\(^{-1}\).

**Azoreductase activity assay.** The azoreductase activity was investigated using as substrate the azo dye methyl red, which is routinely used for the kinetic characterization of azoreductases \(^{13, 21}\). No enzymatic reduction of methyl red with 100 µM NADPH or NADH was observed over 60 seconds.

**Enzyme monitored single turnover.** Time-resolved absorbance spectroscopy was used to monitor the flavin state during the oxidation by oxygen of PA0660 reduced aerobically by stoichiometric NADPH. The first spectrum recorded 1.5 ms after mixing (black line in Figure 8.4 panel A) represents oxidized enzyme in the presence of NADPH. The spectrum recorded at 3 s (red line in Figure 8.4 panel A) corresponds to the flavin hydroquinone, most likely the anionic form, as suggested by the peak at 360 nm \(^{22}\).
Figure 8.4: Time resolved absorption spectra of PA0660 mixed with stoichiometric NADPH at atmospheric oxygen at 25 °C.

The experiment was performed in 20 mM potassium phosphate, pH 7.0, 200 mM sodium chloride, and the concentration of PA0660 and NADPH after mixing was 8 µM. Time-resolved absorption spectra were recorded at 3 s intervals after mixing (2.2 ms) for 600 seconds in the range from 340 to 600 nm. Panel A shows the absorption spectra recorded at 0.0015 (black bold), 3 (red bold), 9 (cyan), 30 (green), 100 (pink), 223 (blue) seconds. Panel B shows the absorption spectra of the three species A (black line), B (red line), and C (blue line) calculated by the global analysis by SPECFIT/32. The inset shows the multivariate analysis of the time-resolved spectra with an A→B→C kinetic model using SPECFIT/32, where A, B, and C represent the oxidized enzyme mixed with NADH, the reduced enzyme, and the oxidized enzyme, respectively.

Time-resolved spectra recorded after 3 s showed the gradual increase of the maxima at 355 and 447, which is due to the formation of oxidized enzyme (Figure 8.4 panel A). The stabilization of the C4a-(hydro)peroxyflavin \(^{23, 24}\), the typical intermediate of flavin-dependent monooxygenases, is characterized by a peak in the 360-380 nm region in the UV-visible absorption spectrum of the reduced enzyme reacting with oxygen. In the conditions tested for PA0660 no accumulation of a species with absorbance at 360-380 nm was detected. Changes in absorbance at 500-600 nm which are indicative of the formation of charge-transfer complexes were not detected. A global analysis performed with SPECFIT/32 identified the model A→B→C (Figure 8.4 panel B) as the best fit for the time-resolved spectra. The extracted spectra for species A, B, and C (Figure 8.4 panel B) correspond to oxidized enzyme in presence of NADH, anionic hydroquinone, and fully oxidized enzyme.
Bioinformatic analysis. A BLAST search and multiple sequence alignment of the protein sequence of PA0660 highlighted the presence of six conserved motifs that are listed in Table 8.1.

**Table 8.1:** Conserved motifs in the protein sequence of PA0660.
The numbering of the residues refers to the protein sequence of PA0660; the brackets identify residues that can alternatively be present in that position, *h* represents position occupied by a hydrophobic residue, while *X* represents a position where any amino acid is accepted.

<table>
<thead>
<tr>
<th>Motif</th>
<th>Consensus sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$^{14}$P-X-h-X-[A/S]-P-[M/L]-F-[L/I]-X-S$^{24}$</td>
</tr>
<tr>
<td>II</td>
<td>$^{33}$C-(X)$_3$-h-[G/A]-[S/T/A]-h-P-A-L-N-X-R$^{47}$</td>
</tr>
<tr>
<td>III</td>
<td>$^{98}$X-(h)$_2$-I-[T/S]-S-[L/V]$^{104}$</td>
</tr>
<tr>
<td>IV</td>
<td>$^{115}$H-X-[Y/W]-G-G-X-V-[F/L]-H-D$^{124}$</td>
</tr>
<tr>
<td>V</td>
<td>$^{139}$[V/A]-D-G-(h)$_3$-V-(X)$_2$-G-A-G-H-A-G-(X)$_3$-P$^{158}$</td>
</tr>
<tr>
<td>VI</td>
<td>$^{192}$G-[A/C]-D-h-X-Y-h-G-[T/S]-X-F-[I/L]-(X)$_3$-E$^{207}$</td>
</tr>
</tbody>
</table>

More than 1000 sequences in the non-redundant protein database belonging to hypothetical nitronate monooxygenases carry these six motifs, which are different from the conserved motifs identified in Pa-NMO and in PA1024 for Class 1 NMO (1) and the new class of NADH:ferricyanide reductases (Salvi, F., Ball, J., and Gadda, G., unpublished data). Figure 8.5 shows the example of seven hypothetical nitronate monooxygenases with the conserved motifs identified in PA0660 boxed in green, while the protein sequences of Pa-NMO and PA1024 show the conserved motifs identified in previous studies (1) (Salvi, F., Ball, J., and Gadda, G., unpublished data) boxed in yellow and orange, respectively. The six conserved motifs identified...
in PA0660 (Table 8.1) identify therefore a third and different class of enzymes with diaphorase activity.

Figure 8.5: Multiple sequence alignment of hypothetical nitronate monooxygenases carrying the conserved motifs identified in the protein sequence of PA0660.

The numbering of the residues refers to PA0660 sequence. The conserved motifs of PA0660, listed in Table 8.1, are boxed in green (sequences 1 to 8). Sequences 9 and 10 belong to PA1024 and to Pa-NMO, with the different motifs described for the class of NADH:ferricyanide reductases similar to PA1024 and for NMO Class 1 boxed in orange and yellow, respectively. The sequences identifiers used are: PA0660, *P. aeruginosa* PAO1 (NP_249351.1);
C. algicola, *Cellulophaga algicola* (WP_013550013.1); B. parapertussis, *Bordetella parapertussis* (WP_015040637.1); S. silvestris, *Solibacillus silvestris* (WP_014823224.1); B. dendrobatidis, *Batrachochytrium dendrobatidis* JAM81 (XP_006680703.1); A. anophageffer, *Aureococcus anophagefferens* (XP_009037650.1); E. huxleyi, *Emiliania huxleyi* CCMP1516 (XP_005792771.1); P. hodgsonii, *Pantholops hodgsonii* (XP_005978881.1); PA1024, *P. aeruginosa* PAO1 (NP_249715.1); Pa-NMO, i.e. PA4202 *P. aeruginosa* PAO1 (NP_252891.1).

**8.6 Discussion**

**PA0660 is not a nitronate monooxygenase.** The current gene function prediction of hypothetical nitronate monooxygenases (more than 5,000 genes in the GenBank) is not reliable, because it is based only on modest overall protein sequence similarity with the fungal enzymes Cs-NMO and Nc-NMO, which represent two separate NMO classes with different substrate specificity. Two previous studies addressed the issue of the gene function prediction of hypothetical NMOs and led to the identification of two different classes of enzymes represented by the enzymes Pa-NMO (PA4202) and PA1024 from *P. aeruginosa* PAO1 (Salvi, F., Ball, J., and Gadda, G., unpublished data). Pa-NMO was characterized as nitronate monooxygenase and the conserved motifs identified in the protein sequence of this enzyme are conserved in 500 hypothetical nitronate monooxygenases, establishing Class 1 NMO (Salvi, F., Ball, J., and Gadda, G., unpublished data). The gene product PA1024, currently annotated as nitronate monooxygenase in the GenBank and in the PDB database, was shown to represent instead a NADH:ferricyanide reductase, specific for NADH (Salvi, F., Ball, J., and Gadda, G., unpublished data). Seven conserved motifs identified in the protein sequence of PA1024 are present in more than 1000 hypothetical nitronate monooxygenases that need therefore to be reclassified as hypothetical NADH:ferricyanide reductases (Salvi, F., Ball, J., and Gadda, G., unpublished data).

In this study we cloned, expressed in *E. coli*, and purified the third hypothetical NMO from *P. aeruginosa* PAO1 encoded by the gene pa0660. Despite its annotation as hypothetical nitronate monooxygenase no enzymatic activity was detected with the physiological substrate propionate
3-nitronate or the neutral form 3-nitropropionic acid. The enzyme PA0660 was not able to use nitronates/nitroalkanes of 2 and 3 carbon chain length as substrates. The lack of activity with neutral nitroalkanes suggests that PA0660 is also not a nitroalkane oxidase \(^{(25)}\).

**PA0660 exhibits diaphorase activity.** Anaerobic reduction experiments of PA0660 with NAD(P)H performed with a stopped-flow spectrophotometer showed that the enzyme is highly efficient in using NADPH or NADH as electron donor, with a preference for NADPH, as the second order rate constant \(k_{\text{red}}/K_d\) with NADPH is 34 times higher than with NADH as substrate. This is remarkably different from the gene product PA1024 from *P. aeruginosa* PAO1, which is specific for NADH only (Salvi, F., Ball, J., and Gadda, G., unpublished data), and suggests a different function of PA0660 and PA1024.

