Crystallographic Determination of Wild Type, Mutant and Substrate-Analogue Inhibited Structures of Bacterial Members of a Family of Superoxide Dismutases

by

Simon Hardie Oakley

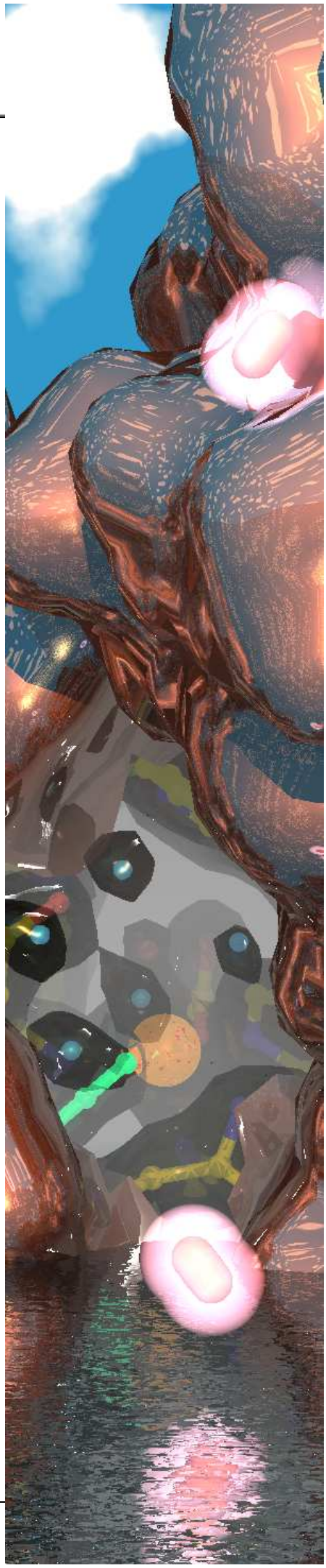
Submitted as part of the requirements for the degree of

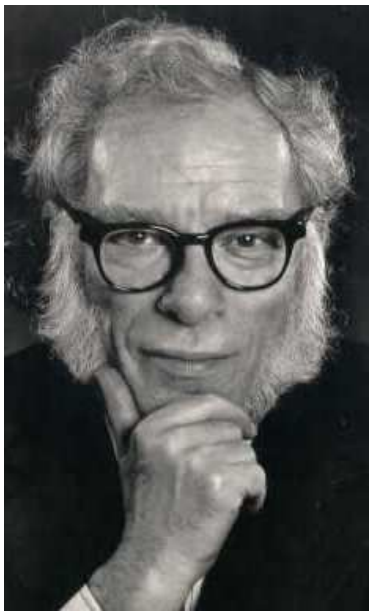
Doctor of Philosophy

Institute of Fundamental Sciences, Chemistry

Massey University, New Zealand

2009





"Part of the inhumanity of the computer is that, once it is competently programmed and working smoothly, it is completely honest."

Isaac Asimov (1920-92), author,

inventor of the word “robotics”.

0.1 Abstract

Crystallographic Determination of Wild Type, Mutant and Substrate-Analogue Inhibited Structures of Bacterial Members of a Family of Superoxide Dismutases

The iron and manganese superoxide dismutases are a family of metallo-enzymes with highly conserved protein folds, active sites and dimer interfaces. They catalyse the elimination of the cytotoxic free radical superoxide to molecular oxygen and hydrogen peroxide by alternate reduction then oxidation of the active-site with the concomitant transfer of protons from the solvent. There are many key aspects of enzymatic function that lack a structural explanation.

The focus of this study is on three crystal structures. The iron-substituted manganese superoxide dismutase from *Escherichia coli* complexed with azide, a substrate-mimicking inhibitor, was solved to 2.2 Å. This "wrong" metal form shows a binding pattern seen previously in the manganese superoxide dismutase from *Thermus thermophilus.* Wild-type manganese specific superoxide dismutase from the extremophile *Deinococcus radiodurans* was solved to 2.0 Å and has an active site reminiscent of other solved manganese superoxide dismutases despite a lack of product inhibition. The azide-inhibited manganese superoxide dismutase from *Deinococcus radiodurans* was determined to a resolution of 2.0 Å and showed binding of azide, and by inference superoxide, different to that seen in *Thermus thermophilus*, but reminiscent of that seen in azide-inhibited iron superoxide dismutases. These results indicate that the azide ion, and by inference superoxide, bind to the metal centre of manganese superoxide dismutases in two modes, and transition between the two modes may be entropy dependent.

These structures, integrated with knowledge from other structures, known biochemistry and various spectra, provide insight into catalytic function. An outer-sphere mechanism of proton transfer that does not rely on through-peptide proton uptake is proposed and compared to a previously proposed inner-sphere mechanism. This is based on the observation that a water molecule moves into the active site of the manganese superoxide dismutase from *Deinococcus radiodurans* upon azide binding, providing a Grötthus pathway for rapid proton transfer to the active site from the bulk solvent.

Also presented in this study are the partially refined structures of four point mutants (S82T, L83M, L133V, and M164L/L166V) of the manganese superoxide dismutase from *Escherichia coli* all solved to roughly 2 Å resolution, designed to investigate product inhibition which varies across species.

0.2 Acknowledgements

Acknowledgement pages generally read like a mix between an Oscar acceptance speech and a shopping list, so here goes. There are many people who contributed to my time at Massey University both professionally and social. Therefore to quote the bard, “brevity is the sole of wit” I apologise to all those not named specifically in my efforts to be both laconic and pretentious.

Firstly, would like to acknowledge my supervisor Geoff for his support, encouragement and vast knowledge on all things chemical, X-ray, and crystal.

Secondly, the post-docs; my co supervisor Bryan for answering questions about various computer software and crystallography, and Jim, whose brief foray into protein crystallography from chemistry was very fruitful.

Thirdly, Gill, Trev, and the other members of the X-lab past and present for providing lab space, office space, computing, stolen pens, interesting talks, moral encouragement, cakes, parties, beer, wine, gossip *et cetera*. These were all gladly accepted but not always reciprocated.

Penultimately, I would like to thank my various flat-mates and friends over my time in Palmerston North who put up with my bad jokes and stories over the past few years.

Lastly, but most importantly, I would like to thank my family for their long-distance encouragement and support.

-Simon

0.3 Table of contents

0.4 List of figures

Colour images occur on the following pages

5,15,22,24,26,27,31,70,72,75,76,77,79,80,81,82,82,83,84,87,88,90,91,91,92,93,100,105
,

107,128,129,132,138,139,139,140,143,145,147,151,154,155,156,158,159,162,163,209,

209,211,211,212,213,227,228.

0.5 List of equations

0.6 List of tables

0.7 Abbreviations

α helix	secondary structure element of proteins
ATP	adenosine-5'-triphosphate
β-sheet	secondary structure element of proteins
CD	circular dichroism
DNA	2'-deoxyribonucleic acid
$E^{o'}$	standard reduction potential at pH 7
EPR	electron paramagnetic resonance
HFEPR	high-magnetic field electron paramagnetic resonance
HPLC	high pressure liquid chromatography
K	kelvin, temperature scale
k_{cat}	the maximum number of enzymatic reactions catalysed per second
K_m	Michaelis constant, approximate substrate concentration, where enzyme activity is half the maximum
L	Litre, measure of volume
LB	Luria broth, media for bacterial growth
M	moles/litre, a measure of concentration
MCD	magnetic circular dichroism
MR	molecular replacement
MIR	multiple isomorphous replacement
mRNA	messenger RNA
mV	millivolt, a thousandth of a volt

NAD(P)H	nicotinamide adenine dinucleotide (phosphate) –reduced form
NCS	non-crystallographic symmetry
NMR	nuclear magnetic resonance
PCR	polymerase chain reaction
PDB	protein data bank
pH	activity of the hydrogen ion, measure of acidity, where $pH = -\log a_{H+}$
p*K*	disassociation constant of a proton or a hydroxide, defined as $pK = -\log K$
redox	reduction/oxidation
R_{free}	set of reflections, excluded from refinement and (electron) density calculations
RMS	root-mean-square, measure of the magnitude of a varying quantity. As RMS deviation, measure of scatter of values about a target value. Relevant to geometric parameters in protein structures.
RNA	ribonucleic acid
ROS	reactive oxygen species, including O_2^-, H_2O_2, O_2NO^- (peroxynitrate), OCl^- (hypochlorite)
RPM	revolutions per minute
SIR	single isomorphous replacement
SOD	superoxide dismutase
SD	standard deviation
TLS	translation-libration-screw

μL	microlitre, measure of volume
UV	ultra violet
V	volt, potential difference across a conductor

0.8 Naming conventions of superoxide dismutases.

Throughout the literature concerning superoxide dismutases different naming conventions have been used for both genes and proteins. Some known names are Cu,ZnSOD, Cu-Zn superoxide dismutase, FeSOD, MnSOD, SOD, SOD-1, SOD-2, SOD-3, SOD-4, SODF, SODS, copper-zinc superoxide dismutase, cuprein, cytocuprein, erythrocuprein, ferrisuperoxide dismutase, hemocuprein, hepatocuprein, superoxidase dismutase, superoxide dismutase I, superoxide dismutase II, *et cetera.*

The most common abbreviations of superoxide dismutase are "SOD" pertaining to protein and "sod" for dealing with genes, but some have used "SD". Superoxide dismutase activity has been detected in a number of polypeptides and subsequently potential superoxide dismutase activity has been inferred for many genes. Many organisms have multiple superoxide dismutases and a variety of naming schemes have been used.

sodA, sodB, sodC…	alphabetically based.
sod1, sod2, sod3…	numerically based.
sodA, sod1..	italicised to indicate gene.
sodM, sodF, sodN…	where the letter refers to the obligate metal cofactor Mn, Fe, Ni.