PA0660 has a very low oxidase activity and is able to turnover with the electron acceptor DCIP and menadione. The high \(k_{\text{red}}\) of 470 s\(^{-1}\) value compared to the \(k_{\text{cat}}\) of ~8.3 s\(^{-1}\) with DCIP and menadione implies that a kinetic step different than flavin reduction is rate limiting for the turnover of the enzyme. The rate limiting step could be product release or an isomerization of the enzyme-product complex. The steady-state mechanism could not be established, as all experimental points at different concentrations of NADPH were in the saturation area, due to low \(K_m\) values for NADPH with the electron acceptors DCIP and menadione. Indeed the lower concentration of NADPH that could be used was 29 and 44 µM with menadione and DCIP and was still saturating. The fact that the Michaelis-Menten constant \(K_m\) for NADPH is lower that the binding constant \(K_d\) suggests that product release is slow compared to the chemical step \(^{(26)}\), which would be consistent with product release being the rate-limiting step in turnover.

Due to the ability of using both NADH and NADPH as electron donors PA0660 can be classified as diaphorase (E.C. 1.6.99.-) \(^{(27)}\). Examples of FMN-containing enzymes with diaphorase activity
are azoreductases \(^{(21)}\), oxidoreductases classified as old yellow enzymes (OYE) \(^{(28)}\), and flavin-dependent monooxygenases \(^{(29)}\).

No accumulation of the C4a-(hydro)peroxyflavin intermediate was detected which would support monooxygenase activity. The lack of detection of the C4a-(hydro)peroxyflavin intermediate suggests that PA0660 most likely is not a flavin-dependent monooxygenase, even if it is also possible that the C4a-(hydro)peroxyflavin intermediate formed but did not accumulate enough to be detected in the conditions used.

PA0660 did not show any enzymatic activity with the azo dye methyl red, which is used to kinetically characterize azoreductases \(^{(21)}\), and this suggests that PA0660 is not an azoreductase.

The bacterial OYE (YqiM) from *Bacillus subtilis* does not carry the six conserved motifs identified in this study in the protein sequence of PA0660 \(^{(28)}\). However, there are different classes of OYE enzymes with different substrate specificities \(^{(30)}\) and it is possible that PA0660 represents a different prototype of OYE. Different potential substrates will be tested for the ene-reductase activity typical of OYE enzymes \(^{(30)}\). Other flavin-dependent enzymes able to use NADPH and to a lesser extent NADH as electron donors are ferredoxin-NADP\(^+\) reductases (FNRs) \(^{(31)}\) but only FAD has so far been described as cofactors of FNRs from different sources \(^{(31,32)}\).

**More than 1000 hypothetical nitronate monooxygenases possess the conserved motifs identified in PA0660.** The six conserved motifs identified in the protein sequence of PA0660 (Table 8.1) are conserved in more than 1000 sequences of hypothetical nitronate monooxygenases in the non-redundant protein database (Figure 8.5). These sequences belong to bacteria, with the notable exceptions from eukaryotic sources of the fungus *Batrachochytrium dendrobatidis* JAM81, the alga *Aureococcus anophagefferens*, the unicellular phytoplankton
Emiliania huxleyi CCMP1516, the starlet sea anemone *Nematostella vectensis*, and the Tibetan antelope *Pantholops hodgsonii*. This is remarkably different from the two classes of enzymes represented by Pa-NMO and PA1024, which are conserved in bacteria and fungi. Diverse species of bacteria possess hypothetical proteins carrying the six motifs identified in PA0660 and some of them are known pathogens such as *Bordetella parapertussis*, *Bordetella bronchiseptica*, and *Leptospira santarosai*. The fact that this class of enzymes appears to be so widely represented in bacteria and not strictly conserved in fungi and other eukaryotes can be interpreted in two ways: either the function of the enzymes represented by PA0660 is involved in a cellular pathway specific to bacteria metabolism \(^{(33)}\), such as for example peptidoglycan biosynthesis, or the corresponding fungal enzymes evolved to a different subclass of enzymes with different conserved motifs.

There is no crystal structure available for PA0660 and due to the protein sequence identity lower than 30% with templates in the PDB database a homology model of this protein would not be reliable. It is not possible therefore to locate the six conserved motifs identified in this study on the three dimensional structure of PA0660. However the multiple sequence alignment of hypothetical proteins carrying the motifs identified in PA0660, Pa-NMO (PA4202) \(^{(1)}\), and PA1024 \(^{(2)}\) (Salvi, F., Ball, J., and Gadda, G., unpublished data), boxed in green, yellow, and orange in Figure 8.5, highlights that three conserved motifs occupy the same position in the alignment. Specifically motif I corresponds to motif I in PA1024 and Pa-NMO, the end of motif III corresponds to motif VII in PA1024 and II in Pa-NMO, and motif V corresponds to motif IV in PA1024 and III in Pa-NMO. These three locations, that will be referred as locations A, B, and C, with different consensus sequences in the multiple sequence alignment are marked with a red asterisks and labeled in Figure 8.5. In the case of PA1024 and Pa-NMO the comparison of the
two crystal structures (PDB codes 4Q4K \(^{(1)}\) and 2GJL \(^{(2)}\)) shows that the three locations A, B, and C overlay also in the crystal structure, specifically in the TIM barrel domain (FMN-binding domain). Location A contacts the face of the cofactor not exposed to the active site cavity, location B is near the pyrimidine moiety of the isoalloxazine ring, and location C is on the surface above the active site with a fully conserved histidine residue (H183 in Pa-NMO and H152 in PA1024) projecting in the active site cavity. Despite the lack of structural information on PA0660 it is possible to hypothesize that it possesses a TIM barrel domain and that motifs I, III, and V identified in this study are located similarly to Pa-NMO and PA1024 (locations A, B, and C). In this case the scaffold of the TIM barrel domain appears to be recycled by nature for three FMN-containing enzymes catalyzing different reactions and the side chains of residues in the positions A, B, and C are important for protein function. The consensus sequences identified for location A, B, and C in the three different classes of enzymes represented by PA0660, PA1024, and Pa-NMO can be relevant for both uses in multiple sequence alignments for gene function prediction and for enzyme design to highlight structural elements to alter enzyme function.

8.7 Conclusions

The biochemical characterization of the hypothetical NMOs from *P. aeruginosa* PAO1 has been instrumental for the improvement of the gene function prediction of more than 5000 hypothetical NMOs in the GenBank. According to the current annotation in the GenBank many organisms possess multiple genes annotated as NMO, which could be either the consequence of unreliable gene function prediction or of the presence of different isoforms of NMO enzymes in the same organism. In the case of *P. aeruginosa* PAO1 the three genes annotated as NMO are *pa4202*, *pa1024*, and *pa0660*. In this study we cloned, expressed in *E. coli*, and purified the gene product
PA0660, one of the hypothetical NMOs from \textit{P. aeruginosa} PAO1. The kinetic characterization of the recombinant enzyme showed the lack of nitronate monooxygenase activity. PA0660 exhibited instead diaphorase activity with preference for NADPH over NADH. Six conserved motifs were identified in the protein sequence of PA0660, which are conserved in more than 1000 hypothetical NMOs belonging mostly to bacteria that need to be reclassified. A previous study on the hypothetical NMOs from \textit{P. aeruginosa} PAO1 confirmed that \textit{pa4202} encodes for a nitronate monooxygenase, Pa-NMO, and identified four conserved motifs that are not present in PA1024 and PA0660 and identify Class 1 NMO, with 500 protein sequences in the non-redundant protein database \textsuperscript{(1)}. The gene product PA1024 was reclassified as NADH:DCIP reductase specific only for NADH, with seven motifs in the protein sequence conserved in more than 1000 hypothetical NMOs from bacteria and fungi (Salvi, F., Ball, J., and Gadda, G., unpublished data).

This study complements the two previous studies on the improvement of the gene function prediction of hypothetical NMOs \textsuperscript{(1)} (Salvi, F., Ball, J., and Gadda, G., unpublished data) and identifies a new class of bacterial enzymes with diaphorase activity. Furthermore the consensus sequence in three locations in the multiple sequence alignment is suggested to be important for enzyme function.

\subsection*{8.8 References}

\textsuperscript{a} The use of the extinction coefficient of free FMN may result in an inaccuracy of up to 25\% in the determination of steady-state kinetic parameters. This does not alter the main conclusion that PA0660 exhibits diaphorase activity instead of nitronate monooxygenase activity.


CHAPTER IX: EXPRESSION, PURIFICATION AND INITIAL CHARACTERIZATION OF THE ESSENTIAL GENE PRODUCT HP0773 FROM *Helicobacter pylori*

### 9.1 Abbreviations

PCR, polymerase chain reaction; IPTG, isopropyl β-D-1-thiogalactopyranoside; PMSF, phenylmethanesulfonylfluoride; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; PMS, phenazine methosulfate; DCIP, 2,6-dichlorophenolindophenol; P3N, propionate 3-nitronate.

### 9.2 Abstract

The infection by the human pathogen *Helicobacter pylori* affects more than 50% of the world population and leads to severe consequences ranging from chronic gastritis and ulcers to an increased risk for gastric cancer. The increased antibiotic resistance of this pathogen towards the current treatments sets the need for the characterization of new drug targets for *H. pylori*. The gene *hp0773* has been identified as a candidate essential gene for this pathogen and the predicted gene function is nitronate monooxygenase, an enzyme involved in the detoxification of the lethal toxin propionate 3-nitronate. The gene *hp0773* was cloned, and the recombinant HP0773 expressed in *Escherichia coli*, and purified. The initial characterization of the purified protein showed that the enzyme contains FMN and iron and is not a nitronate monooxygenase. Furthermore other potential substrates were tested and a homology model generated. Further studies are necessary to establish the function of HP0773.
9.3 Introduction

*Helicobacter pylori*, previously known as *Campylobacter pylori*, is a human pathogen found mainly in the stomach, which affects more than 50% of the world population\(^{(1)}\). The infection by *H. pylori* has been recognized as contributing to chronic gastritis and gastric ulcers, and recently associated to an increased risk of stomach cancer\(^{(2)}\). In some cases the infection by *H. pylori* is asymptomatic\(^{(1)}\), but it still needs to be eradicated to avoid the increased risk for gastritis and gastric cancer\(^{(3)}\). The current therapy consists of an association of proton pump inhibitors with antibiotics containing metronidazole and tetracycline, but antibiotic resistance of *H. pylori* towards these compounds is increasing dramatically\(^{(2, 4)}\). It is therefore essential to identify new drug targets for the eradication of this important pathogen\(^{(2)}\). A study on the entire genome of *H. pylori* with transposon analysis identified 344 candidate essential genes (23\% of the entire genome)\(^{(5)}\). One of these candidate essential genes is *hp0773*, which is present also in the target list of the Center for Structural Genomics of Infectious Diseases for drug targets against *H. pylori*. According to the NCBI gene database\(^{(6)}\), *hp0773* is predicted to encode for a nitronate monooxygenase (NMO; E.C. 1.13.12.16). Nitronate monooxygenases, known as 2-nitropropane dioxygenases until 2010, are FMN-dependent enzymes involved in the detoxification of the metabolic poison propionate 3-nitronate (P3N; Scheme 9.1)\(^{(7)}\). P3N is a potent irreversible inhibitor of succinate dehydrogenase, which hampers energy production in the mitochondria, leading to neurological disorders and possibly death\(^{(7)}\).

\begin{align*}
\text{P3N} & \xrightarrow{\text{NMO}} \text{MSA} \rightarrow \text{Acetyl-CoA} \\
\text{C} & \text{N} \\
\text{O} & \text{O} \\
\text{H} & \\
\text{C} & \text{C} \quad \text{O} \\
\text{H}_2 & \\
\text{O} & \text{O} \\
\text{C} & \text{H}_2 \\
\text{O} & \text{S} \\
\text{C} & \text{O} \\
\text{A} & \text{CoA} \\
\end{align*}

**Scheme 9.1:** Detoxification of the mitochondrial toxin P3N by the FMN-dependent NMO to malonic semialdehyde (MSA), which can be converted to acetyl-CoA\(^{(8)}\).
This lethal toxin is produced by plants and fungi as a defense mechanism against herbivores, and many intoxication cases of humans and livestock have been reported \(^7\). Fungal nitronate monooxygenases have been kinetically characterized and oxidize P3N via a single electron transfer, \(^9\) but very little is known on the bacterial enzymes. Some bacteria, among which \textit{Pseudomonas aeruginosa}, can use P3N as a sole carbon and nitrogen source \(^8\), and a hypothetical NMO from \textit{Mycobacterium tuberculosis} has been associated with the virulence of this pathogen in guinea pigs \(^12\).

In this study we cloned the gene \textit{hp0773} from the genomic DNA of \textit{H. pylori} 26695, we expressed the recombinant His-tagged protein HP0773 in \textit{Escherichia coli} and we purified it by affinity chromatography. The initial characterization of purified HP0773 showed the presence of FMN and iron as cofactors and does not detect nitronate monooxygenase activity. Other putative substrates were tested and a homology model of HP0773 is presented.

### 9.4 Materials and methods

**Materials.** The genomic DNA of \textit{Helicobacter pylori} strain 26695 was from ATCC (Manassas, VA) and \textit{Escherichia coli} strain DH5α was purchased from Invitrogen Life Technologies (Grand Island, NY). The enzymes \textit{BamHI}, \textit{NdeI}, \textit{DpnI}, and T4 DNA ligase were from New England BioLabs (Ipswich, MA), \textit{Pfu} DNA polymerase from Stratagene (La Jolla, CA), and oligonucleotides from Sigma Genosys (The Woodlands, TX). Calf intestinal alkaline phosphatase (CIAP) was from Promega (Madison, WI), and HiTrapTM Chelating HP 5 mL affinity column was from GE Healthcare (Piscataway, NJ). QIAprep Spin Miniprep kit and QIAquick PCR purification kit were from Qiagen (Valencia, CA). BugBuster Master Mix and \textit{E. coli} strains Rosetta(DE3)pLysS were from Novagen (Madison, WI). All other reagents used were of the highest purity commercially available.
Cloning of HP0773 from genomic DNA. All purification steps in the cloning procedures were performed with the QIAGen PCR purification kit, unless specified otherwise. The gene hp0773 was amplified by PCR with an initial denaturation step at 95 °C, followed by 16 cycles of denaturation for 30 seconds at 95 °C, annealing for 1 minute at 55 °C, and extension for 10 minutes at 68 °C. The primers containing the restriction sites for the enzymes NdeI and BamHI are listed in Table 9.1.

Table 9.1: Oligonucleotide primers used for PCR amplification of hp0773. The restriction sites for NdeI and BamHI enzymes underlined, and oligonucleotide primers for site-directed mutagenesis of R196R with the mutation underlined.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP0773for</td>
<td>GGAATTCCATATGGTATCAACACTCAAACCG</td>
</tr>
<tr>
<td>HP0773rev</td>
<td>CGCGGATCCCTAACCCTCTGTAAGCTCTTT</td>
</tr>
<tr>
<td>R196Rfor</td>
<td>GTTTCAAAGAAGAATTCCGTGTAGAAAAAAGTTAGTG</td>
</tr>
<tr>
<td>R196Rrev</td>
<td>GGCACTAAGTTTCTAAACCGAATTCTCTTTGAAAC</td>
</tr>
</tbody>
</table>

The amplification of the primers by Pfu DNA polymerase was checked by agarose gel electrophoresis and the PCR products were purified and subjected to a double digestion by NdeI and BamHI at 37°C for 2 h and further purified. The ampicillin-resistant vector pET15b was digested with NdeI and BamHI at 37°C for 2 h and purified prior to be processed with 1 unit of calf intestinal alkaline phosphatase (CIAP) at 37 °C for 30 min. After a purification step with the PCR purification kit the cut and dephosphorylated vector pET15b was ligated to the insert with incubation for 15 h at 16 °C with T4 DNA ligase. A volume of 3 μL of the ligation mixtures was used to transform E. coli strain DH5α and the resulting colonies on Luria-Bertani agar plates containing 50 μg/mL ampicillin were screened for the presence of the desired insert by DNA
sequencing at the GSU core facility. The DNA sequencing confirmed the correct insertion of the gene in the vector pET15b and the absence of undesired mutations in the gene hp0773.

**Site-directed mutagenesis of the rare codon cga.** The mutagenic primers R196Rfor and R196Rrev listed in Table 9.1 were amplified by PCR in presence of 2% v/v of DMSO using as template the construct pET15b/hp0773 prepared as described above. After an initial denaturation step at 95 °C for 2 min, 25 cycles were performed of denaturation at 95 °C for 45 s, annealing at 55 °C for 45 s (the annealing temperature decreases of 0.2 °C each cycle), and extension at 72 °C for 10 min, followed by a final step at 72 °C for 10 min. The PCR product was purified and digested with DpnI for 2 h to ensure complete removal of the template pET15b/hp0773, prior to transformation into *E. coli* strain DH5α. The resulting colonies on Luria-Bertani agar plates containing 50 μg/mL ampicillin were screened for the presence of the desired mutation by DNA sequencing at the GSU core facility.

**Recombinant protein expression and purification.** The construct pET15b/hp0773/R196R designed to insert an His-tag at the N terminal of the protein was used to transform *E. coli* strain Rosetta(DE3)pLysS and single colonies were used to inoculate 1 L of Terrific Broth containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol. The culture was incubated at 25 °C until it reached an optical density at 600 nm of 0.6. IPTG was then added to a final concentration of 100 μM and the culture was incubated at 25 °C for 20 h. The wet cell paste obtained by centrifugation was resuspended in the lysis buffer (50 mM TRIS-Cl, pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5mM β-mercaptoethanol, 1 mM PMSF, 5 mM MgCl2, 2 mg/mL lysozyme, 5 μg/mL RNase and 5 μg/mL DNase) and subjected to several cycles of sonication. The cell free extract obtained by centrifugation at 10,000 g for 20 min at 4 °C was loaded onto a HiTrapTM Chelating HP 5 mL affinity column equilibrated in buffer A (50 mM TRIS-Cl, pH 8.0, 300 mM
NaCl, 25 mM imidazole, 10% v/v glycerol). After washing the column with 50 mL of buffer A the recombinant protein was eluted with 20% of buffer B (buffer A + 500 mM imidazole). All purification steps were carried out at 4 °C. Samples of the different steps (crude extract, cell free extract, flow through, wash and elution) for SDS-PAGE analysis were prepared by boiling them in sample buffer (60 mM TRIS-Cl, pH 6.6, 10% v/v glycerol, 5% w/v SDS, 5% v/v β-mercaptoethanol, 0.01% w/v bromophenol blue). The fractions eluted from the affinity column were pooled based on the purity estimated by SDS-PAGE analysis and dialyzed against 20 mM HEPES-Na, pH 8.0, 300 mM NaCl, 10% v/v glycerol. UV-visible absorbance was recorded with an Agilent Technologies diode-array spectrophotometer Model HP 8453 PC equipped with a thermostated water bath.