In this thesis a functional and hierarchical nomenclature will be used to describe the enzyme function, the obligate metal cofactor, oxidation state, the species of origin, inhibitors and mutation. Throughout this work "SOD" will used as an abbreviation for superoxide dismutase. Members will be identified first by the element symbol, then abbreviation, for example "NiSOD" for a nickel-specific superoxide dismutase. Hierarchical naming systems have been used in several papers: the system used here is an extension, albeit with minor modifications, of that used by Vance and Miller (Vance & Miller, 1998b).

1. There are at present three major families of superoxide dismutases: nickel dependent (NiSOD), those which require both copper and zinc (Cu/ZnSOD), and the third family that are active with either manganese (MnSOD) or iron

(FeSOD). The abbreviations of “MnSOD” and “FeSOD” are treated as words beginning with vowels; e.g. “FeSOD” should be pronounced “eff-ee-sod” rather than as “i-arn-sod” and should be prefixed by the indefinite object, “an”, rather than “a”.

2. If it is unknown or ambiguous whether a superoxide dismutase is active with manganese or iron it is designated “Fe/MnSOD”. There are some members that have significant activity with both iron and manganese and these are commonly termed “cambialistic” These are referred to as “camb-Fe/MnSOD”, and hence the manganese-only form of a cambialistic enzyme would be “camb-MnSOD”.

3. If the redox state of the active-site metal is known it is denoted as a superscript of the elemental symbol. For example, “camb-Fe^{3+}SOD” and “camb-Fe^{2+}SOD” denote two different oxidation states of the iron in a cambialistic enzyme.

4. It is possible to generate superoxide dismutases with incorrect metal ions substituted for the preferred metals. The preferred metal is indicated immediately adjacent to the SOD expression, preceded by the metal that has been substituted, separated by a hyphen. For example, “Mn-FeSOD” represents an iron superoxide dismutase in which active-site iron is replaced with manganese.

5. By default the enzyme is assumed to have the native sequence. If increased clarity is required, then the prefix “WT” will be used as an abbreviation for “wild type”. If the protein has a point mutation it will be indicated by a prefix. The prefix will be the single letter code of the native sequence, followed by the residue number and the single letter code of the mutant. For example “Y34F-MnSOD” represents a mutation of a tyrosine to a phenyalanine at position 34 in an MnSOD. This is in contrast to the non-mutated native or “WT-MnSOD”.

6. The species of origin is indicated by an italicised expression between the indication of mutation and the metal designation. The basis of the italicised expression may either be the common or the scientific name. For example, “Y34F-*Ec*-MnSOD” and “Y34F-*mito*-MnSOD” represent the same point

mutation of the *Escherichia coli* and the human mitochondrial manganese superoxide dismutases respectively.

7. Superoxide dismutases interact with a number of small-molecule inhibitors to form stable complexes; this is indicated by the chemical name (and charge) of the inhibitor, written after the "SOD" expression. The azide-inhibited form of the cattle Cu/Zn superoxide dismutase would be indicated as "*bovine*-Cu/ZnSOD-N_3^-".

8. Other experimental techniques and variables can be included into this naming convention. These will be introduced after any description of inhibitors. For example "RT" is used to abbreviate for room temperature and "110 K" for an experiment at cryo-conditions of 110 K. For example, there exist spectral studies for Y34F-*Ec*-MnSOD-N_3^- at ambient and cryogenic conditions designated "Y34F-*Ec*-MnSOD-N_3^--RT" and "Y34F-*Ec*-MnSOD-N_3^--110 K" respectively.

9. This nomenclature lends itself to the identification of individual atoms within a structure. This follows the hierarchy of protein name, underscore, chain identifier (italicised), underscore, one-letter amino acid code, residue number, underscore, and atom name. Atom names are capitalised forms, following the naming conventions of the protein data bank (Berman *et al.*, 2000). This makes distinguishing atoms that are closely located spatially unambiguous. The expressions "*Ec*-MnSOD_*A*_D146_OD1" and "*Ec*-MnSOD_*A*_D146_OD2" distinguishes between two atoms of the same amino-acid side chain. If there is a known PDB code for a protein, then this can be substituted for the protein name, for example "1VEW_*A*_D146_OD1".

10. If an atom or side chain in a crystal structure has split or multiple conformations then this will be indicated by a subscript after the atom name. For example, individual atoms within the two conformations of lysine 29 in the A chain of Y174F-*Ec*-MnSOD are described as "1IXB_*A*_K29_NZ_A" and "1IXB_*A*_K29_NZ_B"

11. The naming of water molecules will follow the hierarchy of protein name, underscore, an italicised chain number (if present), underscore, “WAT” followed by the number of the water, for example 1EN5_WAT72.

The naming and abbreviation of amino acids generally follow IUPAC conventions for the use of trivial names, three-letter system and one-letter system (Cornish-Bowden, 1984). The naming and use of elements and chemicals will follow IUPAC conventions, but some trivial names will also be used.