Identification of cofactors. The flavin cofactor of purified HP0773, passed through a PD-10 column equilibrated in water, was extracted by treatment with trichloroacetic acid at a final concentration of 10% v/v for 1 h and 40 min on ice. The protein was removed by centrifugation and the pH of the supernatant was adjusted to 7.0 by addition of NaOH. The extracted cofactor was purified by HPLC using a Shimadzu SPD-M10A system with a C18 reversed-phase column (5µ, 15 cm x 4.6 mm) equilibrated in 5% acetonitrile in water with 0.1% trifluoroacetic acid. A gradient from 5% to 100% acetonitrile in 1 h at 0.5 mL/min was applied, and the flavin cofactor eluted with a retention time of 21.8 min. The identification of the flavin cofactor was performed by mass spectrometry using a Waters Micromass Q-TOF micro (ESI-Q-TOF) in negative ion mode at the Mass Spectrometry Facility of Georgia State University.

For the metal content determination the purified HP0773 was incubated with EDTA at a final concentration of 18 mM on ice in the dark for 1 h in order to remove loosely bound metals, as described for dihydropyrimidinedehydrogenase (13). The protein solution was passed through a
PD-10 column equilibrated in 50 mM HEPES-Na, pH 8.0, 100 mM NaCl, 10% v/v glycerol, diluted to a final concentration of 3.6 μM and filtered. Total protein concentration was determined with the Bradford method using bovine serum albumin as standard\(^{(14)}\). Metal content determination was carried out by inductively coupled plasma optical emission spectrometry (ICP-OES) analysis at the core facility in the Chemistry Department of Georgia State University. A sample of the buffer was also processed in the same way and analyzed in order to subtract the metal background of the buffer.

**Activity assays.** In all cases the enzymatic activity was normalized by the concentration of the enzyme-bound flavin calculated using the extinction coefficient of free FMN \(\epsilon_{450}=12,500\) M\(^{-1}\) cm\(^{-1}\) \(^{(15)}\). For the nitronate monooxygenase activity assay the enzymatic activity was monitored by following the initial rate \(^{(16)}\) of oxygen consumption with a Hansatech Instruments computer-interfaced Oxy-32 oxygen-monitoring system after addition of the enzyme and of the nitroalkane or nitronate stock solution, as previously described \(^{(10, 11)}\). Stock solutions of nitronates and nitroalkanes were prepared as previously described \(^{(10, 11)}\) and the assay was carried out at 30 °C and atmospheric oxygen (i.e. 0.23 mM oxygen). The activity of the cell free extract and of the purified enzyme (with a final concentration of 0.4 μM) was tested with 10 mM 3-nitropropionic acid, 10 mM propionate 3-nitronate, 20 mM nitroethane, ethynitronate, 2-nitropropane, or propyl-2-nitronate in 20 mM sodium pyrophosphate, pH 7.0, and in 20 mM piperazine pH 5.5. The dihydroorotate dehydrogenase activity was tested as previously described for Class 1 \(^{(17)}\) and Class 2 \(^{(18)}\) dihydroorotate dehydrogenases. Briefly, purified HP0773 to a final concentration of 1.4 μM was added to a solution of 50 μM dihydroorotate, and 800 μM fumarate, in 50 mM HEPES-Na, pH 7.0, 10% v/v glycerol, 100 mM NaCl at 25 °C and atmospheric oxygen. The formation of orotate was followed spectrophotometrically by monitoring the increase in
absorbance at 300 nm \((\varepsilon = 2.65 \text{ mM}^{-1} \text{ cm}^{-1})\)^{(17)}\). Alternatively, purified HP0773 to a final concentration of 1.4 \(\mu\text{M}\) was added to a solution of 500 \(\mu\text{M}\) dihydroorotate, 10 \(\mu\text{M}\) 2,6-dichlorophenolindophenol (DCIP), 1 \(\mu\text{M}\) methyl viologen in 50 mM HEPES-Na, pH 7.0, 10% v/v glycerol, 100 mM NaCl at 25 °C and atmospheric oxygen. The enzymatic reduction of DCIP was monitored by following the decrease in absorbance at 610 nm \((\varepsilon = 21.5 \text{ mM}^{-1} \text{ cm}^{-1})\)^{(18)}.

Dihydropyrimidinase activity was tested as previously described\(^{(19)}\) by adding purified enzyme to a final concentration of 1.5 \(\mu\text{M}\) to a solution containing 0.05 to 10.0 \(\text{mM}\) dihydrouracil, and 50 to 100 \(\mu\text{M}\) NAD\(^+\) at 25 °C and atmospheric oxygen. The production of NADH was monitored by following the absorbance at 340 nm. The assay was performed in 50 mM sodium pyrophosphate, pH 11.2, and also in 50 mM HEPES-Na, pH 7.0, 10% v/v glycerol, 100 mM NaCl. The same activity assay was repeated by using 50 \(\mu\text{M}\) NADP\(^+\) instead of NAD\(^+\).

The conversion of uracil to dihydrouracil\(^{(19)}\) was tested by following NAD(P)H consumption with the decrease in absorbance at 340 nm in 20 mM HEPES-Na, pH 8.0, 10% v/v glycerol, 100 mM NaCl. The reaction was started by adding purified enzyme (in a concentration range from 0.4 to 3 \(\mu\text{M}\)) to a solution of uracil and NAD(P)H. Concentrations of uracil and NAD(P)H ranged from 75 to 150 \(\mu\text{M}\), and from 35 to 65 \(\mu\text{M}\), respectively. Spectral changes of the enzyme-bound flavin cofactor upon addition of dihydrouracil (1 mM), uracil (90 \(\mu\text{M}\)), dihydroorotate (1 mM), or NAD(P)H (100 \(\mu\text{M}\)) were monitored spectrophotometrically at 25 °C and atmospheric oxygen in 1 mL solution of 36 \(\mu\text{M}\) purified HP0773 in 20 mM HEPES-Na, pH 7.0, 10% v/v glycerol, 100 mM NaCl over 10 min.

Reduction of the flavin cofactor by NAD(P)H was monitored spectrophotometrically by following the decrease at 340 nm at 25 °C and atmospheric oxygen, in 20 mM HEPES-Na, pH 8.0, 150 mM NaCl and using final concentrations of 116 \(\mu\text{M}\) NADH, or 96 \(\mu\text{M}\) NADPH and 1
μM enzyme. This experiment was also repeated in 50 mM TRIS-Cl, pH 8.0, 10% v/v glycerol, 5 mM β-mercaptoethanol with final concentrations of 117 μM NADH, or 64 μM NADPH, and 20 μM enzyme, and in 20 mM piperazine pH 5.3, with final concentrations of 25 μM NADH, or 25 μM NADPH, and 10 μM enzyme.

The ability of purified HP0773 to oxidize xanthine, trimethylamine, L-arginine, glycine, and choline chloride was tested by using PMS as artificial electron acceptor in a coupled assay with oxygen, as previously described for D-arginine dehydrogenase (20). Briefly the enzymatically reduced PMS is oxidized by oxygen and oxygen consumption is monitored with a Hansatech Instruments computer-interfaced Oxy-32 oxygen-monitoring system. The assay was carried out by adding the enzyme to a final concentration of 0.3 μM to a solution of 1 mL containing 1 mM substrate, and 10 mM PMS, in 50 mM potassium phosphate, pH 7.0 at 25 °C and atmospheric oxygen.

Model of the three dimensional structure. A three dimensional model of HP0773 was constructed with the web-based SWISS-MODEL (21), using as template the crystal structure of Streptococcus pneumoniae enoyl-acyl carrier protein reductase (FabK), PDB code 2Z6I.

9.5 Results

Recombinant protein expression and purification. The gene sequence of hp0773 codes for a protein of 363 amino acids with a predicted molecular weight of 40 kDa. The open reading frame for the gene hp0773 contains 19 rare codons for recombinant expression in E. coli, according to the web-based Rare Codon Calculator (RaCC) of the NIH MBI Laboratory for Structural Genomics and Proteomics. The E. coli strain Rosetta(DE3)pLysS is engineered to provide tRNA for the rare codons (22) present in hp0773, with the exception of the rare codon cga for arginine. This rare codon was therefore removed by site-directed mutagenesis and the recombinant
HP0773 carrying an His-tag at the N terminal end of the protein was overexpressed in the E. coli strain Rosetta(DE3)pLysS and purified by affinity chromatography. Figure 9.1 shows the SDS-PAGE analysis of the cell free extract and flow through, followed by the fractions eluted from the affinity column. Fractions 4, 5, 6, and 7 were pooled and dialyzed against 20 mM HEPES-Na, pH 8.0, 300 mM NaCl, 10% v/v glycerol.

Purified HP0773 showed a UV-visible absorption spectrum with maxima at 369 and 422 nm (Figure 9.2), which is consistent with the presence of a flavin cofactor.
Figure 9.2: UV-visible absorption spectrum of purified HP0773 in 50 mM HEPES-Na, pH 8.0, 10% v/v glycerol, 300 mM NaCl at 25 °C.

The cofactor of purified HP0773 was extracted by treatment with trichloroacetic acid, purified by HPLC and analyzed by ESI mass spectrometry in negative ion mode (Figure 9.3). This analysis detected a peak with m/z value of 455.0, which is consistent with the presence of FMN rather than FAD. The FMN cofactor is completely released to solution after denaturation of the enzyme by trichloroacetic acid, indicating it is not covalently bound to the protein.

Figure 9.3: Mass spectrometric analysis of the extracted flavin cofactor of HP0773.
Metal content determination was performed on purified HP0773 by inductively coupled plasma optical emission spectrometry (ICP-OES). This analysis identified iron as cofactor of HP0773, with a ratio of iron to protein of 1.36 (Table 9.2). The other metals tested, namely copper, nickel, zinc, magnesium, cadmium, and cobalt, were present only in negligible amounts (Table 9.2).

Table 9.2: Inductively coupled plasma optical emission spectrometry (ICP-OES) of purified HP0773.

<table>
<thead>
<tr>
<th>Metal</th>
<th>Concentration (μM)</th>
<th>ratio metal/protein concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe++</td>
<td>4.9</td>
<td>1.361</td>
</tr>
<tr>
<td>Cu++</td>
<td>0.13</td>
<td>0.036</td>
</tr>
<tr>
<td>Ni++</td>
<td>0.07</td>
<td>0.019</td>
</tr>
<tr>
<td>Zn++</td>
<td>0.07</td>
<td>0.019</td>
</tr>
<tr>
<td>Mg++</td>
<td>0.08</td>
<td>0.022</td>
</tr>
<tr>
<td>Cd++</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Co++</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>

a nd, not detected, as the metal content in the sample of purified HP0773 is the same as in the buffer.

Nitronate monooxygenase activity assay. The nitronate monooxygenase activity of the cell free extract and of purified HP0773 was tested as previously described for the well characterized nitronate monooxygenases from C. saturnus (10) and Neurospora crassa (23). No enzymatic activity was detected with 10 mM propionate 3-nitronate (P3N) and 3-nitropropionic acid (3NPA) at pH 5.5 and pH 7.0, as shown by the lack of oxygen consumption (Figure 9.4). No enzymatic activity was detected with 20 mM nitroethane, or ethyl nitronate, neither with 20 mM 2-nitropropane, or propyl-2-nitronate.
Other activity assays. The NADPH oxidase activity of purified HP0773 was tested by monitoring the decrease in the absorbance of NAD(P)H at 340 nm. This assay was performed with different buffer conditions at pH values of 5.3 and 8.0, and different final concentrations of enzyme, NADH, and NADPH were tested. No decrease of the absorbance at 340 nm was detected in 20 minutes in any of the conditions tested. Figure 9.5 shows the assay performed with 96 μM NADPH and 1 μM enzyme at 25 °C and pH 8.0.
The black line represents the UV-visible absorption spectrum of a solution of 96 μM NADPH, and the blue and red lines represent the UV-visible absorption spectra of a solution of 96 μM NADPH immediately after mixing with HP0773 to a final concentration of 1 μM and after 10 minutes, respectively.

Flavin reduction by NAD(P)H was also tested aerobically by monitoring the decrease in absorbance at 422 nm of HP0773 36 μM mixed with excess NADH and NADPH (100 uM) at pH 7.0. No decrease in absorbance at 422 nm was observed.

The dihydroorotate dehydrogenase activity of HP0773 was tested by monitoring orotate formation and with a DCIP colorimetric assay, as previously described for dihydroorotate dehydrogenases Class 1 and 2 (17, 18), and no enzymatic activity was detected. Dihydropyrimidine dehydrogenase activity (19, 24) was assayed with dihydrouracil and NAD(P)H by monitoring the production of NAD(P)H with the increase in absorbance at 340 nm and with uracil and NAD(P)H by monitoring NAD(P)H consumption with the decrease in absorbance at 340 nm. Both uracil and dihydrouracil did not act as substrates for purified HP0773.

The ability of HP0773 to oxidize xanthine, trimethylamine, L-arginine, glycine, and choline chloride was tested using PMS as artificial electron acceptor, as previously described for D-arginine dehydrogenase (20). No enzymatic activity was detected with any of these compounds.

**Model of the three dimensional structure.** A homology model of the three dimensional structure of HP0773 was generated with the web-based SWISS-MODEL server (21), using as template the crystal structure of *Streptococcus pneumoniae* enoyl-acyl carrier protein reductase (FabK), PDB code 2Z6I (25). The template chosen displayed the highest protein sequence identity with HP0773 among the proteins available in the PDB database. However, the sequence identity of the template FabK (PDB code 2Z6I) with HP0773 is only 23 % and the global model quality estimation by SWISS-MODEL provides a QMEAN4 score of 0.4 in a scale from 0 to 1.0 with 1.0 representing the maximum reliability of the model (21). Figure 9.6 shows the overall structure
of the homology model of HP0773, which displays a TIM barrel fold, with the FMN cofactor shown from the superposition with the template FabK (PDB code 2Z6I).

**Figure 9.6**: Homology model of HP0773. Panel A represents the overall structure of the homology model of HP0773 shown as gray cartoon, with the FMN cofactor from the template FabK (PDB code 2Z6I) in yellow sticks. Panel B shows a view of the modeled active site of HP0773 with some residues predicted to be located in the protein environment surrounding the FMN cofactor highlighted as gray sticks and labeled.

Based on the homology model H182 is predicted to be in the active site, while M24 is located on the opposite face of FMN. The protein environment of FMN in the homology model is mainly hydrophobic, with the side chains of A126, V26, and V49 constellating the proposed active site.

**9.6 Discussion**

*HP0773 contains FMN and iron.* A transposon analysis of the genome of *H. pylori* has included the gene *hp0773* as potentially essential for the growth of this pathogen on solid media (5). The protein sequence is not conserved in human and is predicted to be a nitronate monooxygenase based on low overall protein sequence similarity to fungal NMOs (6). The fact that *hp0773* is an essential gene for *H. pylori* not conserved in humans renders the gene product
HP0773 an attractive drug target for antimicrobial agents. The gene *hp0773* was cloned from genomic DNA, expressed as a recombinant His-tagged protein in *E. coli*, and purified by affinity chromatography. SDS-PAGE analysis of the purified protein is consistent with the predicted molecular weight of 40 kDa for the gene product HP0773. The recombinant purified HP0773 contains FMN non covalently bound, as confirmed by mass spectrometric analysis of the extracted cofactor. The UV-visible absorption spectrum shows maxima at 359 and 422 nm, which are characteristic of flavin-dependent enzymes. However, the low energy band of the flavin cofactor at 422 nm is significantly blue shifted from the 450 nm of free FMN. This was already observed in the case of dihydropyrimidine dehydrogenase from *Alcaligenes eutrophus*, which contains iron as second cofactor \(^{(24)}\). The metal content determination performed on HP0773 by inductively coupled plasma optical emission spectrometry (ICP-OES) revealed the presence of iron in a 1.3 ratio to the FMN-bound enzyme, which suggests the presence of a 2Fe-2S cluster in HP0773. Interestingly nitronate monooxygenases characterized so far do not contain iron as a cofactor \(^{(7, 9)}\). Other enzymes reported to contain flavin and iron are dihydropyrimidine dehydrogenase \(^{(24)}\), trimethylamine dehydrogenase \(^{(26)}\), xanthine dehydrogenase \(^{(27)}\), glutamate synthase \(^{(28)}\), histamine dehydrogenase \(^{(29)}\), an iron-sulfur flavoprotein (Isf) from *Methanosarcina thermophila* \(^{(30)}\), sulfide dehydrogenase \(^{(31)}\), aldehyde oxidase \(^{(32)}\), fumarate reductase \(^{(33)}\), 2-hydroxyisonicotinate dehydrogenase \(^{(34)}\), adenylylsulfate reductase \(^{(35)}\), \(\Delta 1\)-pyrroline-4-hydroxy-2-carboxylate deaminase \(^{(36)}\), [NiFe] hydrogenase \(^{(37)}\), 2,4-dienoyl-CoA reductase \(^{(38)}\), the tungsten-containing formate dehydrogenase from *Methyllobacterium extorquens* \(^{(39)}\), nitric oxide synthase \(^{(40)}\), the NAD-linked hydrogenase of *Nocardia opaca* \(^{(41)}\), and NADH dehydrogenase \(^{(42)}\).
**HP0773 is not a nitronate monooxygenase.** Despite its annotation as nitronate monooxygenase no oxygen consumption of HP0773 was detected at pH 5.5 or 7.0 with 10 mM 3-nitropropionic acid or propionate 3-nitrate, and with 20 mM nitroethane, ethyl nitronate, 2-nitropropane, or propyl-2-nitronate as substrates, using the enzymatic assay previously described for fungal NMOs (10, 23). This lead to the conclusion that HP0773 is most likely not a nitronate monooxygenase, as already suggested by the presence of iron as a cofactor, and that the gene function prediction in the NCBI gene database is not reliable for this gene.

**Other activity assays.** A protein BLAST search (43) of the amino acid sequence of HP0773 against the Swissprot database returned 86 hits (Figure 9.7) with the alignment covering only a portion of the protein sequence (typically 100 to 200 amino acids) and characterized by a low score (<80).

The first 12 hits, 9 with an alignment score between 50 and 80 and 3 with an alignment score between 40 and 50, belong to hypothetical nitronate monooxygenases from bacteria and fungi. The protein sequence identity of HP0773 with these proteins varies between 23% and 34%. However, as previously discussed, HP0773 does not display nitronate monooxygenase activity.
Figure 9.7: Protein BLAST search of the amino acid sequence of HP0773 against the Swissprot database. As taken from the protein BLAST website; NMO, nitrate monooxygenase; 5-IMP, inosine-5’-monophosphate; DHODH, dihydroorotate dehydrogenase.

The remaining hits, with the alignment score lower than 40 and protein sequence identity ranging from 26% to 47%, include mainly inosine-5’-monophosphate dehydrogenase (E.C. 1.1.1.205), and Class 2 dihydroorotate dehydrogenase (E.C. 1.3.5.2) from different sources. The sequence alignment covers only portions of the protein sequence of HP0773 with a low score, suggesting it is only distantly related with these proteins. While inosine-5’-monophosphate dehydrogenase, which catalyzes the conversion of inosine-5’-monophosphate to xanthosine-5’-monophosphate in the guanine nucleotide biosynthesis does not contain any cofactor \(^{(44)}\), dihydroorotate dehydrogenase Class 1B contains FMN and iron \(^{(45)}\). Dihydroorotate dehydrogenase enzymes
catalyze the conversion of dihydroorotate to orotate in the pyrimidine biosynthetic pathway and are classified in Class 1 and Class 2 (Scheme 9.2) \cite{46}.

\[
\begin{align*}
\text{DHO} & \quad \text{OROTATE} \\
\text{URACIL} & \quad \text{DHU} \\
\text{TMA} & \quad \text{DMA} \quad \text{FMA} \\
\text{XANTHINE} & \quad \text{URATE}
\end{align*}
\]

Scheme 9.2: Potential substrates tested for the enzymatic activity of HP0773. DHO, dihydroorotate; DHU, dihydrouracil; TMA, trimethylamine; DMA, dimethylamine; FMA, formaldehyde.

Class 1 dihydroorotate dehydrogenases are soluble enzymes that utilize fumarate (Class 1A) or NAD$^+$ (Class 1B) as electron acceptor \cite{46}, while Class 2 dihydroorotate dehydrogenases are membrane-bound enzymes that use ubiquinone as electron acceptor \cite{47}. Dihydroorotate dehydrogenase activity assays developed for Class 1 by monitoring orotate formation by the increase in absorbance at 300 nm \cite{17} and for Class 2 by using a colorimetric assay based on DCIP reduction \cite{18} were performed on purified HP0773, but no enzymatic activity was detected.
Other enzymes reported in literature as containing flavin and iron as cofactors include dihydropyrimidine dehydrogenase (E.C. 1.3.1.2), trymethylamine dehydrogenase (E.C. 1.5.8.2), and xanthine dehydrogenase (E.C. 1.17.1.4). Dihydropyrimidine dehydrogenases catalyze the conversion of uracil to dihydrouracil in the pathway of reductive degradation of pyrimidines, but in certain conditions they can also catalyze the reverse reaction of oxidation of dihydrouracil to uracil (Scheme 9.2) \(^{19}\). Dihydropyrimidine dehydrogenase activity of HP0773 was tested as previously described \(^{19}\) by monitoring consumption/production of NAD(P)H with changes in absorbance at 340 nm. The lack of detection of enzymatic activity showed that HP0773 is most likely not a dihydropyrimidinone dehydrogenase.

Trymethylamine dehydrogenase catalyzes the oxidative demethylation of trimethylamine to dimethylamine and formaldehyde (Scheme 9.2), using as electron acceptor an electron-transferring flavoprotein (ETF) \(^{26}\), while xanthine dehydrogenase catalyzes the oxidation of xanthine to urate in the purine metabolism (Scheme 9.2) \(^{27}\). HP0773 was not able to use trimethylamine and xanthine, as well as glycine, L-arginine, and choline, as substrates using PMS as artificial electron acceptor in a coupled assay that monitors oxygen consumption by the enzymatically reduced PMS, previously described for D-arginine dehydrogenase \(^{20}\).

HP0773 did not display NAD(P)H oxidase activity at pH 7.0 and pH 5.3, as shown by the lack of decrease in absorbance of 340 nm of a solution of NAD(P)H and purified HP0773 at 25 °C and atmospheric oxygen. The FMN cofactor of HP0773 was not able to use NAD(P)H as electron source as shown by the lack of flavin reduction upon mixing of the enzyme solution with NAD(P)H at 25 °C and atmospheric oxygen.

In summary it was not possible to assign any enzymatic activity to HP0773 with the potential substrates tested in this study (Scheme 9.2). The non-reliable function prediction and the poor
sequence similarity of this protein do not give any strong lead to hypothetical functions for HP0773.

Furthermore the genomic context of the gene *hp0773* is composed mainly of genes with hypothetical functions, therefore not giving useful insights about the possible role of this protein in cellular pathways. The gene *hp0773* is preceded by the gene *hp0774* and followed by the gene *hp0772*. The gene function prediction of gene *hp0774* is tyrosyl-tRNA synthetase (E.C. 6.1.1.1), which catalyzes the formation of the covalent linkage between tyrosine and its specific tRNA (48), while the gene *hp0772* (*AmiA*), predicted to encode for a N-acetylmuramoyl-L-alanine amidase (48), was shown to be required for the morphological transition of *H. pylori* into the coccoid form (49). The gene *AmiA* in *E. coli* is preceded by the gene *hemF*, coding for the aerobic coproporphyrinogen III oxidase (50). This may suggest a similar rearrangement in *H. pylori*, with the gene *hp0773* possibly coding for the aerobic coproporphyrinogen III oxidase. However, the aerobic coproporphyrinogen III oxidase from different sources was shown not to contain any flavin cofactor (51-55). The gene *tyrS* of *E. coli* was demonstrated to code for a tyrosyl-tRNA synthetase (56), and the following gene in the genomic context was experimentally identified as pyridoxamine kinase (57). This may suggest that HP0773 is a pyridoxamine kinase, but pyridoxamine kinases do not contain a flavin cofactor (57). In conclusion, the genomic context of *hp0773* does not predict a physiological role for the flavin-dependent protein HP0773.

It is also possible that HP0773 constitutes only one subunit of a multienzymatic system and the in vitro system of recombinant HP0773 is not complete. Further in vivo studies on the gene *hp0773* would shed light on the physiological function, and therefore enzymatic activity, of HP0773.
Model of the three dimensional structure. The homology model of HP0773 based on the template of *Streptococcus pneumoniae* enoyl-acyl carrier protein reductase (FabK), PDB code 2Z6l, does not have a high QMEAN4 score for quality estimation. However the low protein sequence similarity of HP0773 with proteins in the PDB database does not give a better alternative to date. The overall folding is TIM barrel, which is commonly found in FMN binding proteins (58). The positions of the side chains of residues in the active site are not very reliable due to the low QMEAN4 score for quality estimation. H182 is predicted to be in the active site, with the protein surrounding of the FMN cofactor, such as M24, V26, V49, and A126, being mainly hydrophobic. The position of a histidine residue in the active site is conserved in many flavin-dependent enzymes performing different reactions, such as nitronate monooxygenase (59), or glycolate oxidase (60). It is also fully conserved in all the hypothetical nitronate monooxygenases in the GenBank, including PA1024 and PA0660 presented in previous chapters, and it is therefore not helpful to make hypotheses on the protein function. Due to the not very high quality of the homology model of HP0773 no further structural analysis was performed.

9.7 Conclusion

The search of new potential drug targets to develop antimicrobial drugs against *H. pylori* represents a high priority due to the increasing antibiotic resistance developing versus the current treatment options (2). A transposon analysis of the entire genome of this pathogen has included the gene *hp0773* in the list of candidate essential genes for the growth of *H. pylori* (5). The gene *hp0773*, which is not conserved in humans, is predicted to encode for a nitronate monooxygenase, a FMN-dependent enzyme involved in the detoxification of the mitochondrial toxin P3N (7). The purified recombinant HP0773, however, did not display nitronate
monooxygenase activity. Furthermore purified recombinant HP0773 was shown to contain tightly bound iron as cofactor, other than FMN. Other potential substrates for HP0773 were tested in this study, but it was not possible to assign any enzymatic activity to this important gene product. The lack of enzymatic activity is most likely due to the fact that the physiological substrate of HP0773 is still unknown and was not among the molecules tested in this study. Alternatively the recombinant protein may be in an inactive state, for example inactivated by oxygen. The histidine tag may affect the enzymatic activity to a certain extent; however it is less likely that it completely suppress the enzymatic activity. A homology model of HP0773 was generated but the low protein sequence similarity leads to a not very high quality of this model. Further studies are necessary in order to give insight in the physiological role of HP0773.

9.8 References

α The use of the extinction coefficient of free FMN may result in an inaccuracy of up to 25% in the determination of steady-state kinetic parameters. This does not alter the main conclusion that no enzymatic activity was detected for HP0773 with the compounds tested in this study. For this reason we did not attempt a more accurate determination of the concentration of enzyme-bound flavin.


CHAPTER X: GENERAL DISCUSSION AND CONCLUSIONS

Biochemical studies on flavoproteins started with the pioneering work on few enzymatic systems \(^{(1)}\) and extended nowadays to scientific teams working all around the world on a vast variety of flavin-dependent enzymes \(^{(2, 3)}\). Hundreds of flavoproteins have been discovered and characterized in many different biological pathways \(^{(3)}\). More than 2200 crystal structures of proteins containing a flavin cofactor are available in the PDB database and major mechanistic breakthroughs have been achieved during the characterization of this family of enzymes. Flavin-dependent enzymes have proved to be excellent mechanistic tools to understand and characterize different reaction mechanisms, such as hydride transfer, single-electron transfer, and covalent catalysis \(^{(2)}\), as well as to investigate the sophisticated control of oxygen reactivity which is unique to this type of cofactor \(^{(4, 5)}\). Artificially modified flavins have been used as active site probes to shed light on the role played by the protein environment in the fine tuning of the redox properties of the cofactor \(^{(6, 7)}\). The mechanistic knowledge gained with the characterization of flavin-dependent enzymes has largely contributed to the general understanding of the strategies of enzyme catalysis. In parallel to the mechanistic studies, structural characterization of a vast number of flavin-dependent enzymes catalyzing different reactions has enabled to gain important information on the different types of overall folds and active site topologies designed by nature to accomplish different tasks in an enzymatic system \(^{(3)}\). Despite the impressive amount of knowledge that has been collectively gained on flavin-dependent enzymes, many aspects of these old and important enzymes still need to be investigated. For example the physiological substrate of the first flavoenzyme discovered \(^{(7)}\), named old yellow enzyme (OYE) is still unknown \(^{(8, 9)}\), despite the fact that enzymes classified as OYE have been extensively characterized from a kinetic and structural point of view for 80 years mainly for biotechnological applications \(^{(9-11)}\).
New flavin-dependent enzymes have been recently discovered in new metabolic pathways, such as lysine-specific demethylase 1 (LSD1) which is involved in post-translational modifications of histones \(^{(12, 13)}\), RebF and RebH \(^{(14)}\) which play a role in the biosynthesis of the antibiotic rebeccamycin, and YUCCA enzymes which are monooxygenases involved in the biosynthesis of auxin, the primary growth hormone of plants \(^{(15, 16)}\).

Exceptions to the conventional classification of flavin-dependent enzymes are being pointed out, such as the unusual mechanism of nitrate monooxygenases which does not proceed via the typical intermediate of monooxygenases C4a-(hydro)peroxyflavin \(^{(17)}\) and the detection of the latter in the oxidase pyranose-2-oxidase \(^{(18)}\). Recent breakthroughs like the description of a new type of redox chemistry of flavin cofactors with the stabilization of a flavin-N5-oxide observed for the first time \(^{(19)}\) remind us that flavoenzymologists are still marveling about the amazing chemical versatility of these cofactors and that the story of this colorful family of enzymes still reserves many surprises.

The work in this thesis has contributed to the characterization of flavoproteins by combining the mechanistic and structural approach on the same enzymatic systems and has provided important insights in the gene function prediction of nitrate monooxygenases, which led to the identification of new classes of enzymes.

Choline oxidase from Arthrobacter globiformis catalyzes the oxidation of choline to glycine betaine and it has been extensively characterized from a kinetic point of view \(^{(20)}\). The crystal structure of choline oxidase in complex with glycine betaine described in chapter II represented a major achievement as it is the first structure of choline oxidase bound to a physiologically relevant ligand, e.g., the reaction product. This allowed the comparison of the interactions observed between glycine betaine and side chain of residues in the active site with the data
previously obtained from mechanistic studies of variant versions of choline oxidase. The analysis of the structures of the free enzyme and of the enzyme-product complex highlighted the different conformations of loop 250-255 at the dimer interface near the proposed entrance to the active site. This analysis suggested an important role of the oligomerization state of choline oxidase for controlling substrate access to the active site which was not observed before.

The mechanistic characterization of choline oxidase variants of the hydrophobic cluster near the proposed entrance to the active site M62A/F357A, M62A and F357A in chapter III complemented previous studies on the oxygen reactivity of flavoproteins \(^{(4, 5, 21)}\). It investigated an isomerization partially rate-limiting occurring during the oxidation of the reduced cofactor by oxygen that was not detected in other oxidases. It exploited the kinetic tool of solvent viscosity studies combined with site-directed mutagenesis to monitor this isomerization in the variant versions M62A/F357A, M62A and F357A. The Mildvan approach of inverse thinking on double variants \(^{(22)}\) was applied to establish the relative contributions of the two residues M62 and F357, and it was possible to assign the partially rate-limiting isomerization to the side chain of F357. Molecular dynamics simulations supported the conclusion that the side chain of F357 is important for the arrangement of the hydrophobic cluster at the proposed entrance of choline oxidase, while the detection of a narrow tunnel gated by F357 and E312 suggested that the isomerization linked to the side chain of F357 may be related to the control of oxygen access to the active site.

The work performed on choline oxidase in this dissertation has provided precious insights in the gating mechanism of this enzyme. It is now clear that the gating system of choline oxidase is more complex than the stochastic movements of the side chains of the hydrophobic cluster (M62, L65, V355, F357, and M359) previously described by molecular dynamics studies \(^{(23)}\). Indeed it
has been shown in this dissertation that the proposed entrance to the active site delimited by the side chains of the hydrophobic cluster (M62, L65, V355, F357, and M359) is covered by F253 of the other subunit only in the closed conformation of loop 250-255 at the dimer interface. In the open conformation of loop 250-255 there is no steric hindrance for the proposed entrance gated by the hydrophobic cluster. Therefore, the site-directed mutagenesis studies of F357 and M62 are probably not enough to significantly affect the gating mechanism for choline and for future studies triple mutants with removal of F253 could be considered. Site-directed mutagenesis of M62, F357, and F253 to a bulkier residue as tryptophan instead of alanine could significantly affect choline access to the active site, but with the risk of completely preventing substrate access and generating a dead enzyme. Furthermore, a second gating mechanism proposed for oxygen access to the active site has been identified, with a narrow hydrophobic tunnel leading to the FAD cofactor and regulated by F357 and E312. The movement of the phenyl ring of F357 seems to determine open and closed states of the narrow tunnel detected and proposed to be for oxygen access to the reduced cofactor. The position of F357 in both gating mechanisms confirms the higher complexity of the control of substrate access in choline oxidase compared with what previously hypothesized for this enzyme and, together with the description of the additional structural element loop 250-255, suggests a sophisticated coordination of substrate access and product release which is not fully understood. This is consistent with the notion that choline oxidase orchestrates four half reactions and is able to retain the intermediate betaine aldehyde in the active site and release the product glycine betaine only after the second reoxidation of the cofactor.

Human choline dehydrogenase is associated to the mitochondrial membrane and catalyzes the oxidation of choline to betaine aldehyde, which is the intermediate compound in the biosynthesis
of glycine betaine\textsuperscript{(24)}. This enzyme has received increasing medical attention for its correlation to severe diseases such as homocysteinuria\textsuperscript{(25)}, breast cancer\textsuperscript{(26, 27)}, and male infertility\textsuperscript{(28)}, but the biochemical characterization of choline dehydrogenase enzymes from different sources was hampered by the \textit{in vitro} instability. Chapter IV presented a detailed review of the medical importance of human choline dehydrogenase and of the biochemical challenges that hampered the study of choline dehydrogenase from various sources by different research groups since 1980\textsuperscript{(29)}. Chapter V described different strategies screened for the expression of human choline dehydrogenase in \textit{Escherichia coli}, among which the design of deletion mutants based on the hypothesis of a cleavable N-terminal presequence and on the use of particular strains of \textit{E. coli}. The challenges encountered in the purification led to the homology model based on the template of choline oxidase and to the bioinformatic analysis described in chapter IV. The work of this dissertation represents an important advancement, despite the fact that it did not lead to a stable preparation of purified hCHD. Indeed, the bioinformatics analysis of hCHD based on the knowledge of the soluble choline oxidase, suggested two regions in the protein sequence of hCHD that are most likely related to the lack of \textit{in vitro} stability observed by different research groups. More important, one purpose of the biochemical characterization of recombinant hCHD was to investigate in the recombinant enzyme the effect of the mutation L78R that has been associated to impaired hCHD function \textit{in vivo}, leading to male infertility and homocysteinuria\textsuperscript{(25, 28)}. The bioinformatic analysis performed in this dissertation suggested that the location of this residue is on the surface of the protein, near a highly conserved pattern in the protein sequence that was not described before and that is present only in mitochondrial CHDs. This location suggests the role of L78 and of the conserved pattern is not related to catalysis, but the residue is possibly involved in interactions with the membrane or possible effectors. Future \textit{in vivo} studies
should focus on the identification of possible complexes of hCHD with other proteins and on the mechanism that leads to the dysmorphic mitochondrial structure described in the case of the variant L78R (28). The role of hCHD for homeostasis of mitochondria and apoptosis is probably more important than currently known, as pointed out by a recent study on the relocation of hCHD to the outer membrane of mitochondria after exposure to a mitochondrial toxin (30).

Nitronate monooxygenases (NMOs) detoxify the metabolic poison propionate 3-nitronate and represent an emerging interest in the flavin field for the still uncharacterized metabolic pathway and the unusual kinetic mechanism (17, 31, 32). Indeed, the kinetic characterization of fungal NMOs established that the detoxification of propionate 3-nitronate is initiated by a single electron transfer from the substrate to the FMN cofactor with the stabilization of the anionic flavosemiquinone (17, 33), which represents a remarkable exception to the classical mechanism of flavin-dependent monooxygenases with the stabilization of the C4a-(hydro)peroxyflavin (16).

There are more than 5000 genes in the GenBank database annotated as hypothetical nitronate monooxygenases, but the gene function prediction is often not reliable as it is based only on modest overall protein sequence similarity with the fungal enzymes from *Cyberlindnera saturnus* (Cs-NMO) and *Neurospora crassa* (Nc-NMO).

The study of the hypothetical NMO PA4202 from *Pseudomonas aeruginosa* PAO1 in chapter VI was aimed to improve the gene function prediction. It was necessary to highlight which parts of the protein sequence of NMOs are significant in a multiple sequence alignment to predict NMO activity. The absence of a crystal structure of NMO was a major obstacle, and the need for structural information for this class of enzymes in order to proceed with site-directed mutagenesis studies was recently highlighted in a review of classification of flavin-dependent monooxygenases (16). Chapter VI presented the kinetic characterization of recombinant PA4202
as Pa-NMO and the first crystal structure of a NMO enzyme at 1.4 Å resolution, which represented a key breakthrough for the identification of active site residues as candidate for site-directed mutagenesis studies, such as the two histidine residues in the active site H133 and H183, the residues Y109, Y299, Y303, which could be involved in substrate binding/catalysis, and the positively charged K307 at the entrance to the active site possibly playing a role in substrate capture. In this study the combined information from the structure and the kinetic data were used to highlight specific residues to use as reference in the bioinformatic analysis. As a result four conserved motifs were described that identify Class 1 NMO and that can be used to predict the ability of hypothetical enzymes to detoxify propionate 3-nitronate with a kinetic behavior similar to Pa-NMO and to Cs-NMO. Two hypothetical NMOs from *Burkholderia phytofirmans* carrying the four motifs identified in this study were cloned from genomic DNA and expressed in *E. coli*. The cell free extracts obtained from these two samples exhibited nitronate monooxygenase activity, confirming the updated gene function prediction. Interestingly, the NMO from *N. crassa* Nc-NMO shows different substrate specificity compared to Cs-NMO and Pa-NMO, and it also does not align with the bioinformatics analysis presented in this dissertation. The differences highlighted between Class 1 NMO and Nc-NMO are probably contributing to the not reliable function prediction based on Cs-NMO and Nc-NMO and should be further investigated to assess if Nc-NMO is the prototype for Class 2 NMO or if it represents an outlier. In case other members of Class 2 NMO can be identified it would be interesting to understand how the two different NMO classes evolved, and relate the different substrate specificities observed to the physiological role. The crystal structure of Nc-NMO is currently not available and it would shed light on differences in the active site between Class 1 NMO and Nc-NMO and suggest structural elements that could play a role for the different substrate specificity.
Indeed, it would be important to assess if Class 1 NMO is not able to deprotonate 3-NPA, with 3-NPA representing therefore a competitive inhibitor of the deprotonated P3N, or if there is a selectivity filter at the entrance that prevents 3-NPA to enter the active site. The identification of the active site residues of Nc-NMO would also help the bioinformatics analysis by highlighting specific regions in the protein sequence of Nc-NMO to be used in the multiple sequence alignment of hypothetical NMOs. Nc-NMO would also be a better candidate than Class 1 NMO for the application an enzymatic sensor for P3N in the environment and food. The toxin P3N in the environment is mainly present in the neutral form of 3-NPA and the enzymatic detection of food contaminated by 3-NPA should be based on Nc-NMO.

A crystal structure in complex with P3N or the product would provide information for the residues involved in substrate specificity and it may shed light on how NMO enzymes are able to prevent the nucleophilic attack of the negatively charged P3N to the flavin cofactor.

Many questions are still unanswered for the metabolism of P3N, among which the role of P3N detoxification in bacteria, and the biosynthetic route for this toxin. Future studies should investigate the role of P3N especially in bacteria and the work presented in this dissertation will represent a precious asset to identify putative bacterial NMOs.

Chapter VII, VIII, and IX focused on the cloning, expression in E. coli, purification and kinetic characterization of hypothetical nitronate monooxygenases that do not carry the four conserved motifs of Class 1 NMO, and therefore are assumed in this dissertation not to be NMO. Alternatively they could represent a different Class of NMO, and it is important to establish if there are different isoforms of NMO and their physiological importance. The studies presented in chapter VII and VIII demonstrated that the other two hypothetical NMOs from P. aeruginosa PAO1, namely PA1024 and PA0660, are not able to use P3N as substrate and that they exhibited
a different enzymatic activity. PA1024, for which a crystal structure is available in the PDB database at 2.0 Å resolution, exhibited a NADH:ferricyanide reductase activity specific for NADH only and possess seven conserved motifs shared by more than 1000 hypothetical nitronate monooxygenases in the non-redundant protein database. PA1024 is therefore currently wrongly annotated as NMO in the GenBank and the PDB databases and needs to be reclassified as NADH:ferricyanide reductase.

PA0660 did not show NMO activity and was able instead to use both NADPH and NADH as electron donor, with a preference for NADPH. Six conserved motifs were identified in the protein sequence of PA0660, which are present in more than 1000 bacterial hypothetical nitronate monooxygenases in the non-redundant protein database, establishing a new class of enzymes with diaphorase activity. The three hypothetical NMOs of *P. aeruginosa* PAO1 possess a different function, and each one carries different conserved motifs that identify three classes of enzymes. Remarkably, all three enzymes have been classified as TIM-barrel proteins and the two crystal structures of Pa-NMO and PA1024, with TIM-barrel fold, have been superposed. The comparison of the similar overall fold of the FMN binding domain and the different function suggested the presence of some regions of the protein sequence, such as locations A, B, and C described in Chapter VIII, that are important for function prediction. On the other hand, some regions of the protein sequence of Pa-NMO, PA1024, and PA0660, appears to be conserved and they may be related to the integrity of the scaffold of the TIM barrel domain and not specific to protein function. The comparison of the three enzymes Pa-NMO, PA1024, and PA0660 in this dissertation has therefore provided important insights to the structural/functional studies of enzymes and it may be relevant for protein engineering. It would be interesting to insert the side chains identified in Pa-NMO as fully conserved and suggested to be important for NMO activity
in the similar scaffold of PA1024 by site-directed mutagenesis and monitor if the enzyme PA1024 starts displaying NMO activity. As the active site of PA1024 seems to be more exposed to the solvent than Pa-NMO, the opposite attempt to turn Pa-NMO into a NADH:ferricyanide reductase would probably require more structural rearrangements of the substrate binding domain which could be obtained with circular permutation \(^{(34)}\). The crystal structure of PA0660 should be pursued as it would provide useful insights into structural determinants of selectivity between NADH and NADPH, and complement the comparison of the three hypothetical NMOs from \(P. \text{aeruginosa} \) PAO1.

The two classes of enzymes represented by the prototype PA1024 and PA0660 were first identified in this dissertation and they must belong to some cellular pathway. It is therefore important to establish substrate specificity for PA1024 and PA0660 and their physiological function by characterizing \textit{in vivo} deletion of these two genes.

Chapter IX focused on the cloning, expression in \(E. \text{coli}, \) purification and initial characterization of the hypothetical NMO HP0773 from \(Helycobacter pylori\). This gene was classified as candidate essential gene for the growth of this pathogen and it is not conserved in humans, representing an ideal candidate for drug target \(^{(35)}\). A protocol of purification was optimized and the presence of iron and FMN as cofactors was established. It was demonstrated that the protein is not able to use P3N, nitroalkanes or nitronates as substrates. Therefore most likely the gene function prediction of HP0773 as NMO is wrong. Other potential substrates for HP0773 were tested but no enzymatic activity was detected with the compounds selected. Furthermore the enzyme is not able to use NADH and NADPH as electron donor. Future studies should use high-throughput screening of potential substrates to address the function of HP0773. The work in this dissertation has set the stage for future studies on the drug target HP0773 by optimizing a
protocol of purification of the recombinant protein, showing it does not exhibit NMO activity, and determining the presence of iron as cofactor which was not predicted by the protein sequence.

Overall the research of this dissertation has shown the benefits of combining kinetic and structural studies in the characterization of enzymatic systems and it has applied the structural/functional information to the identification of specific regions in the amino acid sequence that can be used to predict protein function.

The crystal structure of choline oxidase in complex with glycine betaine has highlighted the importance of the dimer interface in controlling substrate access to the active site, while site-directed mutagenesis studies of F357 and the kinetic investigation of the variant enzymes associated the side chain of F357 to a slow isomerization in the reaction of the reduced enzyme with oxygen.

The bioinformatics analysis of the protein sequence of human choline dehydrogenase identified a putative targeting presequence at the N terminus of the enzyme and a hydrophobic region possibly related to the interaction with the mitochondrial membrane. Furthermore, a homology model of human choline dehydrogenase was generated based on the crystal structure of the soluble choline oxidase.

The crystal structure of the bacterial Pa-NMO represented the first crystal structure of a nitronate monooxygenase enzyme, setting the stage for site-directed mutagenesis studies of this class of enzymes. The four motifs identified in the protein sequence of Pa-NMO significantly improved the gene function prediction of hypothetical nitronate monooxygenases. The kinetic characterization of the hypothetical nitronate monooxygenases PA1024 and PA0660 not carrying
the motifs of Class 1 NMO established that these proteins are wrongly annotated as nitronate monooxygenases.

PA1024 and PA0660 represent the prototype of two different classes of enzymes, specifically a new class of enzymes with NADH:ferricyanide reductase activity belonging to bacteria and fungi, and a new class of bacterial enzymes with diaphorase activity and preference for NADPH.

The comparison of the three TIM barrel proteins Pa-NMO, PA1024, and PA0660 that exhibit different enzymatic activities has provided insights in the regions of the protein sequence that are relevant for protein function.

The work of this dissertation contributed to the advancement in the field of flavin-dependent enzymes in the identification of structural elements important for regulating substrate access to the active site and in the use of experimental data to improve gene function prediction, with the description of two new classes of flavin-dependent enzymes.

10.1 References:


