Structural Studies of Catalysis and Processing in Galactose Oxidase

Nana Akumanyi

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Faculty of Biological Sciences

August 2007

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would like to thank my supervisor, Prof. Simon Phillips, for giving me the opportunity to contribute to field of protein derived cofactors and for valuable support and encouragement, particularly during my final year.

I would also like to thank Prof. Peter Knowles and Dr. Malcolm Halcrow for their input into my project. I would like to particularly acknowledge Prof. Mike McPherson who has allowed me access to his laboratory facilities and advice, which has been greatly appreciated. I would like to thank Sarah Deacon for her patience in introducing me to the various galactose oxidase expression and purification procedures.

My thanks go to Susan Firbank, Joe Jaeger and Cait Dennis for introducing me to crystallography and answering all my constant questions. My appreciation goes out to Mel Rogers for assistance with the W290 mutants. I give thanks to Diane Wilkinson and Christian Kurtis for contributions to preparation of the C383S protein and the W290G variant respectively.

I thank all members of the SEVP/JJ/MRP/MJM labs for teaching me how to use the numerous processing, refinement and molecular graphics programs utilised to produce this work. A big thanks goes to Christian Kurtis and Russell Green for making me feel at home in Leeds from day one, and making the unsuccessful trips to Daresbury that much more bearable.

My thanks go to Noora and my family; to Mum, Dad, Efua and Ekua, your encouragement and support have given me the platform to reach for my goals.

Abstract

The fungal copper containing enzyme galactose oxidase is one of the earliest examples of a protein derived radical cofactor. A covalent bond, formed between the active site cysteine and the Cε of a tyrosine act as a site for radical formation, allowing the enzyme to catalyse the two electron oxidation of primary alcohols to their corresponding aldehydes, with the co-production of hydrogen peroxide.

The formation of the post-translationally modified redox cofactor is an autocatalytic process *in vitro*, requiring only the aerobic addition of copper. A premature form of galactose oxidase, lacking the cofactor has been crystallised and the structure determined. Copper incubation of this protein has been carried out and the structures determined in order to characterise processing intermediates and the mechanism of cofactor biogenesis.

The mature, fully processed and catalytically active enzyme contains a radical which is thought to be stabilised by the active site environment, particularly a tryptophan residue which stacks over the cofactor site. The structure of a galactose oxidase variant in which the tryptophan is mutated to a glycine has been determined and characterised.

Past attempts to obtain an enzyme-substrate complex with sugar molecules have been unsuccessful since space at the active site of the original crystal form was limited. Additionally crystals have not been activated prior to substrate incubation. Investigation of substrate binding using activated crystals and a different crystal form have been carried out. The C383S variant protein, has been shown to have improved activity against galactose, however, the wide substrate specificity of the enzyme has allowed investigation of substrate binding using a smaller substrate. Two enzyme-substrate structures, determined to 1.8 Å and 1.9 Å resolution have be obtained following incubation of C383S with a smaller alcohol substrate.

Investigation of the fully reduced form of galactose oxidase has proved difficult due to the high reactivity of Cu^{1+} with oxygen. Apo-galactose oxidase crystals were anaerobically incubated with Cu^{1+} and the structures determined in order to characterise the fully reduced form of galactose oxidase.

Contents

Wilkinson D, Akumanyi N, Hurtado-Guerrero R, Dawkes H, Knowles PF, Phillips SEV, McPherson MJ (2004): Structural and kinetic studies of a series of mutants of galactose oxidase identified by directed evolution. *Protein Eng Des Sel* **17**:141-8.

Rogers MS, Tyler EM, Akumanyi N, Kurtis CR, Spooner RK, Deacon SE, Tamber S, Firbank SJ, Mahmoud, K, Knowles PF, Phillips SEV, McPherson MJ, Dooley DM (2007): The stacking tryptophan of galactose Oxidase: A second-Coordination sphere residue that has profound effects on tyrosyl radical behaviour and enzyme catalysis. *Biochemistry, **46**, 4606-4618*

List of Figures

Protein structure figures were prepared using the molecular graphics program SPOCK [Christopher, 1998] or PyMOL [DeLano, 2004].

List of Tables

List of Abbreviations

2MPD	2-methylene-1,3-propandiol
ABTS	2,2'-azino-bis(ethylbenzthiaozoline-6-sulphenic acid)
ADSC	Area Detector Systems Corporation
BMGY	Buffered glycerol-complex medium
BMMY	Buffered methanol-complex medium
CCD	Charge coupled device
CD	Circular dichroism
CNS	Crystallography & NMR System
CCP4	Collaborative Computational Project 4
CTQ	Cysteine tryptophylquinone
DDC	Diethylydithiocarbamate
ECAO	*Escerichia coli* Amine Oxidase
EDTA	Ethylenediamine tetraacetic acid
ENDOR	Electron nuclear double resonance
ESRF	European Synchrotron Radiation Facility
EPR	Electron paramagnetic resonance
GO	Galactose oxidase
HEPES	N-2-Hydroxyethylpiperazine-N'-2-Ethanesulphonic acid
KIE	Kinetic Isotope Effect
LTQ	Lysine Tyrosylquinone
MES	2-(N-morpholino)ethanesulphonic acid
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein Databank
PEG	Polyethelene glycol
PIPES	Piperazine-N,N'-bis[ethanesulphonic acid]
PQQ	Pyrroloquinoline quinone
Pro-GO	Pro-sequence form of galactose oxidase
RCC1	Regulator of Chromosome Condensation
rms	Root mean squared
SDS	Sodium dodecyl sulphate
SRS	synchrotron radiation source

TPQ	2,4,5-trihydroxyphenylalanine quinone
TRIS	2-amino-2-hydroxymethyl-1,3-propandiol (also known as trishydroxymethylaminomethane)
TTQ	Tryptophanyltryptopylquinone
UV	Ultra violet
EXAFS	Extended X-ray Absorption Fine Structure
XANES	X-ray Absorption Near Edge Structure
YPD	Yeast extract, peptone and dextrose

1 Introduction

1.1 Copper proteins

One in three proteins require a metal cofactor for function. Metalloproteins play important roles in particular biological processes such as respiration, photosynthesis, nerve transmission and defence against toxic agents. Many transition elements have been found serving as prosthetic groups in the active sites of proteins and in many cases their redox and electronic properties are exploited in order to perform a variety of complex biological functions [Szilagyi and Solomon, 2002]. Transition metals may well be critical to many proteins' function, structure or stability but metal ions in their unbound state, such as copper in biological systems, are often toxic. On average, an adult has about 150 mg of copper in the body, of which about 10 to 20 mg is found in the liver, the rest being ubiquitously distributed [Beers and Berkow, 1999]. Excess copper is excreted through bile. Regulation of the copper content in the body is essential, as many known copper deficiencies are fatal. Inherited copper deficiency such as Menkes' syndrome, occurs in male infants who have inherited a mutant X-linked gene [Beers and Berkow, 1999]. The disorder is characterised by deficiencies of copper in the liver and serum and, more significantly deficiencies of essential copper proteins, including cytochrome-c oxidase, ceruplasmin and lysyl oxidase. Copper concentration inside eukaryotic cells is strictly controlled to prevent damaging free radical formation. Inside the cell copper is distributed through cellular compartments by copper chaperones. The chaperone proteins are divided into three functional groups: the Atx1-like chaperones, the copper chaperones for cytochrome c oxidase and the copper chaperones for superoxide dismutase [Rosenzweig, 2002]. Little is known about how these proteins acquire copper, but one possibility is through interactions with membrane transporters. Almost all of the copper in the body is present as a component of copper proteins, thereby reducing and maintaining the *in vivo* concentration of unbound copper ions to almost zero [Rosenzweig, 2002].

Copper proteins are mainly classified by their spectroscopic characteristics and coordination environments. The different classes of copper proteins exhibit a range of structures and active sites, comprising simple one-copper sites, to four copper centre representatives [Jian *et al.*, 2004]. Analysis of copper compounds suggests the different oxidation states of copper tend to prefer different geometries around the metal ion. In the Cu^{1+} oxidation state, the ion has been reported to occupy a three or four coordinate geometry, with three equatorial ligands and a fourth axial ligand [Szilagyi and Solomon, 2002]. The Cu^{2+} oxidation state is often observed

with four to six ligands, in a tetrahedral, trigonal bi-pyramidal or square pyramidal geometry [Jian *et al.*, 2004; Solomon *et al.*, 2004; Szilagyi and Solomon, 2002].

The blue copper proteins owe their name to the intense blue colouration due to the electronic configuration of the corresponding Cu^{2+} ion coordinated to the enzymes. These metalloenzymes are known as type I (T1) copper centres and are involved in reversible electron transfer. All proteins in this family have a trigonal copper centre normally coordinated to two nitrogens and a sulphur atom (commonly these ligands are provided by two histidines and a cysteine (Figure 1.1a)). In the majority of T1 proteins, such as plastocyanin [Nersissian, 1998], there is typically an additional more distant axial methionine, and in azurin only, a second axial ligand is provided by a main-chain carbonyl [Chang *et al.*, 1991]. The copper geometry in azurin can be thus described as a distorted trigonal bipyramid. A variation of this T1 geometry is observed in stellacyanin, which possess the same equatorial trigonal ligands, but the methionine ligand is replaced by the oxygen from a glutamine side chain [Hart *et al.*, 1996]. In some proteins, there is no axial ligand, such as seen for one of the three T1 copper centres in ceruloplasmin [Nersissian *et al.*, 1998].

Type II copper proteins (T2) are characterised by a square planar, square pyramidal or tetrahedral copper geometry. These proteins are weakly coloured and are often redox enzymes. The copper site coordinates four strong ligands, with a weaker fifth ligand often present (Figure 1.1b) [Szilagyi and Solomon, 2002]. An example of T2 copper proteins are the copper amine oxidases (CAO), a ubiquitous group of quinoenzymes that catalyse the oxidative de-amination of primary amines to the corresponding aldehydes, with the concomitant reduction of molecular oxygen to hydrogen peroxide [Blaschko, 1958]. These enzymes are dimers of identical subunits - each of which contains a single copper ion coordinated by three active site histidines and one or two water molecules in a square pyramidal geometry. CAO's contain a 2,4,5-trihydroxyphenylalanine quinone (TPQ) cofactor, derived from a post-translationally modified tyrosine side chain [Parsons *et al.*, 1995]. Their function in prokaryotes and lower eukaryotes is to utilise amines as a source for carbon and nitrogen. In higher eukaryotes CAO has been identified in many organisms including mammals and plants [Blaschko, 1958; Kumar *et al.*, 1996], their roles are less clear but have been linked to cell signalling, growth and development. Other T2 copper proteins include copper-zinc superoxide dismutase, galactose oxidase, monooxygenases, dioxygenases and nitrite reductase.

Type III copper proteins (T3) are binuclear proteins classified according to their characteristic absorbance spectra. The two copper ions are approximately 3-4 Å apart and couple

diamagnetically to form an Electron Paramagnetic Resonance (EPR) silent protein. Examples of proteins with T3 copper centres are: haemaocyanins, tyrosinases and catechol oxidase. The classic T3 copper centre is found in ascorbate oxidase, in which each copper is coordinated by three histidines in a trigonal pyramidal geometry [Messerschmidt *et al.*, 1993]. The same ligands are found in a similar arrangement in haemocyanin, a protein that is used as an oxygen carrier in the blood of molluscs and arthropods. Tarantula haemocyanin consists of four hexamers, each subunit has a molecular weight of 75 kDa, and contains a binuclear copper site [Troger *et al.*, 2002]. Dioxygen is able to bind between the two copper ions in a side-on coordination (Figure 1.1c). Tyrosinase is closely related to catechol oxidase and both are able to oxidise catechol via a four electron reduction of molecular oxygen. Tyrosinase, however, is also the enzyme responsible for converting tyrosine to the melanine pigment, while catechol oxidase is involved in the production of the benzoquinone pigment from catechol. The pigment product is in some part responsible for the darkening of fruit and vegetables, such as apples and potatoes, after exposure to air.

A second class of binuclear copper proteins, the CuA centre proteins, include cytochrome c oxidase and nitrous oxidase reductase [Wunsch *et al.*, 2003]. The centre is arranged with two copper ions and two bridging cysteines (Figure 1.1d). Nitrous oxide reductase is the terminal enzyme of the respiratory chain in denitrifying bacteria, while cytochrome c oxidase performs the equivalent role in the respiratory chain in aerobic organisms and is found in the inner membranes of mitochondria and bacteria. Cytochrome c oxidase utilises several metal ions, including Fe^{2+} and Cu^{2+} to transfer electrons onto oxygen molecules. Electrons are delivered to the CuA centre of cytochrome oxidase by the molecule Cytochrome C, while the oxygen itself binds to a different part of the oxidase [Yoshikawa *et al.*, 1998]. The oxygen is pinioned between the haem iron and another copper site, denoted as CuB (Figure 1.1e) and the electrons are transferred to the oxygen molecule, ultimately forming water molecules.

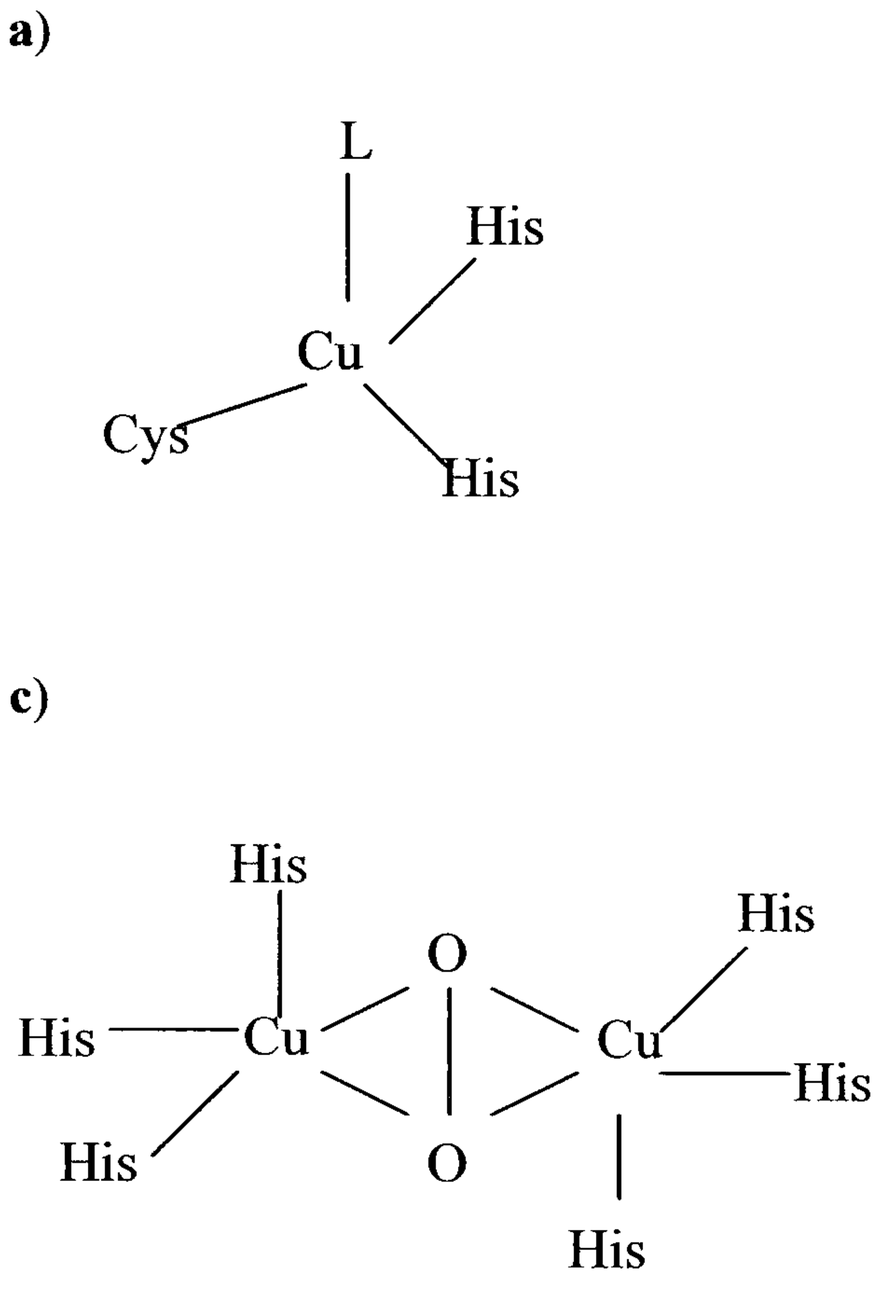


Figure 1.1: Schematic representation of copper sites and geometries found in proteins.

a) Type 1 copper centre, **b**) Type 2 copper centre, **c**) Type 3 copper centre, **d**) Cu A centre, and **e**) Cu B centre. Figure adapted from Firbank [2002].

Many copper proteins are involved in redox reactions, oxygen transport and activation or electron transfer pathways. The properties of copper which make it a valuable protein cofactor are partially derived from its ionic valency (+1, +2, rarely +3), its electronic structure ($1s^2$, $2s^2$, $2p^6$,$3s^2$, $3p^6$, $3d^{10}$, $4s^1$) and its relatively small radius (atomic radius: Cu=0.128 nm, Cu^{+2}=0.071 nm, Cu^{+1}=0.074 nm). In order to reach the Cu^{2+} state, Cu has to break its (stabilising) full 3d shell. In redox active copper enzymes, the metal can only accept electrons from substrates when they act as ligands. When the ligand is coordinated to the metal, the activation energy of the catalysed reaction is lower, thus the reaction rate is faster. One such redox active copper protein is the type II copper enzyme galactose oxidase.

1.2 Galactose oxidase

1.2.1 Introduction

Galactose oxidase is a copper containing enzyme that is secreted by the filamentous wheat-rot fungus *Fusarium graminearum*. The 68.5 kDa enzyme is monomeric and comprises 639 amino acids. The enzyme was discovered in the late 1950s and soon became the subject of much interest, due to its ability to catalyse a two electron reaction with apparently only a single, one electron, copper redox site [Cooper *et al.*, 1959]. The biological role of the enzyme is unknown, yet it has shown an ability to catalyse the oxidation of a large range of primary alcohols leading to the release of hydrogen peroxide, somewhat similar to the reaction carried out by the fungal flavoenzyme glucose oxidase [Keilin and Hartree, 1948]. The role of glucose oxidase is thought to be bacteriostatic, allowing the fungal organism to compete more effectively for nutrients [Whittaker *et al.*, 1998].

The reaction catalysed by galactose oxidase, as mentioned above, is the oxidation of primary alcohols to their corresponding aldehydes, accompanied by the reduction of molecular oxygen to hydrogen peroxide:

$$\mathbf{R\text{-}CH_2OH + O_2 \rightarrow R\text{-}CHO + H_2O_2}$$

The lack of an identifiable secondary cofactor was a major obstacle in understanding the catalytic mechanism. Initially there were three proposals which attempted to address this problem:

a) the redox process involves a Cu^{3+}·Cu^{1+} couple [Hamilton *et al.*, 1973];
b) covalently bound pyrroloquinoline quinone (PQQ) is a secondary cofactor [van der Meer *et al.*, 1989]; and
c) the enzyme has a tyrosine free radical at the active site [Whittaker and Whittaker, 1988].

1.2.2 Hamilton Theory

Hamilton *et al.* carried out extensive studies on the increased enzymatic activity observed when galactose oxidase is assayed against oxidants such as horseradish peroxidase, ferricyanide or EDTA-Mn^{3+} complex [Hamilton *et al.*, 1978]. As a result of this work, Hamilton established that the native enzyme was purified as a mixture of catalytically active and inactive galactose oxidase. The majority of native enzyme was proposed to be in a catalytically inactive Cu^{2+} oxidation state. Oxidation with ferricyanide activated the enzyme, and reduction with ferrocyanide inhibited the substrate oxidative ability of the enzyme. It was thus suggested that the catalytically oxidised state contained a trivalent copper species. The proposed model for the mechanism suggested both electrons from the alcohol substrate would be transferred to galactose oxidase by reducing Cu^{3+} to Cu^{1+} and molecular oxygen would then reoxidase Cu^{1+} to Cu^{3+} (Figure 1.2). The idea of the alcohol oxidation step involving a transient Cu^{2+}-superoxide intermediate was also suggested. The authors were not sure about the electronic configuration of Cu^{3+}, however, as they could not experimentally distinguish Cu^{3+}-X^{-} from its free radical resonance hybrid partner, Cu^{2+}-$X^{\bullet}$, where X was proposed to be an enzymatic group complexed to the copper.

1.2.3 PQQ theory

Van der Meer suggested the presence of PQQ in galactose oxidase [van der Meer *et al.*, 1989]. It was believed to have been extracted as a hexanol adduct under strong acidic conditions and then characterised by comparing its spectroscopic properties and elution profile during liquid chromatography against the authentic synthetic PQQ adduct. Quantitatively one PQQ per protein molecule was identified and hence a model for the catalytic mechanism was proposed (Figure 1.3). The procedure used to extract the PQQ in the form of a hexanol adduct, however, was later shown not to be specific for PQQ [Thomson, 1991] and the methods used to identify the isolated PQQ derivative were not considered rigorous.

1.2.4 Tyrosine free radical theory

Whittaker and Whittaker demonstrated that native galactose oxidase exists as a mixture of two species: a Cu^{2+}-EPR-silent, active form and a Cu^{2+}-EPR-observable, inactive species. Limited exposure of the enzyme to the oxidant potassium ferricyanide produced a stable, homogeneous active species of the enzyme. As the copper stayed in its Cu^{2+} valency state after oxidation of galactose oxidase and no other dissociable prosthetic groups were present,

the protein itself was proposed to be involved in catalysis [Whittaker and Whittaker, 1988]. Whittaker and Whittaker predicted an endogenous protein free radical to be situated in the active site, which was EPR silent due to antiferromagnetic coupling with Cu^{2+}. Spectroscopy confirmed the presence of the free radical, while EPR and ENDOR studies on galactose oxidase associated with deuterated tyrosines suggested the free radical was closely associated with a tyrosine residue ligated to the copper cofactor [Babcock *et al.*, 1992; Whittaker *et al.*, 1989]. The Whittakers' proposed a catalytic mechanism involving a free radical (Figure 1.4). After the one electron process converting the inactive enzyme to the active form, the radical site abstracted an electron or hydrogen atom from the substrate producing a hydroxyl radical species that could reduce the copper site. The aldehyde product was proposed to remain tightly bound to the enzyme while molecular oxygen interacted with the copper site, leading to its two-electron reduction. The dissociation of the aldehyde and H_2O_2 products marked the completion of the turnover cycles.

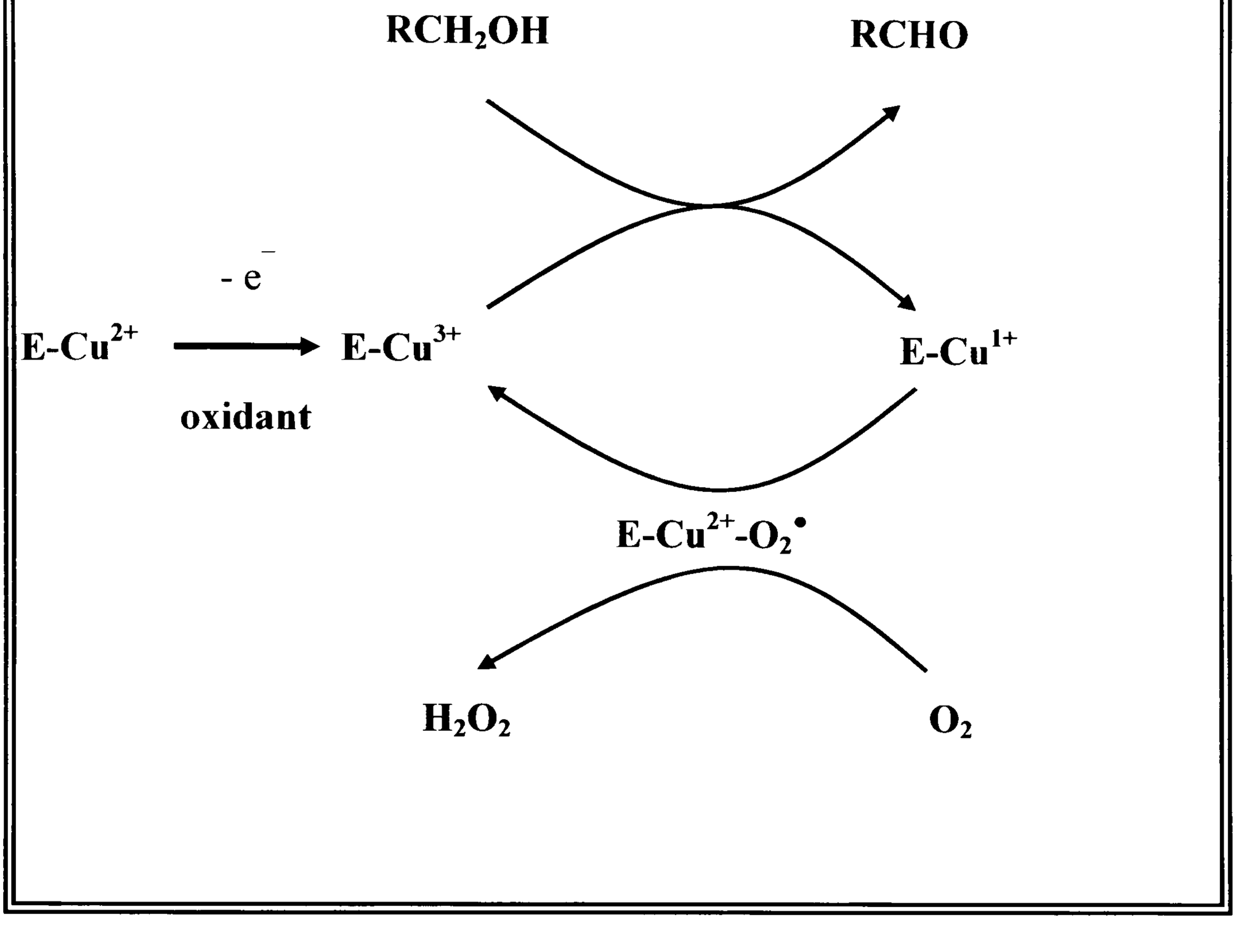


Figure 1.2: Cu^{3+}·Cu^{1+} coupling model for galactose oxidase.

An oxidant is needed to convert the enzyme from the inactive Cu^{2+} form to the catalytically active trivalent copper species. The two redox equivalents would be accommodated into the $Cu^{3+}\leftrightarrow Cu^{1+}$ couple. Adapted from Hamilton *et al.*, [1976].

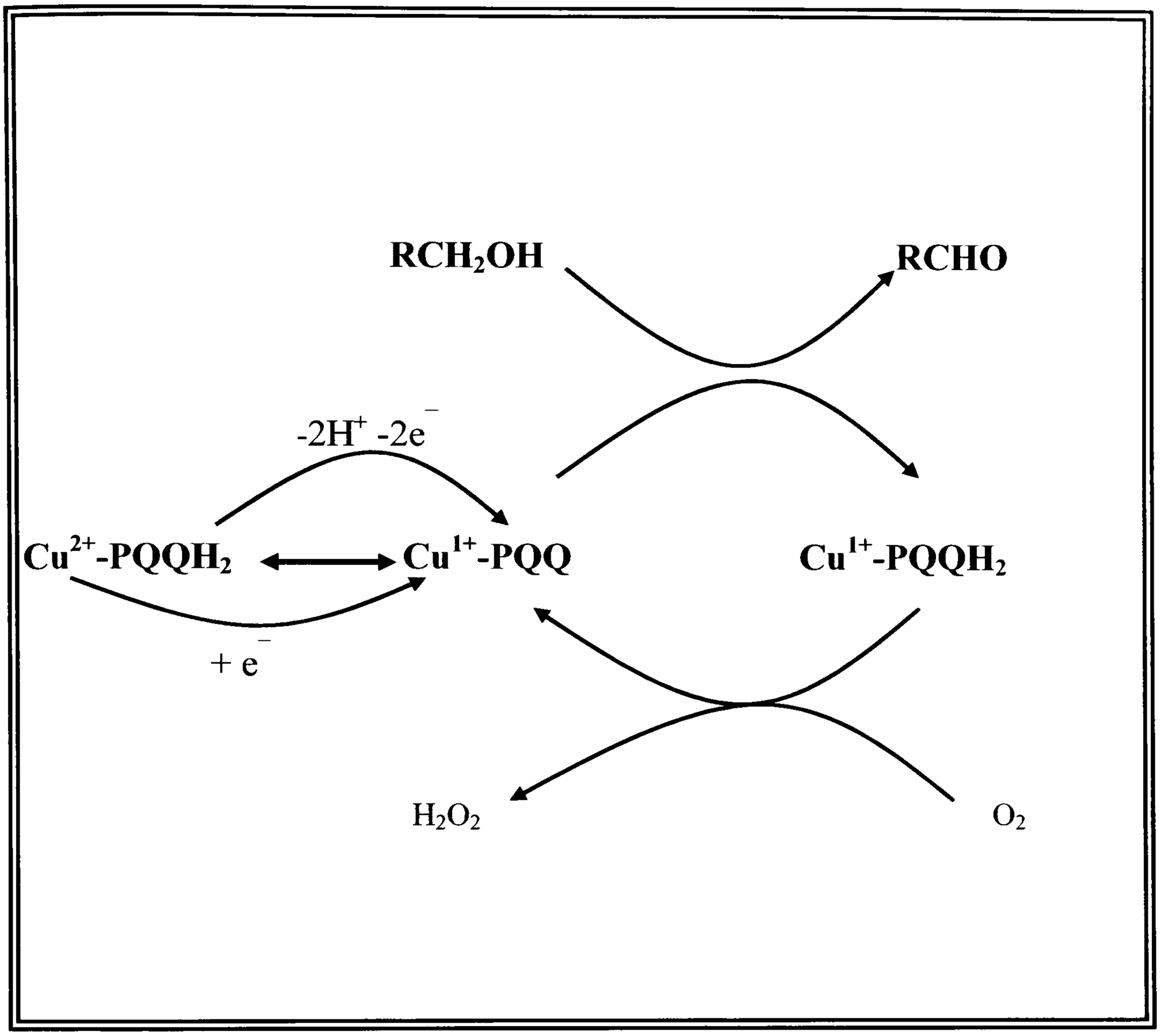


Figure 1.3: Catalytic model for galactose oxidase involving PQQ and copper as cofactors.

In this model the two redox cofactors would be in PQQ while the copper remains as Cu^{1+}. Two hydrogen ions from the substrate would be accepted by PQQ, thus oxidising the alcohol to aldehyde. Reoxidation of $PQQH_2$ to PQQ occurs via Cu^{1+} ion with oxygen reduced to hydrogen peroxide. Adapted from van der Meer *et al.*, [1989].

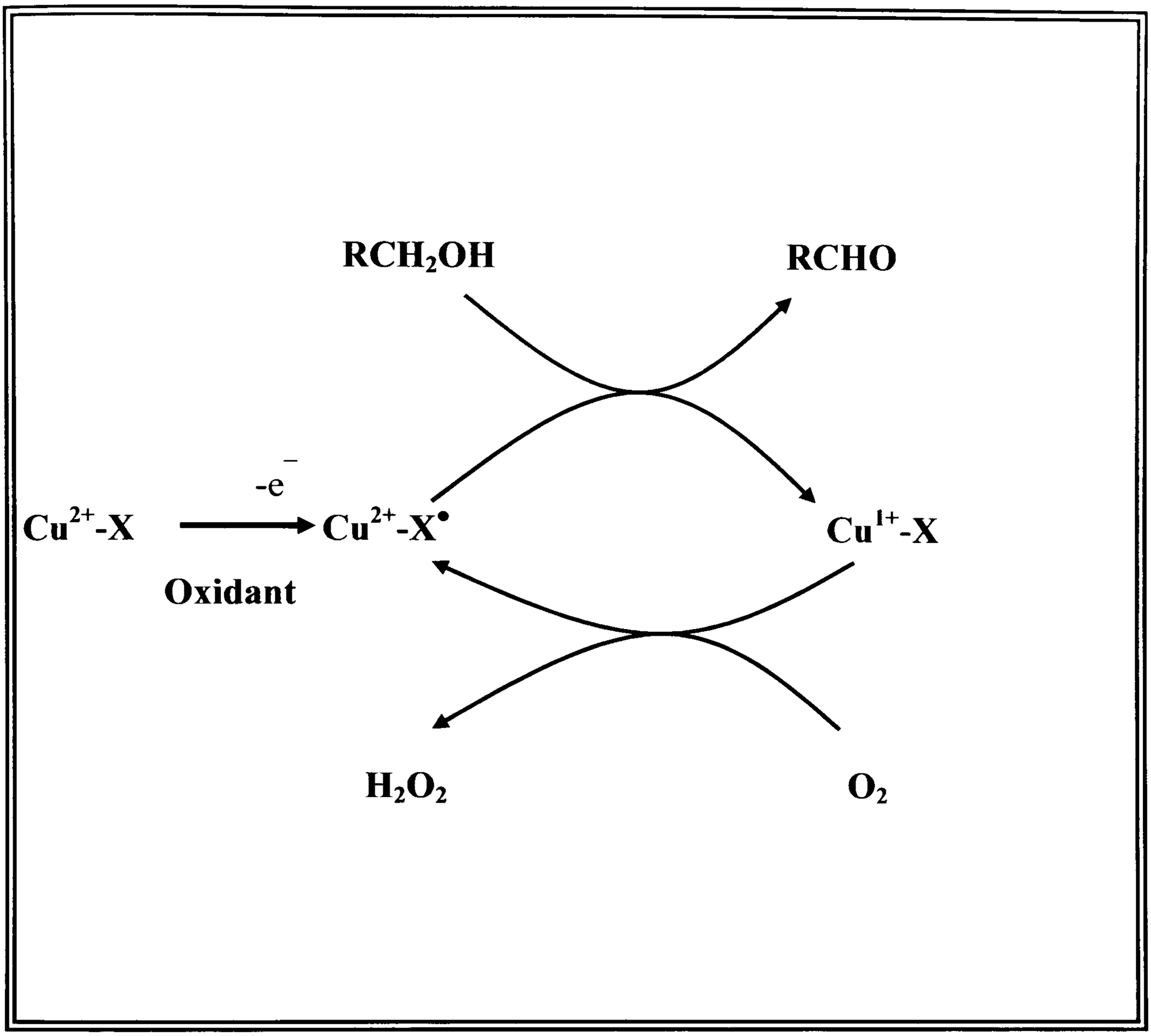


Figure 1.4: Catalytic mechanism of galactose oxidase involving a protein free radical.

X indicates a protein group that undergoes redox modification to create a free radical. Adapted from Whittaker and Whittaker, [1988].

1.3 The crystal structure of galactose oxidase

1.3.1 Introduction

The crystal structure of galactose oxidase revealed insights into the nature of copper binding and helped confirm that the catalytic mechanism did involve free radical chemistry [Ito *et al.*, 1991]. The structure of native enzyme revealed what was at the time a unique fold, comprising three distinct predominantly β-sheet domains (Figure 1.5). Two disulphide bonds were identified between Cys 18-Cys 27, at the bottom of a surface loop and between Cys 515-Cys 518, present in a β-turn. There was no indication of a large, exogenous molecule covalently bound to the enzyme, hence dismissing the theory that galactose oxidase contained a PQQ cofactor.

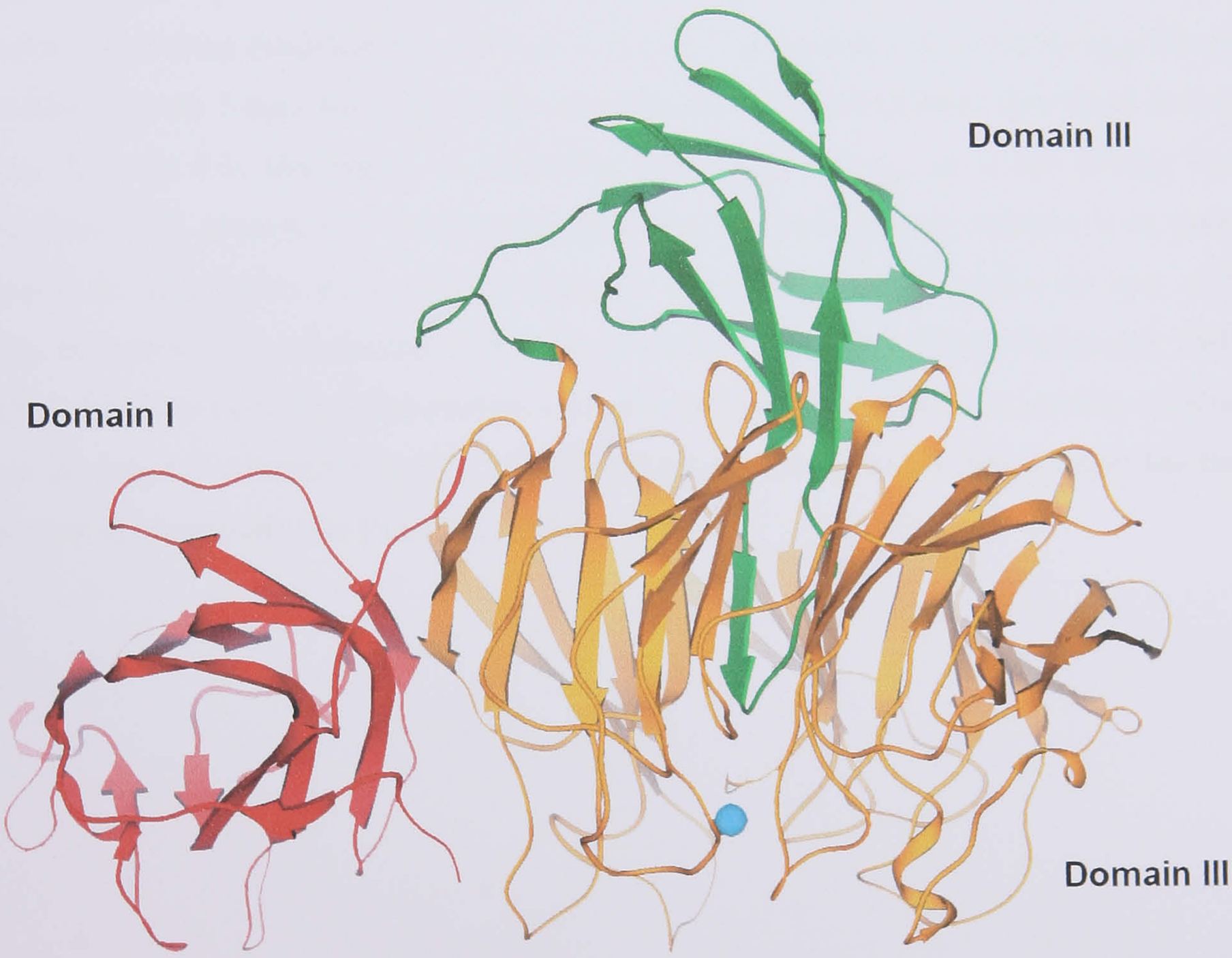


Figure 1.5: The overall structure of galactose oxidase.
Domain I is red, domain II is yellow, and domain III is green. The copper is represented as a cyan sphere.

1.3.2 Domain I

The first domain (residues 1-155) comprises 8 β-strands in a jelly-roll motif with a five-stranded antiparallel β-sheet facing a three-stranded antiparallel β-sheet (Figure 1.6). A galactose binding site was identified in this domain after galactose was soaked into galactose oxidase crystals, although it is not the active site (Figure 1.7), as it does not bind all substrates and is approximately 40 Å from the copper site [Ito, 1991]. The function of this second sugar binding site is unknown, however, it has been suggested that this site may serve to attach the protein to cell walls [Mendonca and Zancan, 1987]. Glycosylation of the enzyme with three to six sugars per polypeptide was reported [Mendonca and Zancan, 1988, 1989], even though galactose oxidase does not contain any glycosylation motifs [McPherson *et al.*, 1993]. The crystal structure revealed no covalently attached carbohydrates, however, enzyme purified and crystallised from *Fusarium* has not been reported. The presence of a second sugar binding site within domain I may thus explain the identification of carbohydrate associated with the protein. Domain I is thought to be important for protein folding, as it has several close interactions with domain II. Recombinant galactose oxidase, lacking domain I, is poorly expressed in *Aspergillus niger*. Protein cleavage separating domain I from the rest of the protein, however, revealed the rest of the protein retained catalytic activity [Mahmoud, 2001]. Domain I has also been shown to contain a metal binding site, identified as binding a sodium ion, in the original structure [Ito *et al.*, 1991] and calcium [Ito *et al.*, 1994]. Copper has been shown not to bind to this site [Ito *et al.*, 1991].

Figure 1.6: Representation of galactose oxidase domain I.

Domain I (coloured red) adopts a jelly roll motif, containing five antiparallel β-strands facing three antiparallel β-strands. The relative orientation of galactose oxidase enzyme is shown below the image of domain I (rotated 90° into the plane of the page compared with Figure 1.5). The blue arrow depicts the location of the sodium ion, and the black arrow indicates the secondary galactose binding site.

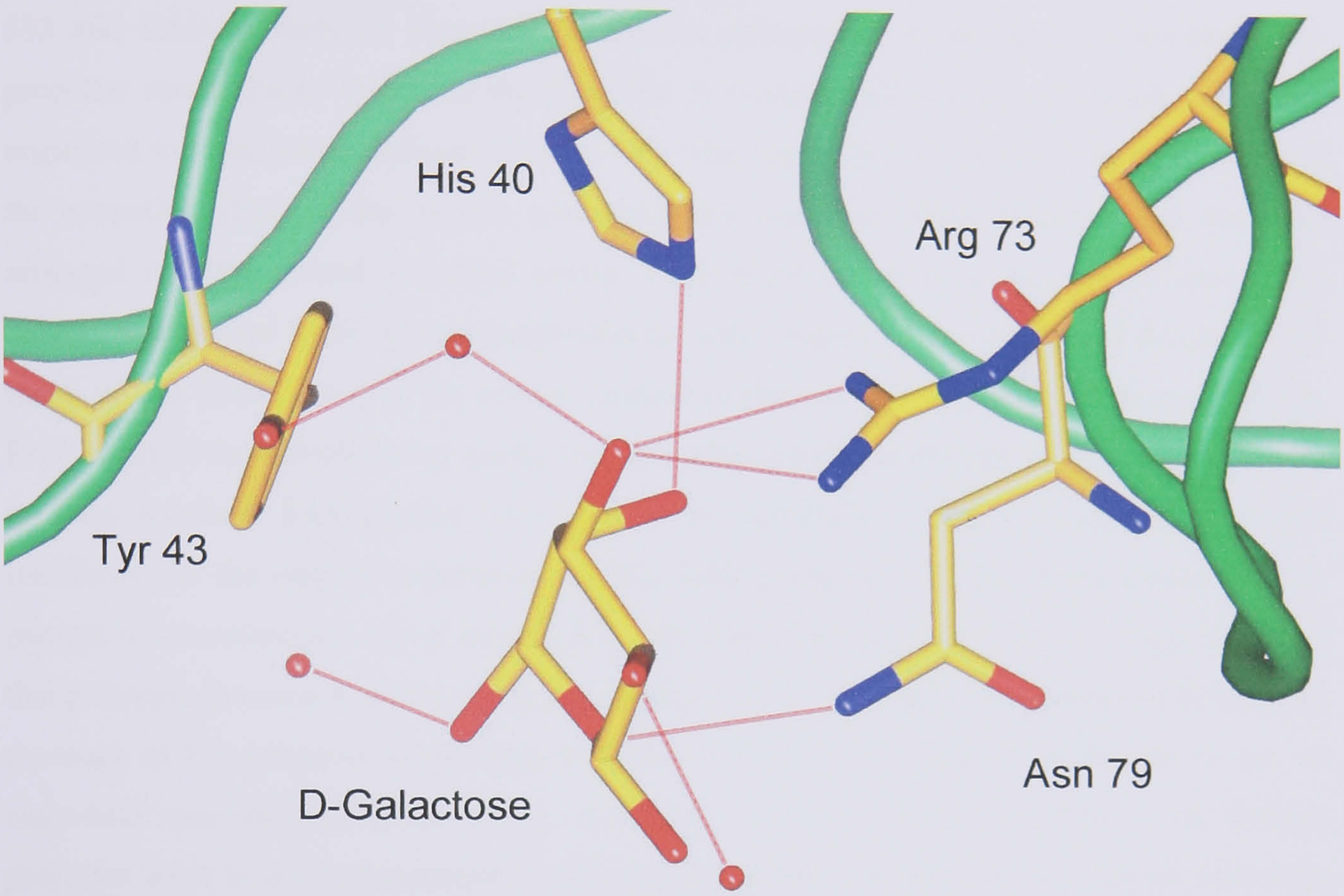


Figure 1.7: The secondary sugar binding site identified in domain I of galactose oxidase.

Sugar soak experiments have identified a binding site in domain I with an affinity for galactose, xylose and fucose [Ito *et al.*, 1994, Firbank, 2002]. Smaller primary alcohols and inhibitors have been shown not to bind to the site.

1.3.3 Domain II

Domain II is the largest of the three domains and contributes three of the four protein derived active site copper ligands, Tyr 272, Tyr 495 and His 496. The domain contains residues 156-532 and folds to form 28 β-strands which are arranged in a distinctive seven-bladed β-propeller motif. Each 'blade' of the propeller is made up from four anti-parallel β-strands organised to form seven β-sheets (Figure 1.8). The copper ion is located near the surface of the propeller, close to the pseudo sevenfold axis with the seven blades of the propeller arranged radially around a central cavity. Each blade twists in a way which causes the innermost β-strand to be almost perpendicular with respect to the outermost β-strand. The result of this twist is to give the central channel of domain II a slightly conical morphology. Each sheet of the circular array packs onto the adjacent sheets through hydrophobic contacts to form localised hydrophobic cores, hence the non-polar residues at the sheet-to-sheet interfaces are the biggest constraints on this folding topology. The central cavity is large enough to accommodate two β-strands from the third domain, with a residue from the tip of this β-hairpin forming a ligand to the active site copper. An early investigation of the folding topology of the propeller motif suggested the seven-fold symmetry is preferable to six- or eight-fold symmetry [Murzin, 1992]. According to the modelling studies, a six β-sheet propeller must have predominantly small residues in the β-strands close to its six-fold axis, but no strong sequence constraints are necessary for a seven-fold β-propeller.

The β-propeller is regarded as a very rigid structure. The closing of the propeller is thought to play an important part in stabilising the propeller structure by holding the blades in a fixed position. The most common method of closure occurs when either the first or last of the blades in a propeller is created from β-strands from both termini of the domain. In galactose oxidase, for example, the first β-strand of the domain forms the outermost β-strand of the last propeller (β-sheet) (Figure 1.8). This arrangement ensures the tight interactions that hold the ring in place and prevents the motif from unfolding. This closure mechanism is referred to as "velcro" or "molecular clasp". Galactose oxidase adopts the N-terminal β-strand closure, while other propeller proteins are known to adopt a C-terminal closure. There are exceptions to this method of propeller closure. The three four-bladed β-propellers that do not exhibit "velcro" closure mechanisms, instead using disulphide bridging between the N and C-termini [Gohlke *et al.*, 1996, Li *et al.*, 1995, Faber *et al.*, 1995].

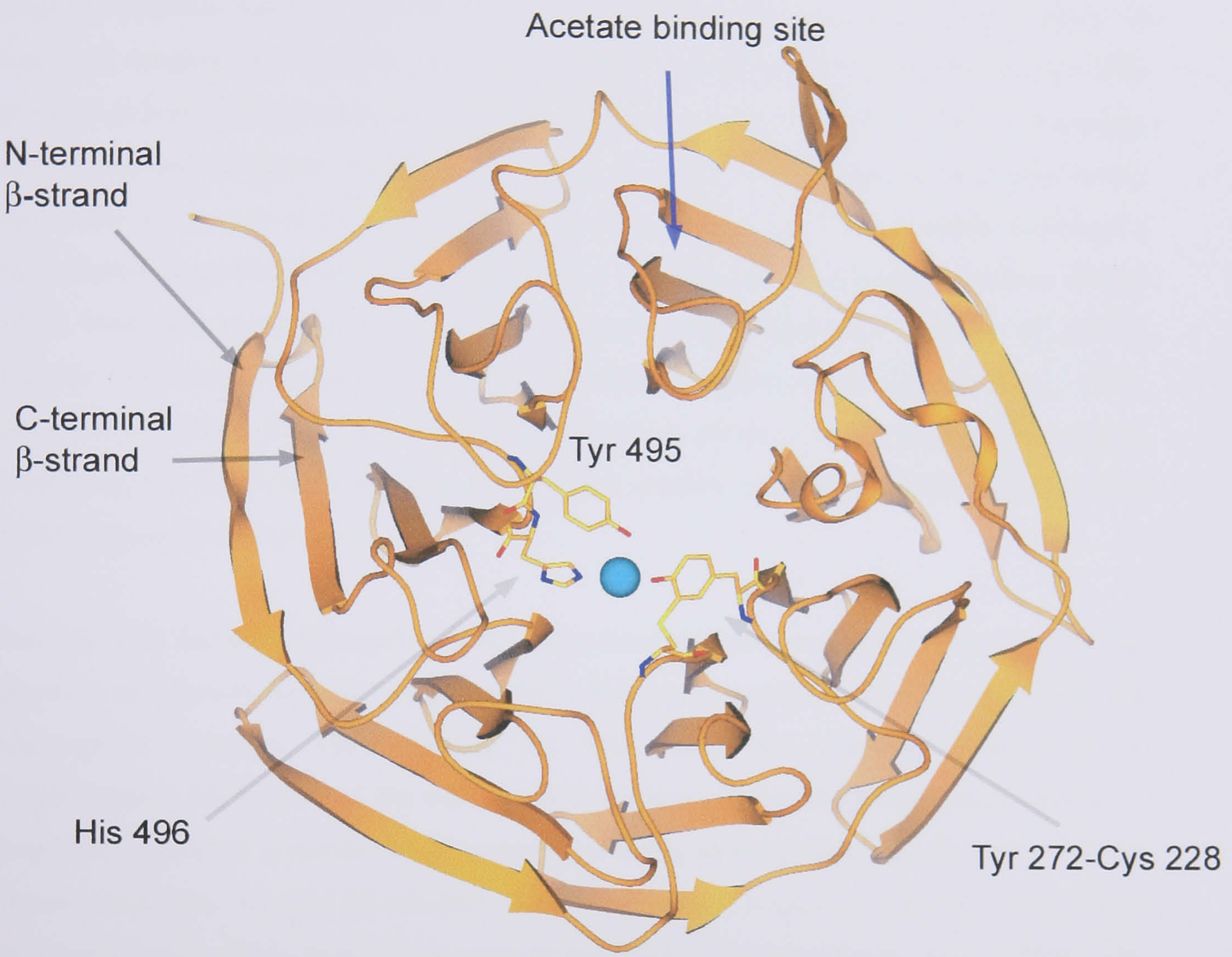


Figure 1.8: The structure of domain II, looking down the pseudo seven-fold axis.

Galactose oxidase domain II is a rare kelch motif protein, considering it contains seven 'blades' and adopts an N-terminal β-propeller closure mechanism. The first strand of domain II is used to complete the final blade of the propeller. The active site copper is represented by a blue sphere and coordinates three ligands contributed by domain II residues. The blue arrow indicates an acetate binding site located on blade 5 [Ito *et al.*, 1991].

The first β-propeller structure reported was the six-bladed propeller of neuramidase from the influenza virus [Varghese *et al.*, 1983]. The next propeller structure reported was the seven-bladed structure of methylamine dehydrogenase [Vellieux *et al.*, 1989], followed closely by the structure of galactose oxidase [Ito *et al.*, 1991]. Currently, over 20 non-homologous but structurally related propeller domains have been described (Table 1.1). Even though β-propeller proteins have similar topology and structures, they perform a huge variety of functions, ranging from enzyme catalysis to scaffold and signalling molecules. Almost 50% of reported β-propeller proteins are enzymes and their catalytic activities cover a broad range of reactions and substrates. Prosthetic groups are generally found bound to the central cavity that forms as a result of the β-sheet packing. For example, the haem of nitrite reductase's eight-bladed propeller [Fulop *et al.*, 1995], the cofactor pyrrolo quinoline quinone (PQQ) which binds non-covalently to the six-bladed glucose dehydrogenase [Oubrie *et al.*, 1999], and the eight-bladed methanol, alcohol, and ethanol deydrogenases [Ghosh *et al.*, 1995; Oubrie *et al.*, 2002; Xia *et al.*, 1992]. The ubiquitous presence of the β-propeller motif in nature suggests the β-sheets acts as a scaffold that acquires specific activities once decorated with functional loop regions.

Proteins with the same β-propeller motif often have different sequences, although, several sequence motifs have been identified in relation to known crystal structures that are predictors of propeller domains. These include the WD motif, the regulator of chromosome condensation 1 (RCC1) motif, the tachylectin-2 repeat, as well as the kelch motif (Table 1.1). Sequence repeats are generally 40-50 residues in length, although loop insertions can result in protein sequences of over 80 residues. The tachylectin-2 sequence comprises five nearly identical repeats which form a symmetrical five-bladed domain [Beisel *et al.*, 1999]. The haemopexin domains have a strong consensus sequence where the first and second strands of each β-sheet are relatively highly conserved [Paoli, 2001]. The RCC1 repeat motif Val-Tyr-x-Trp-Gly is associated with clustering of hydrophobic side chains [Renault *et al.*, 1998]. Sequence repeats thought similar to the RCC1 repeat have been identified in the β-lactamase inhibitor protein II [Lim *et al.*, 2001]. The methanol dehydrogenase sequence repeats provide an example of the "tryptophan docking motif", characterised by the consensus Ala-x-Asp/Asn-x-xThr-Gly-Asp/Glu-x-x-Trp [Ghosh *et al.*, 1995].

Table 1.1: Proteins with known β-propeller motifs. Adapted from Jawad *et al.*, [2002].

No. of β-sheets	Protein	Function	Organism	Reference
*4	Haemopoxin	Haem binding and transport	Mammalian	[Faber *et al.*, 1995]
4	Collagenase	Substrate recognition and binding	Mammalian	[Li *et al.*, 1995]
4	Gelatinase	Substrate recognition and binding	Mammalian	[Gohlke *et al.*, 1996]
*5	Tachylectin-2	Carbohydrate binding in immune response	Invertebrate	[Beisel *et al.*, 1999]
*6	Sialidase	Removal of sialic acid residues	Bacterial	[Crennell *et al.*, 1993]
*6	Neuramidase	Removal of sialic acid residues	Viral	[Varghese *et al.*, 1983]
*6	Glucose dehydrogenase	Conversion of pentose and hexose sugars	Bacterial	[Oubrie *et al.*, 1999]
*6	Phytase	Phytic and acid hydrolysis	Bacterial	[Ha *et al.*, 2000]
*6	TolB	Complex component involved in cell invasion	Bacterial	[Abergel *et al.*, 1999]
*6	Low-density lipoprotein receptor	Receptor signalling/associations	Mammalian	[Jeon *et al.*, 2001]
*6	Keap1	Substrate adaptor protein for transcription factor degradation complex	Mammalian	[Li *et al.*, 2004]

The * symbol denotes proteins with distinct sequence repeats

No. of β-sheets	protein	Function	Organism	Reference
*6	Dispropyl flouro phosphatase	Organophosphatase compounds hydrolysis	Invertebrate	[Scharff *et al.*, 2001]
*6 & 7	Tricorn protease	Cytosolic protein degradation	Archea	[Brandstetter *et al.*, 2001]
*7	Methylamine dehydrogenase	Oxidation of primary amines	Bacterial	[Vellieux *et al.*, 1989]
*7	Galactose oxidase	Oxidation of primary alcohols	Fungal	[Ito *et al.*, 1991]
*7	Transducin G β-subunit	Scaffold in signal transduction complex	Mammalian	[Sondek *et al.*, 1996]
7	Tup1	Transcriptional repression	Yeast	[Sprague *et al.*, 2000]
7	ARPC1	Initiation of actin polymerisation	Mammalian	[Robinson *et al.*, 2001]
*7	Regulator of chromosome condensation	Control of nucleo-cytoplasmic transport	Mammalian	[Renault *et al.*, 1998]
7	β-lactamase inhibitor protein-II	Inhibition of β-lactamase	Prokaryotic	[Lim *et al.*, 2001]
*7	Prolyl oligopeptidase	Hormonal and neural peptides metabolism	Mammalian	[Fulop *et al.*, 1998]
*7	Clathrin head N-domain	Vesicle coating for intracellular transport	Mammalian	[ter Haar *et al.*, 1998]
*7	Nitrous oxide reductase	Reduction of nitrous oxide to nitrogen	Bacterial	[Brown *et al.*, 2000b]

The * symbol denotes proteins with distinct sequence repeats

No. of β-sheets	protein	Function	Organism	Reference
*7	Intergrin extracellular segment	Cell-cell and cell-matrix adhesion	Mammalian	[Xiong *et al.*, 2001]
*7	Quino-haemo amine dehydrogenase	Amines oxidation	Bacterial	[Datta *et al.*, 2001]
*8	Methanol dehydrogenase	Oxidation of primary alcohols	Bacterial	[Xia *et al.*, 1992]
8	Ethanol dehydrogenase	Oxidation of primary alcohols	Bacterial	[Keitel *et al.*, 2000]
8	Alcohol dehydrogenase	Oxidation of primary alcohols	Bacterial	[Oubrie *et al.*, 2002]
*8	Nitrite reductase	Reduction of nitrite and oxygen	Bacterial	[Fulop *et al.*, 1995]

The * symbol denotes proteins with distinct sequence repeats

While the proteins mentioned above have relatively high sequence similarity between β-sheets, other cases, such as the WD and kelch repeats show much lower similarities [Neer and Smith, 1996]. Galactose oxidase was the first member of the kelch repeat to be crystallographically characterised. The founding member of the kelch family was identified in *Drosophila melanogaster*, where the Kelch protein regulates the organisation of actin fibres in the developing embryo [Xue and Cooley, 1993]. The repeats are an ancient and evolutionarily wide-spread sequence motif of 44-55 amino acids in length and have been identified in a diverse number of organisms such as viruses, plants, fungi, and mammals including humans. The consensus kelch repeat motif contains several highly conserved residues, including two adjacent glycine residues and a tyrosine/tryptophan pair separated by exactly seven residues (Figure 1.9).

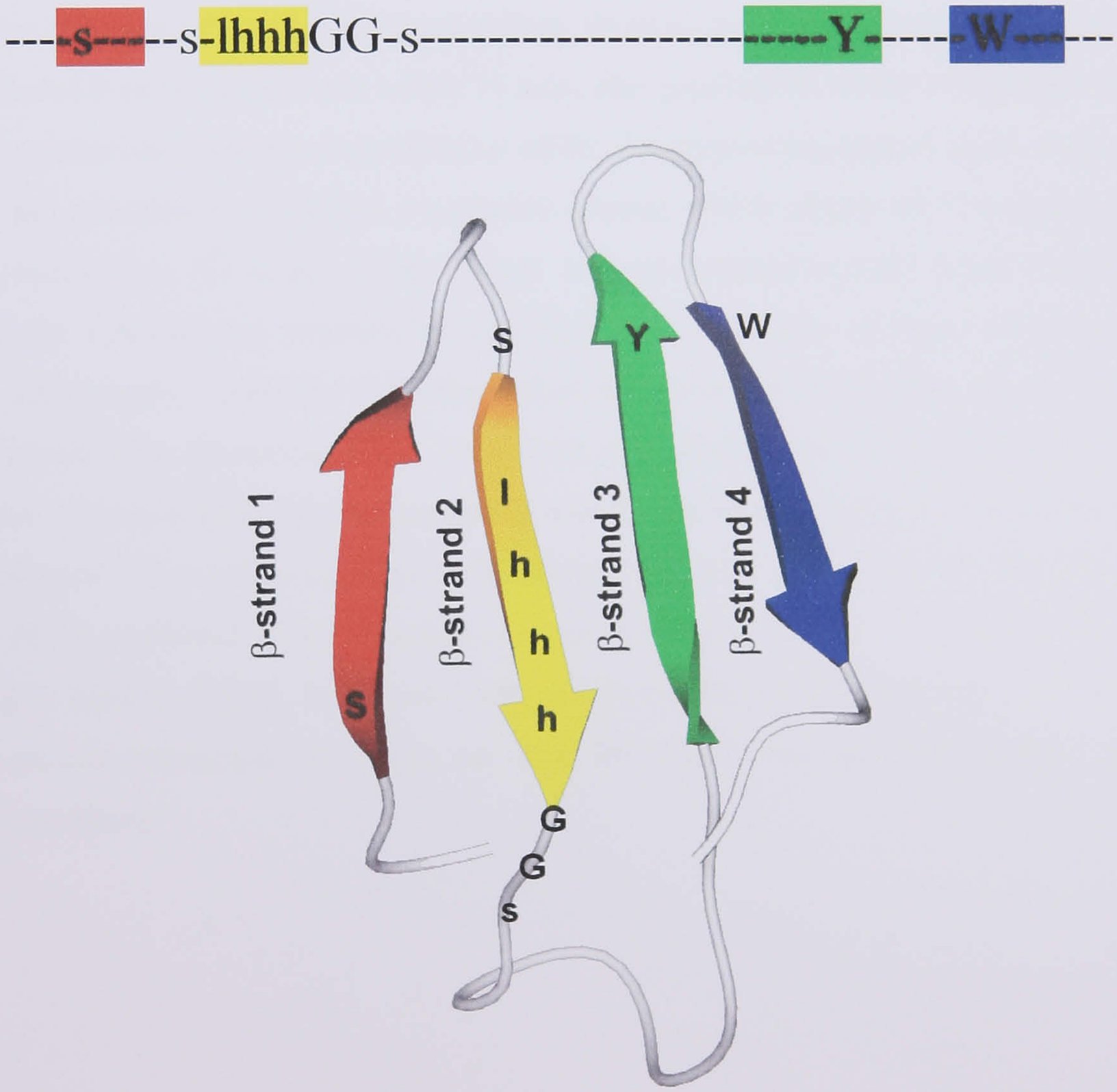


Figure 1.9: The consensus sequence of the kelch motif.

The sequence motif is shown in relation to the four stranded β-sheets that form the blades of the propeller structure. The colours red to blue correspond to the β-strands 1 to 4. In the consensus, G= glycine, Y= tyrosine, W = tryptophan, s = small residue, l = large residue, and h = hydrophobic residue. The most highly conserved residues are located on the β-strands. Adapted from Prag and Adams, [2003].

Based on the knowledge of structures containing sequence repeats, predictions of similar protein folds were made for novel proteins. For example, the bacterial periplasmic protein TolB contained Ala-X-Ser-Pro-Asp motifs and high predicted β-strand content which led to the suggestion of a six-bladed β-propeller domain [Ponting and Pallen, 1999]. This prediction was later substantiated by the X-ray crystal structure [Abergel *et al.*, 1999; Carr *et al.*, 2000]. Similarly, analysis of sequence and secondary structure data, threading methods and experimental data predicted the Tyr-Trp-Thr-Asp (YWTD) motif to assume a six-bladed β-propeller fold [Springer, 1998]. The YWTD motif has been identified in over 60 types of extracellular domains and was confirmed to contain a six-bladed β-propeller domain in the crystal structure of the low-density lipoprotein receptor [Jeon *et al.*, 2001].

An examination of the predicted protein sequences derived from the Human genome revealed at least 71 potential kelch-repeat proteins [Prag and Adams, 2003]. Further analysis suggested 9 of these represented five-bladed β-propellers, 60 were six-bladed and the remaining 2 were seven-bladed β-propellers. 55 out of the 71 were also predicted to adopt a C-terminal β–strand closure mechanism. The recent publication of the structure of the human kelch motif protein, Keap1, has revealed a six-bladed β-propeller domain which adopts the C-terminal β-strand closure mechanism [Li *et al.*, 2004]. Keap1 is a cytoplasmic protein which binds to actin forming the cytoskeleton required to sequester Nrf2 (regulator of stress response genes). Keap1, additionally, binds the Cul3-dependent ubiquitin ligase complex via its N-terminal BTB domain (for Broad-complex, Tramtrack and Bric-a-brac). The BTB domain is a conserved protein-protein interaction motif, usually found at the N-terminus [Ahmad *et al.*, 1998]. Several other kelch-repeat proteins contain the BTB domain. In the Human genome, 51 out of the predicted 71 kelch motif proteins are also predicted to have a BTB domains [Prag and Adams, 2003]. It appears that many BTB/kelch proteins are associated with protein-protein interaction, however, no such interaction has been recorded in regard to galactose oxidase.

1.3.4 Domain III

The third domain (residues 533-639) is located at the opposite end of domain II from the copper site and is made up of a bundle of seven β-strands surrounding a hydrophobic core (Figure 1.10). One long antiparallel β-hairpin penetrates into the central cavity of the domain II propeller motif and at the tip of this loop is His 581, the fourth protein ligand to the active site copper. Domain III may play a role in the stability of the protein, by providing a ligand for the copper, as studies indicated removal of copper facilitated the denaturation of the protein in urea [Kosman *et al.*, 1974]

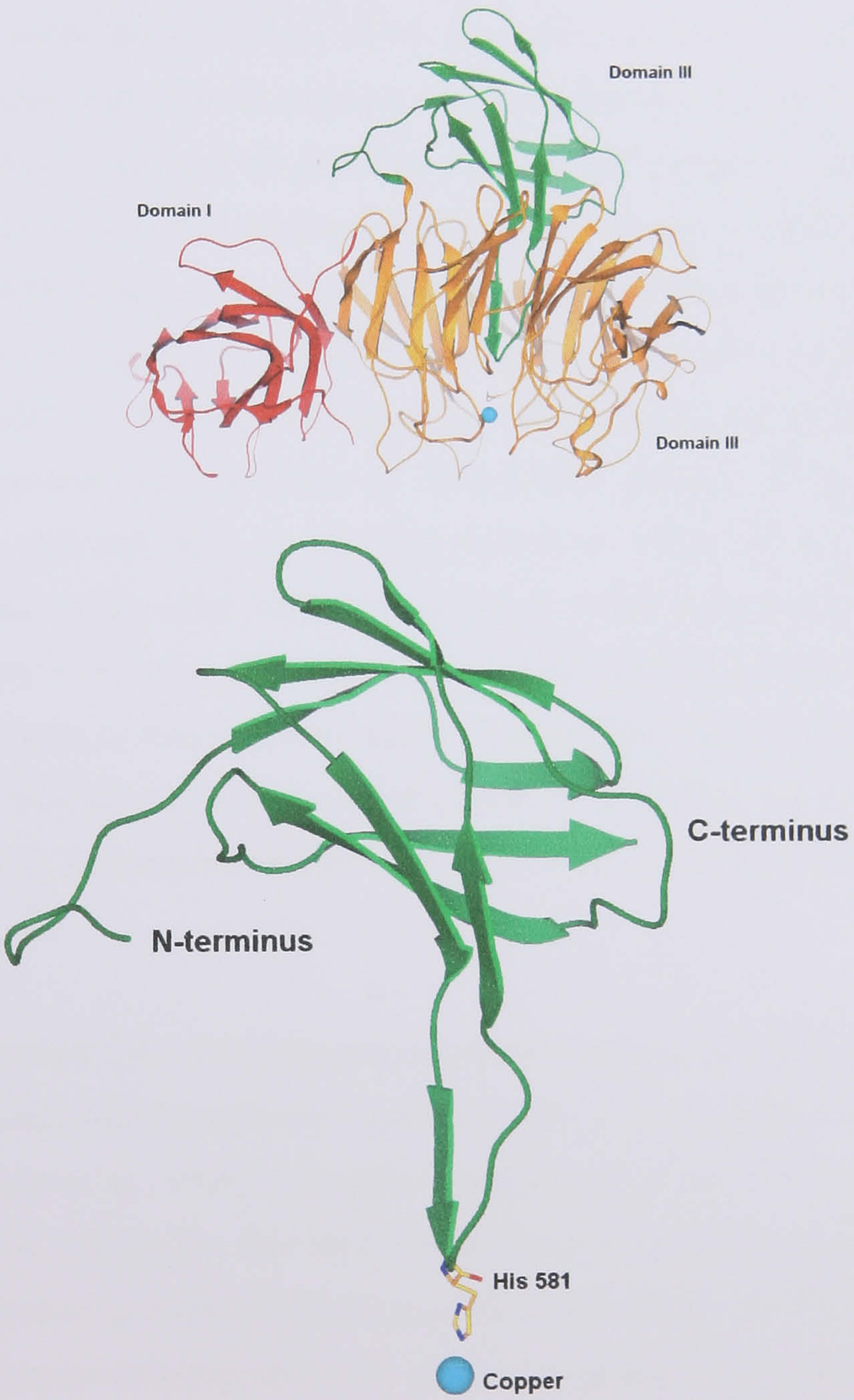


Figure 1.10: Secondary structure of galactose oxidase domain III.

A β-loop from domain III (green) protrudes into the central cavity of Domain II, providing the active site copper with a fourth protein ligand, His 581.

1.4 The thioether bond

Enzymatic cofactors have traditionally been described as low molecular weight structures that are separate from their cognate protein. The first crystallographically characterised cofactors were the haem structures visualised in myoglobin and haemoglobin [Kendrew *et al.*, 1960; Muirhead and Perutz, 1963]. Work over recent years has forced an expansion of the definition of cofactors to include structures that are derived from the protein itself. The publication of the crystal structure of galactose oxidase was the first structural observation of a protein derived cofactor. The structure revealed a covalent bond between the Sγ of Cys 228 and a Cε of Tyr 272 in the active site [Ito *et al.*, 1991]. Prior to the structural confirmation, previous experimental evidence had isolated unusual behaviour associated with the two residues in question. The primary sequence of galactose oxidase, for example, contains six cysteine residues [McPherson *et al.*, 1992] whereas titration of the reduced unfolded enzyme identified only five cysteines [Kosman *et al.*, 1974]. Amino acid sequencing failed to identify residue 272 [McPherson *et al.*, 1992], but use of resonance Raman evidence isolated the proposed radical site to a modified tyrosine [Whittaker *et al.*, 1989]. Supporting evidence for a covalent link within the protein was provided by SDS-PAGE analysis. During electrophoresis galactose oxidase migrated with an apparent molecular weight of 65.5 kDa, despite its calculated molecular weight of 68.5 kDa. This anomaly is due to the presence of the thioether bond, which creates a 44 amino acid loop that is not disrupted by protein denaturation. The loop allows the protein to migrate more rapidly because it is more compact than denatured galactose oxidase lacking the thioether bond [Baron *et al.*, 1994; McPherson *et al.*, 1993; Rogers *et al.*, 2000]. Subsequent research has uncovered a host of similar protein derived cofactors.

The covalently attached 2,4,5-trihydroxyphenylalanine quinone (TPQ) of amine oxidase was the only other protein derived cofactor reported at the time the galactose oxidase structure was described [Janes *et al.*, 1990]. Structural confirmation of the cofactor was presented in 1995 [Parsons *et al.*, 1995]. By that time the cofactor of methylamine dehydrogenase had already been confirmed as tryptophanyltryptophylquinone (TTQ) [McIntire *et al.*, 1991]. In the case of methylamine dehydrogenase, the cofactor is derived from a crosslink between two tryptophan residues. In 1996 another modified tyrosine cofactor was identified in the form of the lysine tyrosylquinone (LTQ) in lysyl oxidase, comprising a crosslink between lysine and a tyrosine [Wang *et al.*, 1996]. In 2001 the structure of quinohemoprotein amino dehydrogenase revealed the presence of cysteine tryptophylquinone (CTQ). As well as three thioether bonds between cysteine sulphurs and the Cβ of aspartate residues, the structure also revealed one

thioether bond between a cysteine sulphur and the Cγ of a glutamate residue [Datta *et al.*, 2001].

Outside of the quinone cofactors, other posttranslational modifications include the cytochrome c oxidase tyrosine to histidine cross link [Yoshikawa *et al.*, 1998], the proposed histidine to cysteine cross link at the CuA sites of catechol oxidase [Klabunde *et al.*, 1998], haemocyanin and tyrosinase. One of the active site histidines has been shown to cross link to a cysteine in the structures of catechol oxidase and haemocyanin. A similar link has been biochemically identified in tyrosinase, however, the enzyme structure has not been determined yet [Lerch, 1982].

The tyrosine to cysteine thioether bond in galactose oxidase is predicted to exist in glyoxal oxidase, a galactose oxidase homologue. The fungal copper enzyme glyoxal oxidase was isolated from *Phanerochaete chrysosporium* and has similar spectral features to galactose oxidase [Whittaker *et al.*, 1996b]. Glyoxal oxidase catalyses the oxidation of aldehydes to the corresponding carboxylic acids with the utilisation of molecular oxygen and co-production of hydrogen peroxide. The active site is predicted to closely resemble that of galactose oxidase, apart from a histidine residue which is thought to replace the tryptophan positioned over the thioether bond. This predicted single change appears to account for the different properties of the two proteins, including; redox stability, pH sensitivity, anion binding reactivity and differing redox potentials.

Since its description, the nature of thioether bond has been extensively investigated. The bond is absolutely necessary for enzyme activity, mutation of the cysteine to a glycine results in an enzyme with negligible activity, although a 76% reduction of copper content is also recorded in the C228G mutant structure [Baron *et al.*, 1993]. The role of the thioether bond is postulated to be the lowering of the tyrosine redox potential, hence allowing easier accessibility for catalysis. The redox potential of free tyrosine is 0.93 V [Harriman, 1987] but in galactose oxidase it is only 0.4-0.5 V [Hamilton *et al.*, 1978]. When the protein is oxidised, an unpaired electron radical resides on the cysteine-tyrosine cofactor [Whittaker *et al.*, 1989].

The active site configuration of galactose oxidase is thought to greatly increase the stability of the radical. One of the main ways of stabilisation is proposed to be the distribution of the radical across Tyr-Cys-Cu metalloradical complex [Gerfen *et al.*, 1996; Whittaker *et al.*, 1996a]. This free radical-coupled copper complex is extremely stable and in the absence of reducing agents has been shown to persist for several weeks at room temperature [Whittaker *et al.*, 1998]. The radical has been suggested to be distributed over the cysteine and tyrosine.

Computational calculations suggested that only 12% of the unpaired spin is situated on the cysteine [Engstrom *et al.*, 2000], although a 28% figure has also been reported [Whittaker *et al.*, 2000]. Additional radical stability may be derived from the active site residues immediately surrounding the radical complex. The active site tryptophan stacks over the thioether bond, and by protecting the bond from solvent is proposed to provide an important role in radical stability [Baron *et al.*, 1993; Ito *et al.*, 1991; Ito *et al.*, 1992]. Whittaker has proposed that with the presence of the copper ligand, Tyr 495, there remains a possibility for the unpaired electron to be delocalised over both tyrosine residues via the copper [Whittaker, 2003; Whittaker, 2005]. Resonance stability, as a consequence may provide the explanation for the unusually low redox potential of the galactose oxidase free radical site (Figure 1.11).

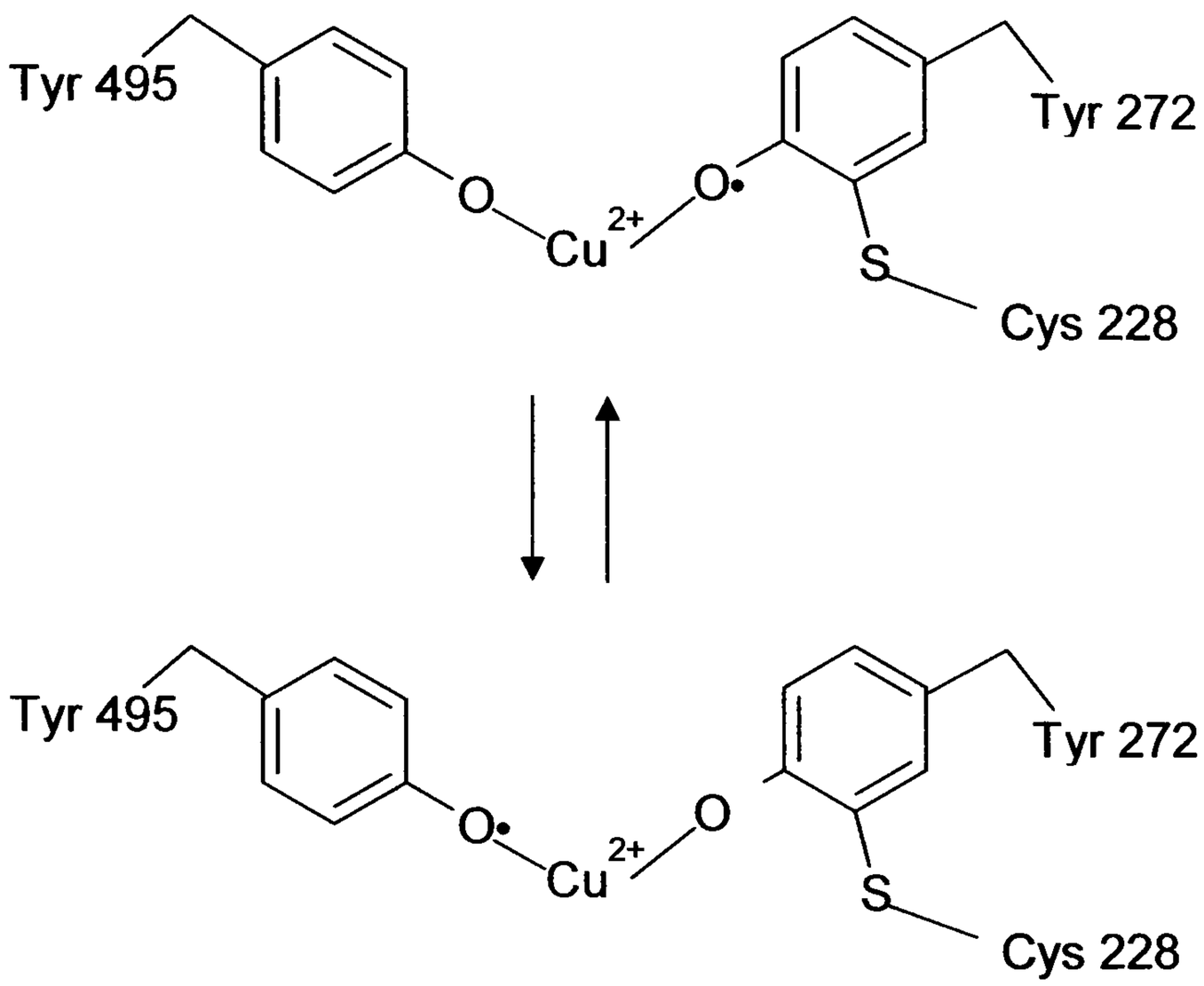


Figure 1.11: Schematic representation of the resonance stabilisation complex.

Schematic representation of the resonance stabilisation complex proposed to promote radical stability. The radical is delocalised across both tyrosine ligands, the copper ion serves as a bridge between the residues. Adapted from Whittaker [2002]

1.5 Biochemistry of galactose oxidase

1.5.1 Active site structure

The active site copper, as mentioned earlier, is coordinated by three ligands from domain II: Tyr 272, Tyr 495 and His 496, with one ligand, His 581, provided by Domain III. A fifth ligand originating from the surrounding solution is also coordinated. Tyr 495 acts as the axial ligand in the square pyramidal geometry around copper (Figure 1.12). The original galactose oxidase structure was crystallised at pH 4.5 in acetate buffer, the fourth equatorial ligand was found to be an acetate ion while a water molecule was found at this position when crystals were transferred to an acetate-free buffer at pH 7.0 [Ito *et al.*, 1994]. Prior to the determination of the protein structure, spectral analysis had predicted a square pyramidal active site [Kosman *et al.*, 1980; Whittaker and Whittaker, 1988], although the axial ligand had been suggested to be a second exogenous ligand rather than Tyr 495. Examination of the copper distance from the axial ligand suggests the weak binding of Tyr 495 may have resulted in the initial conclusions. Tyr 495 is found at a distance 2.7 Å from copper in both the pH 4.5 and 7.0 structures (PDB code 1GOF and 1GOG respectively). The equatorial water molecule is also observed at unusually long distance 2.8 Å while acetate is located at the more reasonable 2.5 Å distance.

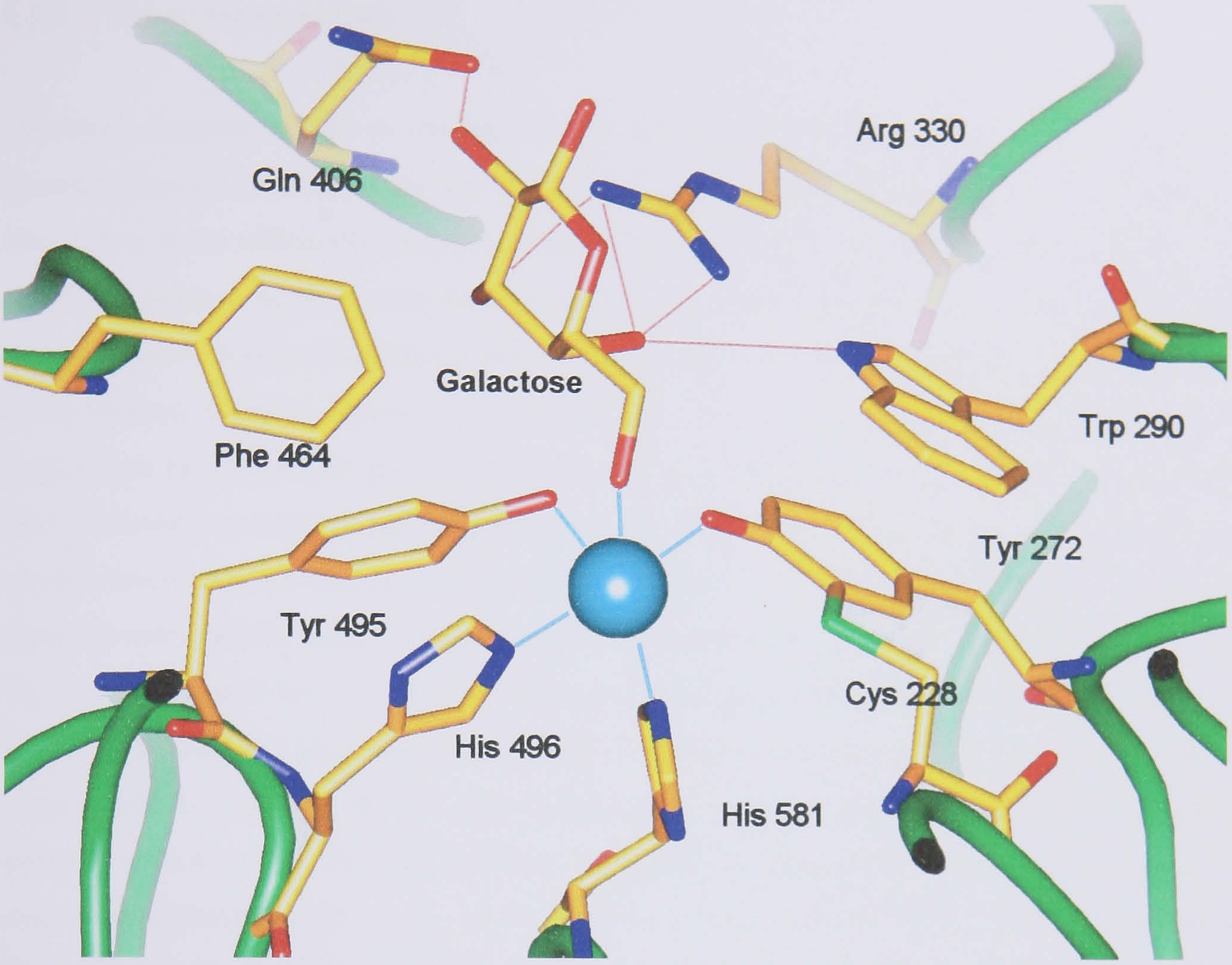
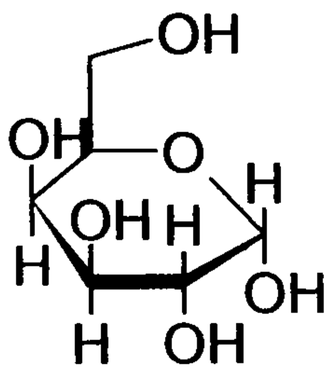


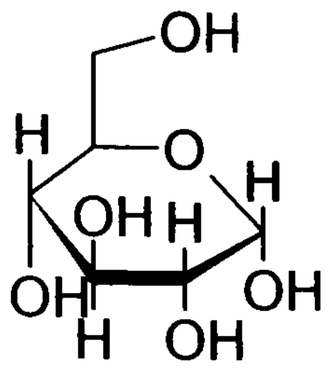
Figure 1.12: Active site of galactose oxidase, with a manually docked D-galactose molecule. H-bonding interactions with Arg 330, Gln 406 and Trp 290 are shown, along with copper to ligand bonds. The hydrophobic Phe 464 also makes favourable hydrophobic interactions with D-galactose. The coordinates for galactose oxidase in this model are derived from the galactose-free structure (PDB code 1GOF). The sugar molecule has been manually docked into the active site with the O6 of galactose in the fourth equatorial ligand site normally occupied by acetate in the native structure [Ito *et al.*, 1995]. Galactose has never been experimentally observed at the active site.

1.5.2 Substrate specificity

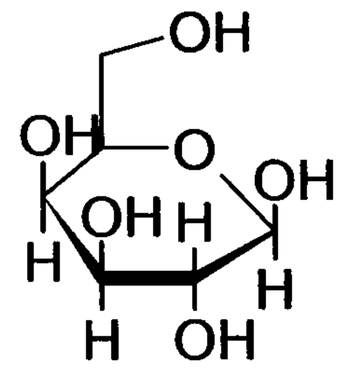
Galactose oxidase is able to oxidise a large range of substrates. Although the enzyme is named galactose oxidase, galactose is not its best substrate, dihydroxyacetone is catalysed three times more efficiently than galactose [Kosman, 1984]. Galactose oxidase is capable of oxidising primary alcohol substrates, ranging from small molecules such as ethanol and simple sugars, to large polysaccharides [Avigad, 1985; Avigad *et al.*, 1962; Kosman, 1984]. Some primary alcohols, however, are not substrates. D-galactose acts as a good substrate, but L-galactose and D-glucose are not turned over (Figure 1.13). The reason for this became clear after molecular modelling of substrates into the active site of the enzyme revealed steric clashes between some sugars and the enzyme [Ito *et al.*, 1994]. Studies of the efficiency of turnover with various hexose sugars suggests that the enzyme is very sensitive to changes in the configuration at the C6, C5, and C4 positions but less affected by changes at C1, C2 and C3 of the hexose ring [Schlegel *et al.*, 1968]. Computer modelling studies have consequently shown that D-galactose in its chair conformation adopts favourable hydrogen bonding partners with Arg 330, Gln 406, and Trp 290 (Figure 1.12), whereas L-galactose and glucose cause steric hindrance [Ito *et al.*, 1994; Wachter and Branchaud, 1996a]. Information about the substrate binding site is derived from observation of solvent binding at the fourth equatorial copper ligand site. The galactose oxidase inhibitor azide has also been shown to bind copper in crystal structures, coordinating at the fourth equatorial site at a distance of 2.5 Å [Ito *et al.*, 1995]. Azide is thought to exert inhibition by occupying the substrate binding site, thus preventing alcohol coordination to the copper [Whittaker and Whittaker, 1993]. Thus the fourth equatorial site has been designated the substrate binding site. Despite numerous attempts to obtain an active site enzyme-substrate complex, however, there is no experimentally derived structure describing substrate binding to the galactose oxidase active site.



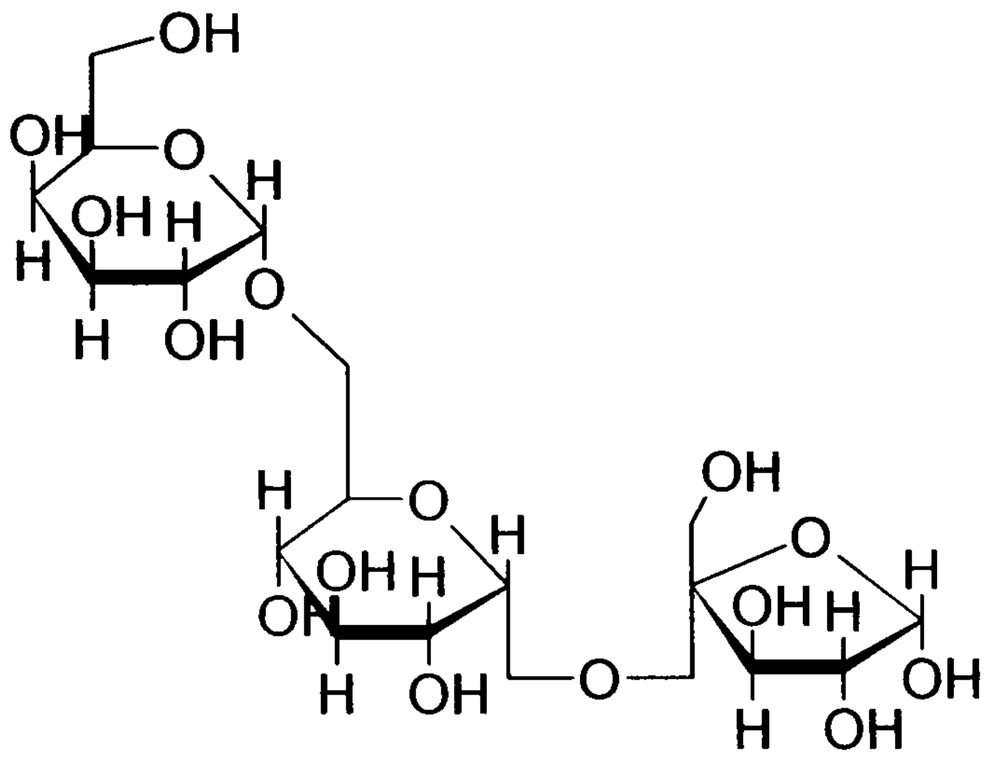
D-galactose



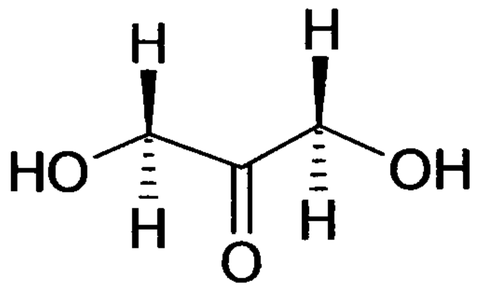
D-glucose



L-galactose



Raffinose



Dihydroxyacetone

Figure 1.13: Primary alcohol structures.

D-galactose, raffinose and dihydroxyacetone are substrates for galactose oxidase but L-galactose and glucose are not.

In recent years, the directed-evolution method of protein engineering has been used in order to increase galactose oxidase specificity to certain sugars. The directed-evolution process comprises two parts. The first part involves introducing random mutations into a parent protein sequence followed by cloning into a plasmid for expression in a suitable host organism (bacteria or yeast). In the next part of the strategy clones expressing enzymes with altered traits (such as improved thermostability, substrate specificity or catalysis) are identified in a high-throughput screen or selection and the genes encoding these improved proteins are isolated and recycled to the next round of directed evolution. This method was used to introduce glucose-6 oxidase activity to galactose oxidase [Sun *et al.*, 2002].

Two independent directed evolution experiments identified several galactose oxidase mutations that improved the enzyme's catalytic activity towards galactose. Since the objective of these studies was to identify improved enzymatic activity in crude extract, there was no measurement of relative expression levels of the variant proteins. The trial carried out by Sun *et al.*, [2001] selected a variant which displayed a 30-fold increase in activity, due to 18-fold higher expression and 1.7-fold greater catalytic efficiency. This contained a silent mutation of Pro 136 and five amino acid substitutions S10P, M70V, G195E, V494A and N535D. The S10P, M70V substitutions were thought to contribute to enhanced expression, while the V494A was claimed to increase expression and thermostability. Another study by Delagrave *et al.*, [2001] identified V494A as enhancing V_{max} while N535D was reported to contribute a modest increase in activity. Recurring substitutions were identified at position Cys 383 and Tyr 436, with a triple mutant combining C383S/Y436H/V494A substitutions showed a 15.9 fold greater catalytic efficiency compared to wild type. This assay determining K_M and V_{max} values was based on crude lysates of bacterial cultures. Enhanced activity to substrate in this assay system may therefore be due to contributions from increased levels of protein expression and/or enhanced catalytic activity.

The variants identified by Delagrave *et al.* were further investigated by generation of a series of single and double mutants of V494A, C383S and Y436H for comparison with the triple mutant (Figure 1.14) [Wilkinson *et al.*, 2004]. Two additional mutants, C383A and Y436A variants were also studied. This investigation allowed the relative contributions of each mutation to be studied for activity against 1-methylgalactose and galactose. The results of this study clearly indicated that the substitution of cysteine with a serine in the 383 position is an important one, which reduces the K_M for Galactose (Table 1.2).

Table 1.2: Kinetic properties of galactose Oxidase mutants with D-Galactose [Wilkinson *et al.*, 2004].

	D-galactose			
Mutant	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat}/K_M WT/variant
Wild-Type	68 ± 5	1090 ± 40	16000 ± 1300	1
C383S	15 ± 1	834 ± 18	55600 ± 3900	3.5
C383A	62 ± 6	639 ± 30	10300 ± 1100	0.6
V494A	53 ± 8	1119 ± 66	21100 ± 3400	1.3
Y436A	61 ± 9	1100 ± 71	18000 ± 2900	1.1
Y436H	53 ± 6	756 ± 36	14300 ± 1800	0.9
C383S/V494A	16 ± 2	915 ± 33	57200 ± 7400	3.6
C383S/Y436A	17 ± 2	1266 ± 63	74500 ± 9500	4.7
C383S/Y436H	22 ± 3	1871 ± 96	85000 ± 12400	5.3
Y436A/V494A	30 ± 8	1398 ± 148	46600 ± 13400	2.9
Y436H/V494A	68 ± 4	1181 ± 36	17400 ± 1200	1.1
C383S/Y436H/V494A	19 ± 2	1301 ± 46	68500 ± 7600	4.3

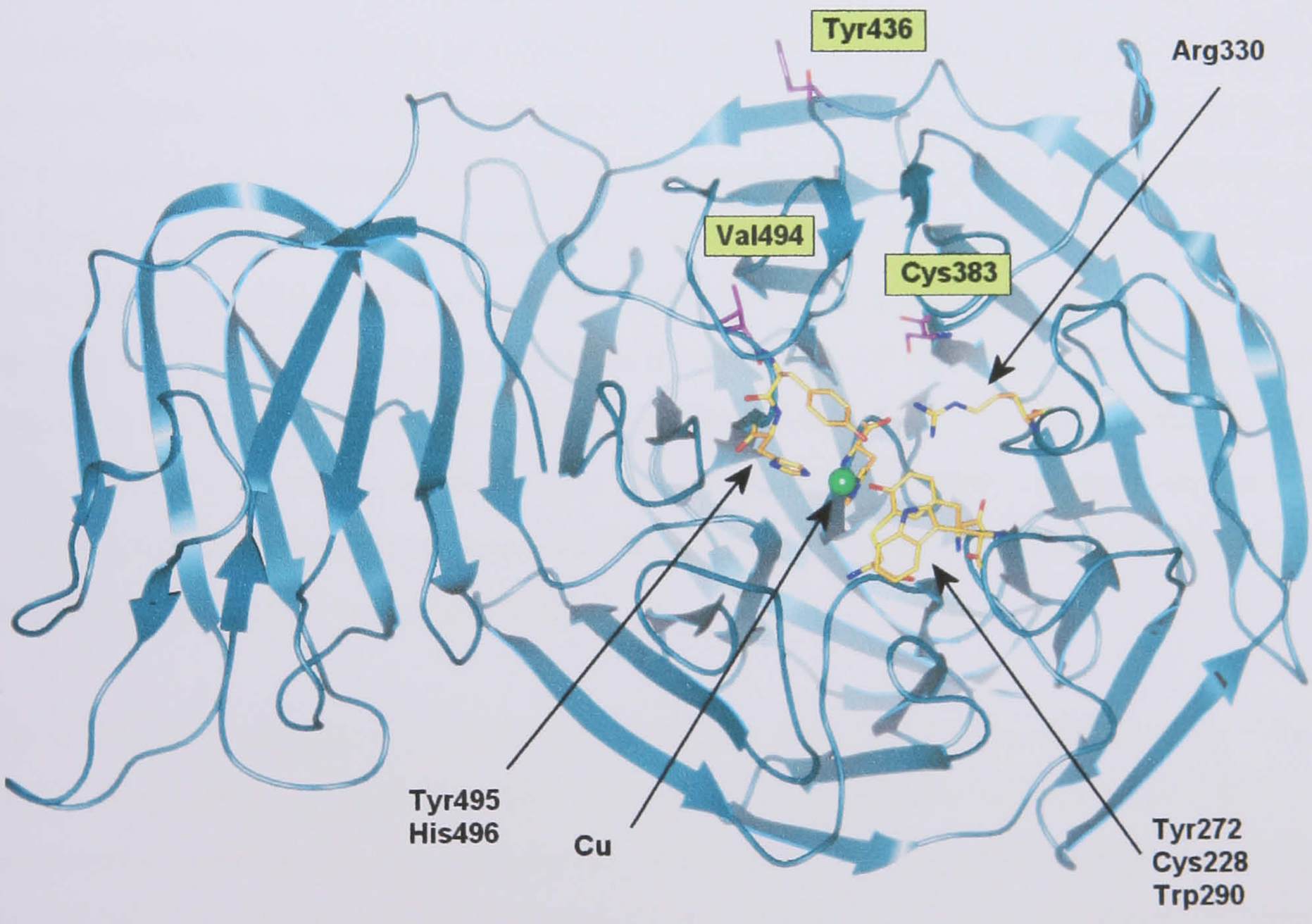


Figure 1.14: Schematic representation of the 383, 436, and 494 galactose oxidase mutations sites

Looking down the pseudo seven-fold axis of the beta propeller. The residues at positions 383, 436, and 494 are highlighted in magenta, copper is coloured green.

The work by Wilkinson *et al*, confirmed that the C383S/V494A/Y436H displayed enhanced k_{cat}/K_M. Although Y436H and V494A appeared to enhance catalytic turnover, the predominant cause of the enzymatic improvement was suggested to be the reduced K_M imparted by C383S. C383S and V494A are found relatively close to the enzyme active site, but Y436H is 30 Å away. It is therefore difficult to say how Y436H causes an increase in enzyme catalysis.

The results from enzyme assay experiments suggested it was not the loss of the sulphur of Cys 383 that was responsible for lowering the K_M, as the variant C383A did not display this reduced K_M. This implied that it was the property of a serine that enhanced substrate binding. C383S had a slightly reduced k_{cat}, but in terms of catalytic efficiency the dominant effect of the reduced K_M gave an enzyme with 3 to 4-fold enhanced k_{cat}/K_M. In contrast, the C383A mutation proved detrimental to enzymatic activity.

The location of the cysteine at position 383, places it close to the domain III loop protruding into the central cavity of the ß-propeller motif (Figure 1.14). The interface between this loop and the ß-propeller comprises an extensive hydrogen bonding network between the protein and water molecules. The role of cysteine in extracellular proteins is frequently structural, via the formation of a disulphide bond with another nearby cysteine residue when conditions are oxidising. Cysteine is also important for metal binding, such as in the zinc finger motif, commonly observed in nucleic acid binding proteins [Betts and Russell, 2003]. Another role reported for cysteine is that of a nucleophile, when part of the catalytic triad comprising asparagine and histidine in cysteine proteases such as papains and caspases. There are also cases, however, where a cysteine residue has no apparent function. As there are no other cysteine residues or catalytic residues in the immediate vicinity of Cys 383, the role of this residue was not thought to be significant.

The mutation of cysteine to serine is effectively the replacement of a sulphur atom with an oxygen atom. Both elements are found in group VI of the periodic table, where Sulphur has and electronic configuration comprising 16 electrons ($1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^4$) while oxygen has 8 ($1s^2\ 2s^2\ 2p^4$). Despite an increased number of nuclear protons in sulphur compared to oxygen, the electron affinity of oxygen is greater than that of sulphur. This is partially due to the increased number of sulphur inner electron shells, which act to screen incoming electrons from the nuclear attraction. Furthermore, the reduced distance between the oxygen nucleus and any incoming electrons also act to increase the electron affinity. The higher electron affinity of oxygen results in stronger hydrogen bonding properties for serine residues

compared to cysteine. How this change in property could affect the catalytic property of the C383S mutant is unclear.

1.5.3 Enzyme radicals and galactose oxidase oxidation states

Free radicals have been regarded as rare species in enzymatic reactions. The semiquinone forms of flavins and quinone cofactors in electron transport were probably the most studied radical species in enzymatic catalysis, with the exception of the tyrosine radical in ribonucleotide reductases [Reichard and Ehrenberg, 1983]. The electron transport process in biological systems, such as the respiratory chain in plant photosystem II, involve metal-based cofactors spaced 10-15 Å apart. In the case of the class I ribonucleotide reductases, there are no metal-based cofactors involved in radical transfer processes, and instead tyrosine, tryptophan and cysteine residues along the pathway perform this task [Chang *et al.*, 2004]. This radical transfer pathway is proposed to occur over a distance of 35 Å, comprising radical "hopping" across three tyrosine residues to form a cysteine radical.

An increasing number of S-adenosylmethionine (SAM)-dependent enzymes are known to function by mechanisms involving free radicals as reaction intermediates. In the cases of lysine 2,3-aminomutase [Chirpich *et al.*, 1970], pyruvate formate-lyase activase [Brush *et al.*, 1988; Knappe and Sawers, 1990], anaerobic ribonucleotide reductase [Reichard, 1993; Sun *et al.*, 1996], and benzylsuccinate synthase [Leuthner *et al.*, 1998], SAM serves as the free radical initiator. SAM undergoes transient cleavage to the 5'-deoxyadenosyl radical, which in turn propagates radical formation by abstracting hydrogen atoms, either from substrate molecules to form radical intermediates, or from glycyl residues of enzymes to activate them for radical-based biochemistry [Frey, 2001].

Although many radical enzymes have now been characterised, obtaining definitive proof for the involvement of radicals has been difficult due to the short lived nature of many radicals. The most convincing evidence for the existence of the galactose oxidase radical was obtained when the radical was observed spectroscopically. The UV-visible spectrum of galactose oxidase is the result of several overlapping contributions, which are difficult to assign (Figure 1.15). electron-paramagnetic resonance (EPR) and resonance Raman experiments, however, have aided in the interpretation of such data. The intense 445 nm peak is assigned to a combination of two elements: a ligand-metal charge transfer between Tyr 272 and copper; and π-π^* ring transitions in the tyrosine radical [Whittaker *et al.*, 1989]. The 800 nm peak was assigned to an intra-ligand charge transfer between Tyr 495 and Tyr 272 mediated by the copper [McGleshen *et al.*, 1995]. Spectroscopic analysis of the metal-free enzyme

conclusively showed that the apo-enzyme radical was located on the Tyr-Cys cofactor [Babcock *et al.*, 1992; Gerfen *et al.*, 1996; Whittaker *et al.*, 1989; Whittaker and Whittaker, 1988]

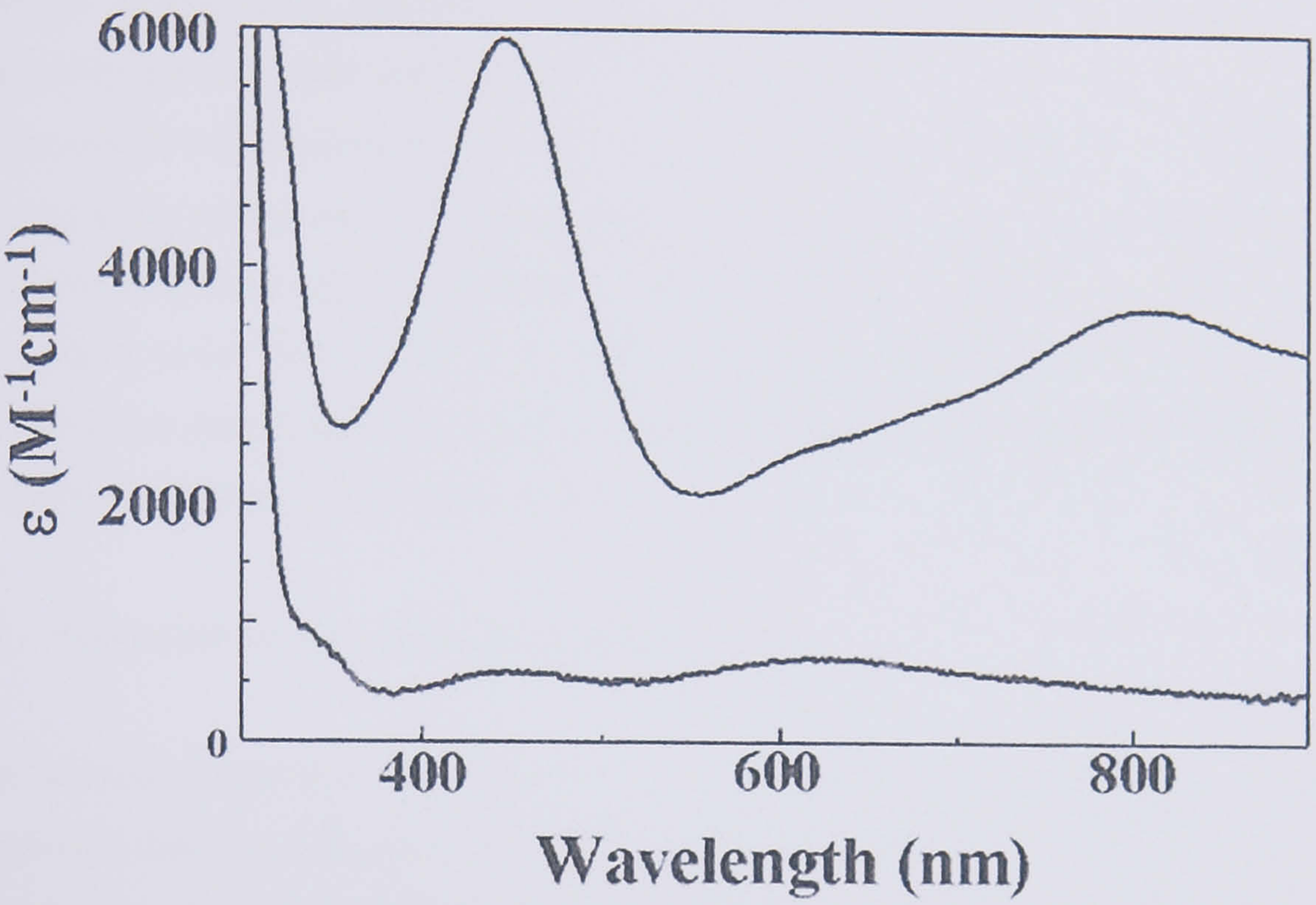


Figure 1.15: UV-visible spectrum of galactose oxidase.

The top line represents the spectrum of the oxidised, radical-containing, active enzyme. The lower line shows the inactive, radical-free, semi-reduced form. The 445 nm peak was assigned to Tyr 272-copper charge transfer and the presence of the radical. The 800 nm peak was interpreted as an inter-ligand charge transfer between Tyr 272 and Tyr 495. Figure from Firbank, [2002].

Although the catalytic properties of galactose oxidase have been studied for over 30 years the UV-visible spectral features of the enzyme were not definitively characterised until the late 1980's. This was mainly due to the fact that the enzyme is able to adopt three different oxidation states:

$$\underset{\text{Reduced}}{Cu^{1+} - Tyr} \underset{+1e^-}{\overset{-1e^-}{\rightleftharpoons}} \underset{\text{Semi-reduced}}{Cu^{2+} - Tyr} \underset{+1e^-}{\overset{-1e^-}{\rightleftharpoons}} \underset{\text{Active}}{Cu^{2+} - Tyr^{\bullet}}$$

In 1988 galactose oxidase was isolated in three distinct redox forms [Whittaker and Whittaker, 1988]. Mild oxidation of the protein with potassium ferricyanide produced the active form of the enzyme, in which the radical was present and the metal ion existed in the Cu^{2+} form. This form of the enzyme produced the characteristic spectrum of galactose oxidase when analysed by UV-visible absorption (Figure 1.15). One electron reduction with potassium ferrocyanide produced a semi-reduced form of galactose oxidase that was catalytically inactive, and contained Cu^{2+} but no radical. The third redox state was formed after anaerobic two electron reduction of the active enzyme by substrate, resulting in the Cu^{1+} oxidation state, and no radical. Restoration of the enzyme to the active form occurs after the addition of molecular oxygen, which is proposed to bind to the reduced copper, oxidising the enzyme back to the radical state (Cu^{2+}-$Tyr^{\bullet}$). Prior to the revelation that the enzyme exists in three oxidation states, galactose oxidase experiments had been carried out on the native non-homogeneous mixture of oxidation states.

1.5.4 Galactose oxidase redox inter-conversion

In the absence of oxygen, spontaneous loss of the radical occurs leaving the enzyme in a catalytically inactive semi-reduced state. Anaerobic incubation of the oxidised, active enzyme in solution resulted in spontaneous and complete radical decay within 3-4 hours [Wright and Sykes, 2001a]. Little is known about why or how this process occurs, but it appeared to involve the Cys 515-Cys 518 disulphide bond lying approximately 10 Å away from the active site. Chemical modification of the disulphide with $HSPO_3^{2-}$ inhibited auto-reduction, and anion binding to the active site copper caused a reduction in the rate of the reaction [Wright and Sykes, 2001a].

The retardation of the auto-reduction process was suggested to be due to the protonation of Tyr 495 or exogenous anion binding to the active site copper. Anion binding is known to reduce the 800 nm peak in the absorbance spectrum, an effect similar to the mutation of Tyr 495 to phenylalanine. Thus spontaneous radical decay was suggested to occur via Tyr 495, with dissociation of the residue from copper reducing the rate of decay from Tyr 272 [Wright and Sykes, 2001a].

Alternatively, spontaneous auto-oxidation was reported to occur in the presence of oxygen, producing a radical containing catalytically active enzyme [Wright and Sykes, 2001a]. This process is unaffected by the chemical modification of the Cys 515-Cys 518 disulphide and in the presence of the inhibitors cyanide or azide, auto-oxidation is 100% complete [Wright and

Sykes, 2001a]. Under aerobic conditions the two competing reactions establish a native galactose oxidase ratio of 5% active to 95% semi-reduced.

The mechanism of the auto-redox reactions is unknown, however, auto-oxidation does not appear to be the reverse of auto-reduction, since disulphide modification or anion presence has no effect on the rate of oxidation. The failure of anion binding to affect oxidation has led to the suggestion that Tyr 272 interacts directly with dioxygen to allow oxidation to occur [Wright and Sykes, 2001a]. There is currently no fully characterised crystallographic description of galactose oxidase in the fully oxidised, the semi-reduced or the fully reduced states. The current structures available describe the protein crystallised in the native semi-reduced/oxidised mix. EXAFS data have determined that no significant difference exists between the average copper ligand distances of the semi-reduced and oxidised galactose oxidase [Knowles *et al.*, 1995], and crystals grown in the presence of the oxidant, potassium ferricyanide, have the same structure as those from solutions of the native mix [Ito *et al.*, 1995].

The fully reduced state of galactose oxidase is the least-well known form of the enzyme, and studies of this form have proved difficult due to its high propensity to react with oxygen. EXAFS data have suggested the fully reduced enzyme active site contains copper in the Cu^{1+} oxidation state and adopts a three coordinate geometry, with an average copper to ligand distance of 2 Å [Knowles *et al.*, 1995]. Tyr 272, His 496 and His 581 are presumed to be the copper ligands in the reduced enzyme.

1.5.5 Catalytic cycle

Catalysis commences from the fully oxidised Cu^{2+} containing enzyme, with the radical site providing one redox centre and the copper ion providing the other. The catalytic mechanism is proposed to occur via a ping-pong mechanism which can be seen as two parts, a reductive half-reaction and an oxidative half-reaction (Figure 1.16 and Figure 1.17 respectively) [Whittaker *et al.*, 1998; Whittaker and Whittaker, 1993; Whittaker and Whittaker, 2001]. The reductive half-reaction comprises the reduction of the enzyme by the substrate, while the oxidative half-reaction is where the enzyme is oxidised by molecular oxygen.

The first step of reductive half-reaction is the abstraction of a proton from the hydroxyl group of the alcohol by the axial ligand Tyr 495, leading to the proposed displacement of Tyr 495 from copper. This step has been intensely studied in both wild-type galactose oxidase and the Y495F mutant form of the protein. Azide binding to the enzyme is thought to mimic substrate

binding by replacing the exogenous acetate or water ligand identified in crystal structures [Ito *et al.*, 1991; Ito *et al.*, 1994]. According to Whittaker, anionic ligands (acetate and azide) interact more strongly with the metal ion compared to a neutral water molecule, this increased bond strength is proposed to occur at the expense of Cu-Tyr 495 interactions. When coordinated to a metal, phenols are relatively acidic, however the decrease of the metal interaction is suggested to return the basicity of Tyr 495 to the more typical value (p*K*a=10). It has been shown that the titration of azide with wild-type enzyme is accompanied by the stoichiometric uptake of a proton from solution by a group in the protein with a p*K*a>9 [Whittaker and Whittaker, 1993]. In the absence of a substrate hydroxyl from which to abstract a proton, Tyr 495 is proposed to acquire a proton from solution. Proton uptake is abolished when Tyr 495 is mutated to phenylalanine, the catalytic rate is also reduced by 1000 times [Reynolds *et al.*, 1995]. Tyr 495 has therefore been proposed to be the catalytic base.

Proton abstraction by Tyr 495 activates the substrate for the next stage of the reaction, the reduction of copper and stereospecific abstraction of the pro-S methylene hydrogen from the substrate by Tyr 272$^{\bullet}$ [Whittaker *et al.*, 1998]. These two steps are proposed to generate a free radical intermediate. Hydrogen atom transfer prior to copper reduction is proposed to produce an intermediary ketyl radical, in which an unpaired electron is localised on the Cα carbon of the substrate (B1, Figure 1.16). Hydrogen atom transfer after copper reduction is proposed to form an alkoxyl radical, with an unpaired electron localised on the oxygen atom (B2, Figure 1.16) There has been no clear experimental evidence for the ordering of the hydrogen transfer/copper reduction steps in the reaction mechanism, although they are thought to happen in conjunction to contribute to the rate determining step in catalysis [Whittaker and Whittaker, 1993]. The overall result of copper reduction and hydrogen atom transfer is the oxidation of the substrate intermediate to the aldehyde product.

Galactose oxidase catalysis of stereospecifically deuterated galactose and benzyl alcohol substrates exhibit over 95% stereospecific hydrogen atom abstraction, as measured by H^1 NMR and mass spectrometry of the aldehyde products [Minasian *et al.*, 2004; Whittaker *et al.*, 1998]. Evidence for the specific hydrogen selection, however, has been reported for experiments conducted only on sugars and sugar derivatives. There is no evidence of such discrimination in smaller substrates. Thus preference for abstraction of the pro-S hydrogen is thought to be due to steric constraints imposed on the substrate by the active site environment.

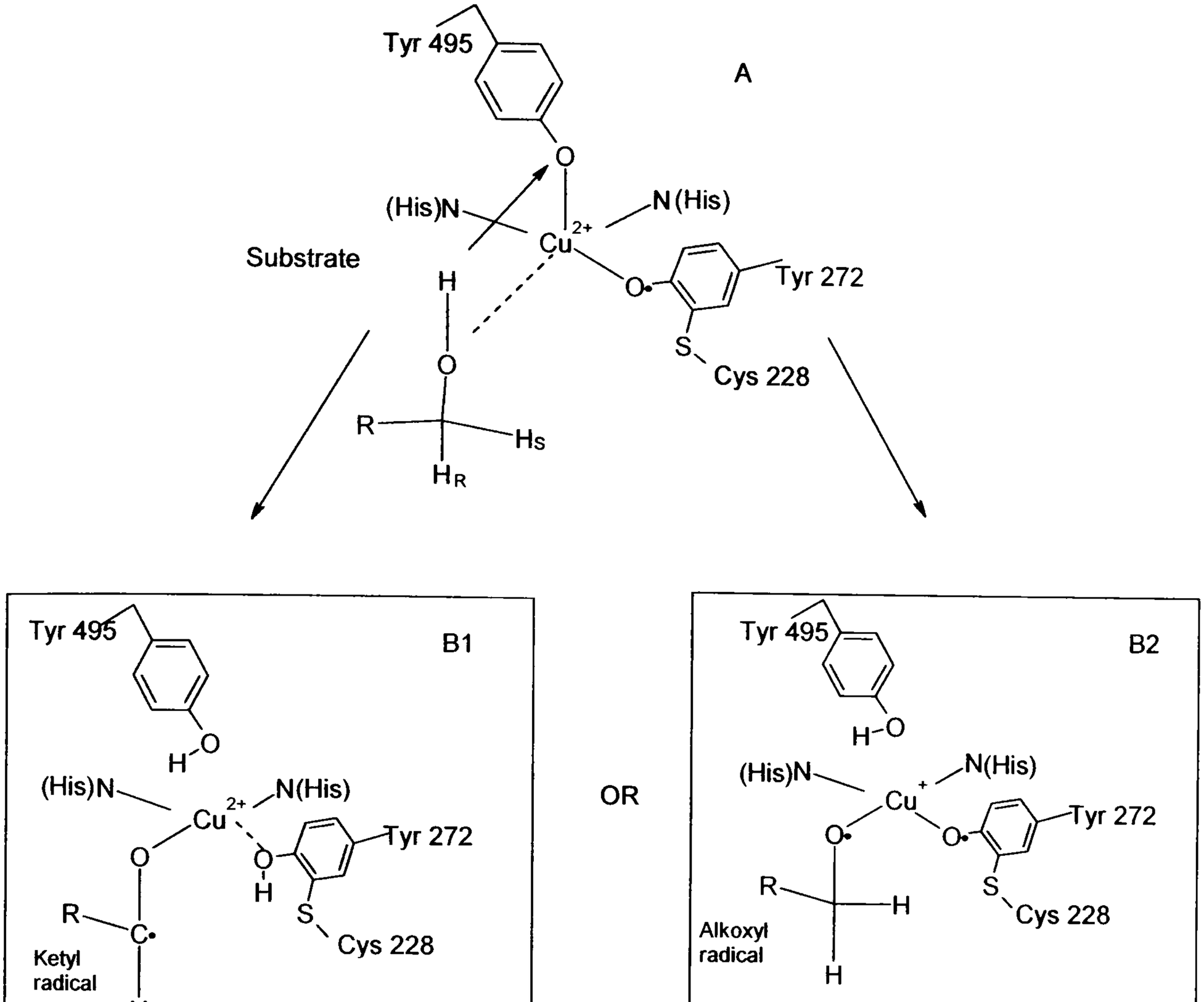


Figure 1.16:The galactose reductive half-reaction.

The galactose oxidase reductive half-reaction proposed by Whittaker *et al.*, [1998]. **A)** Substrate coordination to copper leads to proton abstraction by Tyr 495. Tyr 495 is then displaced from the copper. The reaction is predicted to proceed via one of two radical intermediates. **B1**) A transfer of the pro-S hydrogen atom to Tyr 272, results in the formation of a ketyl radical intermediate. **B2**) An alkoxyl radical intermediate is formed following a single electron transfer from the substrate to Cu^{2+}, thus reducing the metal to Cu^{1+}. **C) From B1**, a single electron oxidation of the ketyl radical intermediate would reduce the metal to Cu^{1+}. **C) From B2**, a hydrogen atom transfer by Tyr 272 results in the reduction of the Tyr-Cys phenoxyl. The overall result of step B and C is the oxidation of the alcohol to the aldehyde product, and the two electron reduction of the enzyme. Dotted lines represent weak ligand-metal interaction.

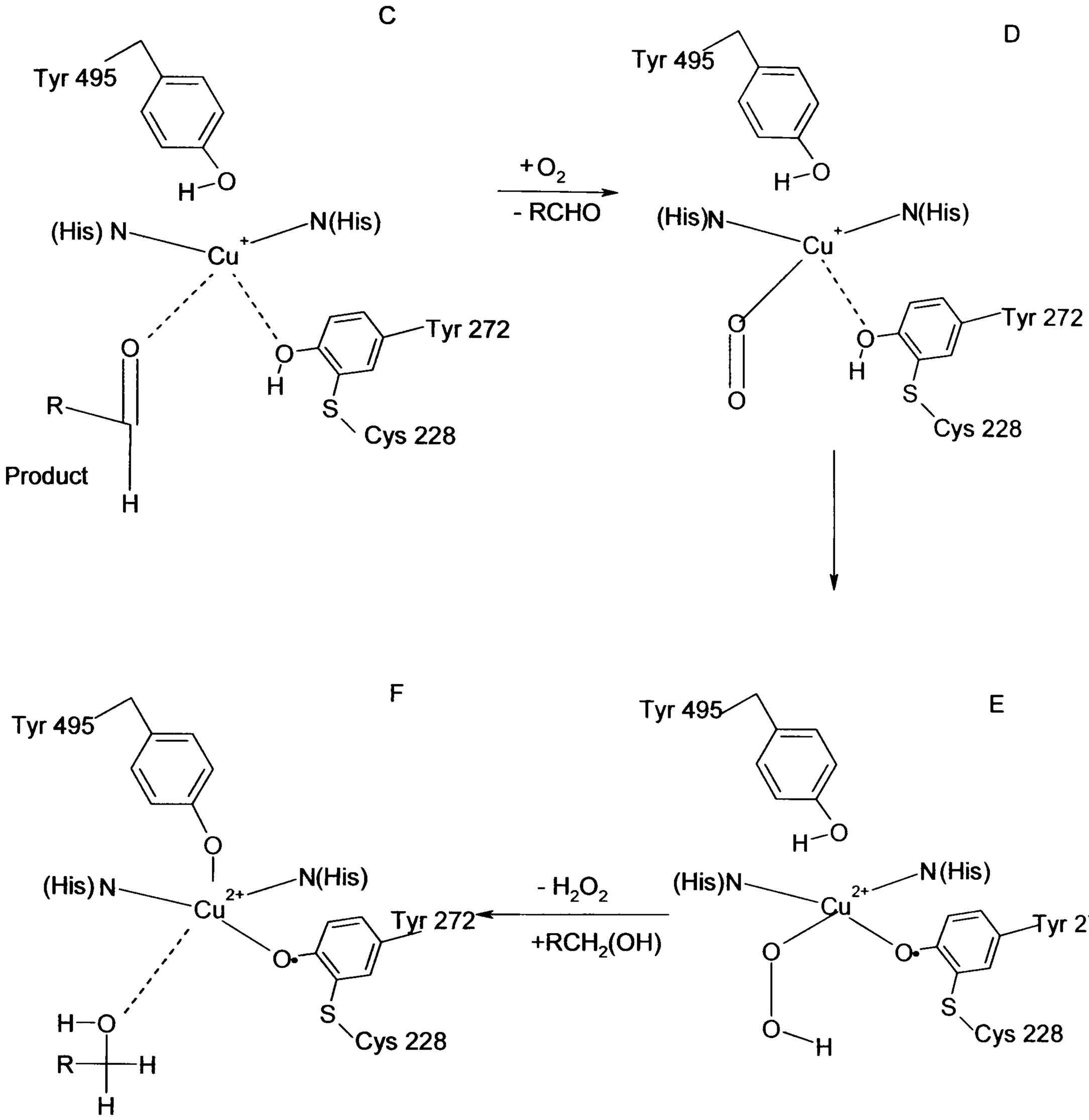


Figure 1.17: The galactose oxidase oxidative half reaction.

The galactose oxidase catalytic oxidative half reaction proposed by Whittaker *et al.*, [1998]. Reduction of dioxygen is proposed to occur essentially via the reverse sequence of steps involved in substrate oxidation (Figure 1.16). **D**) Dioxygen binds to Cu^{1+} displacing the aldehyde product. **E**) The copper is oxidised and the hydrogen atom is transferred from Tyr 272 phenolate to the dioxygen species generating Cu^{2+}-hydroperoxide intermediate. **F**) Proton transfer from Tyr 495 phenolate to the coordinated oxygen atom results in the formation of hydrogen peroxide and regeneration of the fully active enzyme. Dotted lines represent weak ligand-metal interaction.

The fully reduced enzyme is proposed to bind dioxygen via coordination to Cu^{1+}, restoring the enzyme to the active Cu^{2+}, radical state, with the concomitant reduction of dioxygen to hydrogen peroxide. This part of the reaction is not very well characterised, mainly because the extremely fast reaction of Cu^{1+} with dioxygen. Under anaerobic conditions, galactose oxidase has been obtained in the stable fully reduced form, allowing the measurement of a rate constant for reoxidation of Cu^{1+} by oxygen as 1.01×10^7 M^{-1} s^{-1} [Clark *et al.*, 1994]. Studies of the fully reduced form of galactose oxidase rely on the keeping the protein in anaerobic conditions, followed by substrate or chemical reduction [Whittaker, 2005; Wright and Sykes, 2001b]. Substrate reduction of the enzyme occurs at variable rates depending on the substrate used. The rate of re-oxidation, however, is remarkably consistent regardless of which substrate is used to initially reduce galactose oxidase. This is suggested to be supporting evidence for a ping-pong mechanism displaying separable catalytic rates [Whittaker and Whittaker, 1993; Whittaker and Whittaker, 2001], as opposed to the concerted catalytic mechanism initially predicted [Whittaker *et al.*, 1989].

1.5.6 Biogenesis of galactose oxidase

1.5.6.1 Introduction

Galactose oxidase is expressed *in vitro* as a form containing an additional 17 amino acid N-terminal pro-sequence and lacking the thioether bond between Tyr 272 and Cys 228. Sequencing of the galactose oxidase gene from *Fusarium spp* revealed that the start of the mature amino acid sequence [McPherson *et al.*, 1992], as determined by N-terminal sequencing, did not correspond to the cleavage point of the signal peptide predicted from the gene. It was later discovered that expression of galactose oxidase in limited copper conditions resulted in three different forms of the enzyme [Rogers *et al.*, 2000] that could be separated by SDS-PAGE (Figure 1.18, a). The lowest band on the gels corresponds to mature galactose oxidase. The presence of the thioether bond causes mature galactose oxidase to travel faster than expected, with the apparent molecular weight of 65 kDa. The middle band runs with the calculated molecular weight of 68.5 kDa, representing the form of the enzyme lacking the thioether bond. This form also co-migrates with the mutant C228G, which lacks the cysteine involved in thioether bond formation. The upper band runs with the estimated molecular weight of 70.2 kDa and corresponds to a precursor form of the enzyme containing the 17 amino acid pro-sequence. Strict copper-free protein expression and purification allowed the isolation of the 70.2 kDa species (Figure 1.18, b).

The pro-sequence form of galactose oxidase was found to process to the fully mature, active enzyme upon the addition of Cu^{2+} and oxygen [Rogers *et al.*, 2000]. Pro-sequence cleavage, thioether bond and radical formation events all occur despite the presence of a protease inhibitor cocktail, supporting the proposition that pro-sequence cleavage is copper mediated rather than protease dependent. No cleavage was observed under anaerobic conditions, and no other metal ions (Ca, Mg, Mn, Zn, Co and Ni) are able to induce biogenesis.

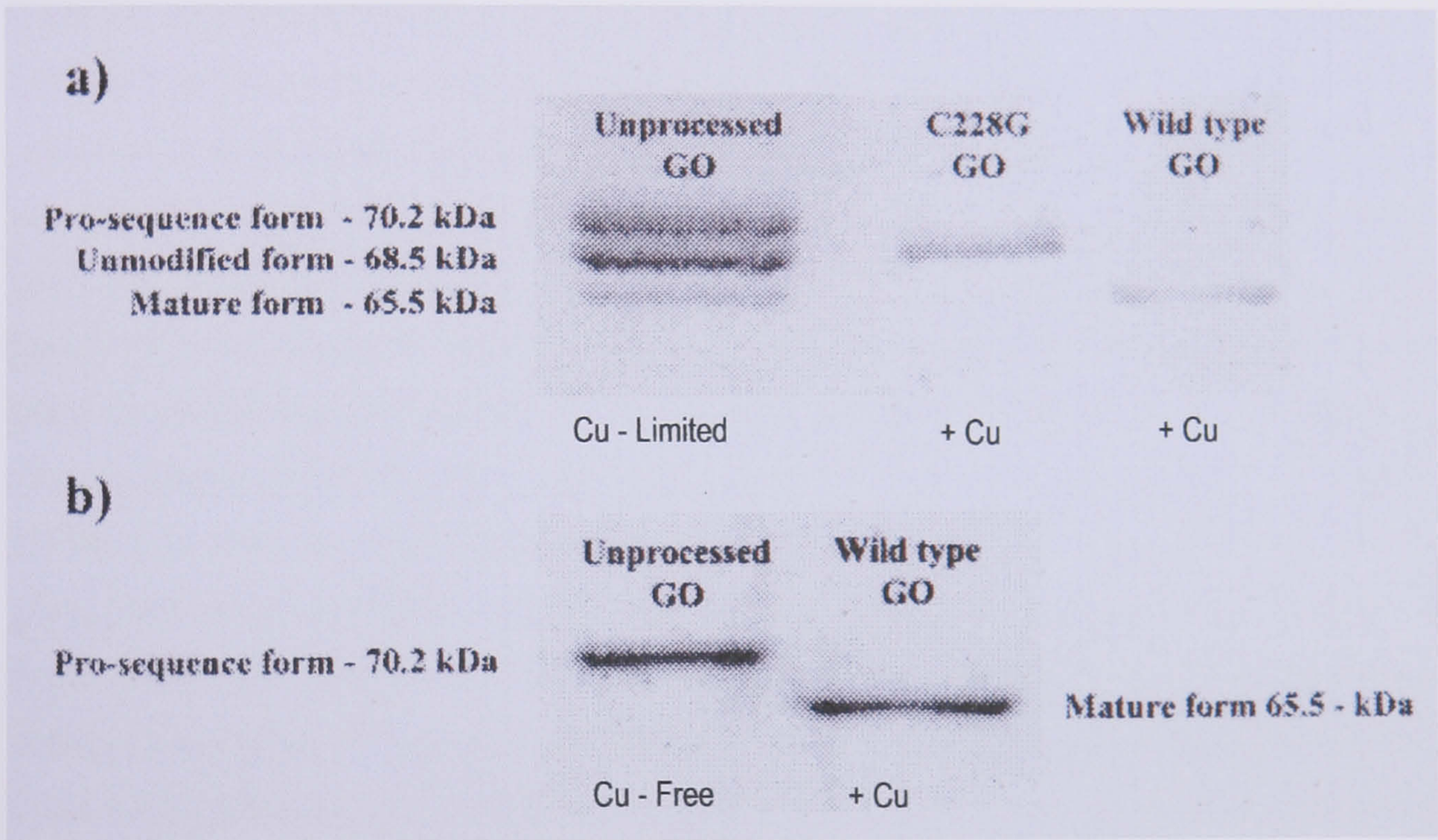


Figure 1.18: SDS-PAGE analysis of galactose oxidase expression.

a) copper limited expression results in the purification of three galactose oxidase species. The lower band is the mature enzyme, while the middle band co-migrates with the thioether-free mutant C228G. The upper band is the pro-Galactose oxidase form containing a 17 amino acid pro-sequence and no thioether bond. **b**) strict copper-free expression and purification of galactose oxidase enabled the isolation of pure pro-galactose oxidase. Figure from Rogers *et al.*, [2000].

1.5.6.2 Structural studies of pro-sequence containing galactose oxidase

The structure of pro-sequence containing galactose oxidase (pro-GO) was reported by Firbank *et al.* [2001]. Copper-free pro-GO (PDB code 1KI3) revealed the pro-sequence to exist as a loop between domains I and II. Analysis of the Cα positions of pro-GO, excluding the pro-sequence, compared with those of the mature enzyme (PDB code 1GOG) revealed that the two structures were very similar overall (root mean squared deviation of 0.7 Å). Five regions of the pro-GO main chain (residues 1-5, 189-200, 216-227, 254-261 and 290-296), however, showed a root mean square deviation of 5.1 Å (Figure 1.19, b). The majority of main chain deviations from the positions in the mature enzyme were due to the presence of the pro-sequence. The presence of the pro-sequence affected the positions of active site residues Cys 228 and Trp 290, with the latter residue showing an 8 Å deviation in the pro-GO structure compared to its positions in the mature enzyme (Figure 1.19, a). Movement of Trp 290 led to an active site in which Cys 228, Tyr 272 and Trp 290 are exposed to solvent. The thioether bond was not present in pro-GO. Tyr 272 and Cys 228, the residues that form the cofactor, were positioned differently from the mature form. The ring of Tyr 272 is rotated almost 90° about $\chi 2$ and the side chain of Cys 228 is pointing away from its position in the active protein. The side chains of the remaining copper ligands, Tyr 495, His 496 and His 581, are in almost the same positions as in the mature enzyme, showing Cα movements of less than 0.4 Å. One unexpected discovery in the copper free pro-GO was the oxidation of the Sγ of Cys 228 to an S-OH group. The presence of S-OH was suggested to represent a possible intermediate in cofactor biogenesis, although an artefact resulting from crystallisation or radiation damage could not be ruled out. The overall effect of the presence of the pro-sequence was to present a more 'open' active site (Figure 1.20). This has led to the proposal that the pro-sequence is responsible for positioning of the active site residues, including copper ligands, to facilitate initial copper binding and the thioether bond formation [Firbank *et al.*, 2001].

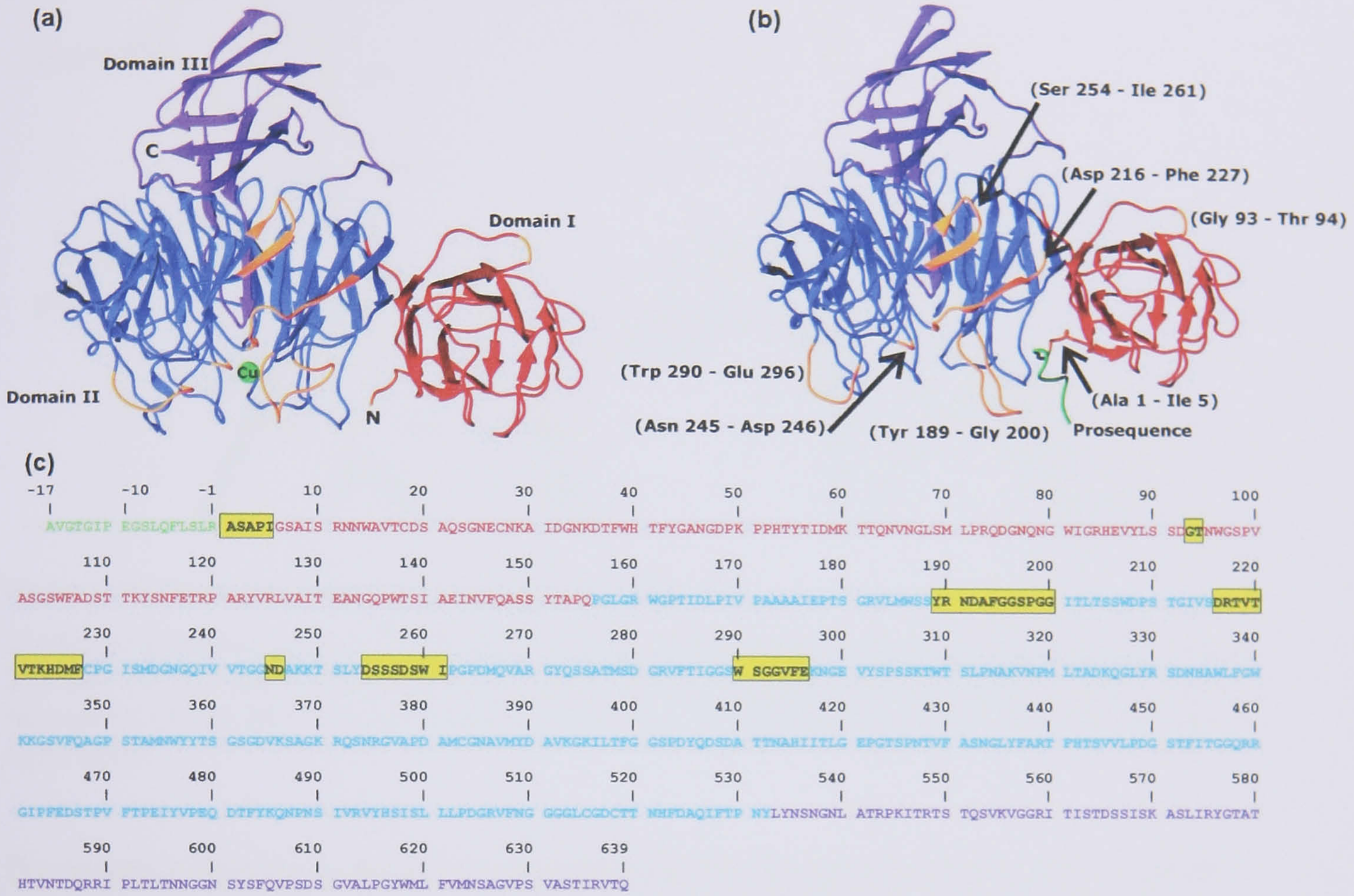


Figure 1.19: A comparison of mature galactose oxidase and pro-galactose oxidase.

Overall structure of galactose oxidase, illustrating similarity in overall fold (domain I red, domain II blue, and domain III purple), and amino acid sequence. **a**) Mature galactose oxidase. **b**) Pro-GO structure, the pro-sequence is coloured green. **c**) Pro-GO amino acid sequence. Regions that differ between pro-GO and the mature enzyme are coloured yellow. From Firbank *et al*., [2001].

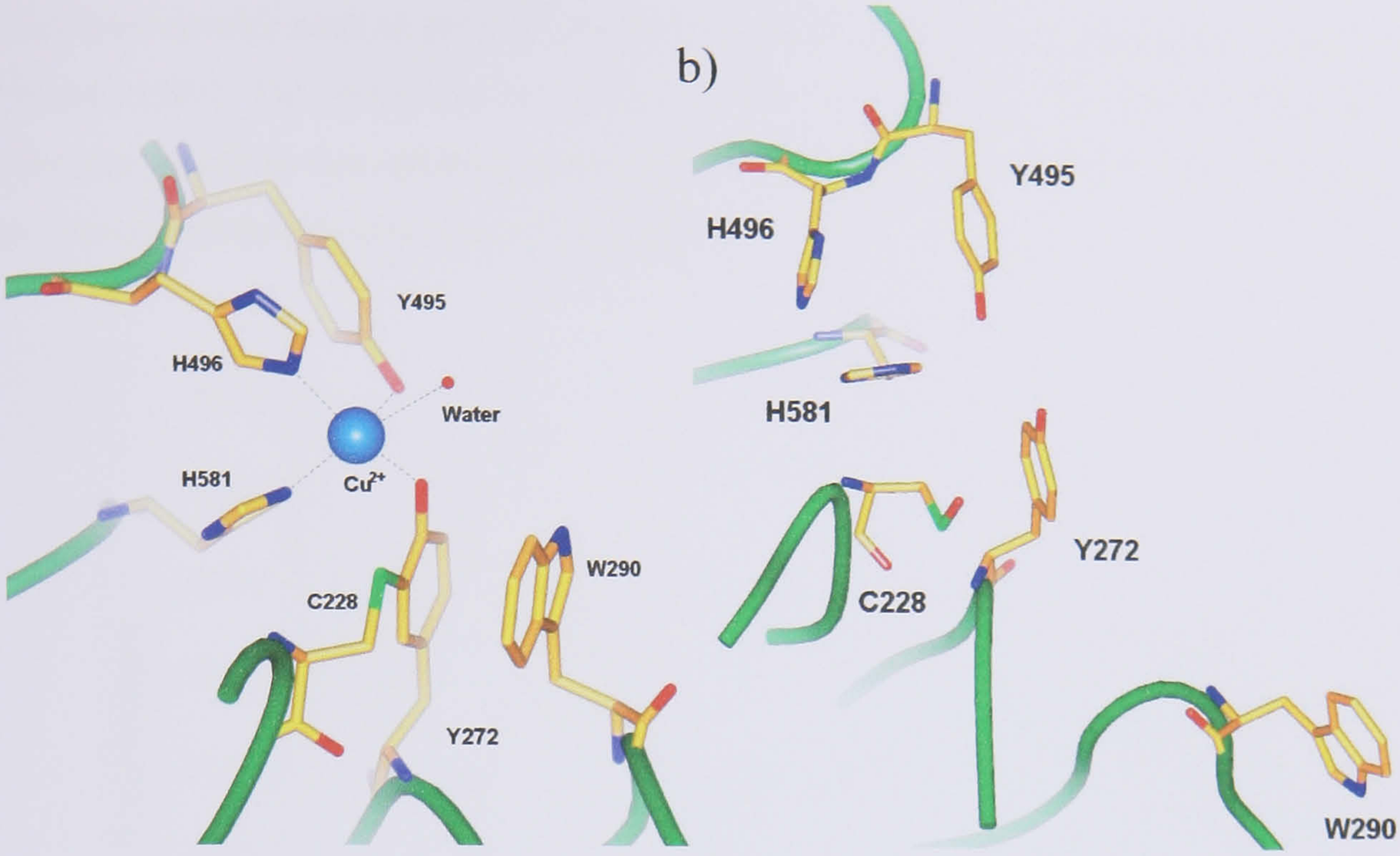


Figure 1.20: Active site structures of pro-GO and mature GO.

a) Mature galactose oxidase (PDB code 1GOG), and b) Pro-GO form (PDB code 1KI3). Cys 228 and Tyr 272 adopt different orientations in pro-GO compared to the mature enzyme. The movement of Trp 290 even more dramatic, the overall effect is to make the pro-GO active site more open.

Pro-sequence cleavage is thought to occur prior to thioether bond formation, since a species that has lost the pro-sequence but not formed the thioether bond has been detected, whereas a species containing both the pro-sequence and the thioether bond has not (as determined by SDS-PAGE, Rogers *et al.,* 2000; Firbank *et al.,* 2001). Solution studies of galactose oxidase biogenesis have shown that the aerobic incubation of pro-GO with Cu^{2+} produces a UV-visible spectrum showing an initial species with a peak at 410 and 750 nm is formed after 10 minutes (Figure 1.21). After 20 minutes these peaks have shifted to 445 and 800 nm, which are characteristic of the mature protein and confirm the presence of the tyrosyl radical [Rogers *et al.*, 2000].

Copper soak experiments were carried out using pro-GO crystals in order to trap processing intermediates [Firbank, 2002]. Anaerobic exposure of pro-GO crystals to 5 mM copper acetate, produced structures with 50% copper occupancy and no thioether bond formation, even after soaks of up to 2 hours. It was reported that the pro-sequence was still present in the crystal structure, although only four of the residues were successfully modelled into the electron density. The active site copper was coordinated by just three ligands: Tyr 495, His 581, and the carbonyl group of Phe 227 [Firbank, 2002].

A two hour aerobic soak of pro-GO resulted in the incorporation of copper into the active site [Firbank, 2002]. The metal was coordinated by Tyr 272, Tyr 495, His 496 and His 581. The active site geometry was reminiscent of the mature protein. The thioether bond, however, was not formed and the pro-sequence still remained.

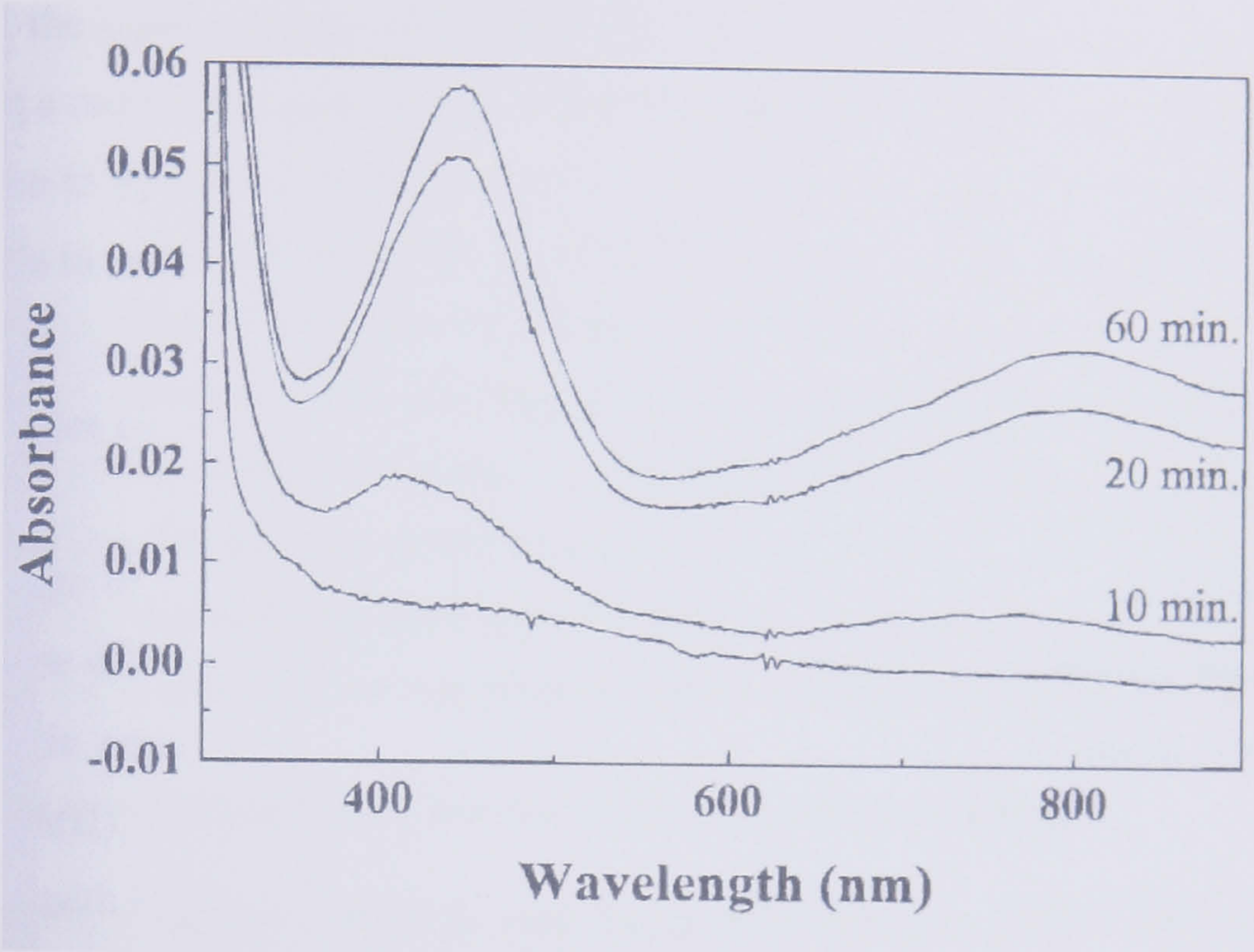


Figure 1.21: UV-visible spectra of pro-go incubated with Cu^{2+}.
UV-visible spectra showing the generation of the active enzyme on the addition of 28 μM of copper to a 28 μM pro-GO. An initial peak is observed at 410 nm, after 20 minutes the characteristic peak at 445 nm is present. Figure from Rogers *et al*., [2000].

A five hour aerobic copper incubation resulted in crystal dissolution. The dissolved crystal solution was subsequent incubated in a water based copper solution and analysed by SDS-PAGE [Firbank, 2002]. Full processing of the whole protein sample was not observed. The results of pro-GO copper soaks suggested the crystallised protein was unable to process. The reason for this is unknown. The apparent oxidation of the Cys 228 residue, however, was proposed to have had an inhibitory effect on cofactor formation.

There have been no formal mechanistic proposals for pro-sequence cleavage. The system requires only oxygen and copper and does not conform to any major family of proteases. Serine and thiol proteases form a catalytic triad with histidine and an aspartate which work to polarise the serine/thiol and thus enable peptide hydrolysis via nucleophilic attack of the peptide carbonyl [Nishihira and Tachikawa, 1999]. Zinc proteases, have been studied in detail, by both crystallographic and kinetic means. Zinc enzymes have been classified according to their binding sites, with many containing the motif HExxHxxGxxH, comprising

three coordinated histidine residues and a glutamate general base [Jiang and Bond, 1992]. Zinc proteases use a zinc ion to polarise the carbonyl carbon-oxygen bond by direct coordination to the carbonyl oxygen, followed by attack of the carbonyl carbon by a coordinated water positioned by an active site glutamate or by a zinc water molecule [McCall *et al.*, 2000]. Copper replacement studies in zinc proteases often result in inactivated enzyme, despite successful catalysis when other divalent metals are substituted. Aspartyl proteases require two active site aspartates in conjunction with a coordinated water molecule, the side chain of one aspartate is proposed to form a hydrogen bond with the carbonyl oxygen of the cleavage site, while the side chain of the other positions a water molecule for nucleophilic attack of the carbonyl [Trylska *et al.*, 2004]. The catalytic residues necessary for the protease mechanisms mentioned have not been identified in galactose oxidase and although other metals have been shown to be able to replace zinc and maintain zinc protease activity, no copper binding site has been identified close to the pro-sequence. The pro-sequence cleavage reaction is not dependent on the residues flanking the site, since mutation of the cleavage site residues (Arg -1 and Ala 1) to phenylalanine did not prevent cleavage [Ogel, 1993].

1.5.6.3 Structural studies of a truncated form of galactose oxidase

Two galactose oxidase constructs were generated [Mahmoud, 2001]. The first of which contained the full protein sequence (AD1), including the signal peptide, 17 amino acid pro-sequence (highlighted green below). A second, truncated galactose oxidase construct (BD1) was generated for expression in *Pichia pastoris* [Mahmoud, 2001].

The mature enzyme sequence (highlighted red below) amplified from the plasmid pGOF102 in *A. nidulans*:

AD1: MKHLLTLALC FSSINAVAVT VPHKAVGTGI PEGSLQFLSL RASAPIGSAIS

A second construct (BD1) was generated containing only the signal peptide and the mature sequence without the pro-sequence:

BD1: MKHLLTLALC FSSINAVAVT VPHK- ASAPIGSAIS

When BD1 was expressed and purified with copper and oxygen present, the protein did not contain the pro-sequence. Expression of the BD1 construct in the absence of copper generated a form of galactose oxidase which contained neither the pro-sequence nor the thioether bond. This species is referred to as premature galactose oxidase and was observed by SDS-PAGE, it co-migrated with the mutant C228G [Rogers *et al.*, 2000]. Incubation of premature galactose oxidase with Cu^{2+} resulted in the formation of the mature protein, as detected by SDS-PAGE

(Figure 1.22) and N-terminal sequencing. Kinetic and spectroscopic analysis of the protein suggests it was at least as active as the protein expressed in *A. nidulans* [Mahmoud, 2001]. Almost all the protein appeared to process within 2 minutes of copper addition [Firbank, 2002].

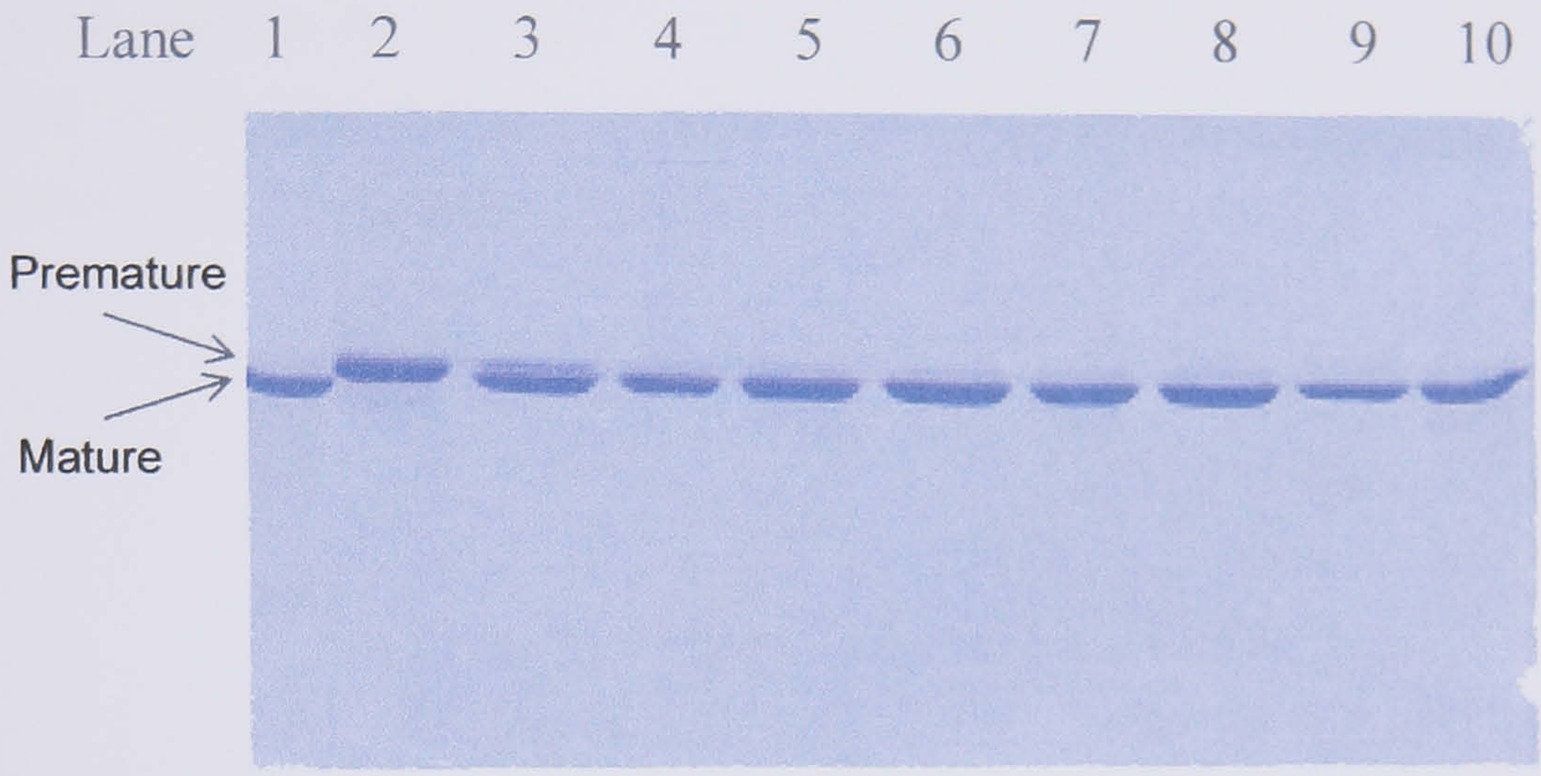


Figure 1.22: SDS-PAGE analysis of Cu^{2+}-dependent biogenesis

SDS-PAGE showing time course of copper addition to premature galactose oxidase [Firbank, 2002]. Lane **1**: mature galactose oxidase, used as standard marker. **2**: purified premature galactose oxidase. **3**: time 0 minutes after addition of copper acetate. **4**: time = 3 minutes. **5**: time = 6 minutes. **6**: time = 9 minutes. **7**: time = 20 minutes. **8**: time = 30 minutes. **9**: time = 2 hours. **10**: time = 3 hours. Although, lane 3 should be considered 0-2 minutes, due to the time taken for sample collection and preparation.

1.5.7 Proposed mechanism for thioether bond formation

A general mechanism for thioether bond formation was proposed [Firbank *et al.*, 2001; Rogers *et al.*, 2000], suggesting the reaction could occur by either a radical or ionic mechanism (Figure 1.23). The copper was proposed to initially bind to two histidines, followed by Tyr 272 coordination, allowing a resonance between Cu^{2+}-Tyr and Cu^{1+}- $Tyr^{\bullet}$ forms. The reaction proceeds via either a radical or ionic mechanism, producing the radical containing fully oxidised enzyme.

In the radical mechanism, radical coupling reaction between the $Tyr^{\bullet}$ and the lone pair of electrons of Cys 228 is proposed to occur. Oxygen is then required to re-oxidise Cys 228 and remove the proton. The ionic mechanism occurs via dioxygen coordination to the Cu^{1+} form to generate a superoxide species. Superoxide is then proposed to attack the Tyr 272 ring, enabling the nucleophilic attack of Cys 228 to tyrosine.

Similar mechanisms were proposed by Schwartz and Klinman [2001] (Figure 1.24). The first proposed the same mechanism as the ionic method, where a dioxygen binds to copper and attacks the tyrosine ring and the Cε position. The cysteine sulphur then carries out nucleophilic attack of the Tyr 272 Cε, forming a bridge between cysteine and tyrosine. Then the remaining proton is removed by either water, or a neighbouring residue acting as a general base. The second mechanism is a variation on the radical coupling mechanism, featuring the binding of dioxygen to the Cu^{1+}-$Tyr^{\bullet}$ form of the enzyme, resulting in a dioxygen radical bound to the copper. This radical is then able to activate the cysteine by producing a radical sulphur which attacks the tyrosine Cε position, thus generating the thioether bond.

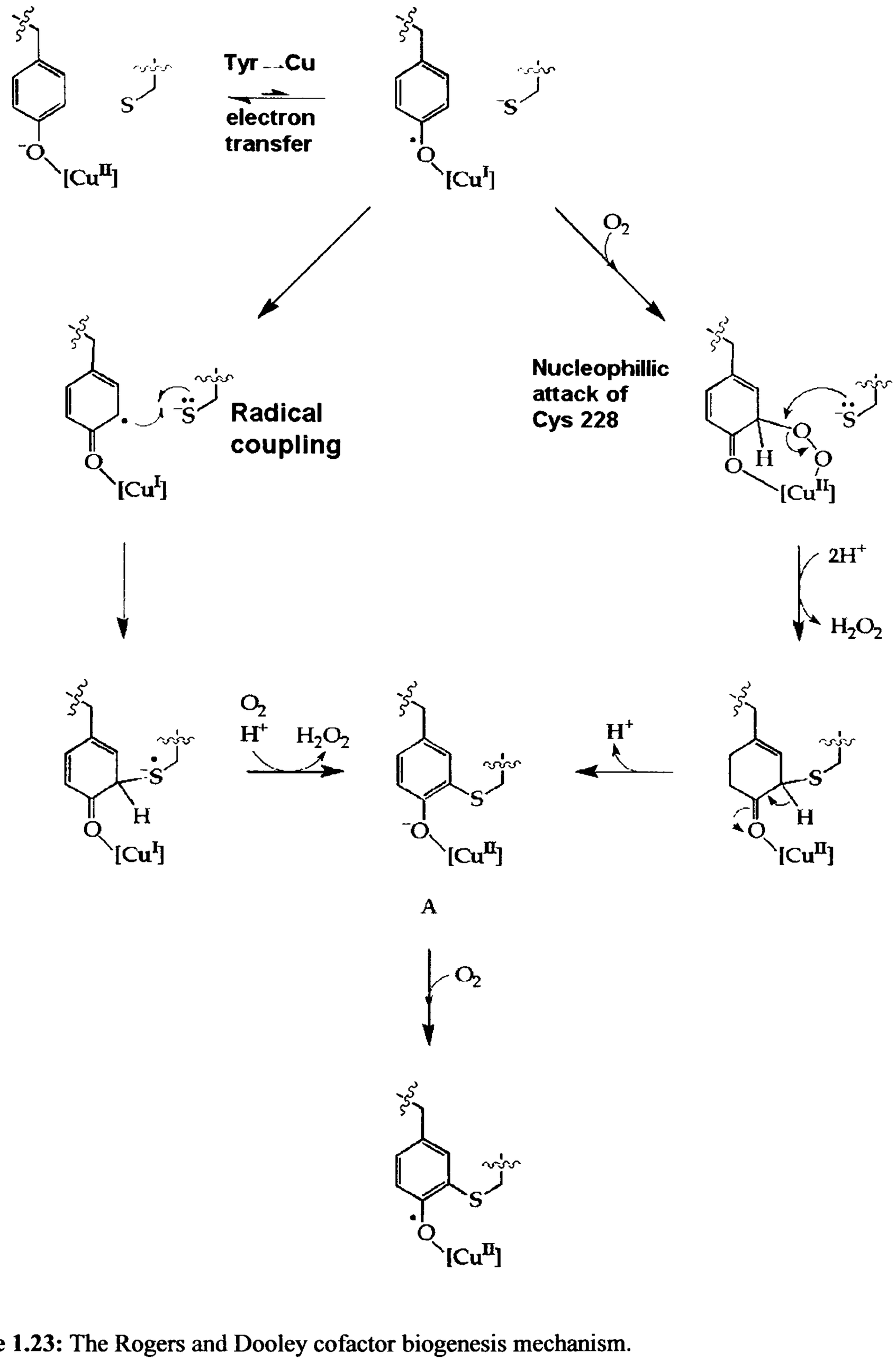


Figure 1.23: The Rogers and Dooley cofactor biogenesis mechanism.

Biogenesis is proposed to occur either via a radical or ionic mechanism, involving the presence of both copper and oxygen. Figure from Firbank *et al.*, [2001].

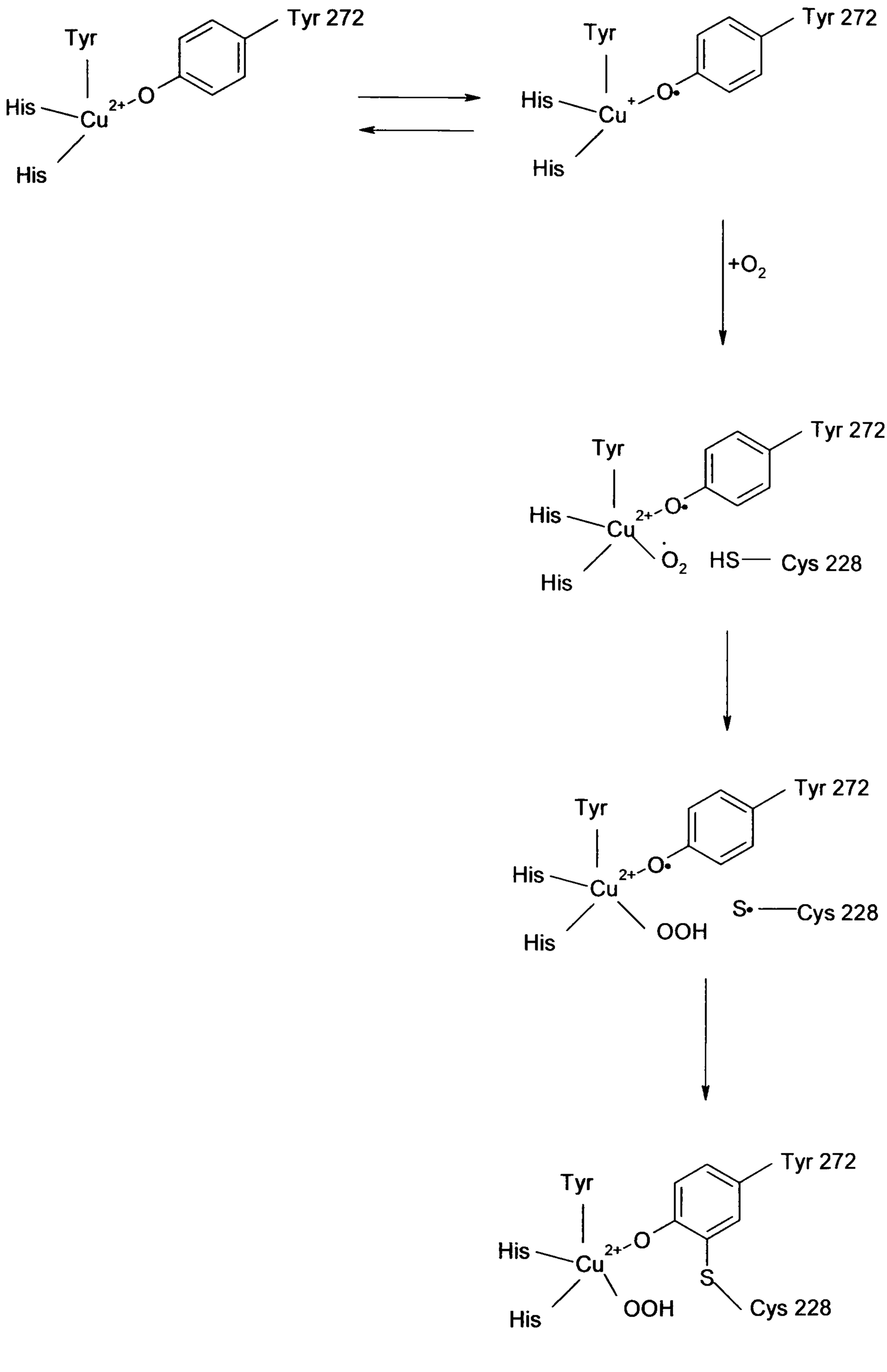


Figure 1.24: The Schwartz and Klinman cofactor biogenesis mechanism.

Thioether bond formation is proposed to occur via radical activation of cysteine via dioxygen binding. Figure adapted from Schwartz and Klinman, [2001].

In all the proposed mechanisms, excluding the Rogers and Dooley radical coupling method, oxygen has been proposed to directly interact with the copper, prior to thioether bond formation. Structural analysis of thioether bond formation using premature galactose oxidase crystals has led to an alternative mechanistic proposal. Anaerobic incubation of the premature galactose oxidase crystals with 2 mM copper did not result in thioether bond formation. His 496 and His 581 were copper ligands and surprisingly, Cys 228 was also interacting directly with the copper ion. It is reported, however, that the copper site was only 20% occupied [Firbank, 2002]. A 24 hour anaerobic copper incubation resulted in thioether bond formation [Firbank, 2002]. Copper was coordinated by His 581, His 496 and Tyr 272, with the axial Tyr 495 ligand at a distance of 2.9 Å. The presence of only three equatorial copper ligands and cofactor formation in anaerobic conditions led to suggestions that copper may have been in its Cu^{1+} oxidation state. Based on the crystallographic results, particularly visualisation of sulphur-bound copper, a mechanistic proposal for thioether bond formation (Figure 1.25) [Firbank, 2002].

In this mechanism Cys 228 coordinates to copper, forming a Cu^{1+}-Cys 228$^{\bullet}$ form in resonance with Cu^{2+}-Cys 228. The radical species would attack the Cε of Tyr 272, oxidation by molecular oxygen would complete the reaction, forming semi-reduced galactose oxidase. A further reaction with molecular oxygen is required to produce the fully oxidised, radical containing protein.

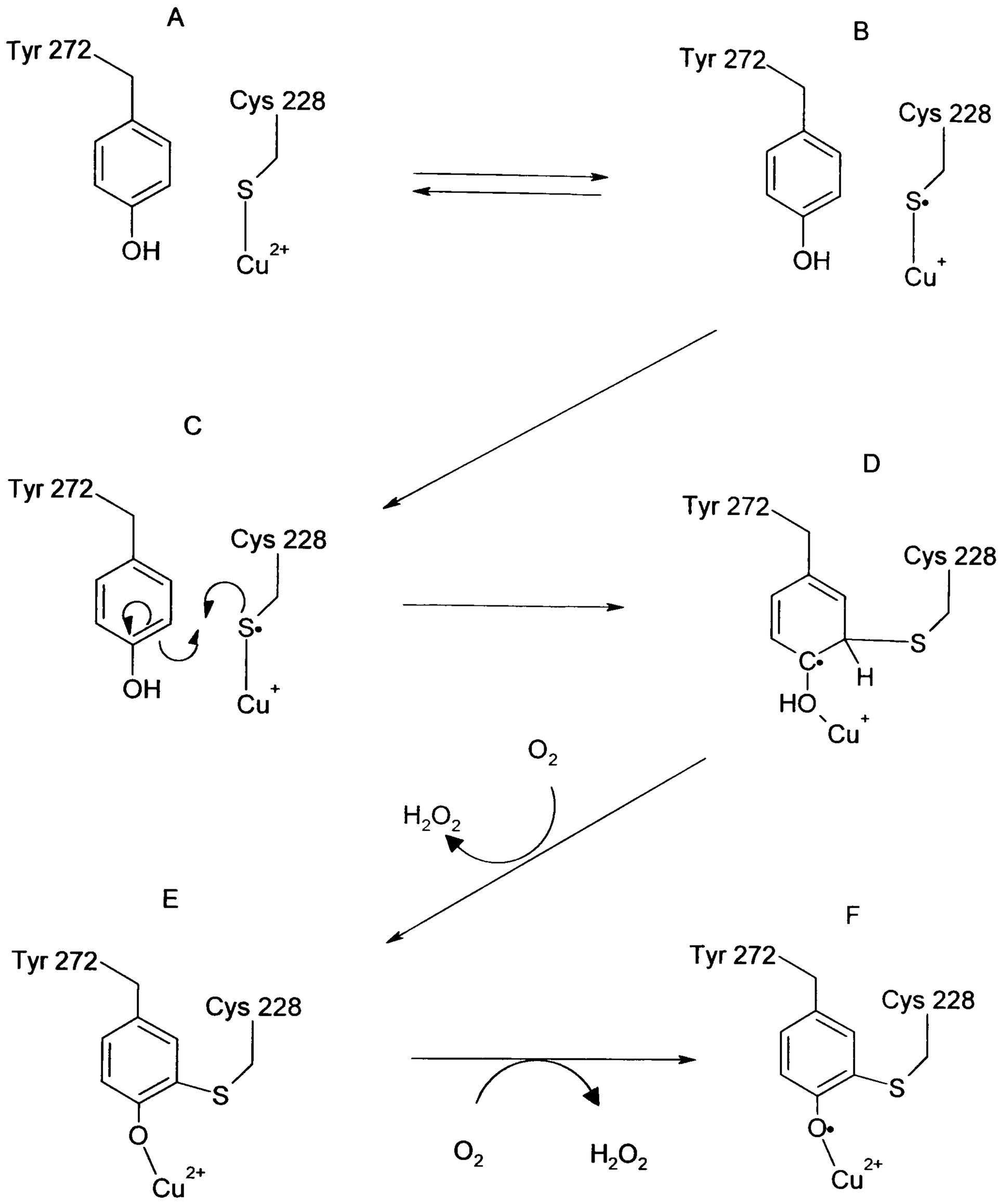


Figure 1.25: Firbank cofactor biogenesis mechanism

Proposed mechanism for the thioether bond formation based on the crystal structures obtained from anaerobic incubation of premature galactose oxidase in the presence of copper. Cys 228 binds to copper, allowing resonance between Cu^{2+}-Cys and Cu^{1+}-$Cys^{\bullet}$ forms (A and B respectively). Radical coupling between Tyr 272 and $Cys^{\bullet}$ initiates thioether bond formation (C and D). Oxidation of the protein by dioxygen results in the removal of two electrons from the protein and the release of hydrogen peroxide. The protein is now in the semi-reduced state (E), reaction with another molecule of dioxygen is required to form the radical containing, catalytically active protein (F). Adapted from Firbank [2002].

There is evidence suggesting that a Cys 228-Cu species could responsible for the transient 406 nm adduct observed during spectroscopic analysis of thioether bond formation. EPR and resonance Raman studies attribute features of a 406 nm complex to a copper-sulphur ligand to metal charge transfer [Whittaker and Whittaker, 2003]. This supports the structural evidence. Furthermore, a 406 nm feature is reported in several low-coordinate type 2 copper-sulphur complexes, such as mutational variants of azurin [van Amsterdam *et al.*, 2002], and Cu/Zn superoxide dismutase [Liu *et al.*, 2002; Lu *et al.*, 1996].

Recent work by Whittaker and Whittaker, investigated the metal oxidation state dependence of thioether bond formation [Whittaker and Whittaker, 2003]. Experiments were carried out using premature galactose oxidase expressed from *Pichia pastoris*. Whittaker reported that addition of Cu^{1+} to premature galactose oxidase under anaerobic conditions followed by exposure to buffer containing oxygen produced fully active enzyme. The results of this study also suggested the rate of Cu^{2+}-dependent thioether bond formation was nearly 10^4 times slower than the Cu^{1+} dependent rate of thioether bond formation. A radical mechanistic proposal was formed based on these results (Figure 1.26).

This reaction begins with the coordination of Cu^{1+} to the active site, followed by dioxygen coordination and oxidisation of Cu^{1+} to Cu^{2+}. The resulting superoxide reacts with the Cys 228 Sγ, creating a thyl free radical. The thyl radical attacks the Tyr 272 Cε, thus generating the thioether bond and inducing the reduction of Cu^{2+} to Cu^{1+}. This form of the enzyme is proposed to be identical to the fully reduced active site formed during substrate turn over and is suggested to undergo rapid reaction with a second oxygen molecule to form the fully oxidised active enzyme.

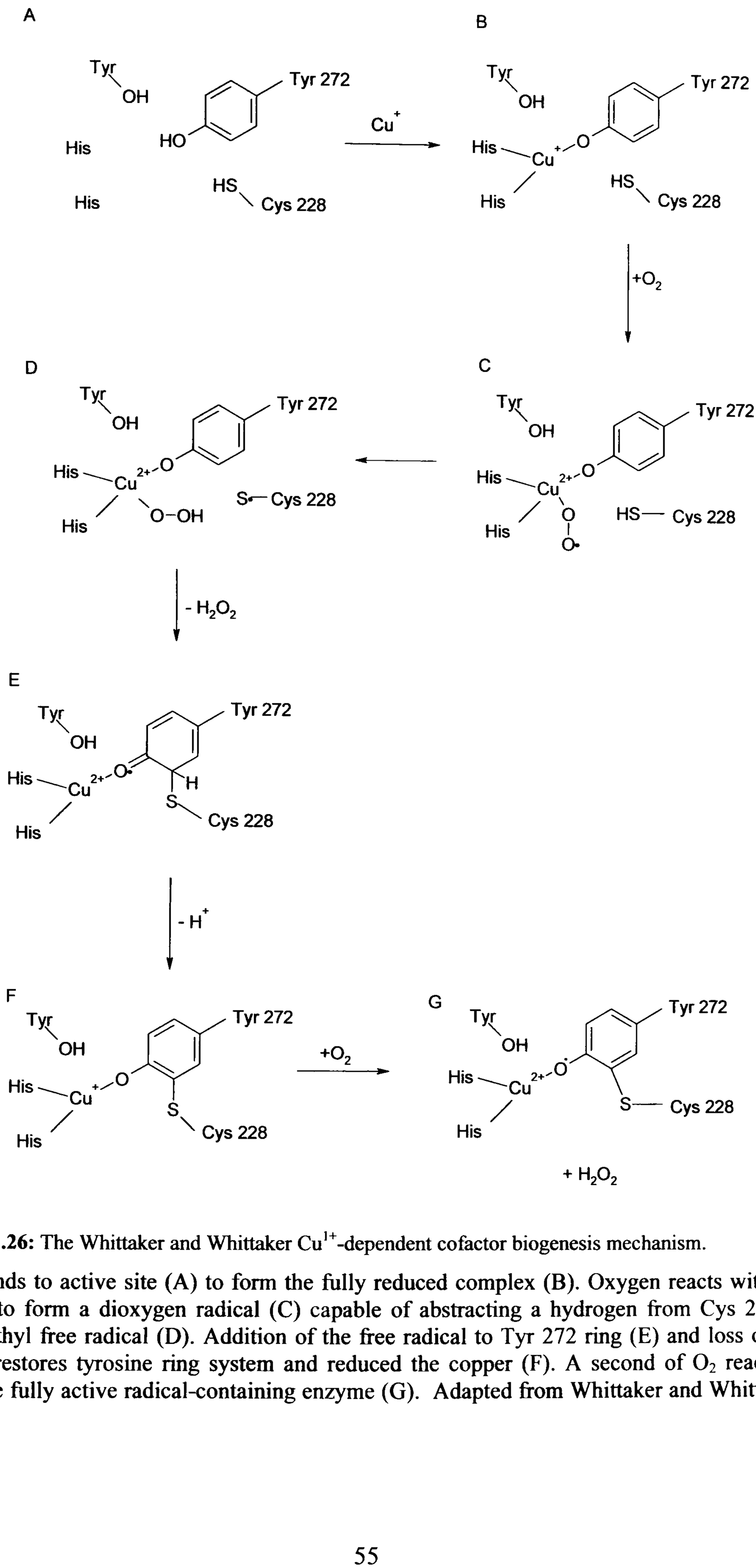


Figure 1.26: The Whittaker and Whittaker Cu^{1+}-dependent cofactor biogenesis mechanism.

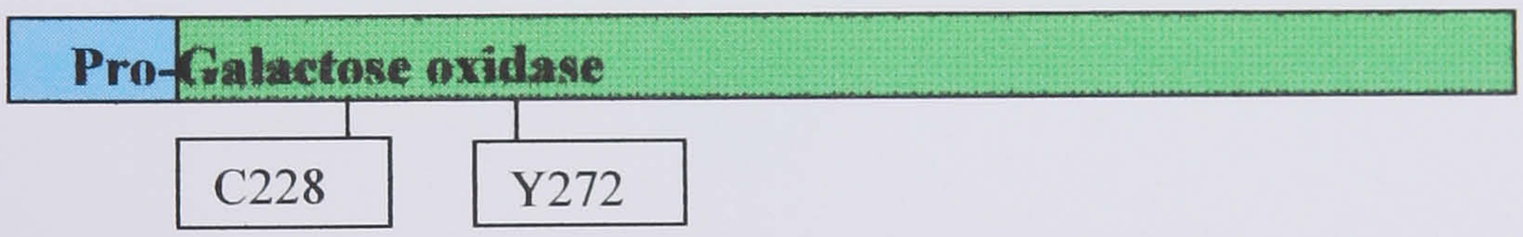
Cu^{1+} binds to active site (A) to form the fully reduced complex (B). Oxygen reacts with the copper to form a dioxygen radical (C) capable of abstracting a hydrogen from Cys 228 to form a thyl free radical (D). Addition of the free radical to Tyr 272 ring (E) and loss of the proton restores tyrosine ring system and reduced the copper (F). A second of O_2 reacts to form the fully active radical-containing enzyme (G). Adapted from Whittaker and Whittaker, [2003]

Whittaker proposed that anaerobic exposure of premature galactose oxidase to Cu^{2+} at low pH conditions (below pH 7.0) produced an abortive Cu^{2+}complex that did not process to form the active enzyme. UV-visible spectra observed the formation of an initial product which slowly converted to a distinct 406 nm absorbing form. Exposure to air resulted in the formation of the mature oxidised galactose oxidase UV-visible spectrum [Whittaker and Whittaker, 2003]. A 400 nm form was identified by single crystal spectrometry after copper addition to the pro-GO form [Firbank, 2002], while a 410 nm form was identified in solution following copper exposure to pro-galactose oxidase [Rogers *et al.*, 2000]. It is possible that all three forms are the same, or they may be representatives of the processing event.

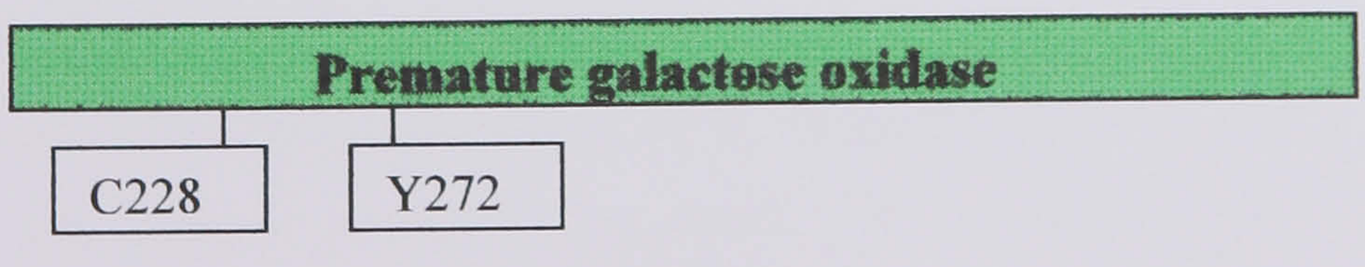
1.6 Forms of galactose oxidase

The existence of three oxidation states of galactose oxidase and the availability of the BD1 and AD1 constructs in *Pichia pastoris* have resulted in several different terms to describe the forms of the enzyme. A brief summary of these terms is provided below.

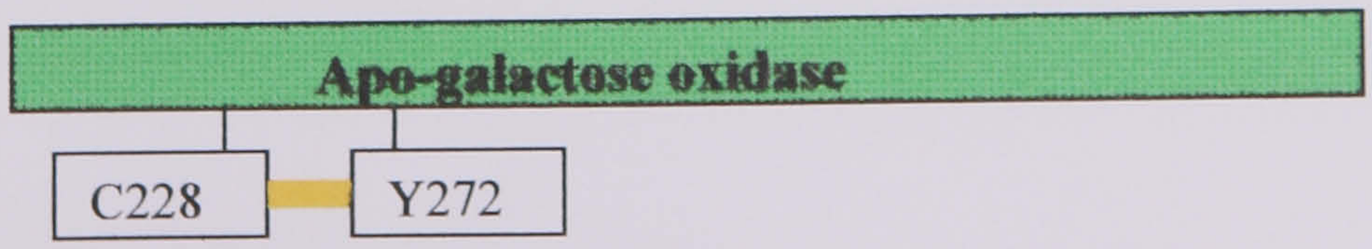
Pro-GO: A form of galactose oxidase that is produced after the cleavage of the signal peptide. This form of the protein does not have a thioether bond and contains a seventeen amino acid N-terminal pro-sequence. It is isolated after strict copper-free expression from *P. pastoris* containing the AD1 construct or copper-free *Aspergillus nidulans* preparations.



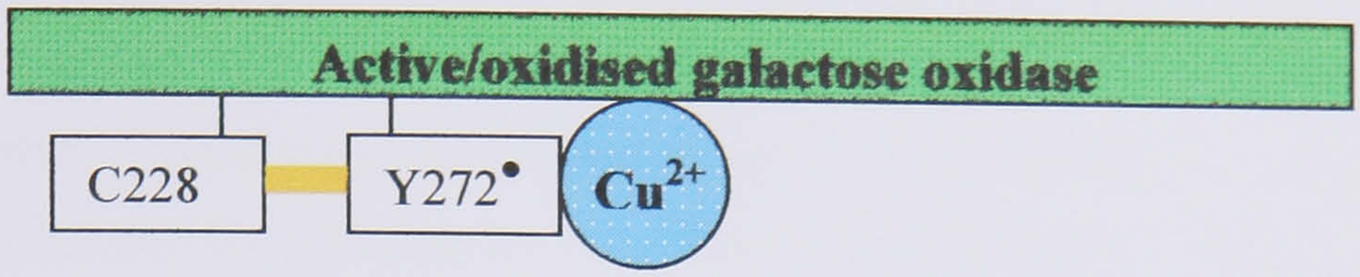
Premature galactose oxidase: A form of galactose oxidase lacking the pro-sequence and thioether bond. This form is also generated from copper-free preparations using the BD1 construct, which does not encode the pro-sequence.



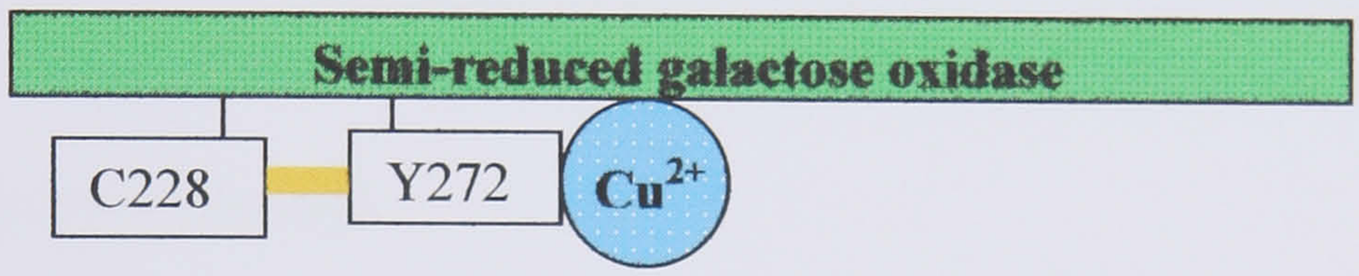
Apo galactose oxidase: After full biogenesis, the protein is said to be in the mature state. When copper is removed, the apo form of galactose oxidase can be crystallised. This form does not contain the pro-sequence and has formed the thioether bond.



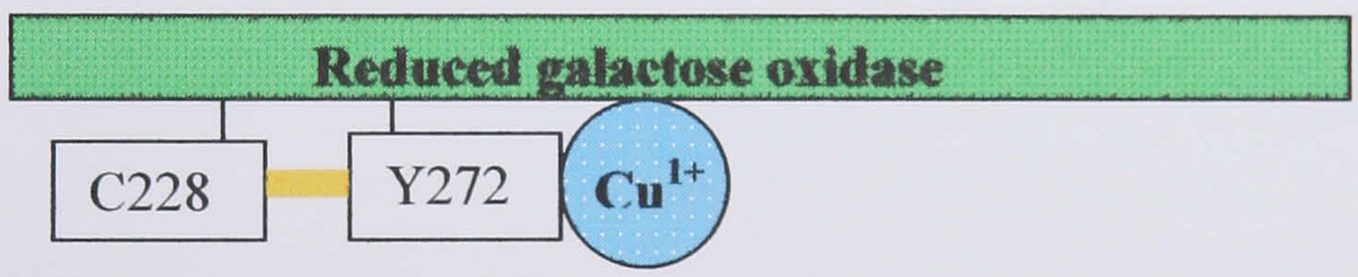
Active/oxidised galactose oxidase: The active enzyme contains the thioether bond and has a radical present at the active site. This form of the enzyme is ready for substrate catalysis.



Semi-reduced: This is an inactive form of the enzyme in which the radical is not present in the active site. Semi-reduced galactose oxidase contains a Cu^{2+} ion but can not catalyse alcohol oxidation. This form is also known as the mature or native form of the enzyme.



Fully/Substrate-reduced: During catalytic turnover the enzyme goes through a two electron reduction involving the loss of the radical and the reduction of Cu^{2+} to a Cu^{1+} ion.



1.7 Structural investigation using flash freezing techniques

Flash freezing techniques involve the rapid cooling of a protein crystal to 100 K. At this cryogenic temperature, catalysis is effectively halted, enabling the trapping of reaction intermediates. An elegant example of the use of this technique to elucidate TPQ formation in amine oxidase was reported by Kim *et al.,* [2002]. TPQ is formed by the self-processed oxidation of a specific tyrosine residue in the presence of copper and molecular oxygen [Cai and Klinman, 1994; Cai *et al.*, 1997; Dove *et al.*, 2000; Schwartz *et al.*, 2000; Schwartz *et al.*, 2001]. As TPQ is generated slowly in the crystal form of amine oxidase, Cu^{2+} incubation of crystals of apo-amine oxidase under strict anaerobic conditions, followed by exposure to oxygen allowed the initiation of the biogenesis reaction. X-ray diffraction analysis of crystals

frozen at 100 K revealed the structures of reaction intermediates, leading to a refined mechanism of TPQ biogenesis.

Elucidation of the catalytic mechanism and the biogenesis of galactose oxidase have advanced greatly since the first crystal structure of the protein was reported. Compelling structural evidence supporting the proposed catalytic mechanism or pro-GO processing, however, has not yet been presented. A complete understanding of these reactions would require the structures of the enzyme at different stages of the catalytic turn over and during the biogenesis process. Both reactions are copper and oxygen dependent, therefore, the initiation of these events in the absence of oxygen should allow the visualisation of reaction intermediates. As mentioned previously, such experiments have been attempted with some success in the case of biogenesis [Firbank *et al.*, 2001], however, trapping of catalytic intermediates during substrate oxidation has been unsuccessful.

1.8 X-ray crystallography

X-ray crystallography is an experimental technique used to elucidate the atomic structure of organic and inorganic molecules based on the fact that X-rays are diffracted by crystals. Based on the diffraction pattern obtained from the regular assembly of molecules or atoms in the crystal, together with the phases of the reflected rays, the electron density can be reconstructed. Pioneering X-ray diffraction studies were carried out in 1912 by the physicist Max Von Laue. He was interested in the way that light and crystals interacted with each other. He was the first to publish work about X-ray diffraction by crystals, for which he was awarded the 1914 Nobel Prize in physics. A year earlier the structure of a diamond was determined [Bragg and Bragg, 1913]. This work identified the length of a carbon-carbon bond to be 1.544 Å, and was awarded the 1915 Nobel prize.

Huge advances in crystallography were made after 1924, when Kathleen Lonsdale and W. T. Astbury, made their first major contribution to the developing field, a profound and systematic study of the theory of space groups, methods for their determination and the possibilities of molecular symmetry that are involved [Lonsdale and Astbury, 1924]. A later continuation and extension of this work consisted of the structure factor tables published in 1936 and the many volumes of the International Tables for X-Ray Crystallography published from 1952 onwards. In 1929, Lonsdale's second major contribution was to determine the structure of the benzene ring in hexamethylbenzene [Lonsdale, 1929], thus describing a structure first postulated by Kekulé in 1865. The description of the benzene ring recognised the molecule as a planar hexagonal molecule, with six carbon-carbon bonds of equal length.

In 1953, James Watson and Francis Crick deciphered and published the structure of DNA using X-ray fibre diffraction data collected by Rosalind Franklin and Maurice Wilkins [Watson and Crick, 1953]. This achievement helped to revolutionalise the field of genetics, as previous theories had suggested proteins and not DNA were the conduit for hereditary information. This Nobel prize winning work was only made possible by the earlier crystallographic studies of small molecules, which allowed the incorporation of bond distances and angles into the macromolecular DNA double helix.

The earliest report of a protein structure came a few years later, when the structure of myoglobin was reported [Kendrew *et al.*, 1958.; Kendrew *et al.*, 1960], shortly followed by the determination of the haemoglobin [Muirhead and Perutz, 1963]. Since the late 1950's, the rate of macromolecular structure determination has hugely increased. In 1971, The Protein Databank, a molecular structure database, was established in Brookhaven National Laboratory, Upton, Long Island, New York. In 1972, the protein databank contained 2 structures, by 1974 there were 12, and more recently there were over 44,800 X-ray diffraction structures of proteins, peptides and viruses in the database (www.rcsb.org/pdb) at the start of July 2007. With the continual advances in synchrotron radiation facilities and crystallographic computational methods, the quantity and quality of macromolecular structures determined each year increase. With the completion of the Human Genome Project [McPherson *et al.*, 2001], crystallographers are faced with the challenge of identifying and investigating genes and proteins that are involved in pathways of therapeutic interest. The use of structural information from biological molecules has the potential to impact heavily on the design of new pharmaceuticals. Structural genomics has taken a place alongside high-throughput chemistry and screening as an integral platform approach underpinning modern drug discovery.

Whilst obtaining the structure of a protein can provide invaluable information about new polypeptide folds or identify enzyme active sites, it is often desirable to observe the protein in a biological complex. For enzyme complexes, such structures can reveal insights into the catalytic mechanisms, protein-substrate or inhibitor interactions and be a guide for structure aided drug designs.

2 Material and Methods

All chemicals and reagents were of the highest quality and were obtained from SIGMA (U.K.) or FISHER (U.K.), unless otherwise stated.

2.1 Expression and purification of galactose oxidase

Mature wild type galactose oxidase was produced by expression and purification from *Aspergillus nidulans*, according to the procedure described by Baron *et al.*, [1994; 1993]. Galactose oxidase was expressed in *Pichia pastoris* following the procedure as described in the product manual (Invitrogen) and was purified by adapting the *A. nidulans* method.

2.1.1 *Aspergillus nidulans* culture growth and expression of galactose oxidase

Stocks of *A. nidulans* transformed with the wild type vector pGOF101 were incubated on a minimal media plate at 37°C for four days [Baron *et al.*, 1994]. The inoculum from the centre of the plate was streaked over the surface of another minimal media plate, which was then incubated at 37°C for a further four days. Six conical flasks (2 litre) containing 500 ml of minimal medium salts ($NaNO_3$ 60 g/L and KH_2PO_4 15.2 g/L) 15 µM $CuSO_4$ and 1.25% glucose were prepared and mixed with the spores. The *A. nidulans* cultures were incubated at 30°C for three days. The cells were induced by adding 2% maltose (final concentration) and 50 µM $CuSO_4$ (final concentration) and the flasks were incubated for a further three days at 30°C before harvesting.

2.1.2 Processing of *Aspergillus nidulans* cultures

The culture medium was filtered to remove mycelia through 15cm diameter discs of Whatman No.1 filter paper under suction. All subsequent steps were carried out at 4°C. 1 M Tris-HCl pH 8.0 was added at 1/40 of the culture volume, to control the pH during salt addition. Ammonium sulphate ($(NH_4)_2SO_4$) was slowly stirred into the medium until 30% saturation. The contaminating proteins were then separated from the medium by decanting. The preparation was brought up to 80% of ammonium sulphate saturation and then left for three hours to allow precipitation of galactose. The mixture was centrifuged at 11000 rpm (Sorvall SLA 3000) for 20 minutes at 4°C. The supernatant was decanted and the pellet re-suspended in 140 ml 10 mM sodium phosphate buffer pH 7.3. The protein solution was dialysed against 5 L of the same buffer for 24 hours, with three buffer changes

2.1.3 Cation exchange chromatography

The protein solution was applied to cellulose phosphate column pre-equilibrated with 10 mM Buffer A (sodium phosphate buffer pH 7.3). Following extensive washing with Buffer A, the bound protein was eluted by applying a linear gradient with buffer B (100 mM sodium phosphate buffer pH 7.3) over a volume of 80 ml. To ensure complete elution of protein, a further 150 ml of Buffer B was used to wash the column. Fractions were analysed by SDS-PAGE and those containing pure protein were pooled. The protein content was estimated by measuring absorbance at 280 nm (see section 2.1.8).

2.1.4 Culture growth and expression of galactose oxidase in *Picha pastoris*

Colonies of *P. pastoris* X-33 cells transformed with the BD1 construct (chapter three and five) or C383S construct (Chapter three) were streaked onto Yeast Peptone Dextrose plates (YPD: 1% yeast extract, 2% peptone, 2% agar, 20% glucose) containing 100 μg/ml of the selection marker Zeocin. After two days, a single colony was picked and used to inoculate a 20 ml overnight starter culture (1.4% yeast extract, 2.9% soytone). This starter culture was used to inoculate 200 ml volumes of sterile buffered glycerol media (BMGY:1.4% Yeast extract, 2.9% soytone, 1 M potassium phosphate pH 6, 13.4% YNB, 0.02% biotin and 10% glycerol) in 2 L flasks and the cultures grown at 30°C, with shaking at 200 rpm, and monitored spectrophotometrically at 600 nm. When the OD_{600} reached 6, the cultures were spun down in pre-sterilised centrifuge pots at 11 000 rpm (Sorvall SLA 3000) for five minutes. The pellets were re-suspended in 200 ml of sterile buffered methanol media (BMMY: 1.4% Yeast extract, 2.9% soytone, 1 M potassium phosphate pH 6, 13.4% YNB, 0.02% biotin, and 5% methanol) and incubated for a further 24-48 hours at 30°C, 200 rpm before harvesting (section 2.1.5). During incubation period additional methanol was added every 24 hours to a final concentration of 0.5% to compensate for losses by evaporation.

2.1.5 Processing of *Picha pastoris* cultures

Cells were pelleted by centrifugation at 11 000 rpm (Sorvall SLA 3000) for 15 minutes and the supernatant decanted. The supernatant was dialysed extensively against 10 mM sodium phosphate buffer, pH 7.3. The dialysed medium was loaded onto a pre-packed cellulose phosphate column that had been previously equilibrated with the same buffer. Washing and eluting the protein from the column was then carried out as described above in section 2.1.3.

The C383S variant was purified from the culture supernatant using a modification of the affinity/size exclusion procedure described by [Sun *et al.*, 2001]. Proteins were concentrated from the 50mM Tris-HCl, pH 8.0 buffered culture supernatant by adding ammonium sulphate to 65% saturation. The solution was centrifuged at 11 000 rpm (Sorvall SLA 3000) for 20 minutes, the supernatant decanted and the pellet re-suspended against 100mM ammonium acetate buffer, pH 7. The solution was dialysed against the same buffer for 24 hours prior to loading onto a pre-equilibrated Sepharose 6B column. The eluted enzyme was copper-loaded according to the dialysis procedure described by [Baron *et al.*, 1994].

2.1.6 Copper loading and concentration of galactose oxidase

The protein was dialysed against 20 mM HEPES pH 6.1, followed by dialysis against the same buffer containing 1 mM copper nitrate, to ensure copper loading at the active site and generation of the fully mature protein. The protein was then concentrated using a spin concentrator with a 30 kDa cut-off, (CENTRICON, UK) to a final concentration of approximately 4 mg/ml.

2.1.7 Measurement of protein concentration

Estimates of protein yields were obtained by measuring the absorbance of a sample at 280 nm using a Pharmacia Ultraspec III or a UV-visible spectrophotometer. The extinction coefficient of the enzyme has been determined as 104 900 M^{-1} cm^{-1} [Kosman *et al.*, 1974]. An absorbance of 1.0 corresponded to 0.65 mg/ml of galactose oxidase.

2.2 Crystallisation of galactose oxidase

2.2.1 Crystallisation of mature galactose oxidase

Galactose oxidase was crystallised by vapour diffusion using the sitting drop method. Sitting drop microbridges (Oxford Crystal Microsystems, U.K.) were placed in 24-well plates (Hampton Research, U.S.A). 1 ml of mother Liquor was placed in the bottom of the wells and 7 µl aliquots of protein solution were placed in the depression of each microbridge and mixed gently with 7 µl of mother liquor by repeat pipetting. Each well was then sealed with a glass coverslip and immersion oil and the trays stored at 18°C.

The C2 space group crystals were obtained from protein crystallised in 0.5-1.0 M ammonium sulphate, and 0.1-0.2 M sodium acetate pH 4.5. The same crystal form was obtained when

copper loaded protein was crystallised in 10-20% PEG 8000, 100 mM MES pH 5.5-6.3 and 200 mM calcium acetate.

2.2.2 Copper free crystallisation

In order to obtain crystals in the space group $P2_12_12_1$, copper was removed from the mature enzyme prior to crystallisation. This was achieved: either by dialysing the enzyme against 2 mM sodium diethyldithiocarbamate (DCC) buffered at pH 7 by 20 mM HEPES at 4 °C overnight; or by adding a concentrated solution of DCC to the buffered enzyme to give a final concentration of 8 mM DCC and incubating at 4°C overnight. Prior to crystallisation, the sample was centrifuged at 13 000 rpm using a Sanyo Microcentaur bench-top centrifuge. Trays were set up with the copper free protein using a mother liquor containing 10 – 20% PEG 8000, 100 mM MES pH 5.5-6.3 and 200 mM calcium acetate. When strict copper free conditions were required, a small amount of CHELEX 100 resin was placed in the base of each well to remove copper from the mother liquor. The trays were mixed gently for at least an hour before the addition of protein, to allow thorough removal of copper from the solutions.

2.3 X-ray data collection

2.3.1 Preparation of crystals for data collection

All data collection was carried out at 100 K and so crystals were flash cooled in liquid nitrogen or a 100 K nitrogen stream prior to exposure to X-rays [Garman and Schneider, 1997]. There are two distinct advantages to data collection at 100 K. The first reason for cooling protein crystals is that they are liable to radiation damage if exposed to X-rays, and this damage can be so serious that the X-ray pattern dies away after a few hours of exposure at room temperature. At 100 K, however, secondary radiation damage is greatly reduced, and virtually disappears. Secondly, data collection at 100 K allows trapping of reaction intermediates in the crystal; for example, crystals can be soaked in substrate anaerobically, frozen, and then data collected without allowing oxygen to enter the crystal lattice.

In order to cause the minimum of damage, crystals were first washed in a cryoprotectant – a solution in which the crystals are stable and are also protected during the cooling process. Cooling of crystals from room temperature to cryogenic temperature was achieved rapidly, to prevent crystallisation of the water in the mother liquor. Ice crystals would disrupt the crystal lattice, thereby reducing the quality of diffraction. Cryoprotectants were made using a

solution similar to the mother liquor, but also containing 10-20% of PEG 400, such that upon cooling to 100 K the solution formed a clear glassy drop. Crystals were washed in this cryoprotectant for a few seconds before being transferred to a cool nitrogen stream, or placed directly into liquid nitrogen.

2.3.2 X-ray sources for data collection

Data were collected both on home sources and using synchrotron radiation. The two home sources were Rigaku rotating anode instruments, with a RU200 rotating anode operating at 50 kV and 100 mA using osmic mirrors to focus the beam. A copper anode resulted in the emission of radiation with a wavelength of 1.5418 Å. Data were collected on image plate detectors (R-AXIS IV++, MSC, U.K.).

Synchrotron radiation was provided by two facilities; The European Synchrotron Radiation Facility (ESRF), Grenoble, France using beamline ID14-4 (wavelength 0.93 Å); and Daresbury Synchrotron Radiation Source (SRS), Daresbury Laboratory, U.K, using beamlines 9.6, 14.1 and 14.2 with wavelengths ranging from 0.87 to 1.488 Å. Synchrotron radiation is produced by circulating electrically charged particles (negatively charged electrons or positively charged positrons) at nearly the speed of light. The particles are injected into a storage ring directly from a linear accelerator or through a booster synchrotron. An advantage of data collection using synchrotron radiation is the intensity of the beam is two orders of magnitude stronger than for conventional X-rays generated by rotating anodes, thus allowing a higher resolution data collection than is possible on home sources.

2.3.3 Data collection strategy

Both at the synchrotron and home sources the strategies employed for data collection were the same. After mounting and aligning the crystal in the beam, an image was taken at 0° and 90°. These images were used to determine the orientation of the crystal, estimate crystal mosaicity, and optimise crystal to detector distance, crystal oscillation angle, and exposure time. The identification of the symmetry and orientation of the crystal can allow data collection to be optimised to ensure that a complete data set is obtained. Programs such as the STRATEGY subroutine of MOSFLM [Leslie, 1992] were used to assist this process.

All crystals have an amount of inherent mosaicity, caused by slight variations in the position of molecules as they attach to the forming crystal lattice. The result is that an X-ray reflection emerges from the crystal as a narrow cone rather than a perfectly linear beam. This not only

causes each reflection to appear slightly diffuse, but also results in spots appearing on more images than expected for a crystal displaying a lower mosaicity. This must be taken into consideration during data collection, to prevent later problems during data processing when it may become impossible to separate the end of one spot from the beginning of another. To prevent this effect, the oscillation angle during data collection can be decreased so that adjacent spots do not overlap from one frame to the next. The crystal to detector distance is adjusted to ensure that spots on the same frame are well enough separated to allow processing. As the crystal to detector distance increases, the spots are more separated but the resolution limit at the edge of the plate is reduced. A balance between adequate spot separation and data resolution limit must be found. The exposure time required for each image changes between different data sets and depends upon several variables, including beam intensity, detector sensitivity, and crystal quality. After all of these parameters have been taken into consideration, data collection can commence.

2.4 Data processing

Data processing was achieved using one of two available programmes; MOSFLM [Leslie, 1992], or the HKL suite [Otwinowski and Minor, 1997]. In order to begin data processing, the data collection parameters should first be input into the processing software. To enable the correct unit cell and orientation of the crystal to be determined, the distance from the crystal to the detector, the wavelength of the radiation used, the angle over which the crystal was oscillated per frame, and beam position must be provided.

2.4.1 Data integration using MOSFLM

Strong reflections from an image were indexed using MOSFLM [Leslie, 1992]. This was usually successful with one image, but in some cases, particularly after soaking crystals in substrate, spots were poor, and more than one image was required. To index the frames, the beam centre and crystal to detector distance were given. Scattering and distance vectors were calculated and matched with the crystal systems. Once the correct space group and unit cell had been identified, the predicted spot pattern at a given crystal orientation was displayed. This allowed verification of indexing, since the predicted pattern should correspond well with the observed reflections. If the two corresponded well, then further refinement of the parameters and cell dimensions were carried out over three sections of 2-5 ° throughout the data. All frames were then processed, allowing the mosaicity to refine throughout the procedure. A summary file produced at the end of processing was used to identify any major problems during processing. If possible the suggestions in this summary file were

incorporated into the input parameters and processing was repeated until the best statistics were obtained.

2.4.2 Data Reduction with CCP4 suite

The integrated reflections were scaled together and an average value of each *hkl* obtained using SCALA [CCP4, 1994]. Several values listed in the output of this programme were used to assess the quality of the data:

R_{sym} is a statistical indication of the agreement between spot intensities, when the same reflection, or its symmetry equivalent, has been measured more than once. A low R_{sym} value is a good indication that the data set has good agreement between equivalent reflections. High values, however, suggest the intensities may not have been accurately measured, or that non-equivalent intensities are being merged, which results from mis-indexing of the diffraction pattern. The R_{sym} is also useful for interpreting the useful resolution limit of the data set.

R_{sym} is defined as:

$$R_{sym}(I) = \frac{\sum_{hkl}\sum_{i} | I_i(hkl) - \overline{I(hkl)} |}{\sum_{hkl}\sum_{i} I_i(hkl)}$$

Where $I_i(hkl)$ is the intensity of an observation for a given reflection, and the $\overline{I(hkl)}$ is the average intensity of the *i* observations of that reflection.

The multiplicity is a record of how many times a reflection, and its symmetry related equivalents have been measured on average. It also indicates the number of observations that have been used to calculate the mean intensity for each reflection. A higher multiplicity shows that more observations have been used to calculate the mean intensity, hence providing a more accurate value for the intensity of a reflection.

A third important statistic is the completeness of the data set. For a given unit cell and resolution, the number of independent reflections can be calculated. The completeness of a data set indicates what proportion of the expected number of reflections at a given resolution

has been collected. Data can be incomplete due to the loss of random reflections throughout the data set, or by the absence of a continuous segment of data.

2.4.3 Data processing and reduction with the HKL suite

Several data sets were processed using the HKL suite [Otwinowski and Minor, 1997]. A single frame from the data set was used for indexing and the predicted diffraction pattern for the image was fitted to the corresponding observed pattern by adjusting various parameters such as beam centre, crystal orientation, spot size, distance, cell dimensions or mosaicity. Once the predicted pattern matched the observed spots, and the statistical estimates of fit (Chi^2 values) indicated a good alignment the frame was processed. The parameters used for this frame were then applied to the remaining frames.

Data were scaled and merged using SCALEPACK [Otwinowski and Minor, 1997]. The quality of data was again assessed by considering the values of R_{sym}, multiplicity and completeness. The output from SCALEPACK was converted to an MTZ format using the CCP4 programme MTZ2VARIOUS [CCP4, 1994].

2.4.4 Conversion from intensities to amplitudes

The MTZ format output from both of these processes containing intensities was then converted to structure factor amplitudes (the square root of the intensities) using CCP4 programme TRUNCATE [CCP4, 1994].

2.5 Structure determination

2.5.1 Introduction

Once the diffraction data have been collected, reduced and scaled, the result is a series of intensities, or amplitudes (the square root of intensities). Unfortunately, this information is not enough to determine the protein structure. When we record diffraction intensities, we lose the phase information that is needed in order to simulate an X-ray-focusing lens. The phase information for each intensity is required to enable the calculation of the electron density within a unit cell. There are several ways to obtain phase information in macromolecular crystallography. If the structure is completely novel, two methods are predominantly utilised. Multiple isomorphous replacement generates phasing information based on the comparison of two crystals which differ because one or more strong scattering elements exist in one crystal (the derivative) that is absent from the other (the native). The technique is called

'isomorphous' because the native and derivative crystals must have the same unit cell dimensions and space group symmetry. Derivative crystals are prepared by soaking of crystals in heavy atom solutions (such as Hg, Pt or Au). The atoms usually bind to one or more specific sites in each unit cell without perturbing its conformation or crystal packing. A data set is collected from each heavy atom derivative, and the changes that occur due to their presence can be used to derive the initial phases for the structure. A second method of generating phases for proteins with an unknown structure is called anomalous dispersion. This method takes advantage of the signal from anomalous scatterers at particular wavelengths. To assist this process, proteins can be expressed containing selenium derivatised methionines, or bound metals in metalloproteins can be used. By collecting data sets at different wavelengths on the same crystal, phases can be obtained by measuring differences due to the changes in anomalous behaviour at the different wavelengths.

There is a third method of phasing that can be employed, molecular replacement. If the new protein is homologous (sequence or structure) to that of a known protein, the known protein may be used as a phasing model and thus solve the phase problem without heavy atom derivatives. Once phases have been obtained, they can be combined with the amplitudes in order to obtain an electron density map into which the structure can be built. The electron density ($\rho(x,y,z)$) is calculated by performing a Fourier transform of the structure factors to convert the crystallographic data into an image of the unit cell and its contents. The formula that relates $\rho(x,y,z)$ to the diffraction pattern is:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h}\sum_{k}\sum_{l} \mathbf{F}(hkl) \exp(-2\pi i(hx + ky + lz))$$

Where $\mathbf{F}(hkl)$ is the structure factor for reflection (hkl), V is the volume of the unit cell and x, y, and z are the coordinate in the unit cell. Each $\mathbf{F}(hkl)$ is a complete description of a diffracted ray recorded as reflection hkl. Being a wave equation, $\mathbf{F}(hkl)$ must specify frequency, amplitude $|F|$, and a phase angle α with the relationship $\mathbf{F}(hkl)=|F|\exp[i\alpha]$. The amplitude is proportional to the square root of the measured intensity of a reflection. The phases of the reflections are obtained as explained earlier.

2.5.2 Molecular replacement

Molecular replacement, can be employed to use phases from a known protein structure to determine the phases of an unknown structure. It involves a systematic trial and error approach to position the model of the known structure in the unit cell of the unknown

structure. This is carried out in two separate stages, first using a rotation function, and then a translation function. This process is not undertaken in real space as this would be very computationally time consuming. Instead Patterson functions, which are independent of phases, are calculated for both the known and unknown structure and used throughout the molecular replacement process. The Patterson map is a vector map, which consists of two parts: intramolecular vectors, or self vectors, which are vectors between atoms within the same molecule; and intermolecular vector, or cross vectors, which are vectors from atoms in one molecule in the crystal to atoms in another molecule. The two vector sets are used to optimally orient and position the search model within the new unit cell.

2.5.2.1 Rotation function

The first step in molecular replacement is to find the correct rotational orientation of the trial model. The self vectors of the search model will overlap with the self vectors of the protein of interest when they are in the same orientation, providing the two structures are sufficiently similar. The search should ideally be limited to self vectors only. To do this, the search is carried out within a given radius that is smaller than the radius of the search model. This will reduce the number of cross vectors within the Patterson map. The rotation search is carried out, and the overlap function (R) calculated for each orientation of the Eulerian angles, α, β, and γ. $P(\mathbf{u})$ is the observed Patterson function, whilst $P_r(\mathbf{u})_r$ is the calculated Patterson function after rotation. The overlap function is defined by:

$$R(\alpha,\beta,\gamma) = \int_U P(u) \times P_r(u)_r du$$

[Rossmann and Blow, 1962] and will have a maximum value for the correct overlap.

2.5.2.2 Translation function

The rotation function establishes the correct orientation for the molecule, but this does not complete the molecular replacement procedure. The position of the molecule in the unit cell must also be established and this is achieved using a translation search. This function may be conceptualised and realised by moving the model through the asymmetric unit, and calculating the correlation between the calculated and observed structure factors. This

process, however, requires extensive computer calculations. An alternative approach is generally taken, again maximising the overlap function of the observed and calculated Patterson functions. Whereas in the rotation search self vectors are used, the translation function is founded on the use of cross-vectors. The overlap of the calculated cross vectors with the observed Patterson map is established as a function of translation vector **t** [Crowther and Blow, 1967]:

$$T(t) = \int_V P_{1,2}(u,t)\, P(u)\, du$$

where $P_{1,2}$ represents the cross vectors between a molecule, 1, and its symmetry related molecule, 2, calculated for different values of the translation vector **t**. The translation function, T, will have a maximum when the two vector sets overlap.

2.5.3 Molecular replacement using AMoRe

Molecular replacement was carried out using the CCP4 supported program AMoRe [Navaza, 1993; Navaza, 1994; Navaza, 2001]. The structure of mature galactose oxidase was used as a search model, after removal of all water molecules and metal ions. The molecular replacement procedure is separated into several stages. The first two stages, SORTING and TABLING are required to prepare the experimental structure factors (SORTFUN and model structure factor, calculated from the model structure file (TABFUN). The structure factors are tabulated to allow interpolation during searches. This makes the procedure more rapid. The rotation function is carried out in the ROTING stage using the Crowther fast rotation function [Crowther, 1972]. The radius of sphere used was approximately half the diameter of the model in the smallest dimension. The best solutions to the rotation search were input into the translation search function (TRAING). The solutions were judged according to their high correlation coefficients and low R-factors. The best solution from the translation search was subjected to rigid body refinement using the FITING procedure and the initial model then re-orientated according to the solution provided by the REORIENTATE routine.

2.6 Electron density map calculations

The solution from molecular replacement provides an initial model for the structure and also approximate phases. The amplitudes have already been measured from the diffraction images and thus sufficient information is available to calculate an initial density map. By combining the observed amplitudes and calculated phases, various maps can by displayed. An F_{obs} map

could be calculated using the observed amplitudes and calculated phases from the model, however, the non-experimental phases would create significant bias to the electron density maps. In order to try and reduce model bias, $2F_{obs}$-F_{calc} maps (2Fo-Fc) are calculated, which amplify the difference between the observed and calculated amplitudes. F_{obs}-F_{calc} maps (Fo-Fc) are also used, which indicate where the calculated amplitudes (model) differ from the observed amplitudes. By changing the model to fit the electron density, the model is improved to represent the target structure. Electron density maps were calculated using the CNS suite of programs [Brunger *et al.*, 1998].

2.7 Model building and refinement

2.7.1 Improving the model

The model obtained from molecular replacement usually contains some differences from the search model. Two methods are utilised to improve this model. The first involves manual rebuilding of the model according to the calculated electron density maps, while the other requires the use of computational methods of model refinement.

Manual model building, allows adjustments to regions of the model that differ significantly from the electron density map. It can, however, result in poor model geometry and does not allow particularly accurate positioning. Model refinement is a method that can vastly improve the model. Protein crystals tend not to diffract to high enough resolution to refine the position of each atom from the observed data. This is taken into account by incorporating restraints during model refinement. Small molecule crystallography has allowed accurate determination of stereochemical information, as well as standard bond distances and angles of organic compounds, these known parameters can be used in macromolecular structure determination. Refinement procedures apply a weighting to improve the geometry of the structure, or to fit the observed X-ray data. A balance must be found to suit the data under consideration, so that a sensible model is obtained that correlates well to the observed data.

The correlation between the model and observed amplitudes is measured by a reliability factor, referred to as the R-factor:

$$R = \frac{\sum_{hkl} \left| \, |F(hkl)_{obs}| - |F(hkl)_{calc}| \, \right|}{\sum |F(hkl)_{obs}|}$$

In the case of a random structure the theoretical values of R are between 0.6-1.0. Thus if the R-factor approach these levels, the structure is not likely to be correct. As the structure improves and the amplitudes calculated from the model approach the observed amplitudes, the R-factor decreases. The progress of the refinement is followed using a *freeR*-factor, or R_{free} as well as the R-factor. The R_{free} is computed with a small set of randomly selected intensities, the "test set", which are set aside from the beginning and not used during refinement. They are used only in the cross-validation or quality control process of assessing the agreement between calculated (from the model) and observed data [Kleywegt and Brunger, 1996]. If the model is built appropriately, both R-values should decrease. For a structure determination with data to 2.0 Å resolution, the final R-factor should be close to 20% with R_{free} a few percent higher. For each model presented in this thesis a *freeR* set containing 10% of reflections was generated using the CPP4 program FREERFLAG [CCP4, 1994]. A different set of *freeR* reflections were inadvertently selected for each structure, this has led to similar R and R_{free} values.

2.7.2 Least squares refinement

Crystallographic refinement can be seen as minimising the difference between the observed and calculated structure factor amplitudes. This can be achieved using "least squares" refinement, by minimising the least squares residual:

$$Q = \sum_{hkl} w_{hkl} \left(|F_{obs}(hkl)| - |F_{calc}(hkl)| \right)^2$$

In words, the function Q is the sum of the squares of differences between observed and calculated structure factors ($F(hkl)$). The sum is taken over all reflections *hkl* currently in use. Each difference is weighted by the term w_{hkl}, a number that depends on the reliability of the corresponding measured intensity. The minimum of Q is found by varying the atomic parameters to make the differentials of Q zero with respect to each atomic parameter. The structure factor, however, is a non-linear function of the positional coordinates in the protein atomic structure (x, y, z) and temperature (B factor) and therefore the least squares residual cannot be minimised directly. Instead an iterative approach must be adopted and the minimisation process is continued until the changes are insignificant as measured by a convergence of the R-values.

2.7.3 Maximum likelihood

Least squares refinement is based on the assumption that errors in the prediction of the observations follow a Gaussian distribution. This, however, is not always the case and it has been shown that errors in structure factor amplitudes can be better described by a Rice distribution. Based on this finding an alternative approach to refinement called maximum likelihood can be used [Pannu and Read, 1996; Read, 1986].

During maximum likelihood refinement, observations are considered as probabilities, where a good model would be consistent with the observations. By incorporation of the Rice distribution, these probabilities include the errors in both the model and observations. The consistency between the model and the observations is measured statistically by considering the probability that the observations would have been made, given the current model. As the errors get smaller, the probabilities improve, and the likelihood increases.

2.7.4 Refinement using CNS

The CNS programs use maximum likelihood refinement methods [Brunger *et al.*, 1998]. The restraints to be used are input and the weight for refinement can be given or determined. Several routines were used during refinement of the galactose oxidase structures.

2.7.4.1 Rigid body refinement

At the start of refinement, the model structure is optimally positioned in the unit cell using rigid body refinement. During this procedure, the whole molecule is defined as a rigid-body, in which atoms are unable to move with respect to one another. This results in six degrees of freedom during refinement, three rotation functions and three translation functions. All exogenous atoms were removed from the initial prior to rigid body refinement.

2.7.4.2 Positional refinement

Once the position of the whole molecule in the unit cell has been determined, the individual atom coordinates can be improved. Although this is partially carried out by manual model building, energy minimisation plays an important role that is not feasibly accomplished by manual methods. The CNS minimisation procedure uses an energy function first proposed by Jack and Levitt [Jack and Levitt, 1978]. The X-ray term is minimised along with a potential energy function that includes terms for bond stretching, bond angle bending, torsion

potentials and van der Waals interaction. As exogenous molecules are added to the model, similar energy functions can be added to the CNS program for each addition molecule. Depending on the resolution of the data, these restraints can be stringently or minimally imposed. The refinement progress was monitored by the values of the R-values (R-factor, and R_{free}).

The thioether bond found in the mature form of galactose oxidase is a covalent link between two amino acids of the active site. The parameters for this bond are not present in the standard refinement dictionaries and must therefore be added prior to refinement. The expected length of the thioether bond has been reported as 1.8 Å in the original galactose oxidase structure. This distance corresponds well to the carbon–sulphur distances determined from small molecule crystallography. Thus a patch was placed in the input file for CNS refinement program files so that, if the thioether bond was present, the restraints were applied, whereas if the bond was not present, the two residues were treated as separate entities.

2.7.4.3 Temperature factor refinement

The protein structure factors are not only affected by the positions of atoms, the atomic vibrations also make a contribution. Thus, in order to obtain the best model, the temperature factors of the model must also be refined. The resolution limit of macromolecular diffraction restricts the number of observations per atom. At low resolution there may only be sufficient information to refine an overall B-factor for the whole molecule, whereas at higher resolutions (> 2.8 Å), the B-factor of side chains, or individual atoms may be refined. Once a reasonable model had been obtained by model building and positional refinement, temperature factor refinement was carried out in between cycles of positional refinement.

2.7.4.4 Simulated annealing

Positional refinement by minimisation methods works by minimising a residual function of observed and calculated amplitudes. If a false minimum is reached, however, it can be difficult to escape and enter the true, or global, minimum. In such cases the procedure of simulated annealing may be used to 'escape' the false minima. The model is allowed to move as if at high temperature, in hopes of lifting it out of local energy minima. Molecular dynamics routines assign each atom an initial velocity, determine the force acting on each atom based on the potential energy of the whole molecule, and calculate the resulting acceleration of each atom. Then the model is cooled slowly to find its preferred conformation at the temperature of diffraction data collection. The advantage of molecular dynamics

refinement is that the radius of convergence is greatly increased, compared to minimisation methods, and thus reduces the need for manual intervention [Brunger *et al.*, 1989].

2.8 Addition of solvent molecules to the model

A significant proportion, typically 50%, of protein crystal unit cells comprise of water molecules. Some of these water molecules are interacting with the protein itself and may show up in electron density maps as ordered molecules. In addition to the water molecules, other components of the mother liquor may be found incorporated into the crystal structure. In galactose oxidase crystals, acetate molecules are often found, as are metal ions. During model building and refinement, solvent derived molecules become apparent in the electron density maps and require incorporation into the protein coordinate file. Structures solved using molecular replacement methods did not contain water molecules in the initial model. Acetate and metal ions, therefore, were added after examining difference maps and considering hydrogen bonding geometry. Water molecules were added using CNS [Kleywegt and Brunger, 1996], being modelled into electron density peaks with values over 3 sigma, and within a distance of 2.6 – 4.0 Å of any atom and 2.0 – 3.2 Å of hydrogen bonding donor/acceptors.

Structural and refinement information about exogenous molecules (such as galactose, hydrogen peroxide, acetate, sulphate and 2-methylene-1,3,propandiol) and atoms were obtained from the Hetero-compound Information Centre Uppsala (HIC-Up) [Kleywegt, 1998]. Once the refinement procedure had been completed, the quality of the model was assessed using SFCHECK [CCP4, 1994] and PROCHECK [Laskowski *et al.*, 1993]. The structures were aligned with one another using LSQMAN [Kleywegt and Jones, 1997].

3 Structural studies of galactose oxidase catalysis

3.1 Introduction

Wild type galactose oxidase has been crystallised into two space groups, C2 and $P2_12_12_1$. In past experiments, exposure of C2 space group crystals to galactose has resulted in crystal dissolution [Firbank, 2002]. Investigation of the packing in the space group revealed a symmetry related molecule abuts the active site pocket, restricting access of galactose to the active site residues. The $P2_12_12_1$ space group crystal form appears to possess a crystal lattice which presents better access to the enzyme active site [Vinecombe, 1999]. It has been shown that galactose oxidase crystallised in the $P2_12_12_1$ space group and activated to the radical form can turn over galactose [Firbank, 2002]. This was shown using a coupled assay which measures the presence hydrogen peroxide produced during catalysis. One of the assay components is the compound 2,2'-azino-bis(ethylbenzthizoline-6-suphonic acid) (ABTS), which is oxidised by hydrogen peroxide, changing from yellow to green. This colour change can be monitored spectroscopically [Baron *et al.*, 1993].

Unfortunately, past attempts to obtain a galactose bound active site structure using $P2_12_12_1$ crystals have failed, furthermore, structures often revealed a loss of the active site copper [Ito, 1991, Vinecombe, 1999, Firbank, 2002]. It was suggested that the loss of copper in the active site may have been caused by chelation of the metal by the high concentration of PEG in the crystal incubation conditions [Firbank, 2002]. In addition to high PEG concentration, soaking experiments were often carried out using crystals that were not oxidised to ensure the presence of the catalytically active enzyme.

Computer modelling studies have suggested galactose makes several favourable contacts with active site residues of galactose oxidase, furthermore, apo-galactose oxidase has been oxidised to the radical form, which has persisted for several weeks in solution studies [Whittaker and Whittaker, 1989]. This chapter explores attempts to achieve active site coordination of galactose using oxidised galactose oxidase (in the absence of PEG) and the oxidised apo-enzyme. Although full oxidation of the substrate would not be expected to occur in the absence of copper, it was hoped that in the presence of excess substrate, strong substrate-enzyme interactions would allow visualisation of an active site bound complex.

Fully reduced galactose oxidase can be obtained in solution studies using substrate or chemical reduction under anaerobic conditions [Whittaker and Whittaker, 1993; Wright and Sykes, 2001b]. Crystallographic visualisation of the Cu^{1+} form of the enzyme has never been reported. Cu^{2+} has been successfully introduced to apo-galactose oxidase crystals, so it was thought that under anaerobic conditions introduction of Cu^{1+} into apo-galactose oxidase crystals could be successful. It is hoped that visualisation of a Cu^{1+} bound active site will provide an insight into the active site structure of fully reduced galactose oxidase and the proposed catalytic mechanism.

Directed evolution experiments identified the C383S mutation as an important alteration which increased catalytic activity towards galactose [Delegrave *et al.*, 2001, Sun *et al*, 2001]. This increased activity was mainly due to a dramatic reduction in K_M, reported in later and more detailed studies [Wilkinson *et al* 2004]. This chapter presents the structure of a C383S mutant and discusses the changes that might be responsible for the dramatic reduction enzymatic K_M (Table 1.2). Residue 383 resides 9 Å from the active site copper (Figure 1.14) and does not interact with any of the active site residues predicted to be important to substrate binding. Investigation of a substrate bound active site could provide valuable insight into how the mutation affects substrate binding. C383S was crystallised in space group C2, however, as mentioned earlier the use of galactose in a substrate soak was likely to cause crystal degradation, so a smaller substrate was selected. 2-methylene-1,3-propandiol is a better substrate than galactose and is analogous to dihydroxyacetone, the best known substrate for the enzyme [Kosman, 1984]. Galactose oxidase shows 3.8 and 1.3 times more activity against dihydroxyacetone and 2-methylene-1,3-propandiol, respectively compared with galactose. 2-methylene-1,3-propandiol was chosen as the substrate for crystal soaks as it was hoped it would be small enough to enter the enzyme active site within the C2 crystal packing. Furthermore, due to the slower catalytic turn-over of 2-methylene-1,3-propandiol, compared to dihydroxyacetone (Figure 3.1, Figure 3.2), it was thought to be a better candidate to allow visualisation of the substrate bound enzyme.

Figure 3.1: Dihydroxyacetone

$(OH)CH_2 - C\ O - CH_2(OH)$

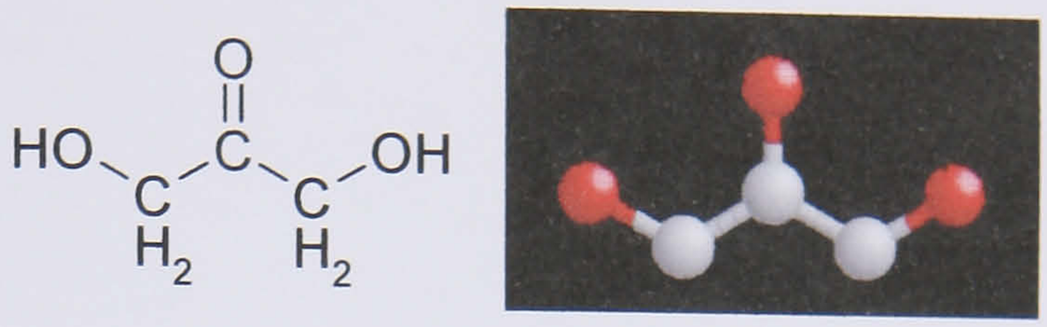
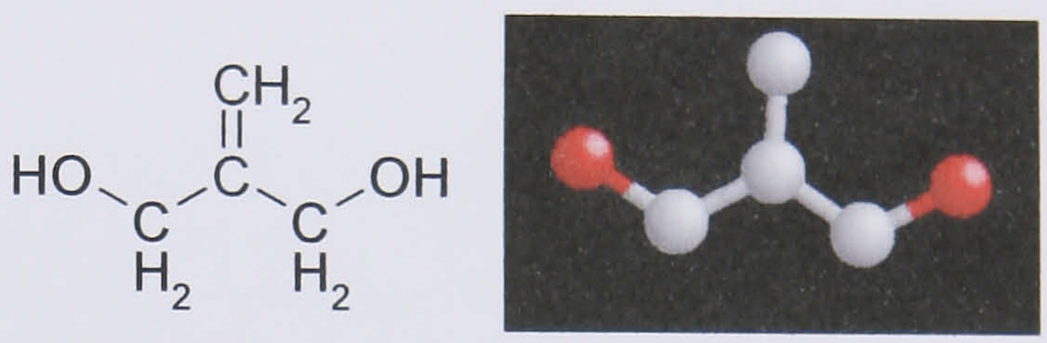


Figure 3.2: 2-methylene-1, 3-propandiol

$(OH)CH_2 - C(CH_2) - CH_2(OH)$



3.2 Results

3.2.1 Apo-galactose oxidase crystallisation

Galactose oxidase was expressed and purified from *Aspergillus nidulans* as described in chapter 2. The protein was purified to a high quality, as judged by SDS-PAGE analysis (Figure 3.3 and Figure 3.4). 151.9 mg of protein was purified from 3 litres of culture medium. Copper was removed prior to crystallisation as described in the standard conditions for copper-free galactose oxidase crystallisation in material and methods section (Chapter two, section 2.2.1).

Galactose oxidase crystals grew after five days and reached a maximum size after three weeks. Previous observations of copper-free galactose oxidase crystals have described their morphology as hexagonal rods [Firbank, 2002; Vinecombe, 1999]. Examination of the crystals confirmed the presence of hexagonal rods (Figure 3.5) as well as rhomboid tipped rods (Figure 3.6). Crystals grew to up lengths of 1.3 mm and were reasonably robust to handling.

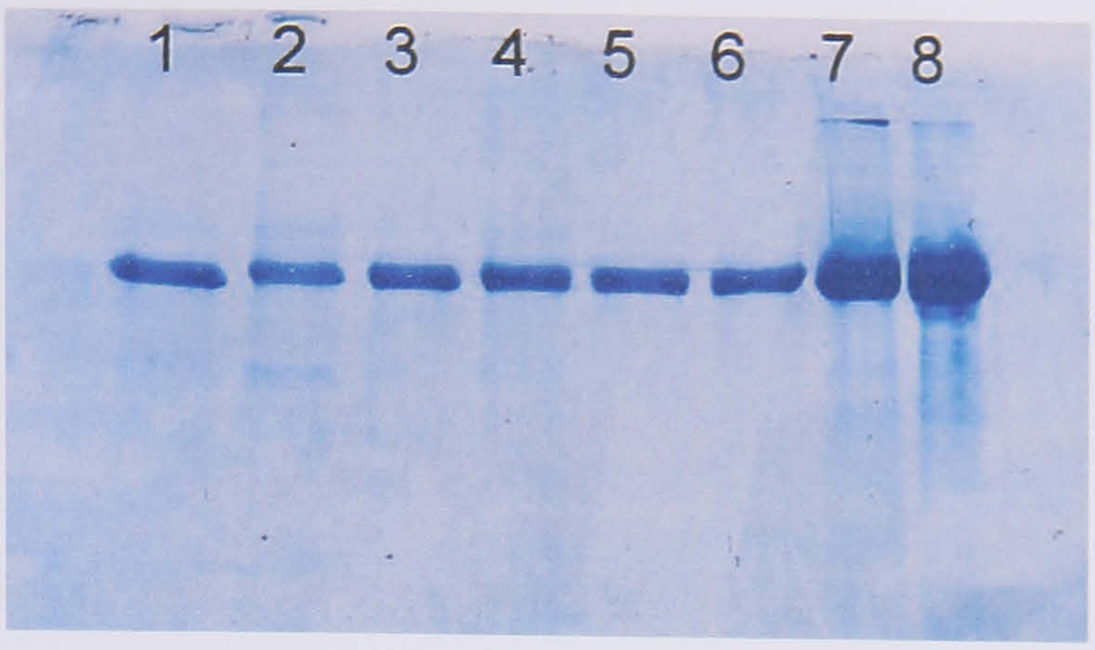


Figure 3.3: 12.5% SDS-PAGE of wild type galactose oxidase expressed from *Aspergillus nidulans*.
Lanes 1-6 represent a sample from each of the six 0.5 litre cultures. Lane 7 is a supernatant sample taken after ammonium sulphate precipitation step, lane 8 is the pooled sample after cation exchange chromatography. Each well contained 7 µl of sample.

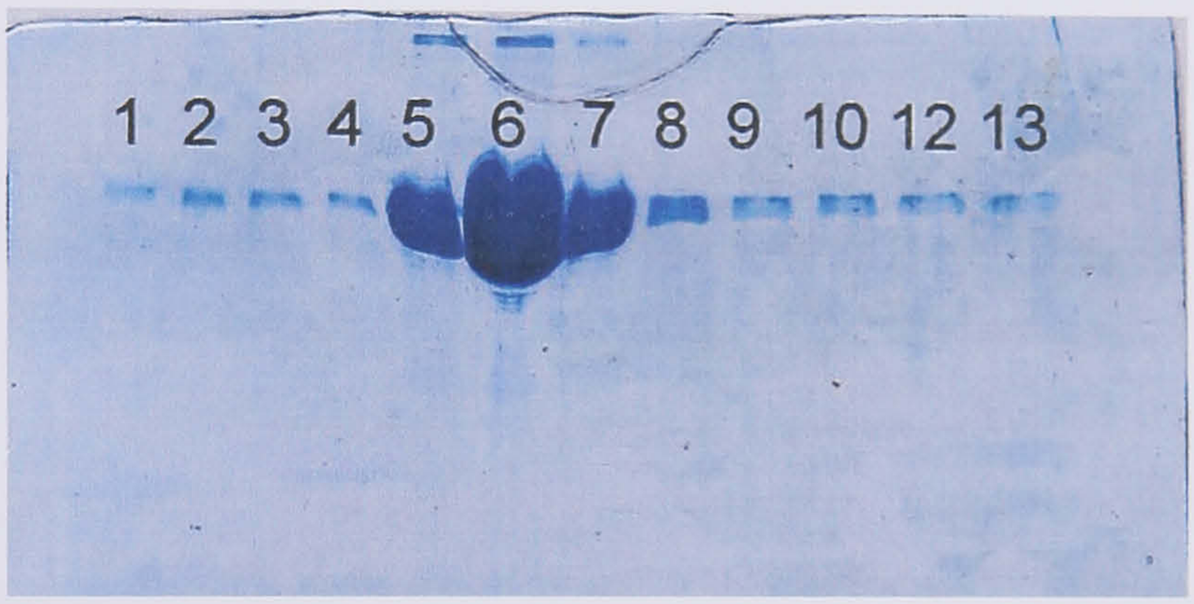


Figure 3.4: 12.5% SDS-PAGE gel displaying elution fractions from the cation exchange column.
Each well contained 7.5 µl of sample.

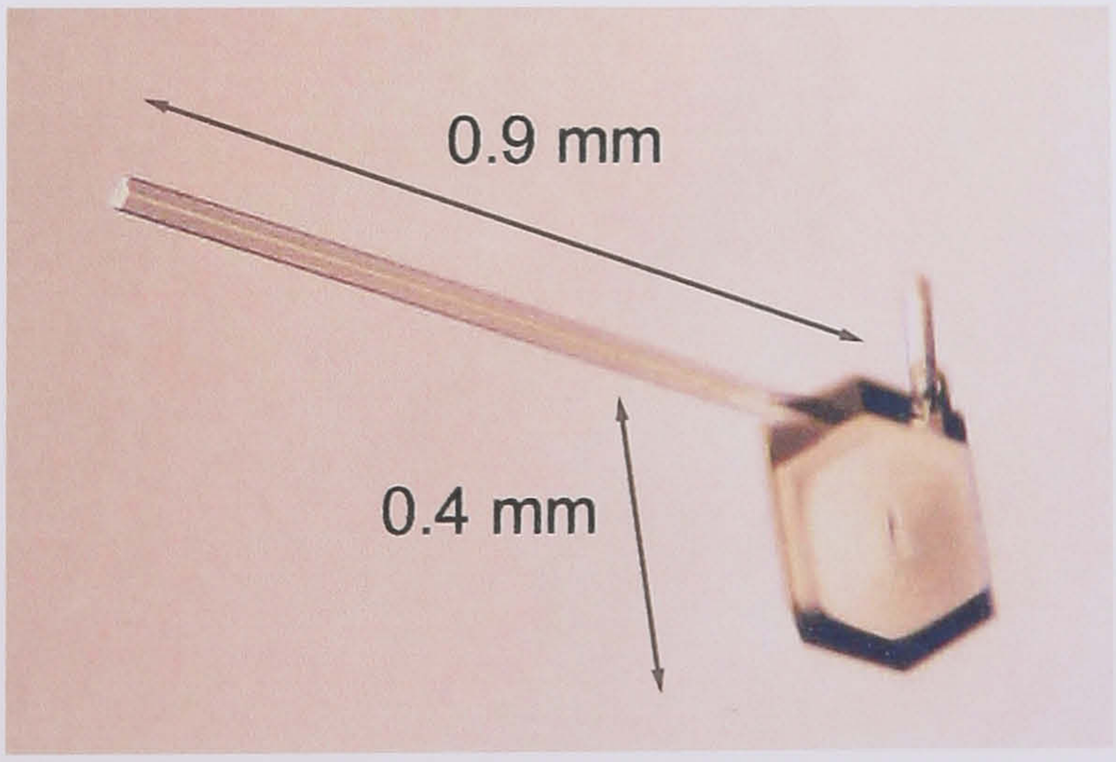


Figure 3.5: Copper-free wild type galactose oxidase hexagonal rod crystal.
The protein crystallised as rods, with hexagonal tips. The crystal on the right (0.4 mm) is growing perpendicular to the bottom of the well, while the crystal on the left (0.9 mm) is parallel to the bottom of the well.

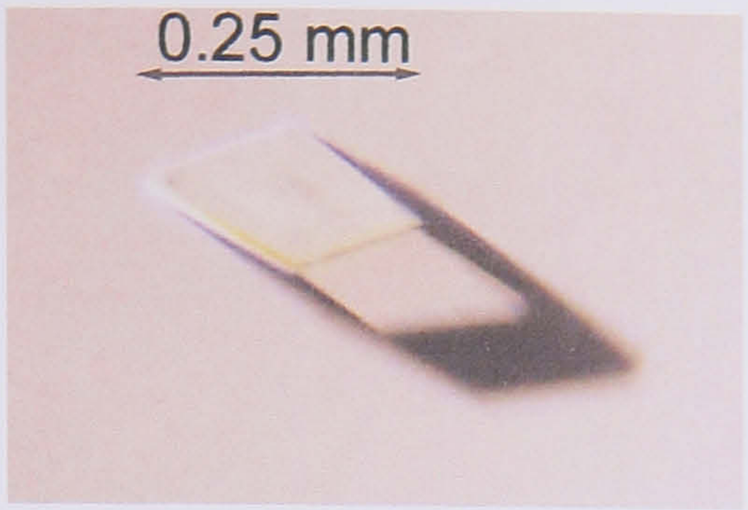


Figure 3.6: Copper-free wild type galactose oxidase rhomboid rod crystal.

The crystal is growing perpendicular to the bottom of the well, and has a rhomboid shaped tip.

3.2.2 Galactose soaks into apo-galactose oxidase crystals

A crystal was incubated in mother liquor (16% PEG 8000, 200 mM calcium acetate and 100 mM MES pH 6.1) comprising 50 mM potassium ferricyanide for five minutes, prior to transfer into mother liquor with 100 mM galactose. After 3 minutes the crystal was placed in cryoprotectant comprising mother liquor and 30% PEG 400, then flash frozen in liquid nitrogen. Data were collected using a laboratory X-ray source as described in the material and methods chapter (Section 2.3). The crystal diffracted to 2.3 Å resolution (Table 3.1), the data were processed using the HKL suite and the structure determined using the standard procedures described in methods section [Ontwinoski and Minor, 1997]. The phases for structure determination were obtained from the $P2_12_12_1$ wild type galactose oxidase structure provided by SJ Firbank [2002]. All non-protein atoms were removed prior to rigid body refinement.

Table 3.1: Data processing and refinement statistics for apo-galactose oxidase and oxidised wild type galactose

Apo-galactose oxidase and oxidised wild type galactose data processing and refinement statistics.		
X-Ray facility	R-AXIS IV++, University of Leeds, UK.	Beamline ID14-4, ESRF, Grenoble, France.
Crystal soak	Apo-galactose oxidase 100 mM galactose soak	3 minute oxidised wild type galactose soak
Cryoprotectant	Mother liquor, 30% PEG 400	Mother liquor, 25% PEG 400
Wavelength (Å)	1.54	0.98
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions	**a**= 59.3 Å, **b**= 88.5 Å, **c**= 133.4 Å **α=β=γ**=90 °	**a**= 59.2 Å, **b**= 88.8 Å, **c**= 132.9 Å **α=β=γ**=90 °
Total no. of reflections	142517	1066959
No. of unique reflections	23186	106042
Resolution limits (Å)	50 – 2.3	100 – 1.7
R_{sym} (%)	5.3 (25.4)	8.0 (22.2)
Completeness of Data (%)	82.2 (67.4)	88.0 (76.2)
Mean I/σI	7.6	9.9
Refinement statistics		
R_{cryst} (%)	21.0	20.4
R_{free} (%)	25.0	22.3
Number of atoms	5220	5230
Number protein of atoms	4839	4830
Number of water atoms	374	382
Average overall B-factor ($Å^2$)	19.2	13.3
Rms bond lengths (Å)	0.004	0.011
Rms bond angles (°)	1.1	1.6

Figures in the parentheses are from the highest resolution shell

Examination of the density maps revealed that neither copper nor galactose were present in the active site (Figure 3.7). Three water molecules are present at the active site, forming hydrogen bonds with Trp 290, Arg 330 and Tyr 272. The protein structure was essentially the same as the reported apo-galactose oxidase structure (PDB code 1GOH) [Ito *et al.*, 1995], except for the presence of a galactose molecule located at the reported domain I sugar binding site (Figure 3.8). Favourable hydrogen bonds are made with Nη1 and Nη2 of Arg 73 (2.8 Å and 3.1 Å), and the Nε2 of His 40 (3.3 Å). Calcium was also bound to domain I, at the previously reported binding site loop (occupied by sodium in original structure), coordinated by six oxygen atoms contributed from a surface loop comprising residues 29 to 37 (Figure 3.9).

Figure 3.7: The active site of apo-galactose oxidase incubated with 100 mM galactose.

No galactose was present in the active site, however, three water molecules were identified hydrogen bonding with residues that are proposed to interact with sugar substrates. Potential hydrogen bond interactions are represented by red lines.

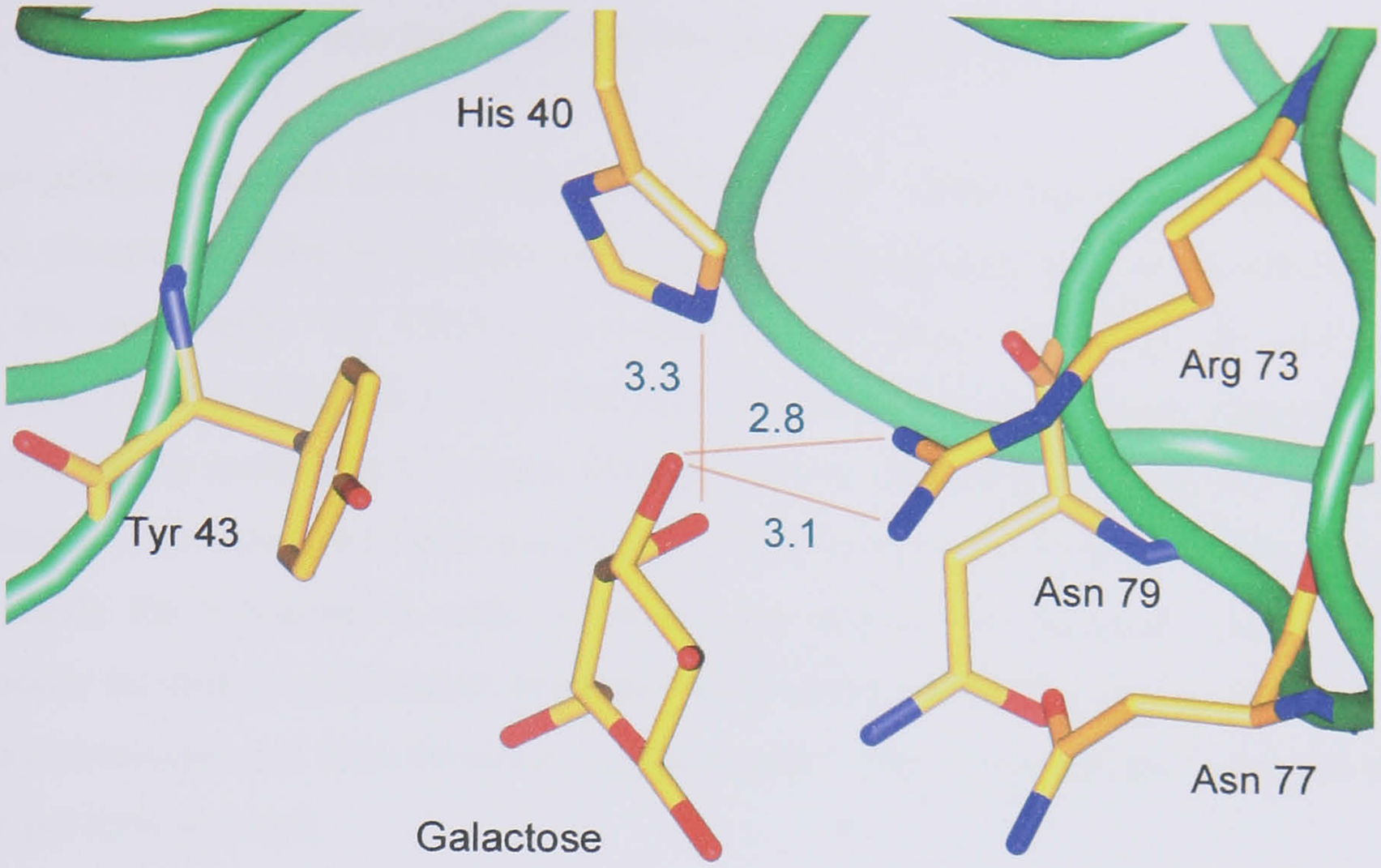


Figure 3.8: Domain I galactose binding.

Galactose was found binding at the domain I sugar binding site. Possible hydrogen bond interactions are represented by red lines.

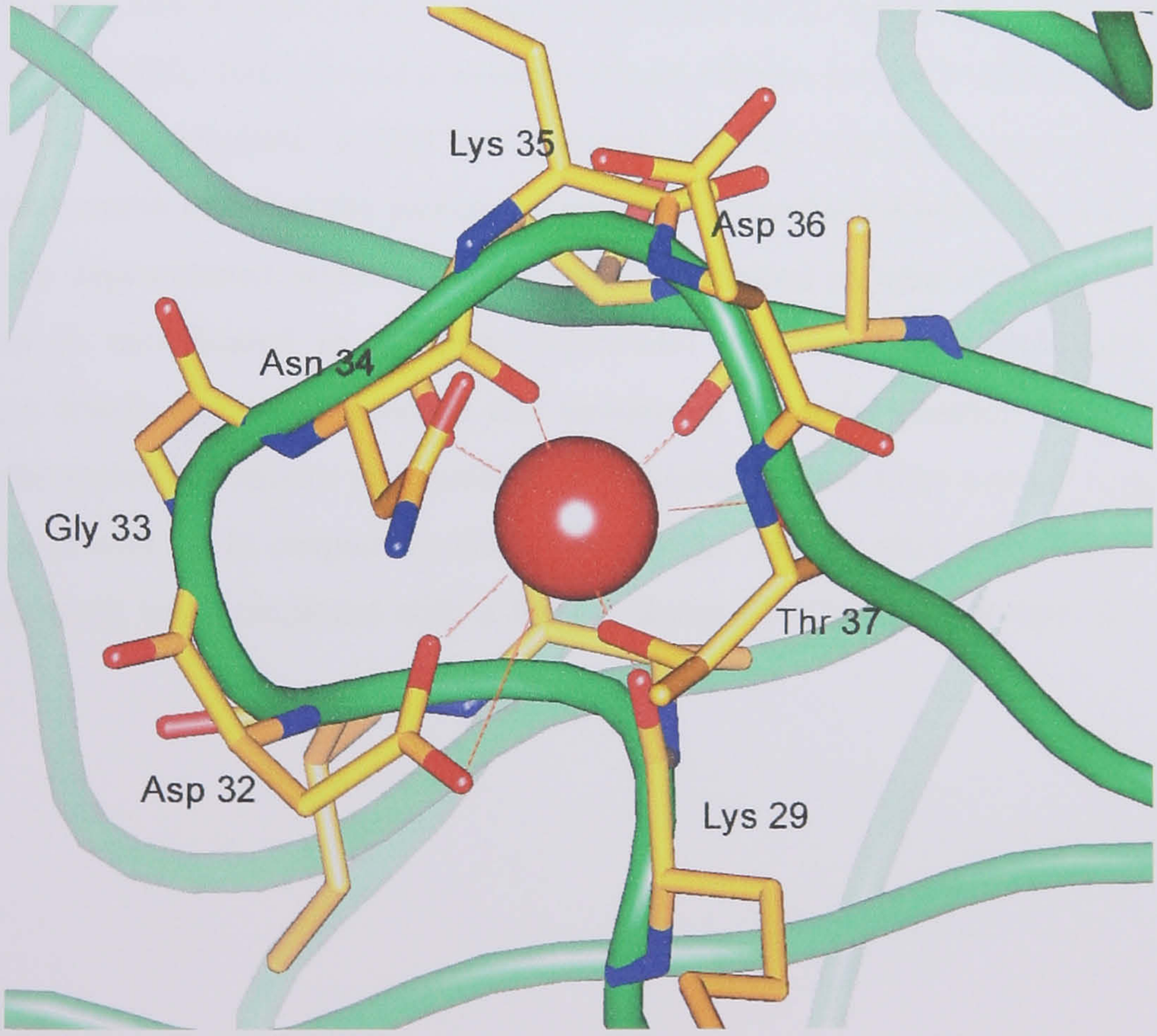


Figure 3.9: Domain I calcium binding.

A calcium ion was modelled into domain I, at the binding site reported in native structure 1GOF [Ito *et al.*, 1991].

3.2.3 Substrate soak into ferricyanide treated galactose oxidase

An apo-galactose oxidase crystal was transferred from the copper-free crystallisation buffer to sodium phosphate buffer by progressively diluting PEG solutions starting at 16% PEG, then 10%, 5% and finally 0% PEG. The crystal mother liquor comprised 50 mM sodium phosphate, 100 mM MES pH 5.7 and 200 mM calcium acetate. Copper was reconstituted into the active site by incubating the crystal in mother liquor containing 20 mM copper acetate for 20 minutes. The crystal was then transferred to mother liquor containing 50 mM potassium ferricyanide for 5 minutes in order to ensure the enzyme was oxidised. The crystal was aerobically incubated with 50 mM galactose for 10 minutes in mother liquor, prior to transfer into cryoprotectant and flash-frozen in liquid nitrogen. The cryoprotectant contained mother liquor and 25% glycerol.

The crystal diffracted to 1.4 Å initially, data were collected on beamline ID14-4, ESRF, Grenoble, France and were 88% complete at 1.7 Å resolution (Table 3.1). The active site density did not indicate galactose binding although the sugar was again identified at the domain I binding site (Figure 3.10). Copper was present in the active site, coordinated by Tyr 272 and 495, His 496, His 581 and a water molecule (Figure 3.11). Overall this structure was very similar to the substrate soaked apo-galactose oxidase structure except the N-terminal region of the protein revealed the presence a metal ion coordinated with the terminal nitrogen of Ala 1, the deprotonated amide nitrogen and the carbonyl oxygen of Ser 2 (Figure 3.12). The copper is coordinated in a square pyramidal geometry, with one water molecule occupying a fourth equatorial position and another in the axial position. This geometry is similar to the active site square pyramidal geometry and appears to be a novel copper binding site (Further discussion in chapter 6). No further copper binding sites were identified and the model refinement was completed with a final R-factor of 20% and a R-free of 22% (Table 3.1).

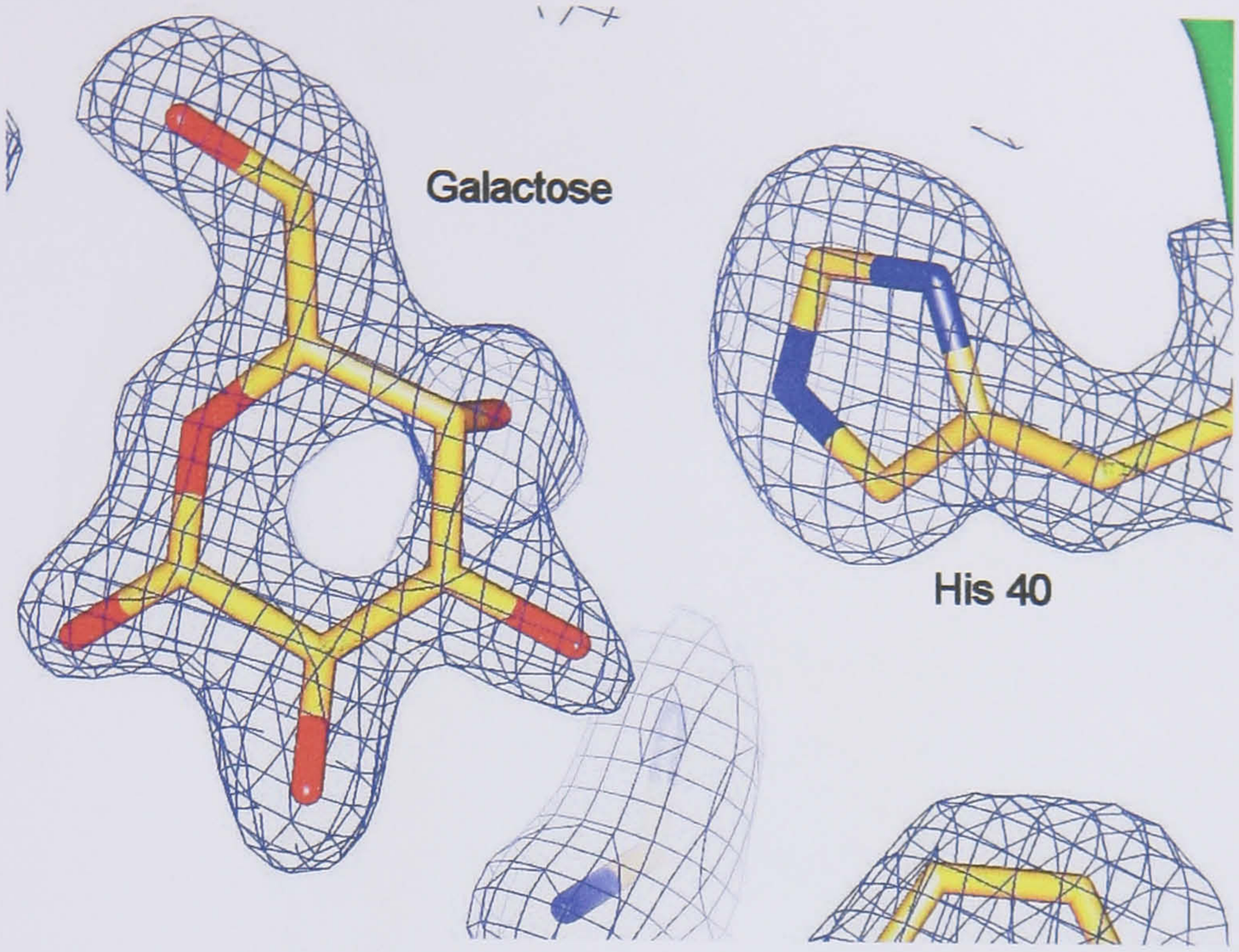


Figure 3.10: 2Fo-Fc map of galactose in domain I of Cu^{2+} reconstituted galactose oxidase

2Fo-Fc electron density map contoured at 1 x rms density, showing D-galactose binding to the domain I sugar binding site.

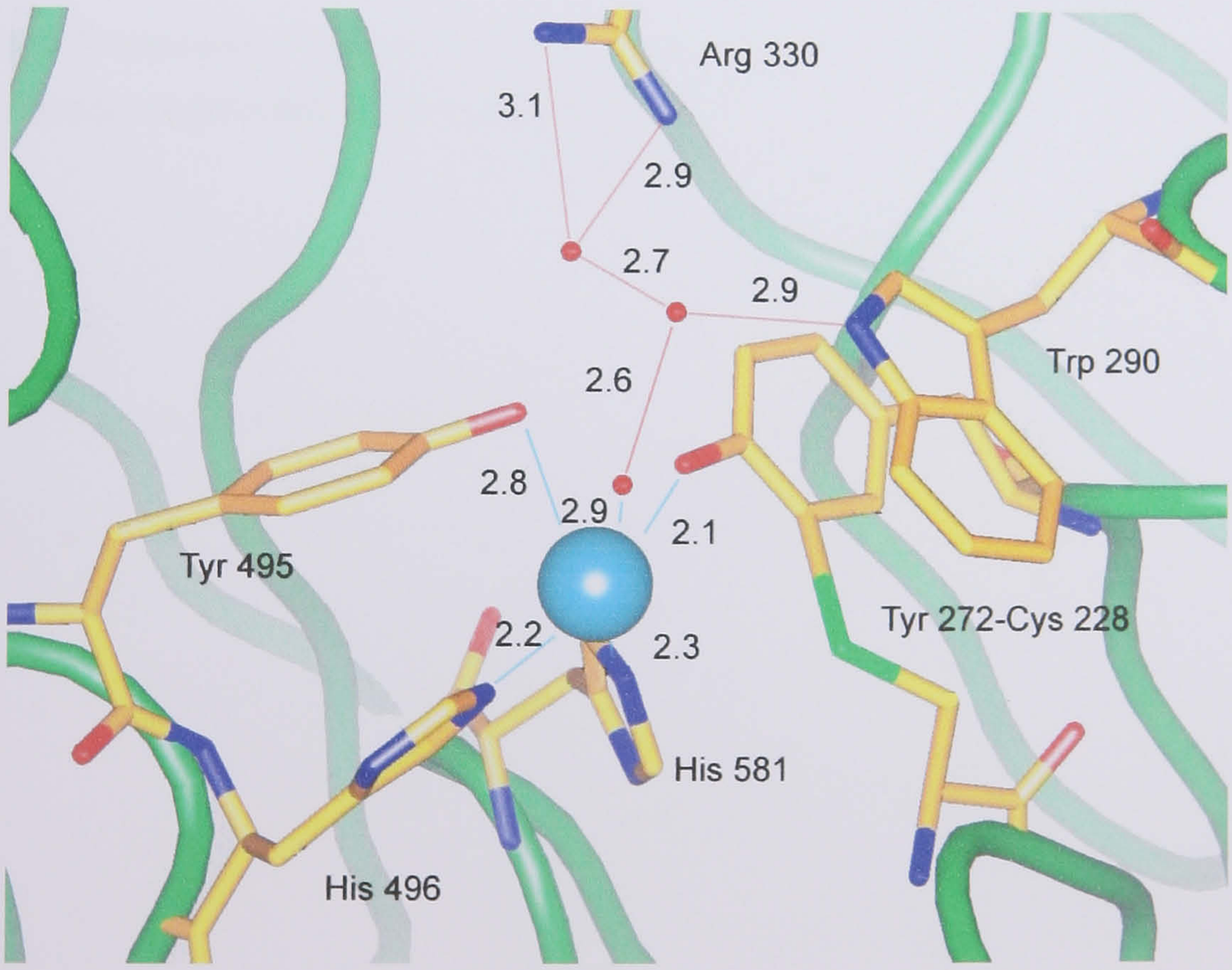


Figure 3.11: Active site of ferricyanide oxidised galactose oxidase.

Active site of ferricyanide oxidised galactose oxidase incubated with galactose for three minutes. No galactose was found bound to the active site. Copper was coordinating five ligands in a distorted square pyramidal geometry, the same geometry reported in native enzyme. Red lines represent hydrogen bond, blue lines a copper ligand bonds.

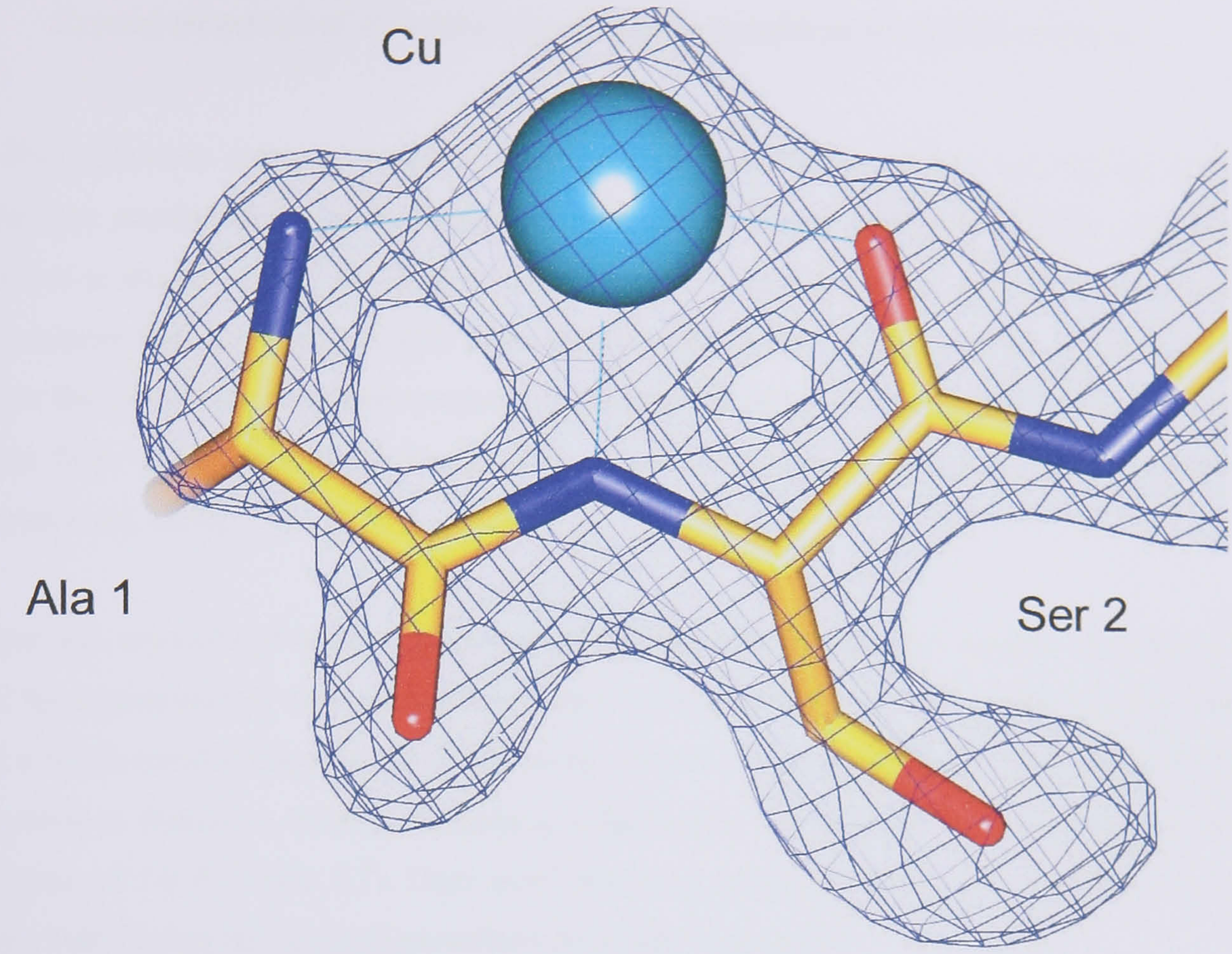


Figure 3.12: N-terminal Cu^{2+} binding after 3 minute soak of wild type galactose oxidase. Copper ligands are provided by three protein atoms.

3.3 Crystal structure of apo-galactose oxidase crystallised from BD1 construct

The BD1 galactose oxidase construct was expressed from *Pichia pastoris* and crystallised in copper-free conditions by S. J. Firbank, University of Leeds [2002], using the procedure described in chapter two. This form of the enzyme is does not possess the pro-sequence and under copper-free conditions it also lacks the thioether bond. In the presence of copper and oxygen the premature form (no copper, no thioether bond) autocatalytically converts to the mature form which essentially identical to galactose oxidase expressed from *A. nidulans* [Deacon *et al*, 2004].

Protein was crystallised under copper-free conditions and remained in aerobic conditions at 18°C for approximately two years prior to data collection. Prior to Cu^{1+} binding studies, data from a single aerobic copper-free crystal were collected using beamline 14.1, Darsebury SRS (Synchrotron Radiation Source, Daresbury laboratories, Cheshire, U.K) and diffracted to a resolution of 2.0 Å (Table 3.2). Data were processed using the HKL suite [Otwinowski and Minor, 1997] using the standard procedure described in chapter 2.

The 2Fo-Fc electron density map revealed the thioether bond was present and there was weak density at the copper binding site, adjacent to the hydroxyl oxygen of Tyr 272 (Figure 3.13). A water molecule was modelled into the density and the model refined well, with a final R-factor of 21%, R-free 24% (Table 3.2). It was a surprise to find the thioether bond present in this structure as copper-free conditions had been applied during the handling of the protein [S.J. Firbank, personal communication]. This raises the possibility of cofactor formation in the absence of copper. The copper binding ligands, residues 495, 496, 581 and 272 are all found within 2.6 Å of the proposed water molecule (Figure 3.14). Tyr 272 is 1.9 Å from the water, a distance more indicative of copper binding. The histidine residues, however, are at bond distances which suggest hydrogen bonding (2.3 Å and 2.5 Å). Attempts to model copper into the density were unsuccessful, even when the copper occupancy was reduced as low as 10%. As a crystal structure is a composite of all the molecules in the crystal population, there is the possibility that a small amount of contaminating copper has been present in the crystallisation conditions, which over time (two years in crystallisation buffer), has converted the enzyme in the crystal to the mature protein.

Table 3.2: Data processing and refinement statistics for wild type galactose copper soaks			
Crystal soak	Copper-free	5 minute 5 mM Cu^{1+}	25 minute 5 mM Cu^{1+}
X-ray facility	14.1 Daresbury SRS	14.2 Daresbury SRS	14.2 Daresbury SRS
Cryoprotectant	16% PEG 8000, 25%PEG 400	16% PEG 8000, 25%PEG 400, 1 mM EDTA	16% PEG 8000, 25% PEG 400, 1 mM EDTA
Wavelength (Å)	0.978	0.978	0.978
Space group	C2	C2	C2
Unit cell dimensions (Å)	**a**= 97.0, **b**= 88.8, **c**= 85.5, **β**= 117.4°	**a**= 97.1, **b**= 89.9, **c**= 85.4, **β**= 116.9°	**a**= 97.1, **b**= 88.9, **c**= 85.4, **β**= 116.8°
Total no. of reflections	152428	253402	194987
No. of unique reflections	42450	27213	17230
Resolution limits (Å)	32.0 – 2.0	29.0 – 2.4	29.0 – 2.5
R_{sym} (%)	4.3 (18.9)	6.7 (35.1)	8.2 (39.3)
Completeness of data (%)	96.4 (72.9)	89.4 (78.3)	76.7 (62.3)
Mean I/σI	21.9	16.6	14.2
Refinement statistics			
R_{cryst} (%)	21.0	24.4	19.0
R_{free} (%)	23.8	28.6	24.4
Number of atoms	5066	4974	4991
Number protein of atoms	4829	4829	4829
Number of water atoms	232	137	155
Average overall B-factor ($Å^2$)	33.42	45.03	36.69
Rms bond lengths (Å)	0.013	0.007	0.007
Rms bond angles (°)	1.9	1.4	1.5

Figures in parentheses indicate values in the highest resolution bin.

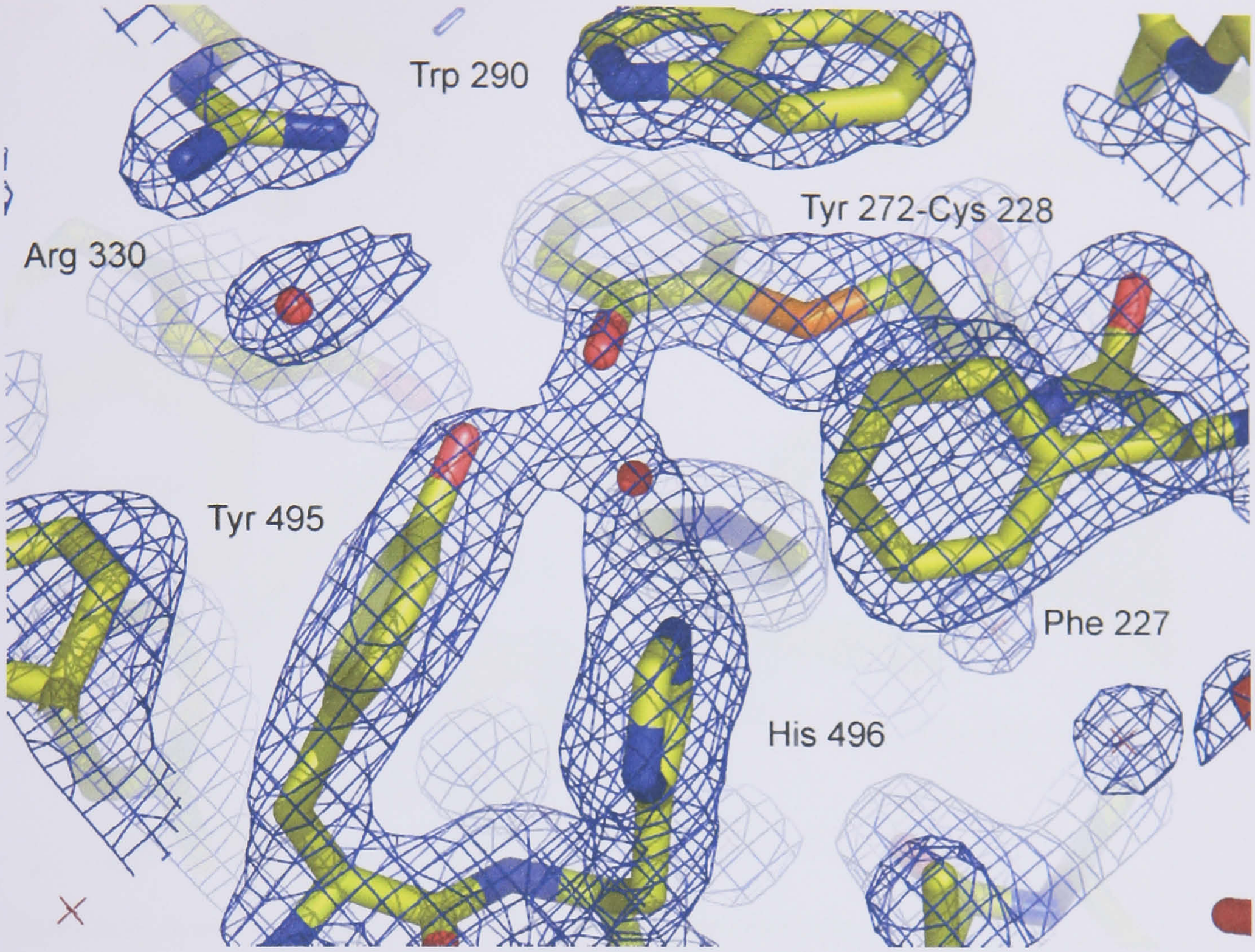


Figure 3.13: 2Fo-Fc electron density map of the of copper-free premature galactose oxidase active site.

The thioether bond is present and a water molecule has been modelled at the mature copper binding site. The density for the active site residues was good, and clearly showed no fourth exogenous equatorial ligand was present. Density map was contoured at 1 x rms density.

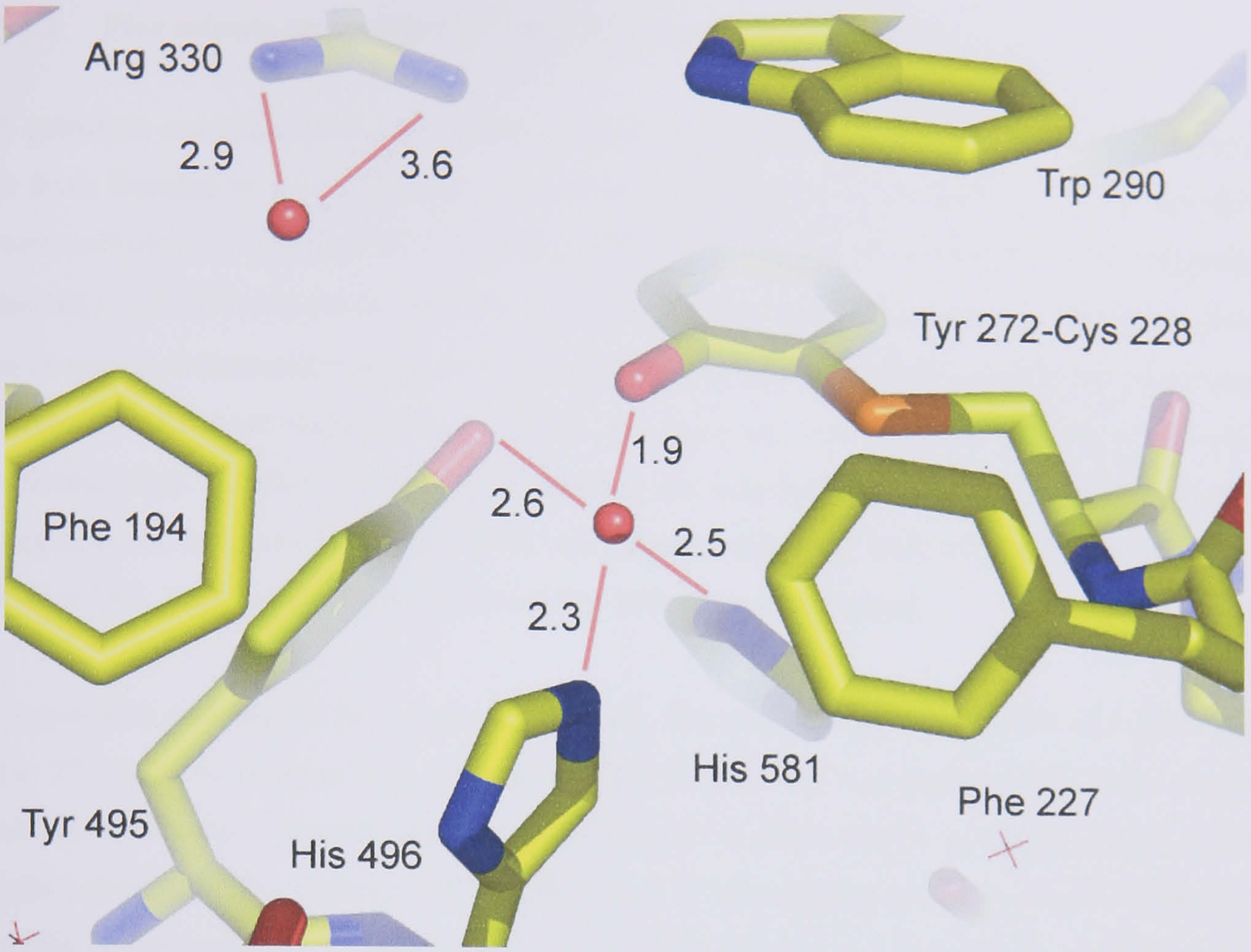


Figure 3.14: Active site of premature galactose oxidase stored aerobically.

Potential hydrogen bonds are represented by red lines. The thioether bond is present and a water molecule has been modelled at the mature copper binding site. Although Tyr 272 is only 1.9 Å away from the water, the two active site histidines and Tyr 495 possess bond lengths indicative of hydrogen bonding.

3.3.1 Anaerobic Cu^{1+} soak into galactose oxidase crystallised from the BD1 construct

Prior to introduction of Cu^{1+}, the crystals and all solutions were transferred to an anaerobic glovebox (Belle Technologies) for a minimum of 24 hours. Nitrogen gas was bubbled through all solutions for 10 minutes and the whole experiment was carried out in the anaerobic glove box. Before the crystal soaks were carried out, Cu^{1+} in the form of $Cu(I)(CH_3CN)_4PF_6$ (Copper (I) hexaflourophosphate) was dissolved in acetonitrile to a final concentration of 50 mM. Acetonitrile copper(I) hexaflourophosphate has been previously used successfully to incorporate Cu^{1+} into galactose oxidase in solution studies [Whittaker and Whittaker, 2003]. The crystal soaking solution was also used as cryoprotectant and contained 16% PEG 8000, 25% PEG 400, 100 mM MES pH 6.1, 200 mM calcium acetate, 1 mM EDTA and 5 mM Cu^{1+}. EDTA was added as a Cu^{2+} chelator, to prevent the interaction of Cu^{2+} with galactose oxidase, as shown in previously studies [Whittaker and Whittaker, 2003]. A crystal was incubated in the Cu^{1+} solution for five minutes and another for 25 minutes, quickly followed by submersion in liquid nitrogen, prior to data collection at Daresbury SRS, UK (Table 3.2).

3.3.2 Five minute anaerobic Cu^{1+} soak into apo-galactose oxidase

A galactose oxidase crystal was incubated anaerobically with 5 mM Cu^{1+} for 5 minutes prior to flash freezing in liquid nitrogen. The crystal diffracted to a resolution of 2.4 Å and data were collected on beamline 14.2, Daresbury SRS, U.K (Table 3.2). Data were processed using the HKL suite [Otwinowski and Minor, 1997] following the standard procedure described in chapter 2. Examination of the density maps revealed a strong density peak at the active site, indicative of copper binding (Figure 3.15). A copper ion was modelled into the active site, however, high B-factors suggested the binding site was only partially copper occupied. The copper occupancy was adjusted to 70% and refined fairly well, with a final R-factor of 24%, R-free 29%. The density for Tyr 272 and Trp 290 was not very good.

Copper was coordinated by two protein ligands, His 496 (1.9 Å) and His 581 (2.1 Å), while Tyr 272 was a weak ligand at a distance of 2.5 Å (Figure 3.15 and Figure 3.16). There was no additional density indicating the equatorial acetate or water ligand commonly found in the native structures. The hydroxyl of Tyr 495 is 3.5 Å from copper, so is not a copper ligand. A nearby water molecule is within hydrogen bonding distance of the hydroxyl oxygen of Tyr 495. A water molecule was located in a similar position in the copper-free galactose oxidase crystal described earlier (Figure 3.14), however, the water molecule in this structure is closer to Tyr 495. Density is observed between the copper ion and the thioether bond, perhaps indicating some disorder in this region. The copper ion is coordinated in a trigonal geometry, although the Tyr 272 coordination is weak. The copper site is only partially occupied, so it is difficult to assign the copper oxidation state unequivocally in this model. The presence of EDTA (a Cu^{2+} chelater) in the soaking conditions, the geometry and number of ligands, however, is suggestive of the Cu^{1+} oxidation state.

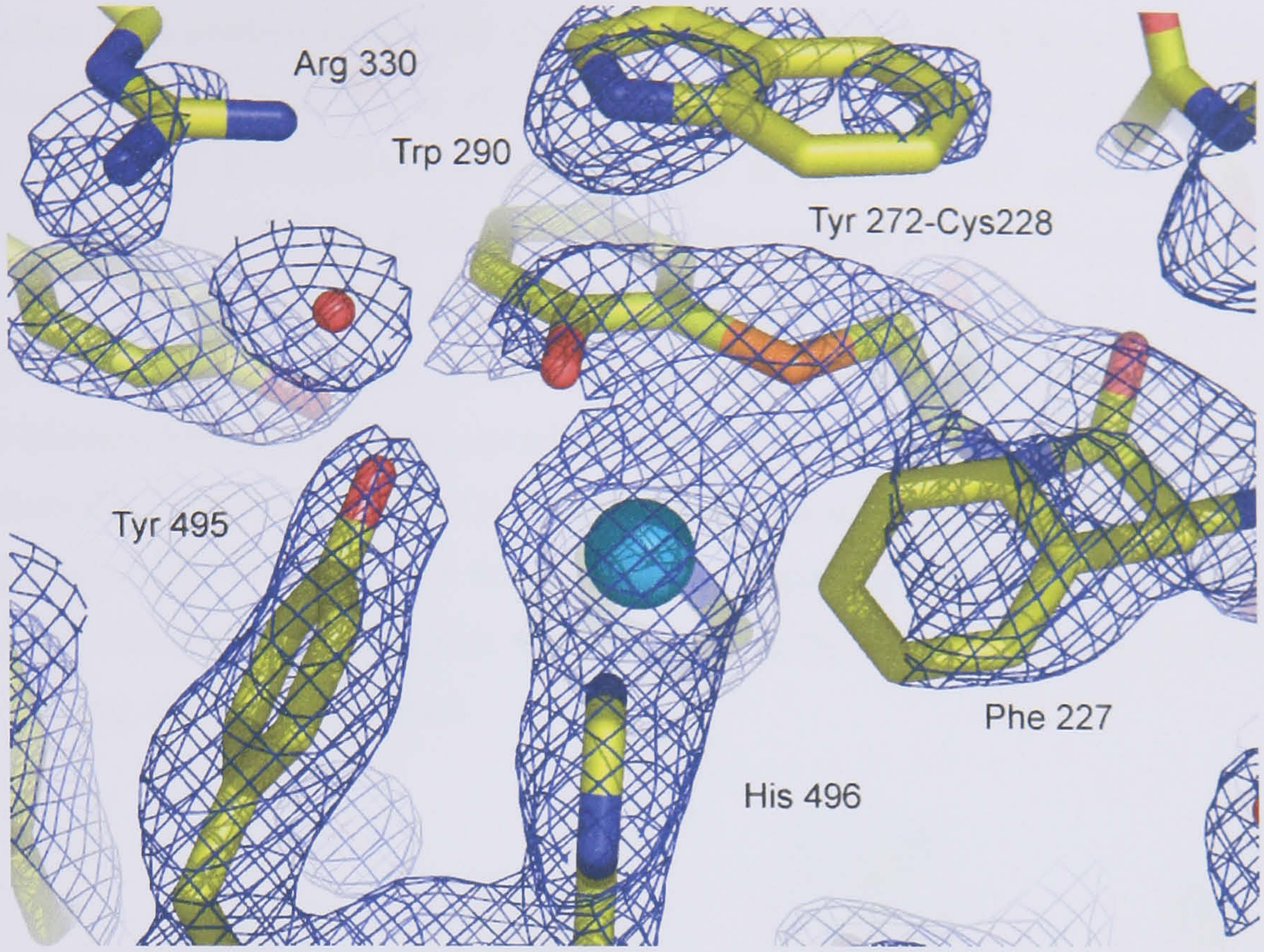


Figure 3.15: 2Fo-Fc electron density map of five minute anaerobic Cu^{1+} soak with premature galactose oxidase crystal.

Electron density at the active site indicated the presence of copper. The density for Tyr 495, His 496 and His 581, were very good. The Tyr 272 and Trp 290 density, however, were less well resolved.

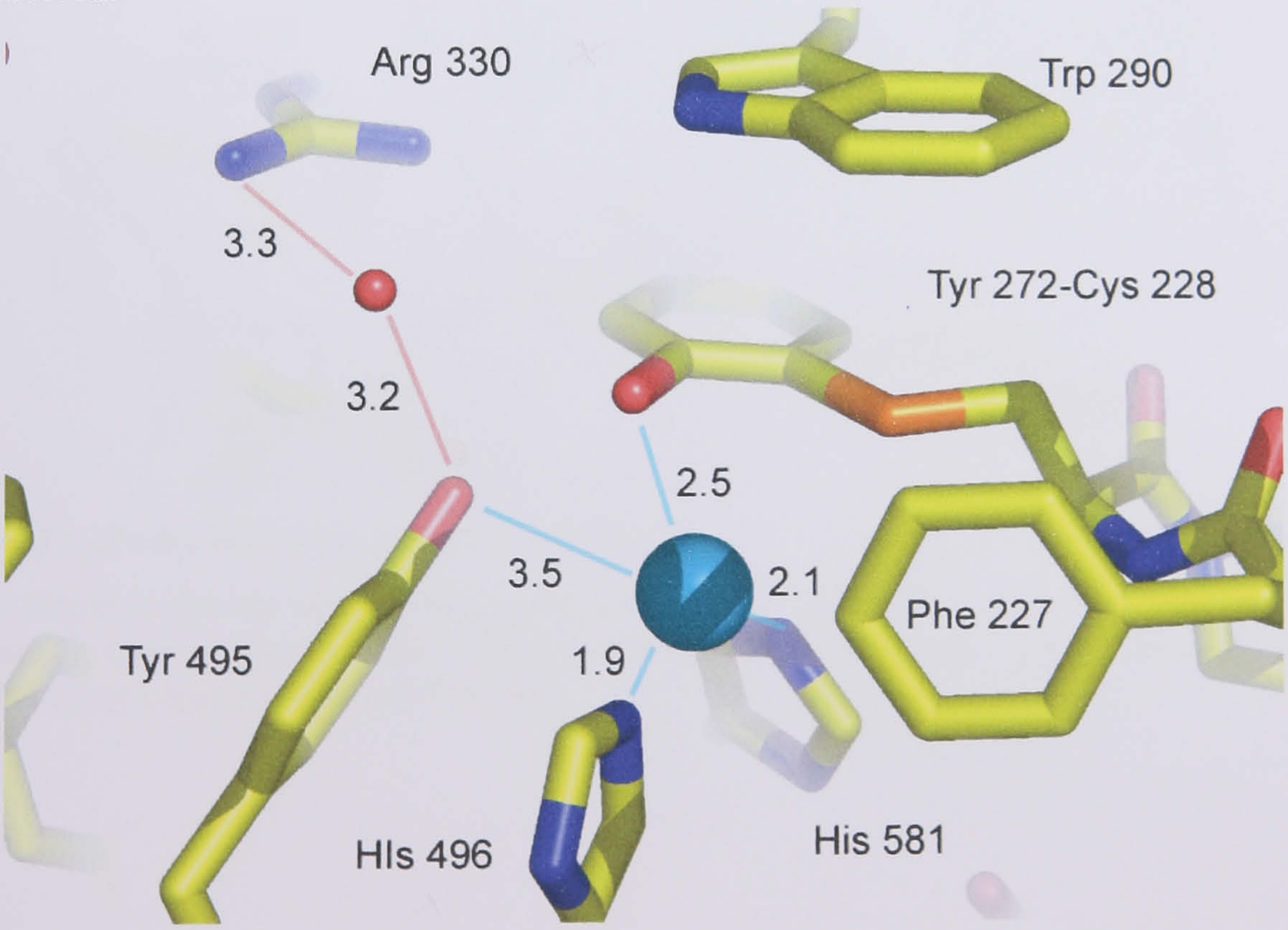


Figure 3.16: Active site of galactose oxidase incubated anaerobically for five minutes with Cu^{1+}.

Cyan lines represent copper distances and red lines are potential hydrogen bond interactions. Copper has strong bonds with His 496, His 581, and weak interaction with Tyr 272. Tyr 495 is not a copper ligand.

Two further large positive peaks were identified adjacent to Cys 383 (Figure 3.17). The peaks were found within 2 Å of the Sγ of Cys 383, but did not refine well as water molecules. Copper binding to this residue has been reported in previous Cu^{2+} soaking experiments [Firbank, 2002], therefore, the peaks may be partially occupied copper. Two copper ions were modelled into density either side of Cys 383 and the imidazole ring of His 334 was rotated so that Nδ1 coordinated one copper. Both metal ions refined well at 80% occupancy. There was a 2.7 Å distance between the to two metal ion sites, whilst one metal was 2.1 Å from His 334, 2.2 Å from the Sγ of Cys 383 and 3.6 Å from the sulphur of Met 382. Although not a ligand, the hydroxyl of Tyr 405 was 3.1 Å from one of the copper ions. Calcium was modelled into the domain I binding site and acetate was present in at the domain II binding site. No other copper binding sites were identified.

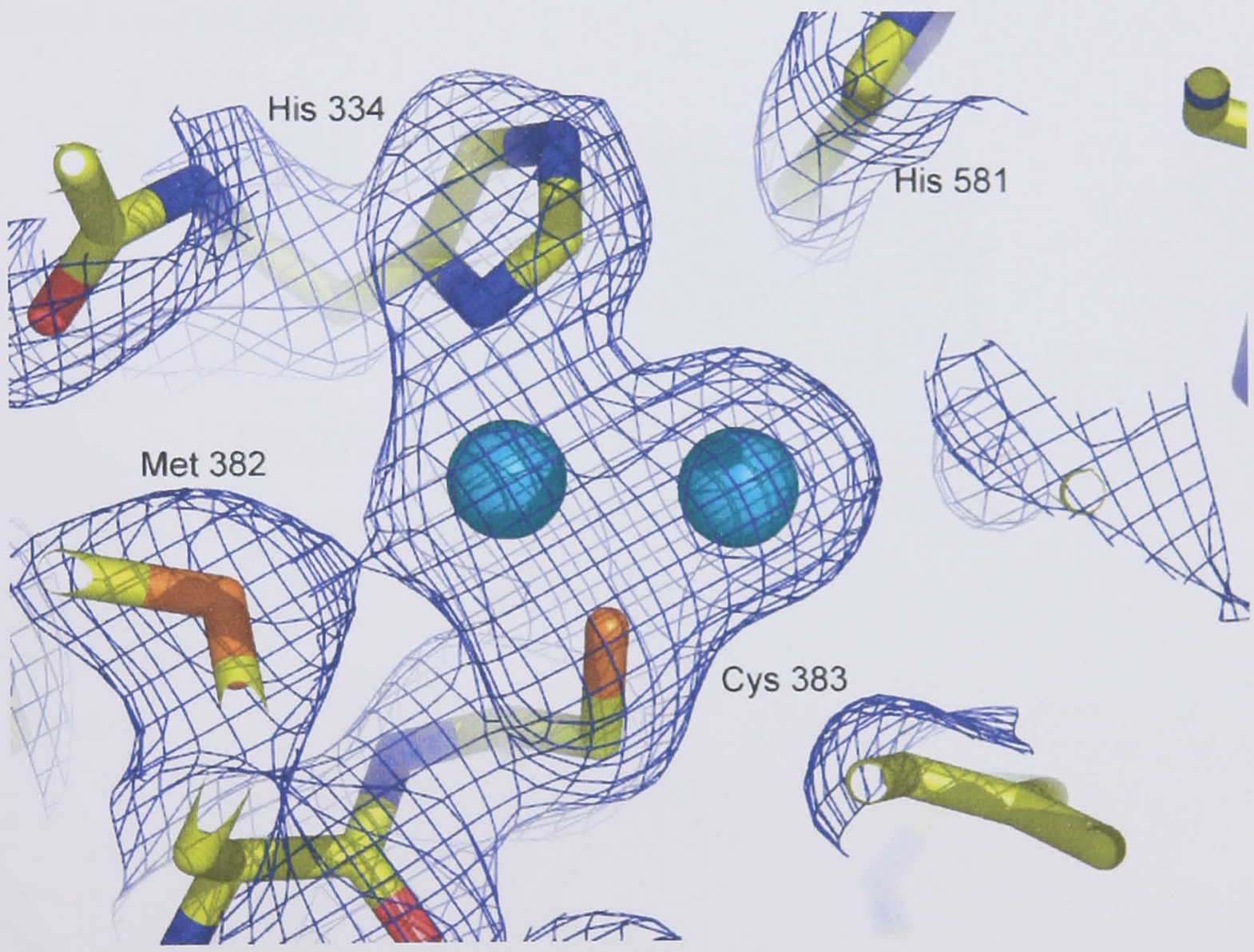


Figure 3.17: 2Fo-Fc map of the region around Cys 383.

2Fo-Fc electron density map of the region around Cys 383 after refinement, contoured at 1 x rms density. One copper density is close to the side chain atoms of Cys 228 met 382 and His 334. The other copper is close to Cys 383.

3.3.3 25 minute anaerobic Cu^{1+} soak into an apo-galactose oxidase crystal

An apo-galactose oxidase crystal was anaerobically incubated for 25 minutes in 5 mM Cu^{1+} solution prior to flash-freezing and data collection. The crystal diffracted to a resolution 2.5 Å and data were collected on station 14.2, Daresbury SRS, UK (Table 3.2). Data were processed using the HKL suite [Otwinowski and Minor, 1997] following the standard procedure described in chapter 2. The active site structure is very similar to the five minute copper soak structure, except no water molecules were within hydrogen bonding distance of Tyr 495 or Arg 330 (Figure 3.18). Copper is again located at the active site at 70% occupancy, forming strong bonds with His 496 (2.1 Å) and His 581 (2.0 Å) (Figure 3.19). Tyr 272 is a weaker ligand at a distance of 2.6 Å from the metal. Tyr 495 is 3.1 Å away from copper and is again not considered a copper ligand.

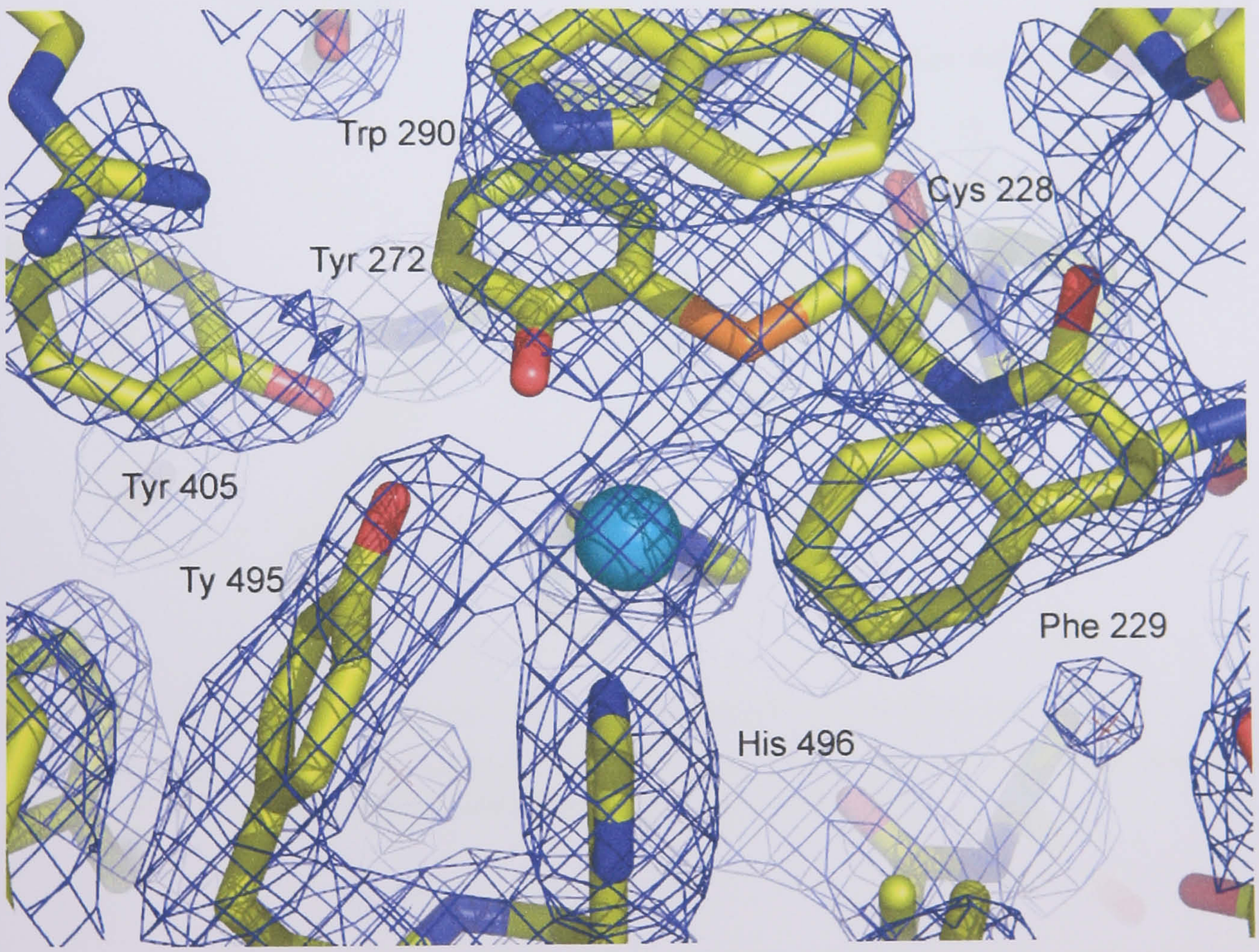


Figure 3.18: 2Fo-Fc electron density map of 25 minute anaerobic Cu^{1+} soak with premature galactose oxidase crystal.

Electron density at the active site indicated the presence of copper. The density for the active site residues was very good. Again there is density between the thioether bond and the copper ion.

An additional copper, also 70% occupied, was identified at the N-terminus of the structure. The copper was coordinated to the terminal nitrogen of Ala 1 (2.6 Å), the peptide nitrogen of Ser 2 (2.4 Å) and the carbonyl oxygen of Ser 2 (2.4 Å) (as in Figure 3.12). There were no other copper ligands, as the nearest water molecule is found 5.3 Å from the copper. Thus the N-terminal copper geometry can be considered as T-shaped, suggesting the N-terminal metal has also maintained the Cu^{1+} state. A calcium and acetate ion were identified and modelled into the known domain I and domain II binding sites respectively.

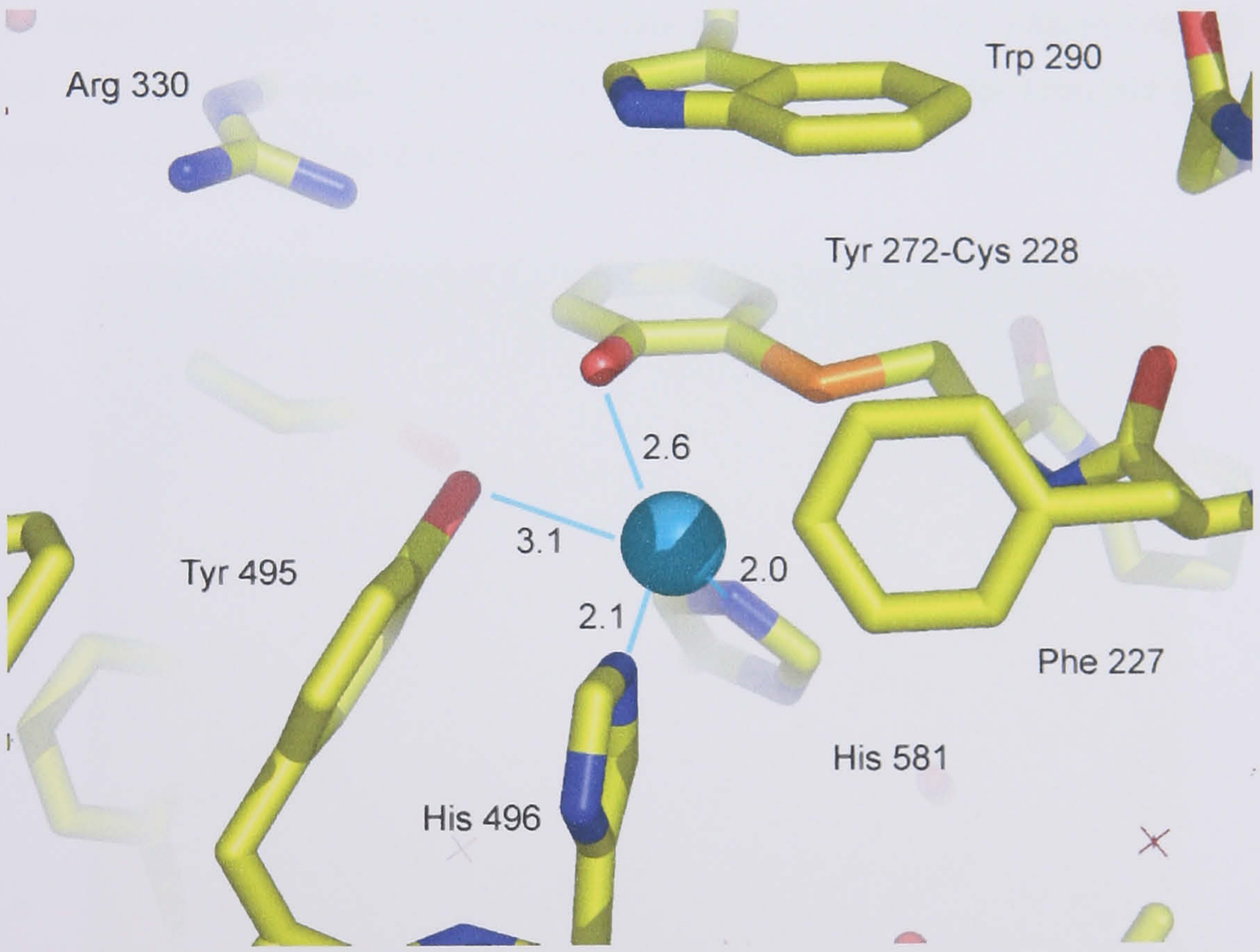


Figure 3.19: Active site of galactose oxidase incubated anaerobically for 25 minutes with Cu^{1+}.

Cyan lines represent copper distances. The active site copper forms strong bonds with His 581, His 496, but weaker interactions with and Tyr 272 and Tyr 495.

3.4 The crystal structure of C383S

The protein expression in *Pichia pastoris* and purification was carried out by D Wilkinson, University of Leeds, using as described in Chapter two [Wilkinson *et al.*, 2004]. The optimal crystallisation conditions were found to be in 0.7 M ammonium sulphate, 0.5 M sodium acetate pH 4.5, 5% PEG 4000. Rectangular plate shaped crystals were formed after one week (Figure 3.20), and grew to 0.4-0.7 mm after three or four weeks at 18ºC (Figure 3.21). A single crystal was transferred to cryoprotectant comprising mother liquor and 20% PEG 400, followed by flash freezing in liquid nitrogen. Data were collected at beam-line 14.1 (wavelength 1.488 Å), SRS (Synchrotron Radiation Source, Daresbury laboratories, Cheshire, U.K.). Data indexing confirmed the crystal was in the C2 space group and were isomorphous with the original galactose oxidase crystals [Ito *et al.*, 1990]. The structure for the native galactose oxidase (PDB code, 1GOF) with all non-protein molecules removed was used to obtain phases and as a starting point for structure determination.

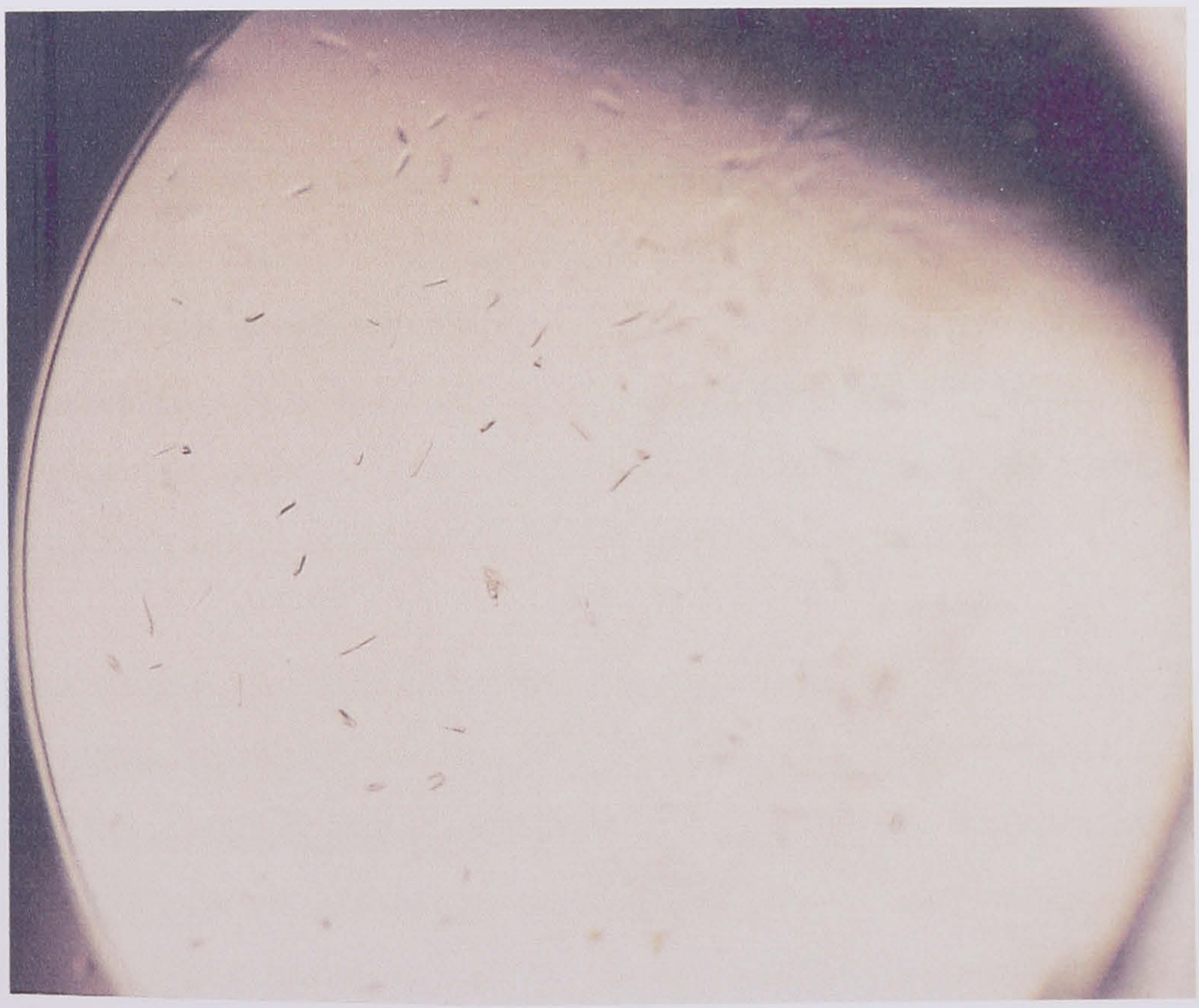


Figure 3.20: C383S crystals after one week

C383S crystallised in mother liquor containing 500 mM sodium acetate pH 4.5, and 700 mM ammonium sulphate, 5% PEG 4000. Small crystals appeared after one week.

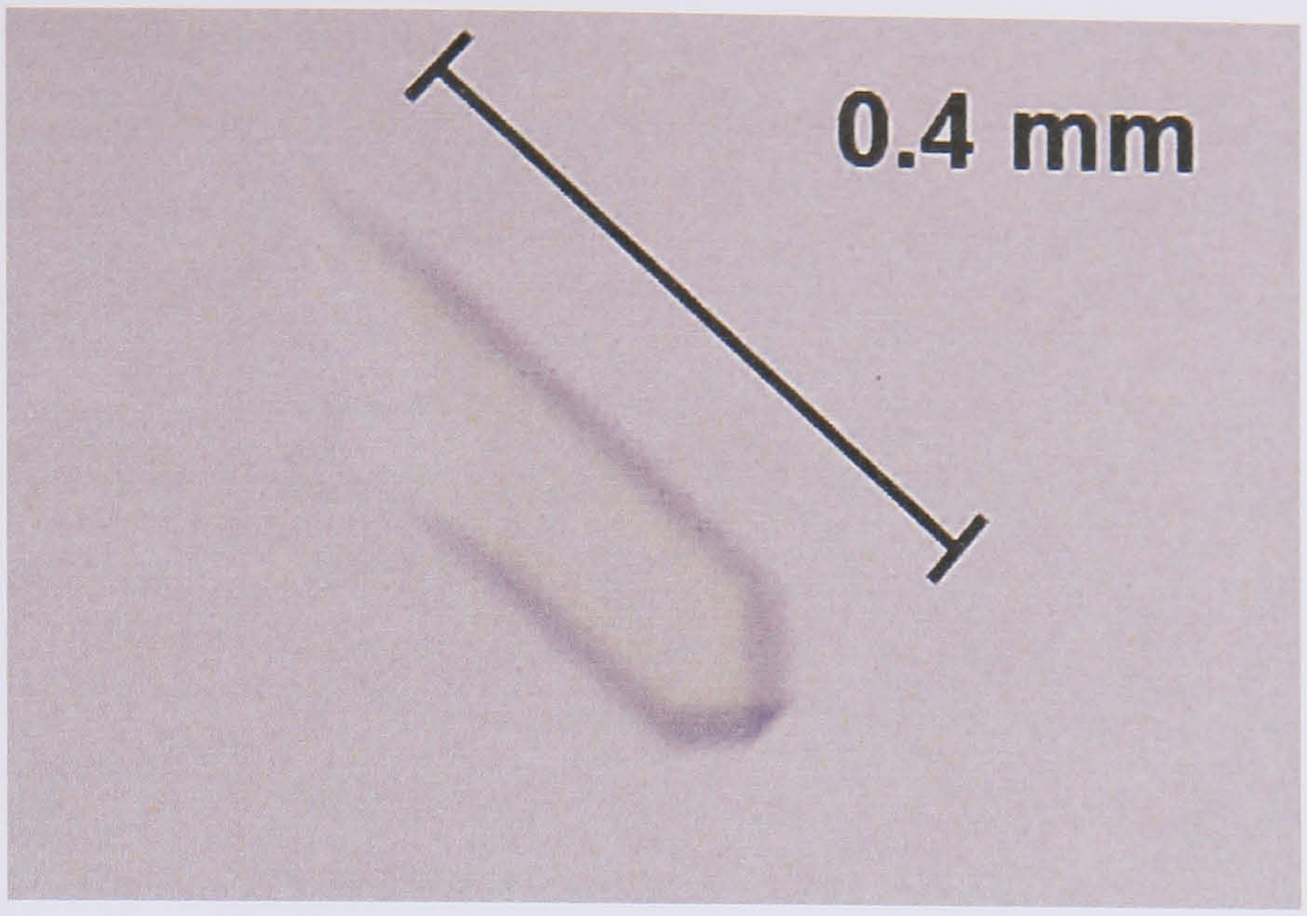


Figure 3.21: Single mutant C383S crystal.

The crystals were approximately 500 μm in length and 100 μm in diameter after four weeks. Typical crystals were plates with triangular tips.

The crystal diffracted to a resolution of 2.3 Å (Table 3.3). Initial electron density maps showed positive Fo-Fc density at the active site indicating that copper was present and the density at the 383 position was consistent with that expected for a serine residue, so a copper ion and serine residue were added to the model. The final refined C383S structure contains all 639 residues, plus 312 water molecules, three acetate ions, a copper and a sodium ion with a final R-factor of 22% and an R-free of 23%, as summarised in Table 3.3.

The active site reveals the presence of the thioether bond between Tyr 272 and Cys 228 (Figure 3.22), whilst copper is coordinated by His 496 (2.0 Å), His 581 (2.1 Å) and Tyr 272 (2.0 Å). Tyr 495 is a weak ligand at a distance of 2.7 Å. The four copper associated residues do not show significant movement compared to wild type. Examination of the active site revealed electron density peaks close to the fourth equatorial copper ligand position, which refined as a water molecule and an acetate ion. Inspection of the Fo-Fc difference density map and the high B-factors resulting from refinement of the two molecules suggested these additional molecules were not fully occupied in the crystal. Occupancies were therefore refined to 60% and 40% for water and acetate respectively.

Table 3.3: Data processing statistics for C383S mutant.

Data processing statistics for C383S mutant galactose oxidase and C383S crystals incubated with 2-methylene-1, 3-propandiol.			
	C383S (1T2X)	60 second soak	3 minute soak
X-ray facility	Station 14.1, Daresbury SRS	ID 14-4, ESRF, Grenoble	ID 14-4, ESRF, Grenoble
Crystallisation conditions	700 mM ammonium sulphate, 500 mM sodium acetate pH 4.5, 5% PEG 4000	700 mM ammonium sulphate, 500 mM sodium acetate pH 4.5, 5% PEG 4000	700 mM ammonium sulphate, 500 mM sodium acetate pH 4.5, 5% PEG 4000
Cryoprotectant	14% PEG 8000, 200 mM sodium acetate, 100 mM MES pH 6.3, 20% PEG 400	Mother liquor, 10% PEG 400, 5% 2-methylene-1, 3-propandiol	Mother liquor, 10% PEG 400, 5% 2-methylene-1, 3-propandiol
Wavelength (Å)	1.49	0.97	0.97
Space group	C2	C2	C2
Unit cell dimensions	**a**= 97.5 Å, **b**= 89.0 Å, **c**= 85.8 Å **β**=117.5 °	**a**= 96.9 Å, **b**= 88.8 Å, **c**= 85.4 Å **β**=117.5 °	**a**= 97.14 Å, **b**= 88.8 Å, **c**= 85.4 Å **β**= 117.7 °
Total no. of reflections	142517	448520	532576
No. of unique reflections	28426	48752	57069
Resolution limits (Å)	62 – 2.3 Å	100 – 1.9 Å	100 – 1.8
R_{sym} (%)	5.3 (25.4)	6.5 (21)	13.2 (34.2)
Completeness of data (%)	99.2 (95.8)	95.9 (76.1)	95.9 (73.4)
Mean I/σI	12.9	13.7	5.9
Refinement statistics			
R_{cryst} (%)	21.8	19.2	20.2
R_{free} (%)	22.8	22.3	22.3
Number of atoms	5157	5244	5364
Number protein of atoms	4829	4829	4829
Number of water atoms	312	383	504
Average overall B-factor ($Å^2$)	31.3	34.3	35.2
Rms bond lengths (Å)	0.012	0.012	0.009
Rms bond angles (°)	1.9	1.6	1.7

Figures in parentheses indicate values in the highest resolution bin.

The acetate and water molecule are only 1.9 Å apart, therefore, it is unlikely they are both present in the active site at the same time. The original galactose oxidase crystals were grown in a high concentration of acetate (over 0.5 M) [Ito, 1991] and the structure revealed acetate coordinated to the active site copper. Crystallisation without acetate, however, or at lower acetate concentrations often results in structures with a water molecule occupying the fourth equatorial copper ligand position [Ito *et al.*, 1994]. In the C383S structure both molecules appear to be present, although the attributed density for acetate suggests partial occupancy. This could be explained by the change of acetate concentration between the crystallisation conditions (0.5 M sodium acetate) and the cryoprotectant (0.2 M sodium acetate). A sodium ion and an additional acetate ion were located at previously reported binding sites (Domain I (Figure 1.6) and domain II (Figure 1.8) respectively) [Ito, 1991]. A third acetate ion was located on the surface of domain III, approximately 40 Å from the active site copper (Figure 3.23).

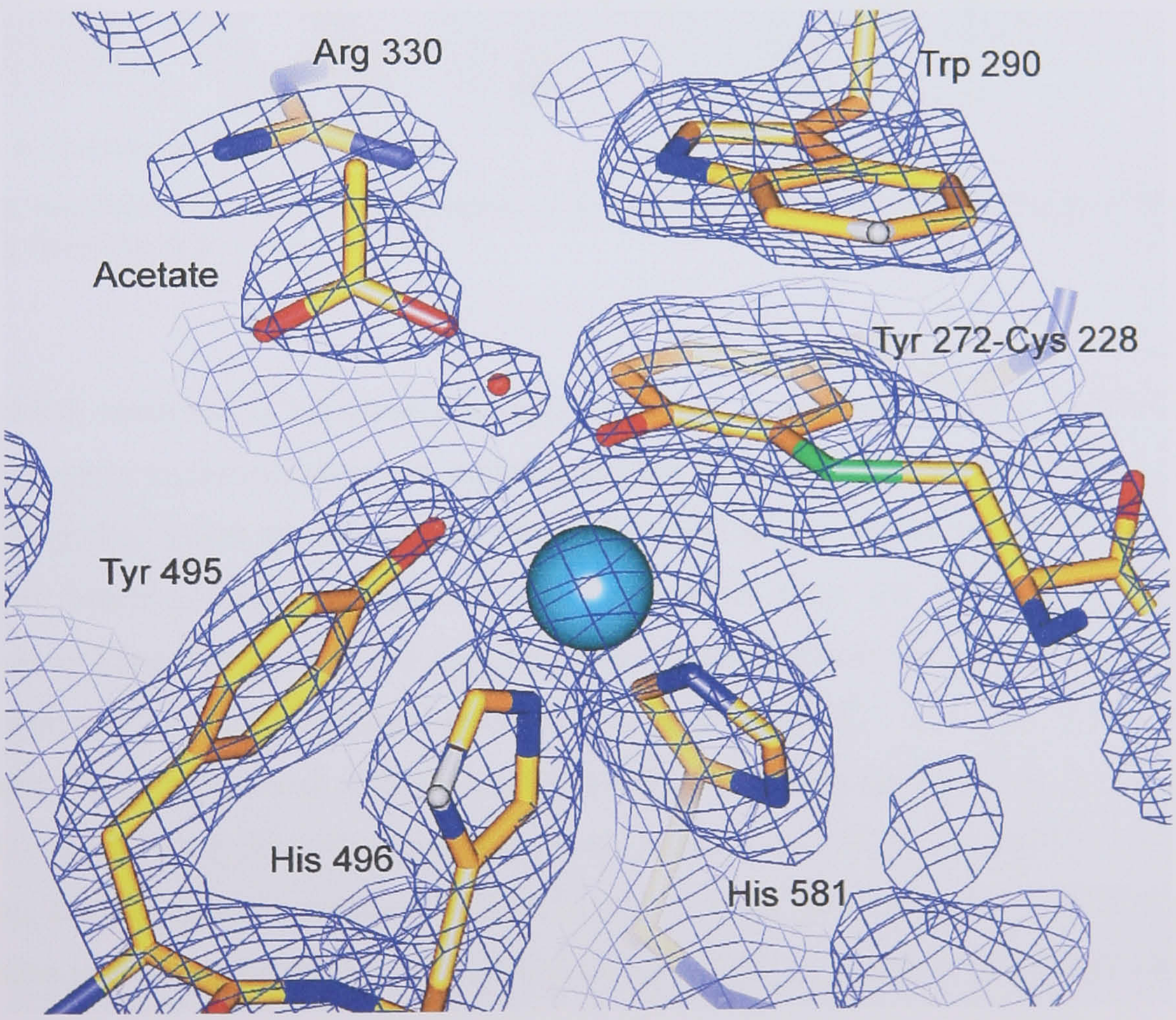


Figure 3.22: 2Fo-Fc density map of the C383S active site.
2Fo-Fc electron density map contoured at 1 x rms density showing the C383S active site. The presence of the mutation did not prevent the thioether bond formation.

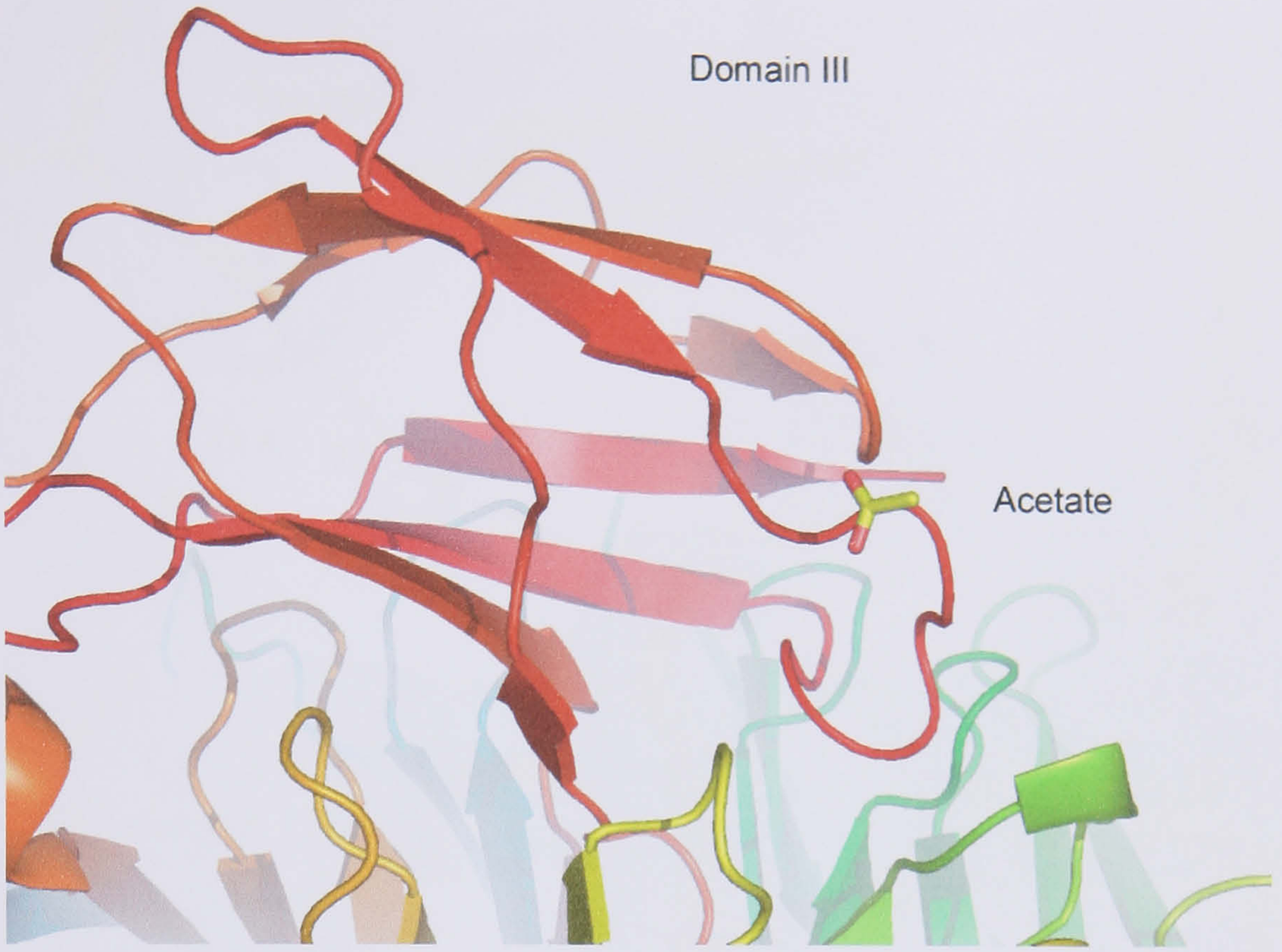


Figure 3.23: Domain III acetate binding

Acetate was found bound to a loop region of domain III. The molecule was approximately 40 Å away from the active site copper.

The C383S mutation is effectively a replacement of a sulphydryl group with hydroxyl, allowing water molecules near the mutated residue to form strong hydrogen bonds to the hydroxyl group. There are three water molecules close to Ser 383 of which two are within hydrogen bonding distance (Figure 3.24). This differs from the situation in wild type galactose oxidase, where there are only two water molecules near residue 383 (Figure 3.25). The additional weakly occupied water molecule, labelled 'C' in Figure 3.26, forms a hydrogen bond to the serine residue and the nearby main chain carboxyl oxygen of Tyr 405 (see discussion). It appears that there may be more space created by the mutation, therefore, allowing the incorporation of an additional water molecule next to residue 383. Sodium and an acetate ion were identified and modelled into positions reported in the original galactose oxidase structure.

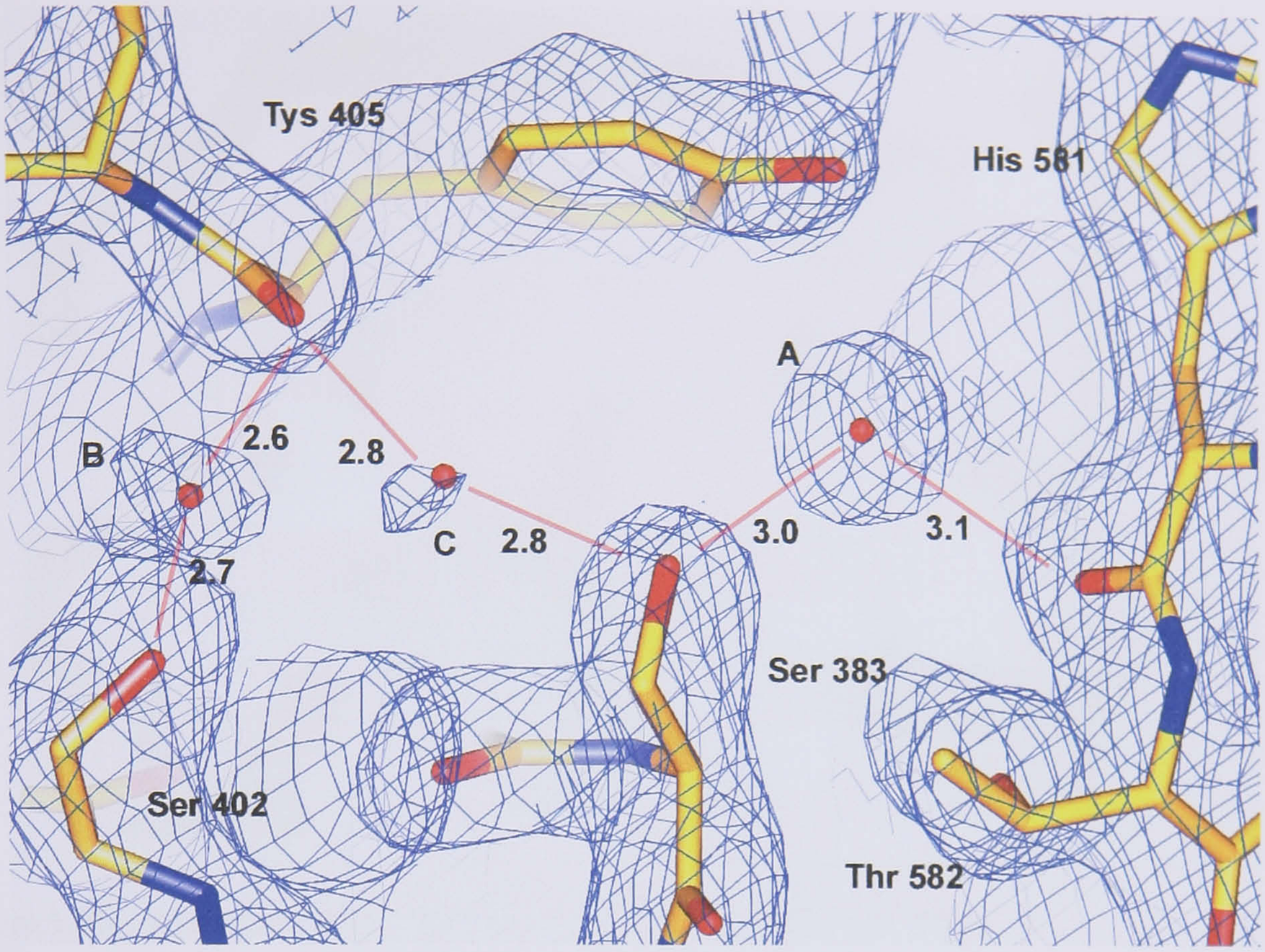


Figure 3.24: The 2Fo-Fc density map of C383S

The 2Fo-Fc map is contoured at 1 rms density. The density around the oxygen atom of residue 383 is smaller than density expected for the Sγ of a cysteine, and refined well as a serine residue.

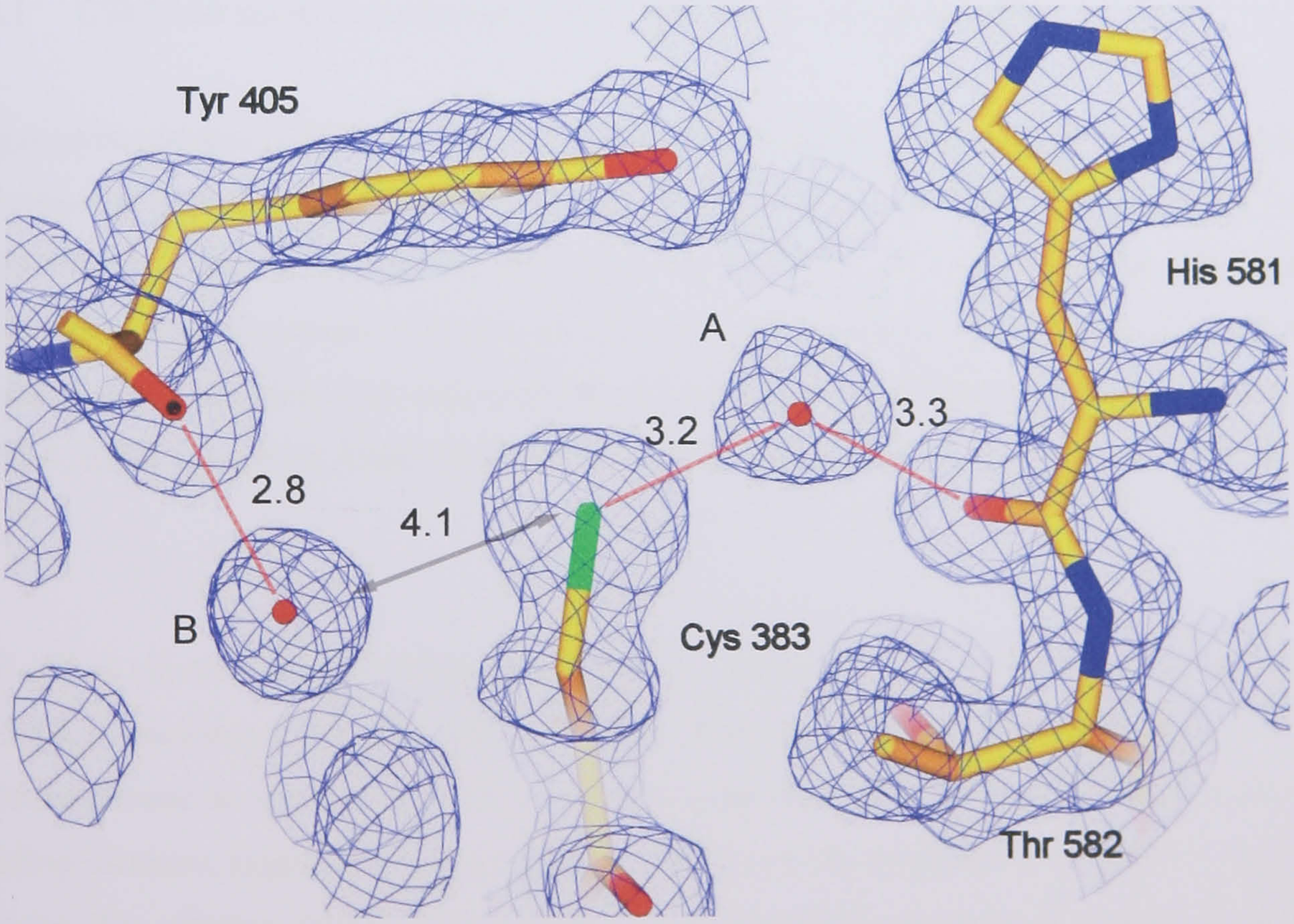


Figure 3.25: 2Fo-Fc density map of the region around Cys 383 in wild type galactose oxidase.

2Fo-Fc electron density map from a 1.7 Å crystal structure of wild type galactose oxidase contoured to 1 x rms density. Map calculated from Data collected from 3 minute galactose soak into wild type galactose oxidase (Table 3.1).

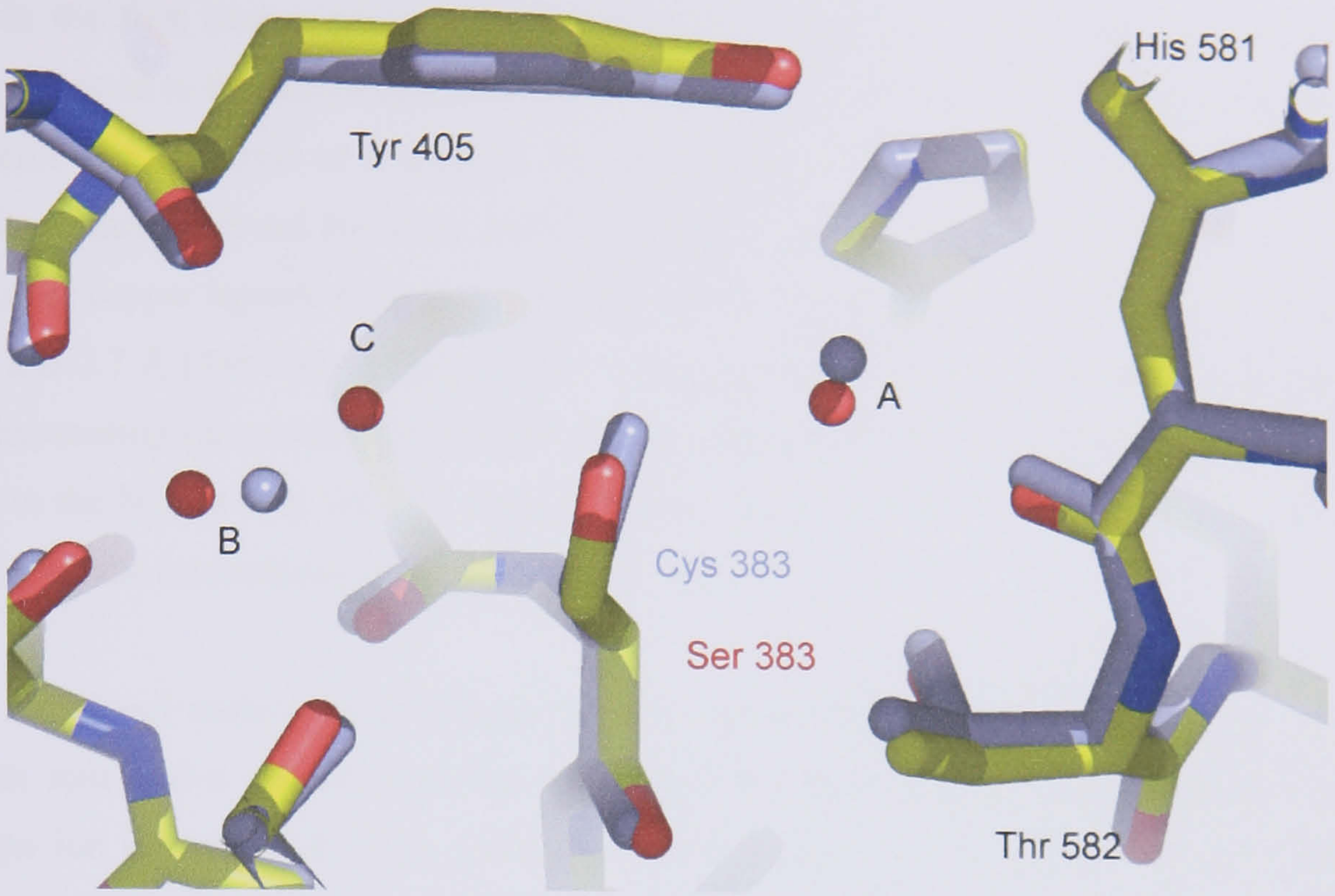


Figure 3.26: Overlay of C383S with wild-type galactose oxidase (blue).

There are three water molecules in close proximity to the residue 383 (labelled A, B and C), water C is only found in the mutant structure.

3.4.1 C383S 60 second incubation with 2-methylene-1,3-propandiol

A galactose oxidase C383S crystal was aerobically soaked for 60 seconds in 10 µl soaking solution comprising mother liquor, 10% PEG 400 and 5% 2-methylene-1,3-propandiol (v/v). Following the soak, the crystal was mounted on the R-AXIS IV++, and frozen at 100 K using a liquid nitrogen cryostream. Initial exposures showed the crystal diffracted to 2.7 Å. It was stored in liquid nitrogen until data could be collected using synchrotron radiation, (beam line ID14-4, ESRF Grenoble, France), where the crystal diffracted to a resolution of 1.9 Å (Table 3.3).

The initial electron density maps confirmed the position of the active site residues were essentially unchanged from the original C383S structure. Once again there were three water molecules close to the hydroxyl of serine 383, with two water molecules within hydrogen bonding distance, suggesting it is a particular feature of the mutation of cysteine to serine at this site. The electron density at the active site suggested the presence of a fourth equatorial copper ligand. This density peak was too large to be water and did not refine well as an acetate molecule. A 2-methylene-1,3-propandiol molecule was modelled into the density giving a good fit and refined well at full occupancy (Figure 3.28).

This is the first time a galactose oxidase structure has been solved with the substrate or product bound to the active site. It is not possible to tell, however, if the exogenous active site molecule is a substrate or a product. The molecule is 3.0 Å away from copper, a distance similar to that observed for water molecule in the original, acetate-free structure [Ito, 1994]. The other copper ligands refined to distances of 2.0 Å (Tyr 272), 2.0 Å (His 496), 2.1 Å (His 581) and 2.7 Å (Tyr 495) (Figure 3.28). Assuming the exogenous molecule is substrate (but not discounting the possibility it is a product), 2-methylene-1,3-propandiol forms a hydrogen bond to the Nη1 of Arg 330 (2.7 Å), and thus it is expected to be stabilised in the active site mainly via the interactions with Arg 330 and copper.

Two additional molecules of 2-methylene-1,3-propandiol were identified, as well as three acetate ions and a sodium ion. One of the acetate ions was located in domain II and the sodium ion in domain I, these positions were identified in the original galactose oxidase structure (Figure 1.6) (PDB code 1GOF). A 2-methylene-1,3-propandiol molecule and the two remaining acetate ions were bound at the surface of domain III (Figure 3.29). The final 2-methylene-1,3-propandiol molecule was located in domain II (Figure 3.30), bound between the second and third blade of the propeller motif. All exogenous ions were modelled into the structure and the model refined to a final R-factor of 19%, R-free 22%.

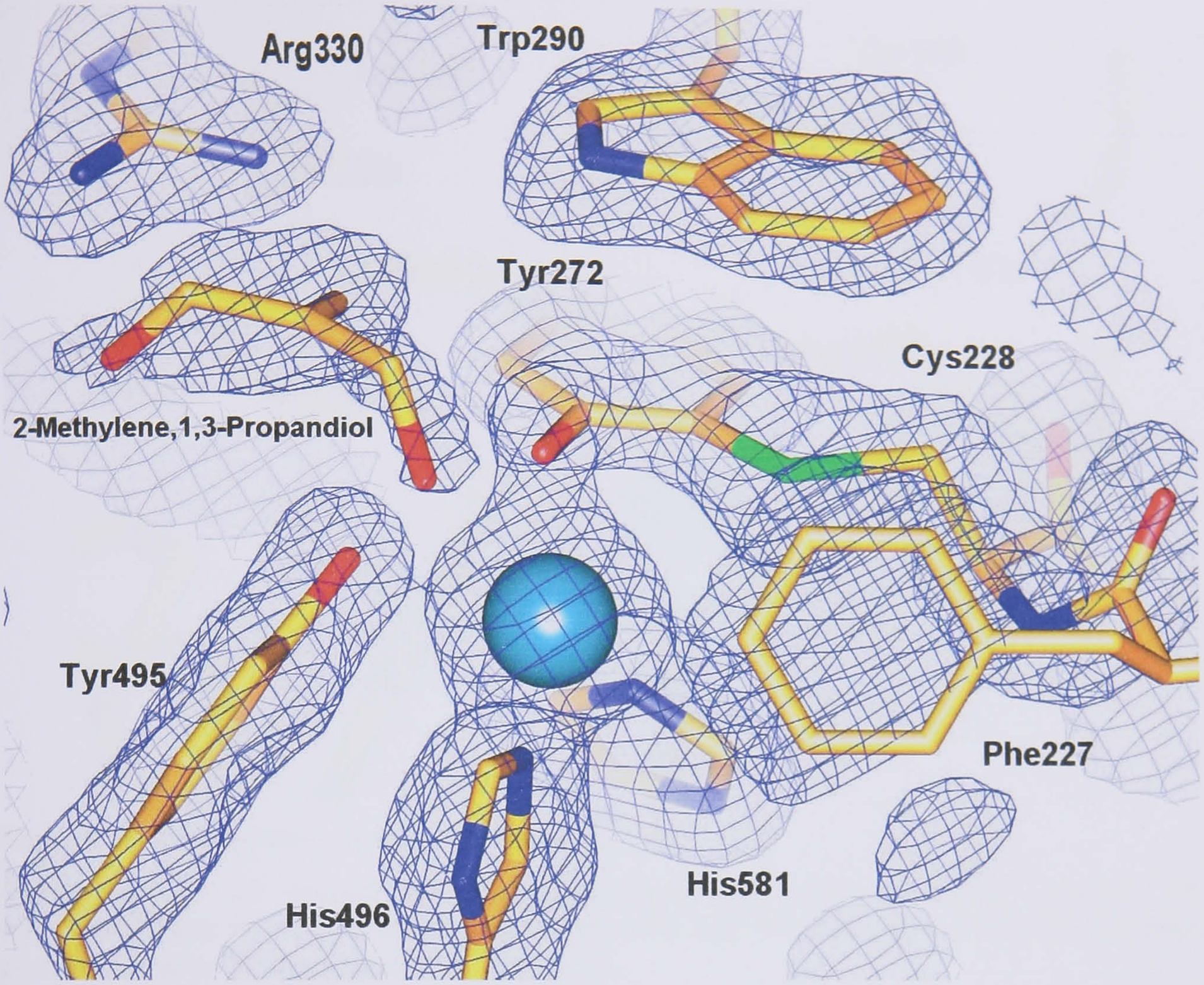


Figure 3.27: 2Fo-Fc density map of 2-methylene-1,3-propandiol bound C383S active site. 2Fo-Fc electron density map showing 2-Methylene-1,3-propandiol bound 3.0 Å from the active site copper after 60 second substrate soak of a C383S galactose oxidase crystal. Contoured at 1 x rms density.

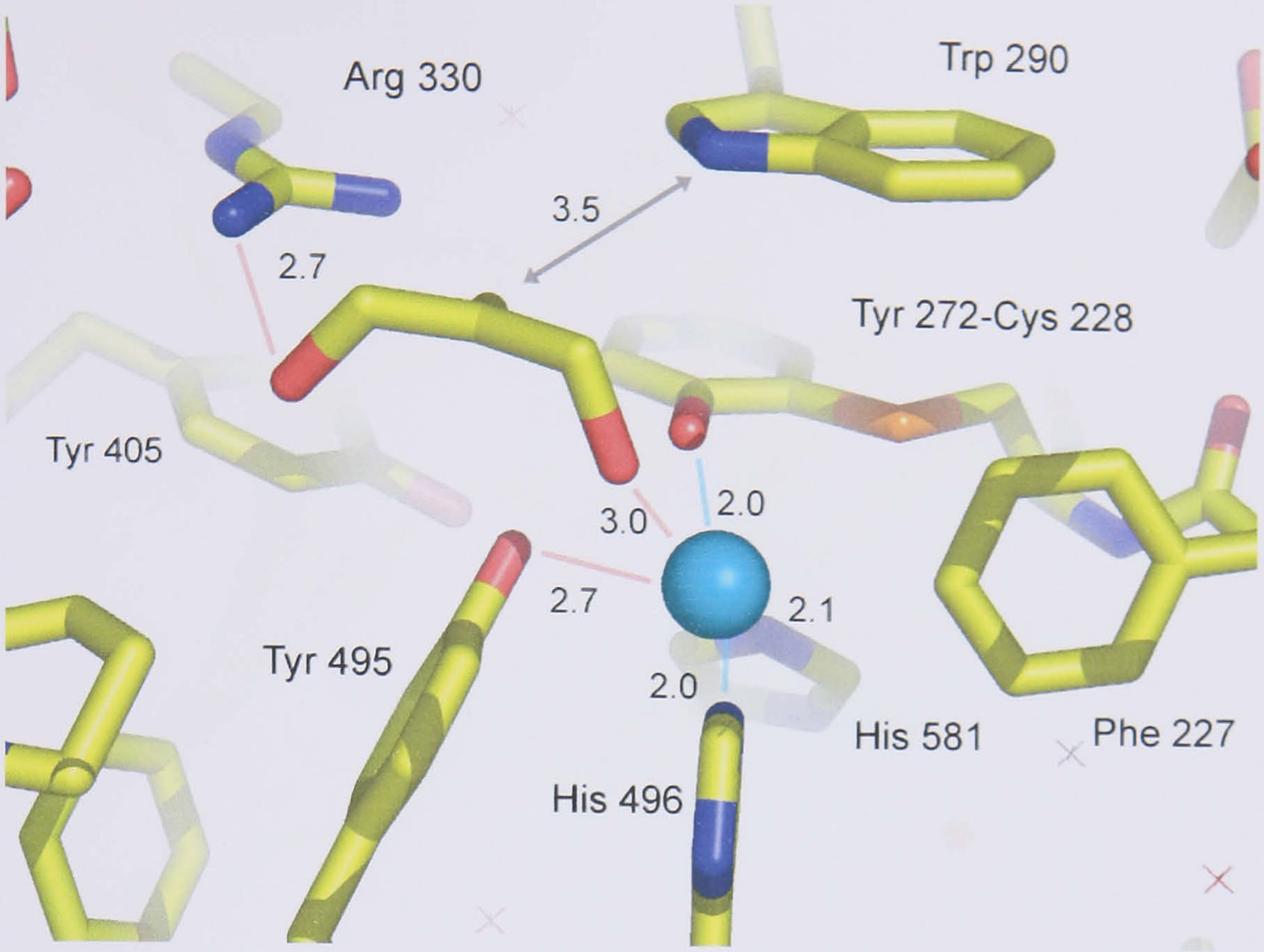


Figure 3.28: 2-methylene-1,3-propandiol bound active site.

Active site of 60 second 2-methylene-1,3-propandiol soak of C383S crystal. The substrate/product forms favourable interactions with Arg 330 and copper, but is too far away from the metal to be considered a strong ligand.

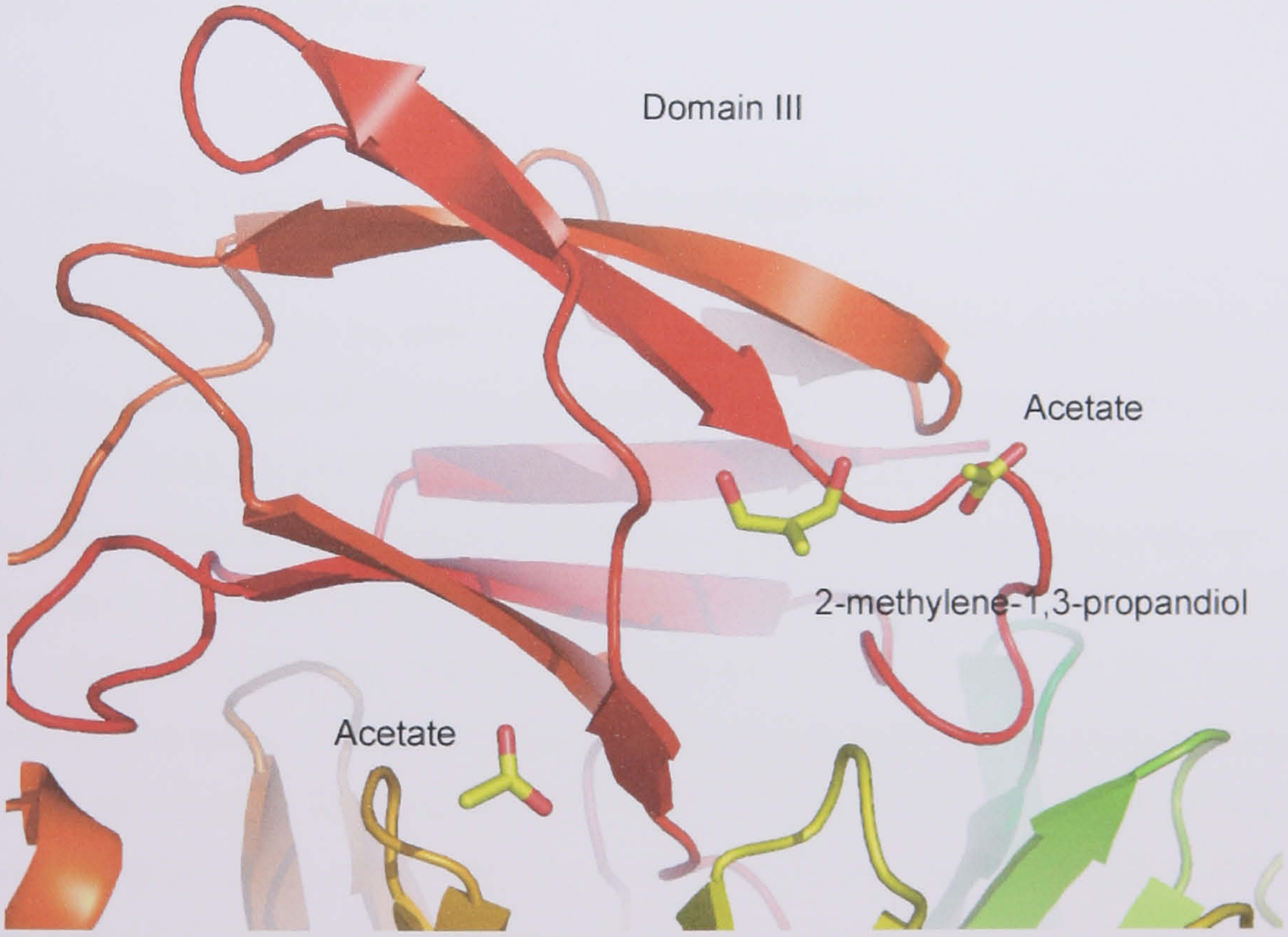


Figure 3.29: Domain III binding sites after 60 second 2-methylene,1,3-propandiol soak.

Two acetate anions and a 2-methylene-1,3-propandiol molecule were identified bound to domain III. One acetate (closest to the substrate) was in a similar position to the acetate identified in substrate-free C383S (Figure 3.23).

Figure 3.30: Domain II binding sites after a 60 second 2-methylene-1,3-propandiol soak.

One of the three 2-methylene-1,3-propandiol molecules identified in the structure was bound on the surface of domain II, coordinated by residues from the inter strand loop of β-sheet/blade of the seven-bladed β-propeller motif. Domain III can be seen in the top left–hand side of the figure and is coloured red.

3.4.2 Aerobic 3 minute 2-methylene-1,3-propandiol soak

A second crystal, soaked in with 5% 2-methylene,1,3-propandiol solution for 3 minutes, diffracted to a resolution of 2.8 Å on the RAXIS IV++, and was improved to a resolution of 1.8 Å at the ESRF, beam-line ID14-4, Grenoble (Table 3.3). Examination of the electron density maps revealed the presence of elongated density within coordination distance of the copper (Figure 3.31). A water molecule was initially modelled into the density coordinated to copper, while an acetate ion was modelled into the density adjacent to this. The model, however, did not refine well, as the acetate and water molecule were too small to account for the active site density. The presence of strong difference electron density suggested the elongated electron density could represent a diatomic molecule, such as hydrogen peroxide or oxygen, both of which would be present in the active site at some point during catalytic turn over.

A peroxide molecule was modelled into the elongated density. It is not possible, however, to determine whether the exogenous molecule is diatomic oxygen, superoxide or hydrogen

peroxide, as the only distinguishing feature would be the oxygen-oxygen bond distance. Hydrogen peroxide has a greater O-O bond length than superoxide and molecular oxygen (1.49 Å, 1.34 Å and 1.21 Å respectively) [Kleywegt, 1998]. With a 1.8 Å resolution electron density map, however, the error in refining atomic bond distances would be too large to accurately determine the identity of the diatomic molecule. Therefore, the molecule was assumed to be hydrogen peroxide and a restraint of 1.49 Å was placed on the O-O bond distance. There are no other molecules that are likely to be present and the density is too strong to be accounted for by two alternate positions of a water molecule. Refinement of the molecule with one of the oxygen atoms at zero occupancy resulted in the presence of strong positive density adjacent to the fully occupied oxygen atom. The peroxide was eventually modelled at 70% occupancy. Although the density was assumed to be peroxide, this does not rule out the possibility of a different diatomic molecule occupying the density. The molecule refined to a distance of distance of 2.6 Å from copper, while the four remaining copper ligand distances refined to 1.9 Å (Tyr 272), 2.2 Å (His 581), 2.1 Å (His 496) and 2.6 Å (Tyr 495) (Figure 3.32).

The density adjacent to the peroxide was assumed to be 2-methylene-1,3-propandiol and the putative molecule was modelled in at 80% occupancy. Hydrogen bonds are formed with Trp 290 (3.2 Å), Tyr 495 (2.5 Å) and Tyr 329 (2.8 Å) (Figure 3.32), residues that have been highlighted as being important for substrate binding [Baron *et al.*, 1994; Deacon *et al.*, 2004; Ito *et al.*, 1990; McPherson *et al.*, 1993; Reynolds *et al.*, 1995; Saysell *et al.*, 1997]. The pattern of hydrogen bonding, however, does not allow assignment of the molecule as the original diol substrate or an aldehyde product. The alcohol substrate was modelled into the density, but this does not rule out the possibility of an aldehyde product occupying the density.

A second 2-methylene-1,3-propandiol and a sulphate molecule are located on the surface of domain II, bound between the first and second blades of the β-propeller motif (Figure 3.33). Sodium and an acetate ion were identified and modelled into positions reported in the original galactose oxidase structure. The structure was refined to a final R-factor of 20% and an R-free of 23% (Table 3.3)

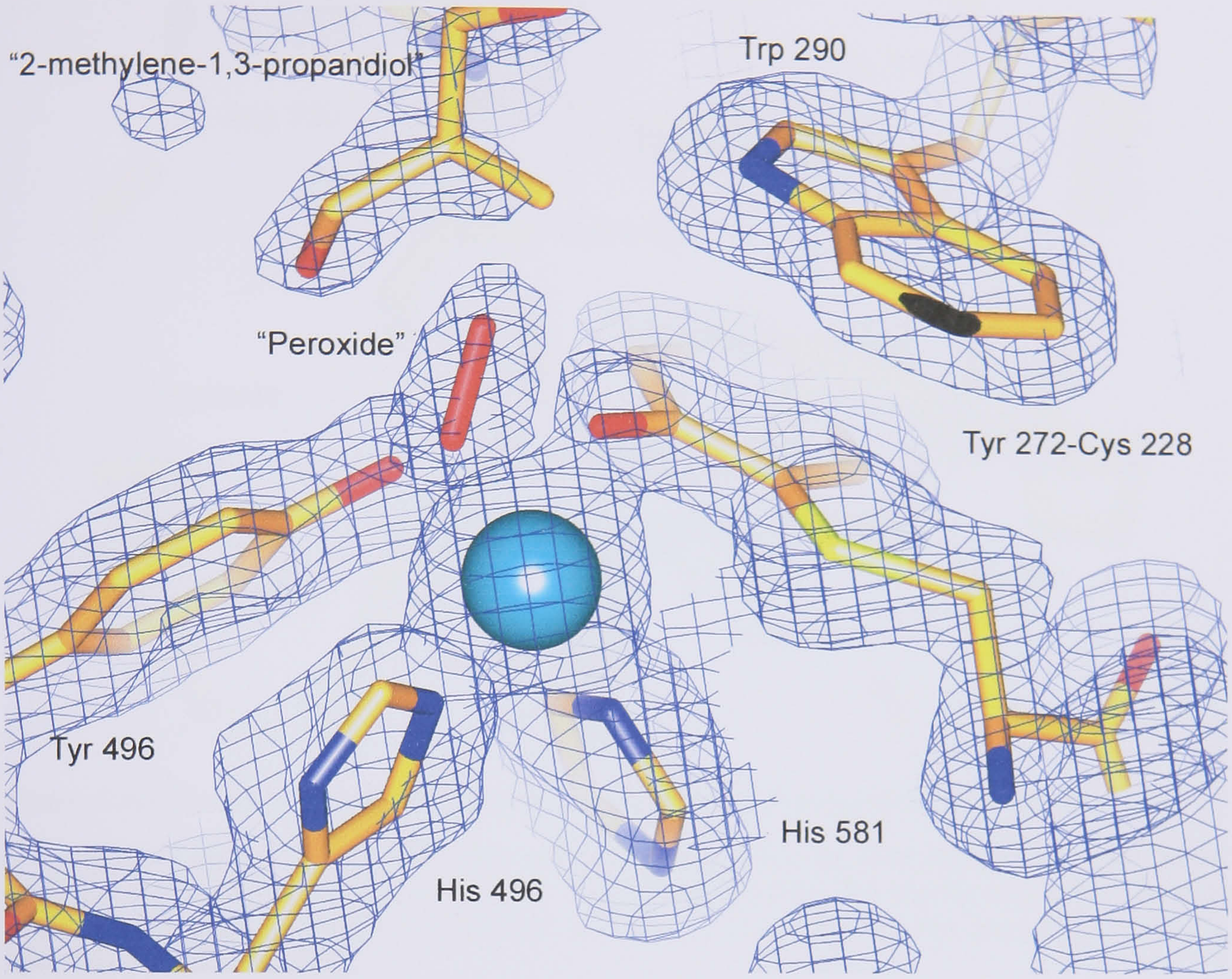


Figure 3.31: 2Fo-Fc density map of the active site of C383S soaked for 3 minutes

The 2Fo-Fc electron density map of the active site of C383S soaked aerobically for 3 minutes in 5% 2-methylene-1,3-propandiol. The diatomic species is bound in the fourth equatorial copper ligand position and modelled as a putative hydrogen peroxide. Close to this site a putative 2-methylene,1,3-propandiol molecule has been modelled.
The map is contoured at 1 x rms density.

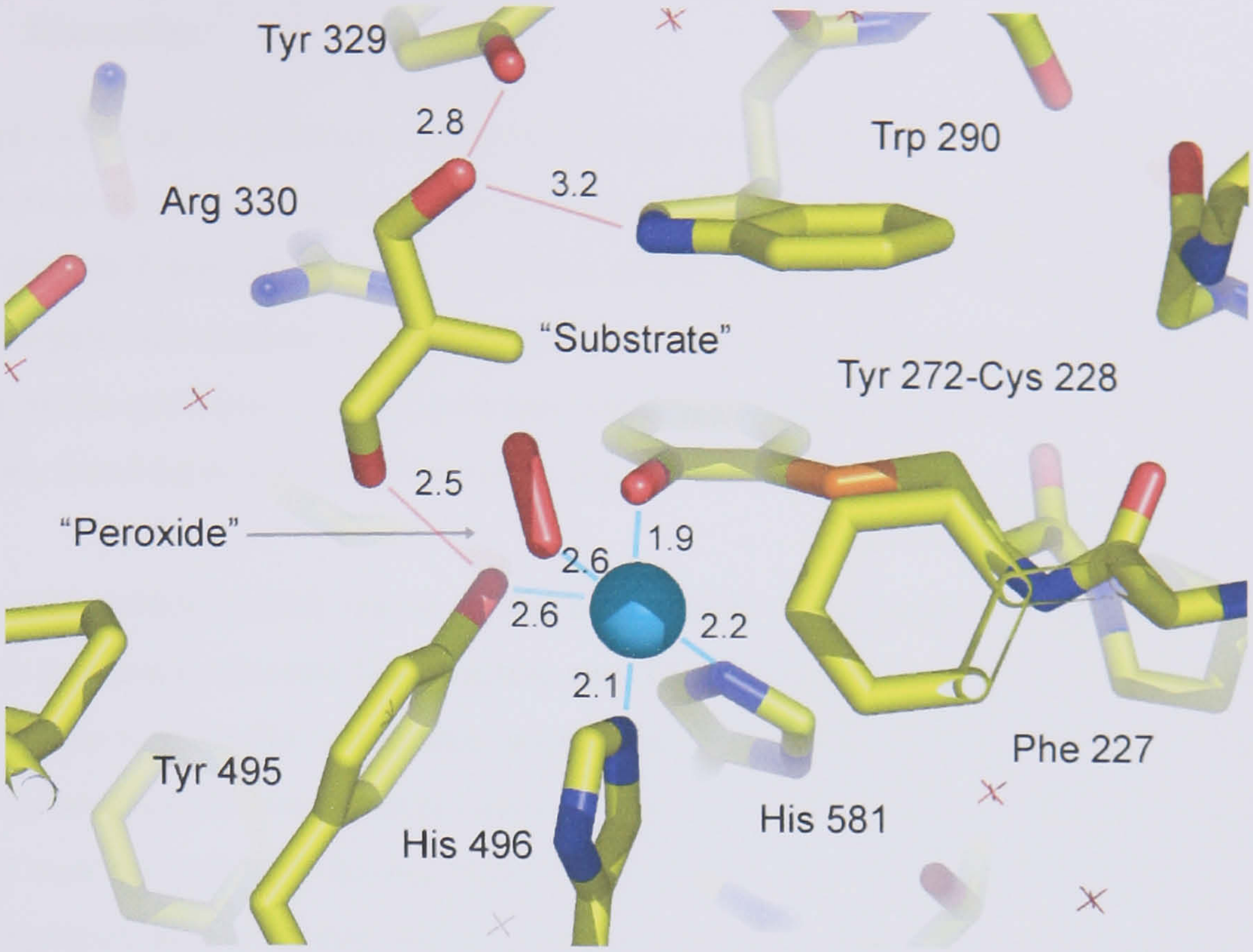


Figure 3.32: The active site of C383S soaked in 2-methylene-1,3-propandiol for 3 minutes.

The red lines represent potential hydrogen bonds and the cyan lines represent copper ligand bonding. A putative peroxide molecule was modelled into the fourth equatorial binding position. The putative alcohol substrate is stabilised in the active site by interactions with Tyr 329, Trp 290 and Tyr 495.

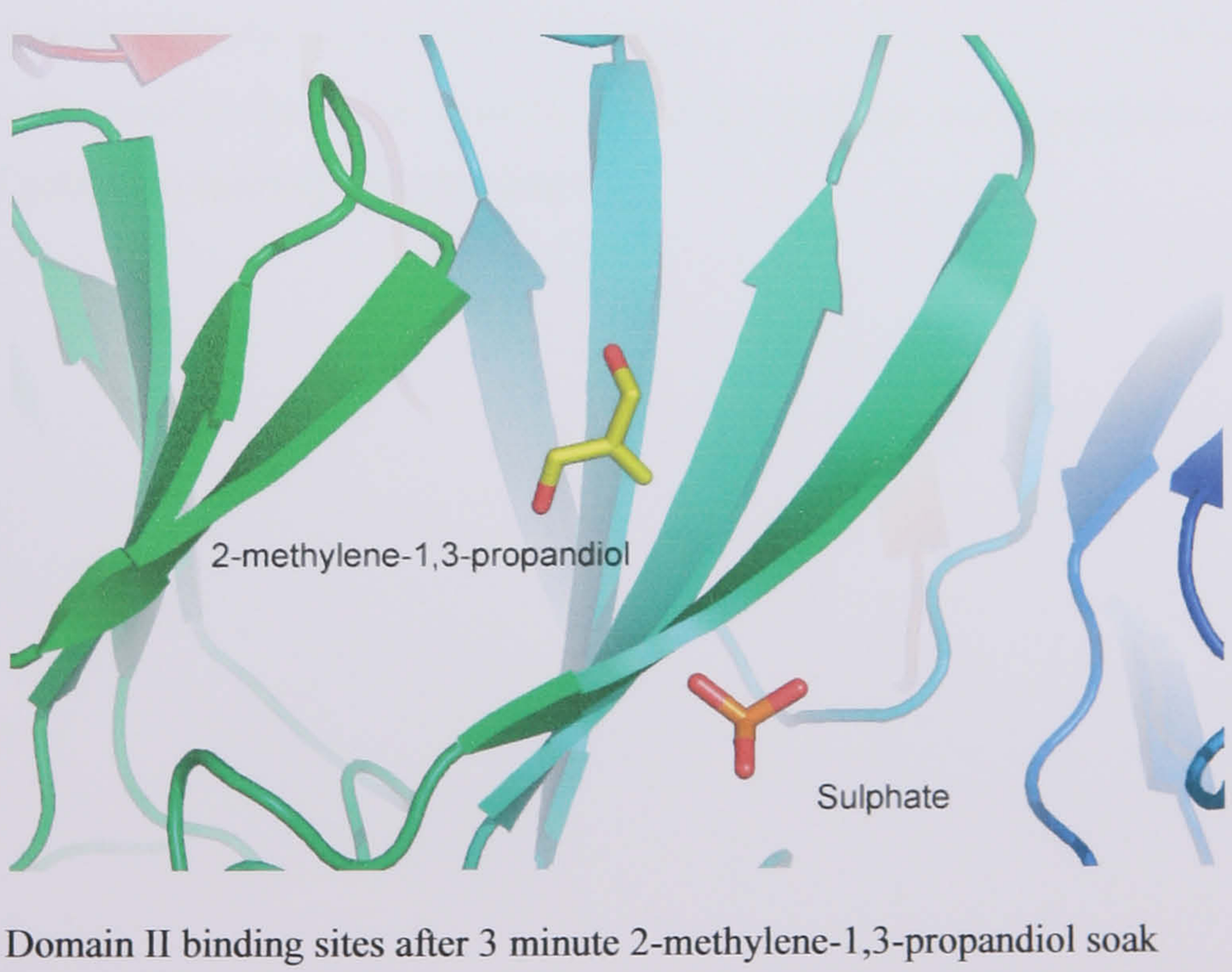


Figure 3.33: Domain II binding sites after 3 minute 2-methylene-1,3-propandiol soak

A sulphate ion and a second molecule of 2-methylene-1,3-propandiol were modelled into the crystal structure. The molecules were located in domain II, bound between the first and second β-sheet/blade of the β-propeller motif (see Figure 1.8 for full propeller motif).

3.5 Discussion

Attempts to obtain a galactose oxidase-substrate complex have proved unsuccessful in the past. Earlier experiments utilised varying sugar concentrations but binding was only observed to the domain I site. In the work presented in this chapter, a 100 mM galactose incubation with apo-galactose oxidase resulted in binding of galactose in domain I, as assessed by strong density at the previously known galactose binding site. There were no galactose molecules, however, found associated with the active site.

A 100 mM galactose soak into a crystal of oxidised, copper loaded galactose oxidase also failed to produce a substrate bound active site. The apo-enzyme and the ferricyanide oxidised enzyme have very similar active sites, containing three water molecules which form hydrogen bonds to catalytically important residues (Figure 3.7 and Figure 3.11). The water molecules interact with the Nη1 (2.9 Å) and Nη2 (3.0 Å) of Arg 330, and the Nε1 (2.8 Å) of Trp 290. These residues, as well as Gln 406 are thought to be important in the stereo-selectivity of the enzyme toward galactose. Protein engineering studies in which residues Arg 330, Trp 290 and Gln 406 were mutated showed an increase in glucose oxidation activity [Sun *et al.*, 2002], and more recently the mutant R330K was shown to have increased activity for fructose oxidation [Deacon *et al.*, 2004]. A comparison of the active site of the galactose soaked enzyme with the proposed model of galactose binding suggests the interactions that may stabilize galactose in the active site could be the same as those stabilising the three water molecules, thus leaving a "ghost" of galactose binding (Figure 3.34).

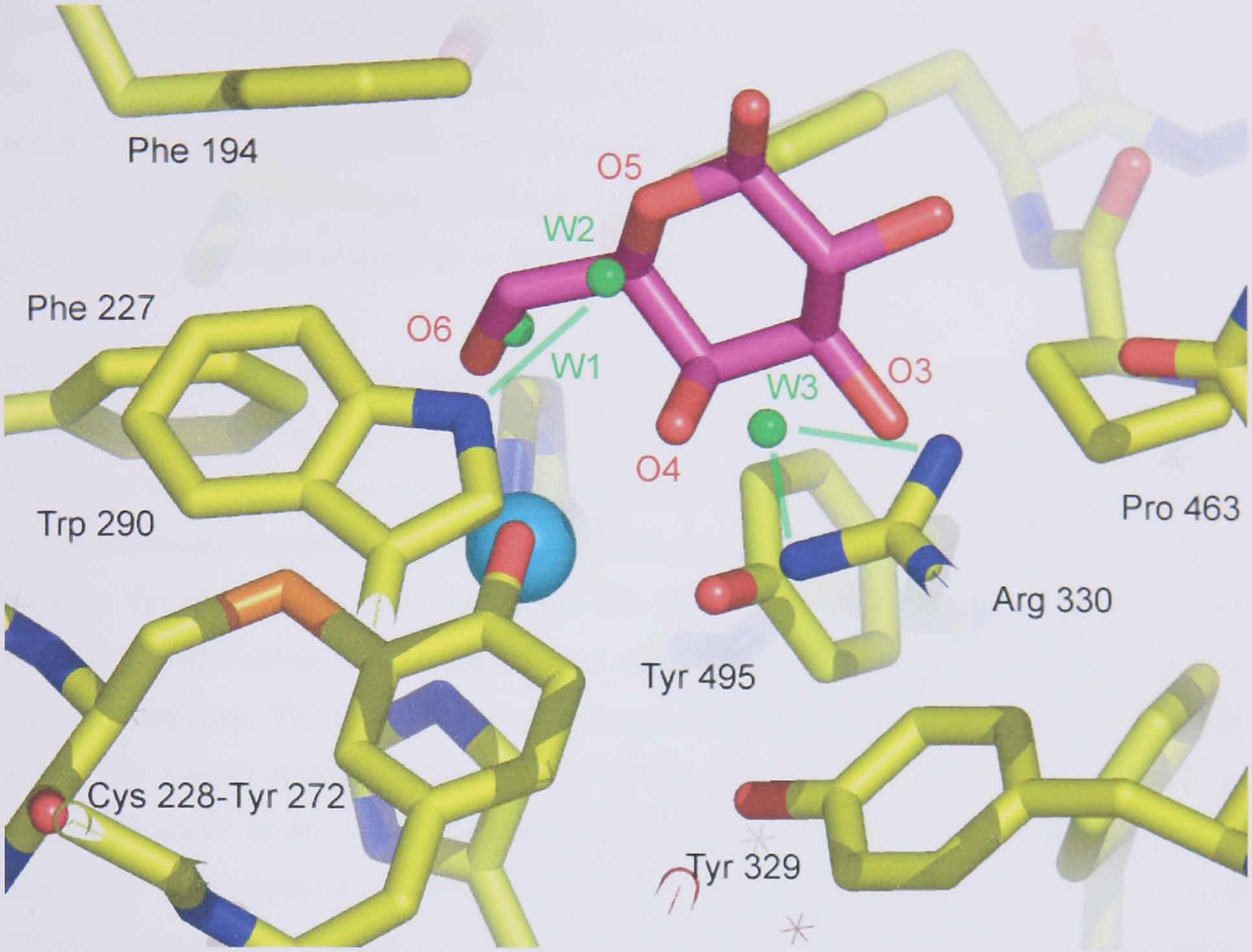


Figure 3.34: Alignment of manually docked galactose onto the activated wild type active site

Galactose (magenta) from manual docking model [Ito *et al.*, 1994], superimposed onto the active site of activated galactose oxidase soaked with 100 mM galactose. O5 and O3 of galactose may correspond to W2 and W3 respectively. The O6 of galactose would bind to copper, and has aligned well with the W1 water molecule in the enzyme active site.

Copper coordination by His 581 is important for the stability of galactose oxidase. In early studies, removal of copper from the protein resulted in an enzyme with reduced resistance to urea denaturation [Kosman *et al.*, 1974]. His 581 is found at the tip of a beta hairpin loop which extends from domain III and into the central cavity of the domain II beta propeller structure (Figure 1.8). An extensive hydrogen bond network comprising 40 conserved water molecules exists at the interface between domain II and the long finger loop of domain III. Two of these water molecules are found close to the Sγ of Cys 383 in all wild type galactose oxidase structures (1GOF, 1GOG, 1GOH) [Ito *et al.*, 1991; Ito *et al.*, 1994]. The water molecules are also conserved in pro-galactose oxidase structure, which was crystallised under different conditions and solved in space group $P2_12_12_1$ (1K3I) [Firbank *et al.*, 2001]. Mutation of Cys 383 to a serine produced an enzyme with improved catalytic activity against galactose, mainly due to an decrease on the K_M for the substrate (wild type K_M=68 mM, C383S K_M=15 mM) [Wilkinson *et al.*, 2004]. The mutation of Cys 383 to serine has a direct effect on the water network around the residue. The two water molecules near the Sγ of wild type 383, are found 3.2 Å away (water-A) and 4.1 Å away (water-B) (Figure 3.25 and Figure

3.26). The carbonyl oxygen of His 581, a residue that is a ligand to the active site copper, forms a hydrogen bond with water-A. There is a hydrogen bond network extending from water-A, which incorporates the main chain amide between the copper ligands Tyr 495 and His 496. Water-B forms a hydrogen bond with the main chain carbonyl of Tyr 405 but is not at an appropriate angle or distance to form a hydrogen bond with the Sγ of Cys 383.

In the C383S mutant crystal structure three water molecules are found close to residue 383 (Figure 3.24 and Figure 3.26). The mutation apparently allows the incorporation of an additional water molecule at this site. The additional water (Water-C) forms a hydrogen bond with the carbonyl oxygen of Tyr 405 and this interaction contributes to the decreased distance between Tyr 405 and the general base Tyr 495 (Table 3.4). The hydroxyl groups of Tyr 405 and Tyr 495 are closer to each other in the mutant crystal structure compared to wild type galactose oxidase. Therefore, the new hydrogen bond network around Ser 383 indirectly affects the copper binding residues His 496, His 581 and Tyr 495. It is unclear how an additional water molecule has any effect on galactose binding. Studies on the role of conserved water molecules in crystal structures of ten aspartic proteinases, however, have reported that conserved water molecules in the crystal structures were strategically located to stabilise the enzyme structure and maintain the active site geometry [Prasad and Suguna, 2002]. Such a role for the water molecules around the C383S mutation site can not be ruled out, but it is difficult to know with certainty.

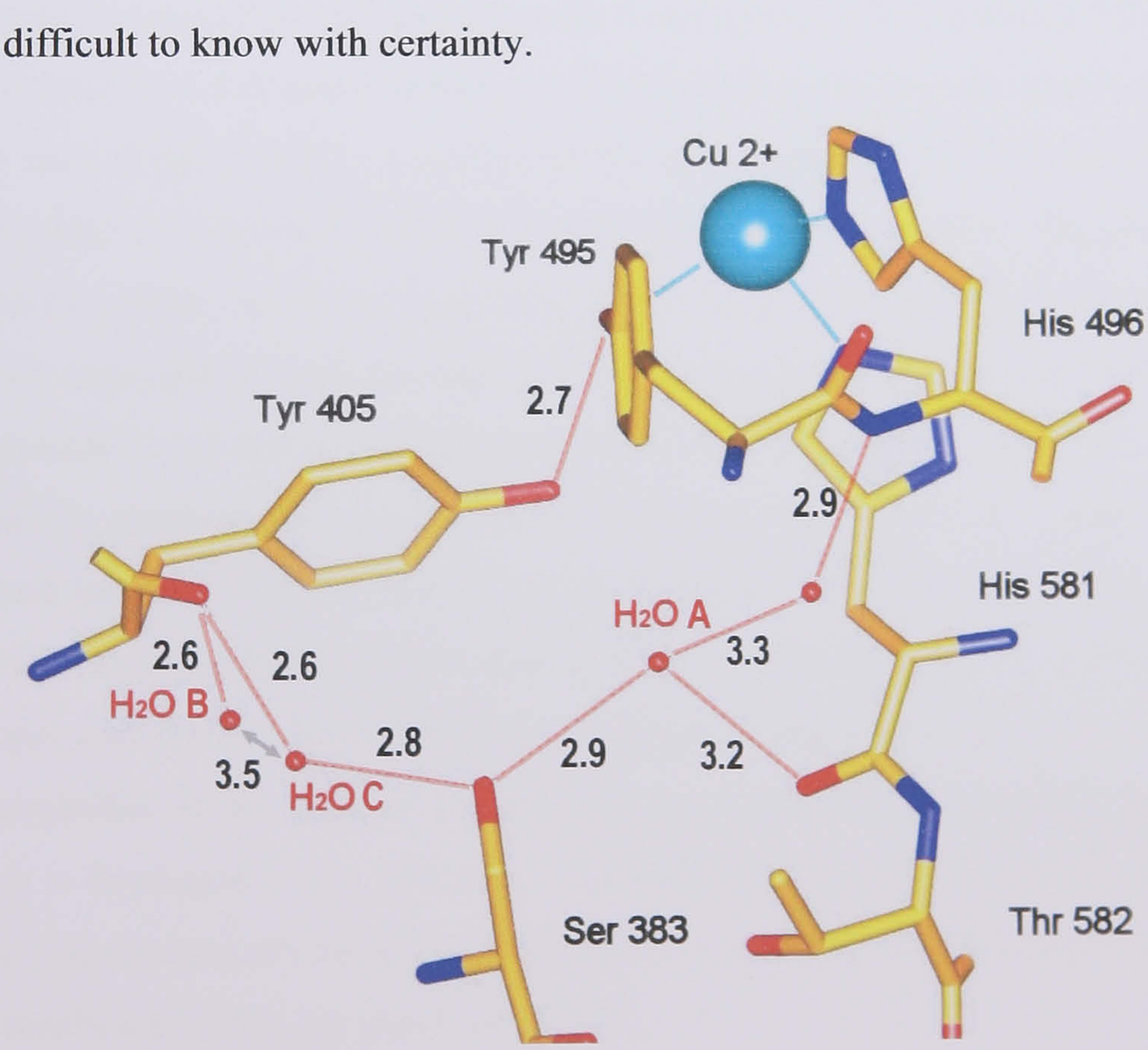


Figure 3.35: Hydrogen bonding network around the C383S mutation site.

Hydrogen bond lengths red and copper coordination is cyan. The water molecules A, B and C form a hydrogen bond network through Ser 383.

Table 3.4: Hydrogen bond distances around the C383S mutation site.

Comparisons of the hydrogen bond distances around the Ser 383 mutation site.				
Hydrogen bond distance (Å)	Galactose Oxidase Structure			
	Wild type (1GOG)	C383S (1T2X)	C383S 60 second substrate soak	C383S 3 minute substrate soak
*Ser 383-Oγ to H_2O A	3.2	2.9	3.1	3.0
H_2O A to His 581- carbonyl	3.3	3.2	3.0	3.1
*Ser 383-Oγ to H_2O B	4.1	5.7	4.9	4.7
H_2O B to Tyr 405 carbonyl	2.8	2.6	2.7	2.6
*Ser 383-Oγ to H_2O C	-	2.8	2.8	2.8
H_2O C to Tyr 405 carbonyl	-	2.6	2.8	2.8
H_2O B to H_2O C	-	3.5	2.7	2.6
Tyr 405-OH to Tyr 495-OH	3.3	2.7	2.8	2.8

*For wild type galactose enzyme distance is from Cys 383-Sγ

A C383S crystal incubated for 60 seconds with 2-methylene-1,3-propandiol (5% v/v of soak solution) diffracted to 1.9 Å and revealed a galactose oxidase active site with bound substrate for the first time (Figure 3.27). 2-methylene-1,3-propandiol forms a favourable hydrogen bond with the Nη2 of Arg 330 and is 3.0 Å from the active site copper. The crystal was not oxidised prior to aerobic substrate incubation, therefore the majority of enzyme molecules in the lattice were expected to be in the catalytically inactive semi-reduced state. Indeed the long copper to substrate bond distance suggests the substrate is not interacting very strongly with the metal. Strictly speaking direct coordination would not be expected at distances over 2.1 Å. A copper bond length of approximately 2.5 Å, however, could be considered indicative of weaker, mostly electrostatic interaction. This is similar to the interaction commonly observed between copper and the axial Tyr 495 ligand. Longer distances would have no effect on the spectra or properties of the copper ion and are considered as a secondary interaction, no stronger than a hydrogen bond [Personal communication Dr M Halcrow, University of Leeds]. Galactose oxidase structures have also been observed with long copper to ligand bond distances at fourth equatorial site [Ito *et al.*, 1995].

It is interesting to note that the methylene group of the substrate is orientated towards Trp 290 and is only 3.5 Å away from the Nε1 of the residue. The Trp 290 imidazole amide and the methylene group are in van der Waals contact, however, if dihydroxyacetone were bound in the same orientation as 2-methylene-1,3-propandiol, a hydrogen bond could be formed between the carbonyl oxygen of dihydroxyacetone and the Nε1 of Trp 290. Such an interaction would be expected to allow both Nη1 and Nη2 of Arg 330 to interact with the hydroxyl group of the substrate molecule, thus perfectly orientating the dihydroxyacetone for oxidation. From analysis of this new structure it is clear why dihydroxyacetone is a good substrate for galactose oxidase, as it could be stabilised by at least three protein-substrate hydrogen bonds in addition to copper binding during catalysis.

A longer crystal soak with 2-methylene,1,3-propandiol revealed an active site containing the substrate molecule some distance from the copper and what appears to be a diatomic copper ligand (Figure 3.32). The electron density for the diatomic molecule and the substrate molecule in this structure is very good, suggesting that a one or both of these molecules were bound to the majority of the galactose oxidase active sites. The substrate diol was modelled into the active site, however, the density could represent the aldehyde product. The molecule is bound to the active site in a different orientation, compared to the shorter soaked structure (Figure 3.36).

The diatomic molecule was modelled as a peroxide, but it could be oxygen or superoxide since the identity could not be determined from the electron density. It was a surprise, however, to find a diatomic molecule bound at the fourth equatorial position, as the C383S crystals were not expected to be fully catalytically active. Observation of a diatomic molecule in the active site at the same time as a substrate/product molecule suggests the enzyme was in the oxidised active form.

There are reported to be two competing redox reactions that alter the oxidation state of the resting enzyme [Wright and Sykes, 2001a]. Auto-reduction is the spontaneous reduction of the oxidised enzyme, whilst in aerobic conditions auto-oxidation regenerates the radical. One method to detect the oxidation state of galactose oxidase in crystallised form, is by single crystal spectroscopy. Attempts to obtain such spectra for substrate and Cu^{1+} soaked crystals failed, often resulting in the loss of the crystal. The redox state of the enzyme in these crystals could not be established experimentally, therefore, auto-oxidation leading to catalytically active enzyme could not be confirmed or ruled out.

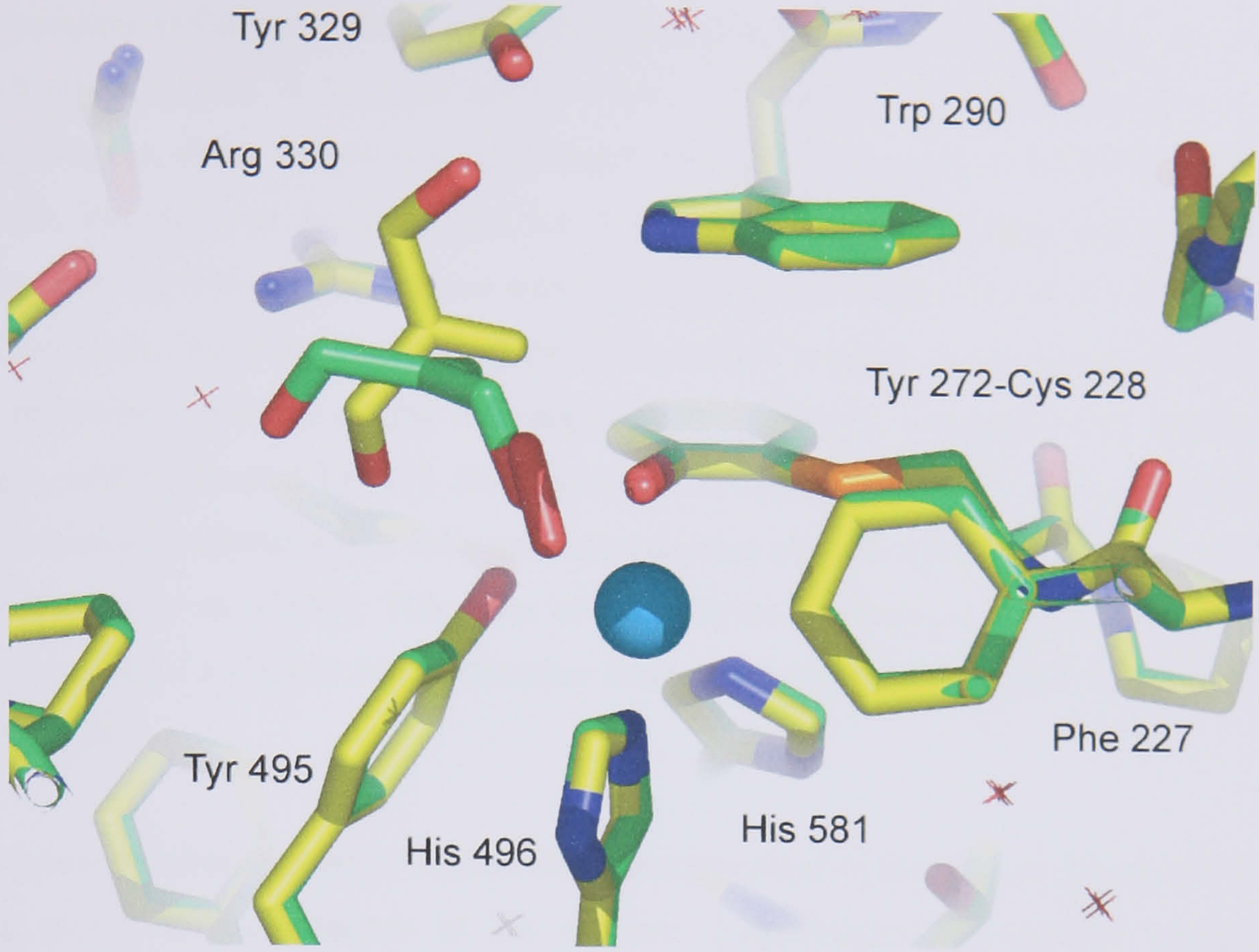


Figure 3.36: Alignment of the two 2-methylene,1,3-propandiol soak active sites structures.

The overlay of the active sites of the 60 second 2-methylene,1,3-propandiol soak (carbons coloured green) and the 3 minute soak. The protein copper ligands are in the same positions in both structures. The diatomic molecule approaches the copper in an end-on orientation and occupies the fourth equatorial ligand position. 2-methylene,1,3-propandiol molecule (green) is slightly further away from the metal but the angle of approach is similar to the diatomic.

A crystal structure has been obtained for the active site of galactose oxidase containing Cu^{1+}, allowing the first observation of the fully reduced enzyme. The galactose oxidase protein was expressed from the BD1 construct, isolated and crystallised under copper free-conditions two years prior to data collection. Structure determination, however, revealed the presence of a thioether (Figure 3.13). This discovery suggests the possibility of thioether bond formation in the absence of copper. There is no evidence, however, supporting this in solution studies and it is possibly that contaminating copper may have entered the protein crystals and initiated thioether biogenesis in all the enzyme molecules over a two year period. A small electron density peak adjacent to Tyr 272 was modelled as a water molecule, although the possibility of the density representing a small proportion of copper contamination in the enzyme population can not be ruled out.

Incorporation of Cu^{1+} into the protein was achieved by anaerobically incubating the enzyme with 5 mM $Cu(I)PF_6$. A 5 minute and 25 minute incubation revealed similar active sites in both structures, where copper was coordinated in a trigonal geometry to Tyr 272, His 581 and His 496. Tyr 495 was found at least 3.1 Å away from copper and thus was not a copper ligand. The Tyr 272-copper distance was the longest of the copper bonds (at least 2.5 Å) in both structures. This was a surprise as Tyr 272 is usually a strong copper ligand. The structure may represent a mixture of Tyr 272 populations comprising protonated and deprotonated species, with protonated Tyr 272 disassociating from copper. The Cu^{1+} containing structures look remarkably similar to the fully reduced galactose oxidase active site structure predicted by Whittaker *et al.* [1993]. This form of the enzyme is proposed to exist after substrate oxidation (Figure 1.17) and during thioether bond formation [Whittaker and Whittaker, 2003] (Figure 1.26).

Determination of the galactose oxidase structures presented in this chapter has provided some "snap shots" of the catalysis of the enzyme. 2-methylene-1,3-propandiol crystal soak experiments have provided the first structural confirmation of an alcohol substrate at the fourth equatorial ligand site and the visualisation of a peroxide-like molecule coordinated to copper. The copper soak experiments have also revealed additional copper binding sites in galactose oxidase, including N-terminal coordination. Whilst previous copper soak experiments have reported copper association with Cys 383 [Firbank, 2002], observation of copper binding at the N-terminus is completely novel for galactose oxidase and could be related to pro-sequence cleavage (for further discussion see chapter 5).

4 Structural studies of galactose oxidase W290G variant

4.1 Introduction

The Trp 290 residue is found in the active site of galactose oxidase, positioned directly over the thioether bond [Ito *et al.*, 1990]. The close proximity of Trp 290 to the thioether bond has led to many speculative theories concerning the role the residue performs for the enzyme. An early theory suggested Trp 290 could act as catalytic base during substrate turnover [Ito *et al.*, 1990; Ito *et al.*, 1991]. Under this proposal, the substrate would bind to the copper in the active site and be activated by the transfer of the hydroxyl proton to the Trp 290 residue. This idea was superseded when mutant galactose oxidase studies showed that proton uptake associated with wild-type catalysis was eliminated when the tyrosine at the 495 position was mutated to a phenylalanine (Y495F) [Baron *et al.*, 1994]. It was clear from this result that Y495 was the catalytic base as previously suggested [Baron *et al.*, 1994; Whittaker *et al.*, 1989].

Manual docking experiments with galactose were carried out upon the release of the first galactose oxidase structure [Ito *et al.*, 1992; Ito *et al.*, 1994] and further modelling and refinement has been achieved since [Wachter and Branchaud, 1996a]. The studies showed that a D-galactose molecule in its chair conformation could be modelled into the active site, producing complementary and favourable interactions between enzyme and substrate (Figure 1.12). One half of the active site pocket provided hydrophobic interactions for the galactose, while the other half suggested possible hydrogen bonding interactions via Arg 330, Gln 406 and the indole ring nitrogen of Trp 290. Hence, these studies suggested Trp 290 could have a primary role in the initial recruitment and coordination of sugar substrates in the active site [Ito *et al.*, 1995].

The presence of the thioether bond between Tyr 272 and Cys 228 causes the reduction of tyrosine redox potential from the 940 mV of a free tyrosyl radical to 430 mV [Harriman, 1987]. There is evidence that the stacking interaction of Trp 290 is responsible for lowering the redox potential required to generate the radical state. Studies where Trp 290 was mutated to a histidine showed that the redox potential increased to 730 mV [Saysell et al., 1997]. Further support for the role of Trp 290 in enabling the thioether bond to lower the Tyr 272 redox potential is provided by the enzyme glyoxal oxidase.

Glyoxal oxidase (GLX), is a copper-enzyme secreted by the wheat-rot fungus *Phanerocheaete chrysosporium*, and has many similarities to galactose oxidase [Kersten and Cullen, 1993]. Although the two proteins share only 28% sequence identity, sequence

alignments suggest the GLX lacks the equivalent of the galactose oxidase domain I, but the key catalytic residues are conserved. Mutational studies have established residues equivalent to Cys 228, Tyr 272, His 496 and Tyr 495 (Cys 70, Tyr 135, His 378, Tyr 379 and His 471 respectively in GLX) [Whittaker *et al.*, 1999]. Like galactose oxidase, GLX also exists in an inactive form which requires one electron oxidation for activity. Spectral studies indicate the GLX active site has structural similarity to galactose oxidase, but with a histidine at the position equivalent to the Trp 290 of galactose oxidase [Whittaker *et al.*, 1999; Whittaker *et al.*, 1996b]. It has a tyrosine that corresponds to Tyr 272 in galactose oxidase, except the GLX tyrosine has a higher redox potential of 640 mv. It is not clear why the tryptophan in galactose oxidase stacking over the thioether bond appears to lower the tyrosyl redox potential more than the histidine in glyoxal oxidase. It has been speculated that as tryptophan is larger than histidine, the residue could provide more protection for the free radical against solvent [Liu et al., 2002]. There has been a suggestion that the free radical that exists on the Tyr-Cys cofactor could be delocalised from Tyr 272 to Trp 290, thus increasing the stability of the radical and increasing the ease of redox access. Indeed the two residues are located only 3.4 Å apart, suggesting the possibility of delocalisation, however, there is no evidence for this from EPR and ENDOR studies [Babcock et al., 1997; Gerfen et al., 1996].

There is currently no direct evidence that Trp 290 has a significant effect on galactose binding or catalytic turnover. This chapter presents a structural study of the W290G mutant and provides a view of the active site without the influence of tryptophan.

4.2 Results

4.2.1 Crystallisation and data collection

W290G mutant was purified and crystallised from *Aspergillus nidulans* by C. Kurtis, University of Leeds, using the methods outlined in chapter two. Crystal trays were set up using a variation of the native crystallisation conditions 1-2 M ammonium sulphate and 0.1-0.5 M sodium acetate, pH 4-5 [Ito *et al.*, 1991]. The crystals were flat plates and were either hexagonal or half hexagonal in morphology. Micro-crystals appeared after 5-7 days growth at 18°C (Figure 4.1). After 3-4 weeks crystals were no longer increasing in size and grew to a maximum of approximately 0.8 mm by 0.3 mm. The crystals were prepared for data collection by freezing to 100 K using mother liquor containing 25% glycerol as a cryoprotectant.

The data were collected using synchrotron radiation at Beamline 9.6, Daresbury SRS, Warrington, UK. Initial test exposures revealed the protein had crystallised into a previously observed $P6_522$ space group, with the unit cell dimensions comprising a long c axis (**a**=89.9, **b**=89.9, **c**=415.3, γ=120°) [Vinecombe, 1999]. The presence of a long c axis resulted in reflections that were extremely close together. The detector was moved further away from the crystal to enable the reflections to be adequately resolved. Although the crystal diffracted beyond 2 Å, data above 2.2 Å resolution could not be processed. Spot resolution was further aided by reducing the crystal oscillation angle to 0.1°. The statistics for data collection are shown in Table 4.1.

4.2.2 Structure determination

The W290G data set was processed using the HKL suite [Otwinowski and Minor, 1997], using the standard method described in chapter two. Overlapping spots along the c axis resulted in some reflections being lost in processing, leading to the lower completeness. Despite this, the six-fold symmetry of the space group allows calculation of electron density maps without significant distortion. The structure of W290G was solved by using the phases from the previously solved structure of a tyrosine 495 to lysine (Y495K) mutant [Vinecombe, 1999]. The phases from the Y495K mutant were applied to the W290G data. This process was possible because Y495K was isomorphous to the W290G $P6_522$ space group. The Y495K structure, with water, copper and active site residues removed was then used to begin the W290 modelling. Active site residues were modelled back into the structure following analysis of the omit map using the program O [Jones *et al.*, 1991]. The resulting structure was put through the standard refinement procedure described in chapter two. The molecular model

for glycerol was obtained from HIC-Up (the hetro-compound information centre-Uppsala) [Kleywegt *et al.*, 2003; Kleywegt, 1998].

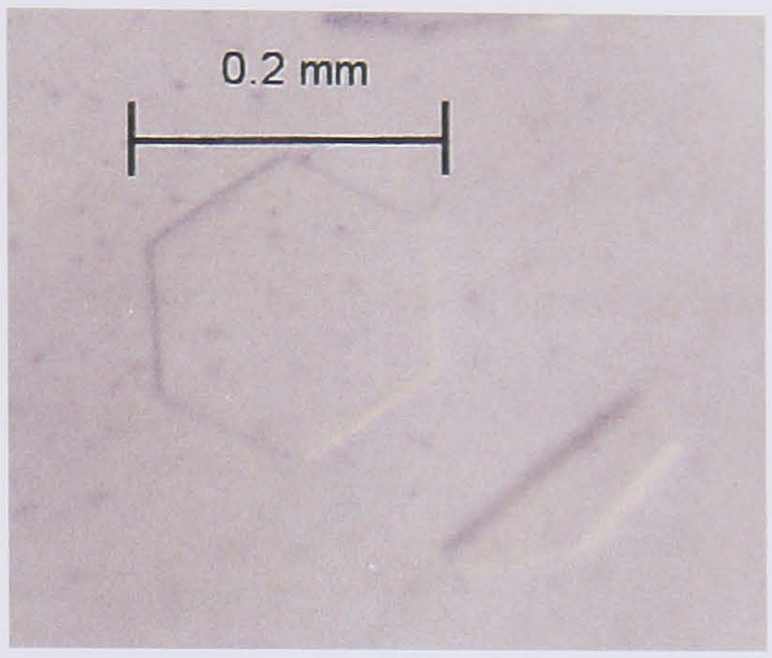


Figure 4.1: W290G crystals.

W290G crystals were grown in Ammonium sulphate and Sodium Acetate. The crystals formed after 5-7 days, and were full sized in up to four weeks.

Table 4.1: W290G data collection and refinement statistics (PDB code 2EID)	
X-Ray facility	Beamline 9.6, Daresbury SRS, UK
Crystallisation conditions	1.5 M Ammonium sulphate, 0.1 M sodium acetate pH 4.5
Cryoprotectant	Mother liquor, 25% glycerol.
Wavelength (Å)	0.86
Space group	$P6_522$
Unit cell dimensions	a= 89.9 Å, b= 89.9 Å, c= 415.3 Å α= 90°, β= 90°, γ= 120°
Number of observed reflections	33161
Number of unique reflections	12672
Resolution limits (Å)	10 - 2.2
R_{sym} (%)	7 (13)
Completeness of data%	64.6 (63.8)
Refinement statistics	
R_{cryst} (%)	19.1
R_{free} (%)	22.7
Number of atoms	5341
Number of protein atoms	4819
Number of water atoms	497
Rms bond lengths (Å)	0.013
Rms bond angles (°)	1.6

Figures in parentheses represent the highest resolution shell

4.3 W290G crystal structure

The mutation of W290 appears to have a minimal effect on overall protein structure, copper content or presence of the thioether bond. The 2Fo-Fc electron density map showed copper and the thioether bond were clearly present and there was an absence of density that could be attributed to a tryptophan residue (Figure 4.2). The absence of Trp 290 in the active site has produced a more open active site and the empty space generated was filled by additional electron density not directly linked to the protein. The majority of the electron density corresponds to water molecules, however, electron density immediately next to Tyr 272 did not refine well as water. Attempts to model acetate into the density were unsuccessful, as the acetate ion was too small. The only other exogenous molecule the crystals were exposed to was glycerol, which was used as a cryoprotectant. A glycerol and a water molecule were

modelled into the electron density over Tyr 272 (Figure 4.2). Two more glycerol molecules were identified in domain I and domain II, as well as the domain I sodium ion and three acetate ions, one in each domain (Figure 4.3). The model refined with a final R-factor of 19% and an R-free of 22%. The active site copper is coordinated by three active site ligands, Tyr 272 (2.1 Å), His 496 (2.2 Å), His 581 (2.2 Å), with Tyr 495 found occupying the axial position at a long bond distance (2.8 Å) (Figure 4.4). A water molecule (W1, Figure 4.4) is found near the fourth equatorial ligand position, but at a distance of 3.3 Å from copper and is probably not a copper ligand. Thus, the active site copper geometry is not square pyramidal, as described for the native enzyme [Ito *et al.*, 1990]. Glycerol is not a copper ligand, but forms hydrogen bonds with three active site water molecules (W1 (2.9 Å), W2 (2.9 Å), W3 (2.8 Å)) and is within hydrogen bonding distance of the hydroxyl oxygen of Tyr 272 (2.6 Å). The Cα of Gly 290 is in the same position as the wild type Trp 290 Cα, but the mutation of a bulky residue like tryptophan to the smaller glycine has serious consequences for access to the active site. Galactose oxidase has an exposed, solvent accessible active site pocket which is perfectly shaped to accommodate a sugar substrate such as galactose (Figure 4.5). The W290G active site pocket, however, leaves copper and the catalytic residues much more exposed, hence leaving the radical site open to solvent attack (Figure 4.6).

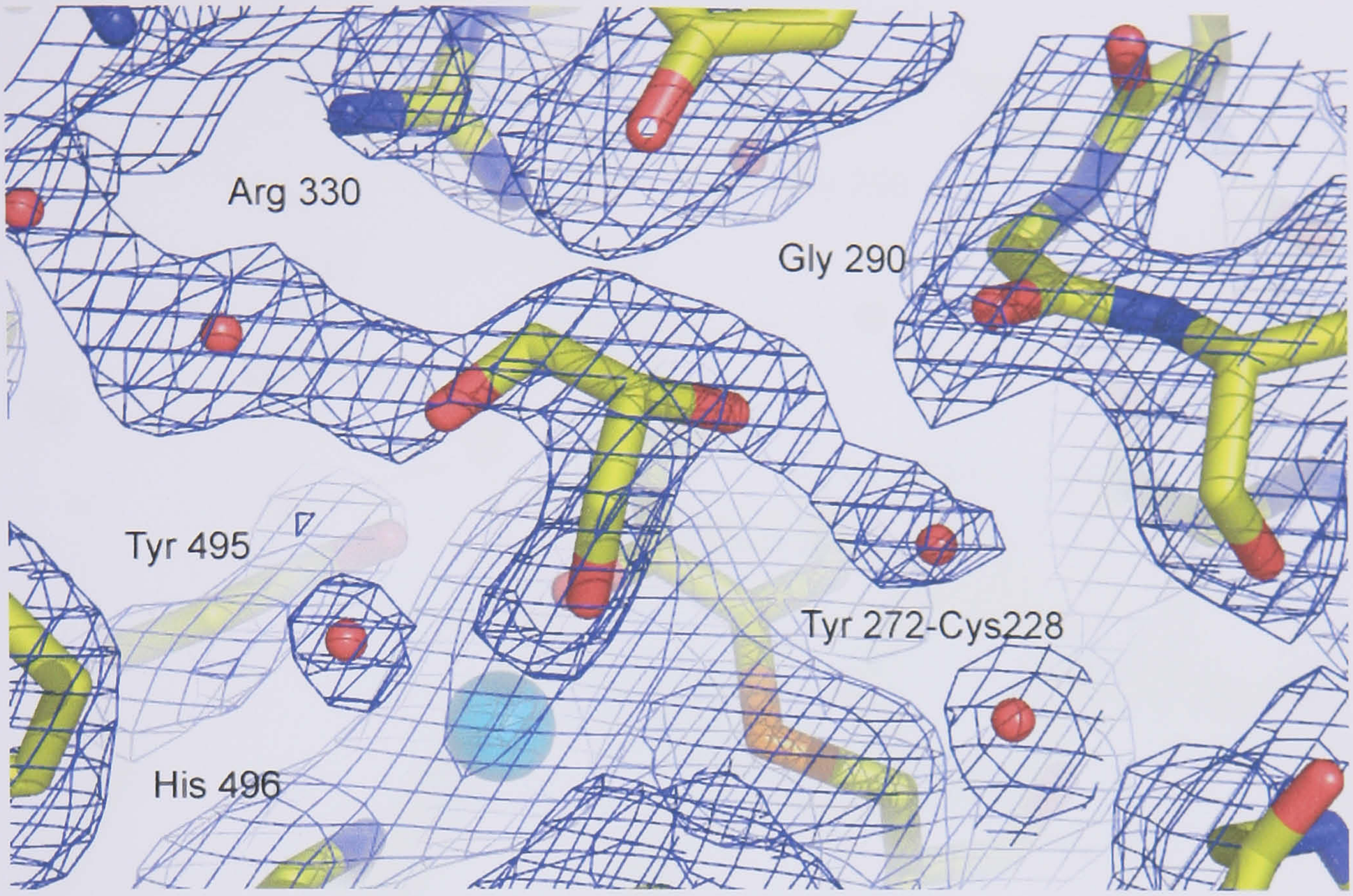


Figure 4.2: 2Fo-Fc map showing the mutation site at residue 290.

The 2Fo-Fc map electron density map is contoured at 1 x rms density, and shows density between Tyr 272 and Cys 228, indicative of thioether bond formation. A large area of density near Tyr 272 did not correspond to any protein residues, but refined reasonably well as glycerol and water.

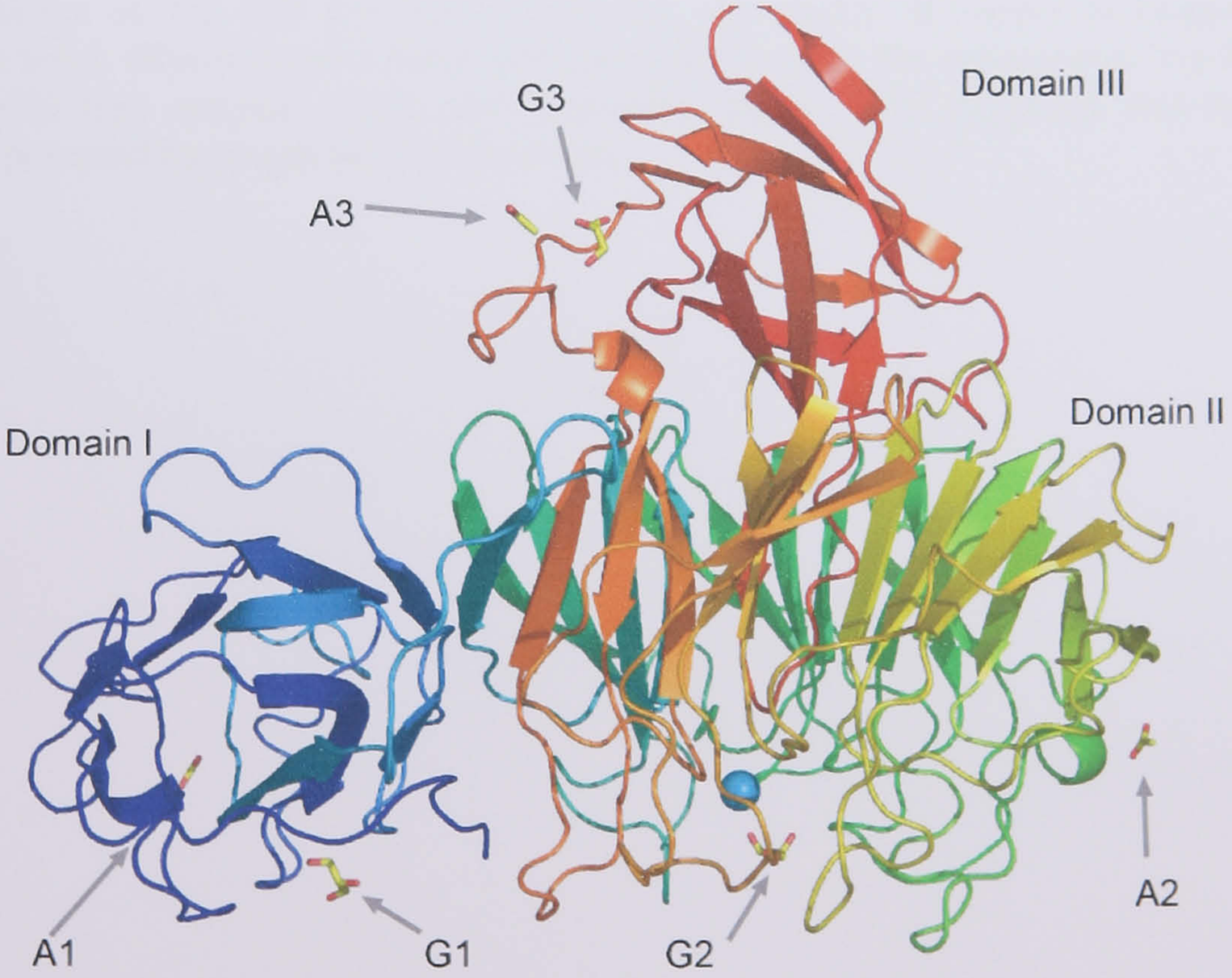


Figure 4.3: W290G glycerol and acetate binding sites.

Overall three acetate ions (A1-3) and three glycerol molecules (G1-3) were modelled into the structure.

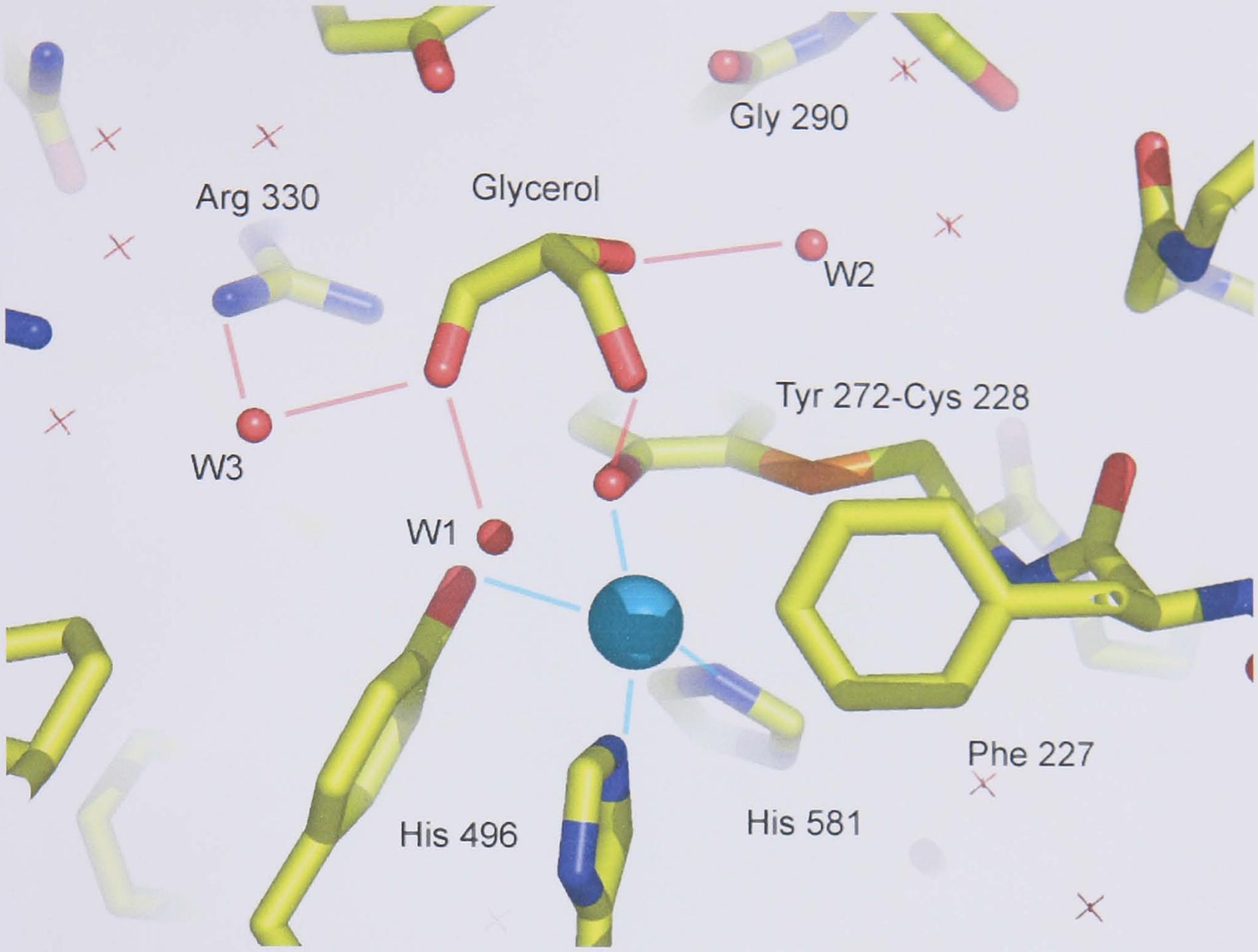


Figure 4.4: Galactose oxidase W290G active site.

The mutation of Trp 290 does not prevent the recruitment of copper or biogenesis of the thioether bond. Glycerol and a water molecule (W2) occupy the space where Trp 290 is found in the wild type enzyme. Cyan lines represent copper bond distances, and the red lines indicate potential hydrogen bond interactions.

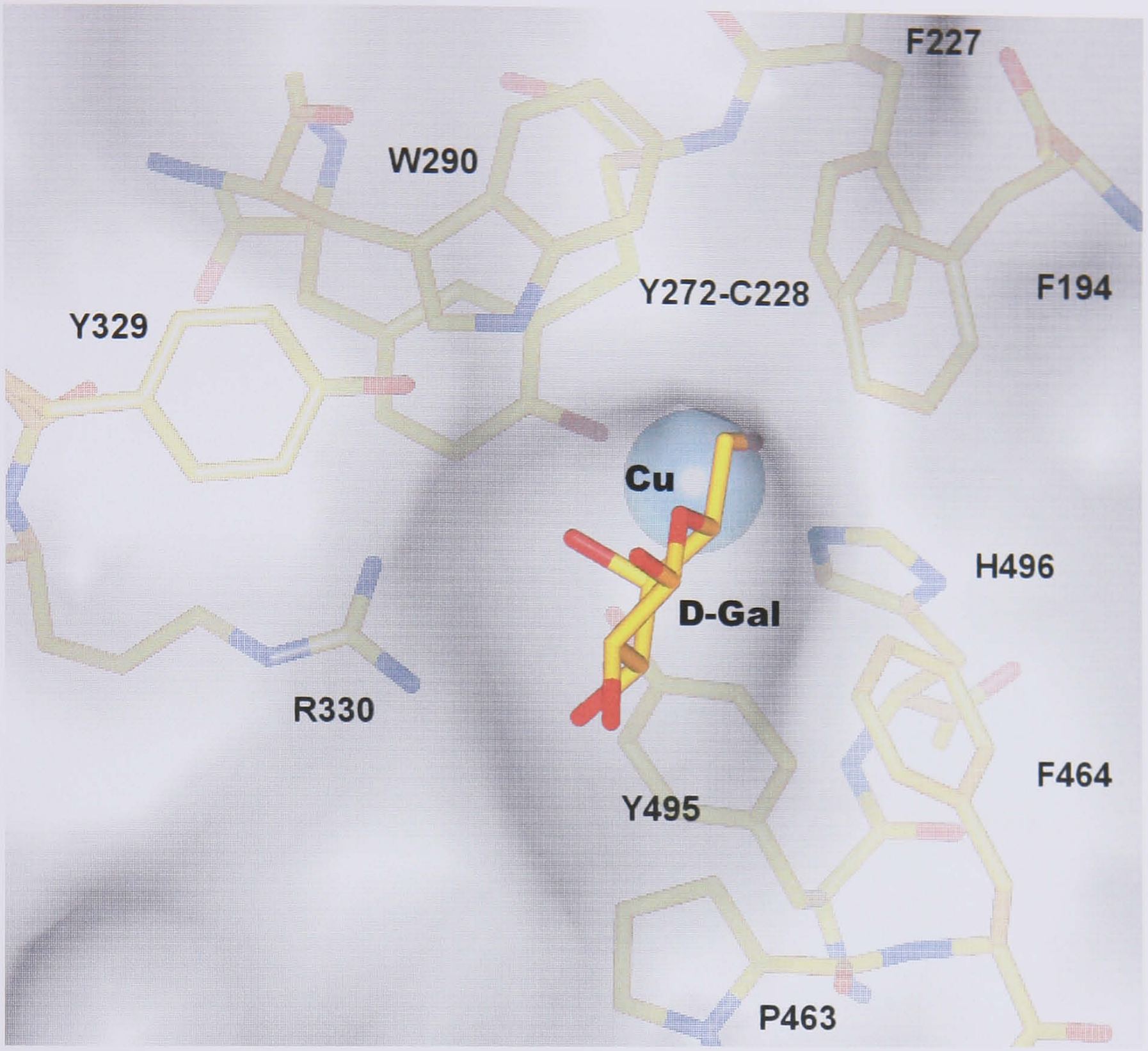


Figure 4.5: Solvent accessible active site pocket of wild-type galactose oxidase.

D-galactose is modelled into the pocket. The active site has a distinctly hydrophilic half and hydrophobic half. The hydrophilic half of galactose is predicted to form hydrogen bonds with W290, Y329, R330, as well as catalytic interactions with copper, Y272 and Y495. The hydrophobic resides F194, F227, F464, P463, and to a certain extent W290 are predicted to contribute hydrophobic interactions.

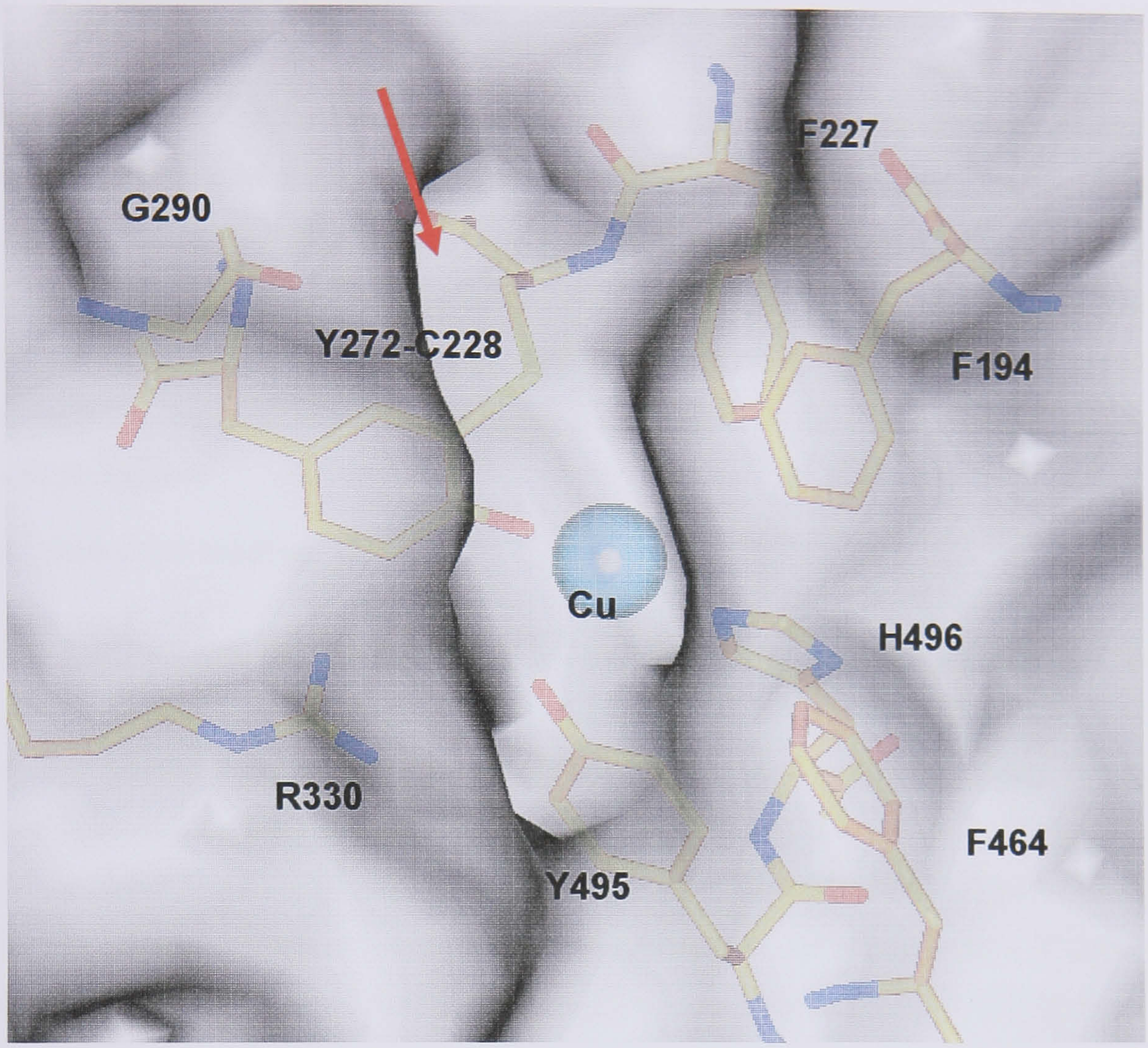


Figure 4.6: Solvent accessible active site of W290G.
The mutation of tryptophan to glycine has caused the shape of the active site pocket to be perturbed (red arrow), exposing the radical cofactor site to solvent attack.

4.4 Discussion

The W290G crystal structure has allowed the visualisation of the galactose oxidase active site without the presence of the tryptophan residue. Although the protein structure is essentially the same as that of the wild type protein, the mutant has a vastly reduced ability to oxidise galactose [Rogers *et al.*, 2007]. Examination of W290G, combined with previously solved crystal structures of two other mutants, W290F and W290H (PDB code 2EIC and 2EIB respectively), has proved useful in providing some insight on the role the tryptophan residue plays in the enzyme.

4.4.1 Substrate binding

In the W290H variant, the Trp 290 indole is conservatively replaced by a histidine imidazole [Rogers *et al.*, 2007]. The structure was solved to a resolution of 2.1 Å and revealed the presence of an active site acetate ion, which forms a hydrogen bond with the His 290 (Figure 4.7) and is coordinated to the active site copper. The wild type Trp 290 is also known to form a hydrogen bond to an active site acetate ion in the original native crystal structure [Ito *et al.*, 1991]. This mimics the interactions predicted for substrates, such as galactose (Figure 1.12). In W290H it is likely that the histidine residue can also aid substrate binding by forming a hydrogen bond between the substrate and the Nε2 of the imidazole ring, since structural studies show the imidazole nitrogen is in the same position as the wild type indol nitrogen (Figure 4.8). Kinetic data support the structural evidence, as it has been observed that the K_M for galactose in W290H is lower than in the wild type enzyme. In contrast to wild type and W290H, however, the variants W290F and W290G exhibit large K_M values, indicative of a greatly reduced ability to bind galactose (Table 4.2) [Rogers *et al.*, 2007].

Table 4.2: Galactose Oxidase mutant kinetics against galactose [Rogers *et al.*, 2007].

Galactose Oxidase structure	K_M (mM)	K_{cat} (s^{-1})
Wild Type	82	503
W290F	2950	371
W290G	1686	1.7
W290H	45	0.2

4.4 Discussion

The W290G crystal structure has allowed the visualisation of the galactose oxidase active site without the presence of the tryptophan residue. Although the protein structure is essentially the same as that of the wild type protein, the mutant has a vastly reduced ability to oxidise galactose [Rogers *et al.*, 2007]. Examination of W290G, combined with previously solved crystal structures of two other mutants, W290F and W290H (PDB code 2EIC and 2EIB respectively), has proved useful in providing some insight on the role the tryptophan residue plays in the enzyme.

4.4.1 Substrate binding

In the W290H variant, the Trp 290 indole is conservatively replaced by a histidine imidazole [Rogers *et al.*, 2007]. The structure was solved to a resolution of 2.1 Å and revealed the presence of an active site acetate ion, which forms a hydrogen bond with the His 290 (Figure 4.7) and is coordinated to the active site copper. The wild type Trp 290 is also known to form a hydrogen bond to an active site acetate ion in the original native crystal structure [Ito *et al.*, 1991]. This mimics the interactions predicted for substrates, such as galactose (Figure 1.12). In W290H it is likely that the histidine residue can also aid substrate binding by forming a hydrogen bond between the substrate and the Nε2 of the imidazole ring, since structural studies show the imidazole nitrogen is in the same position as the wild type indol nitrogen (Figure 4.8). Kinetic data supports the structural evidence, as it has been observed that the K_M for galactose in W290H is lower than in the wild type enzyme. In contrast to wild type and W290H, however, the variants W290F and W290G exhibit large K_M values, indicative of a greatly reduced ability to bind galactose (Table 4.2) [Rogers *et al.*, 2007].

Table 4.2: Galactose Oxidase mutant kenetics against galactose [Rogers *et al.*, 2007].

Galactose Oxidase structure	K_M (mM)	K_{cat} (mM/second)
Wild Type	82	503
W290F	2950	371
W290G	1686	1.7
W290H	45	0.2

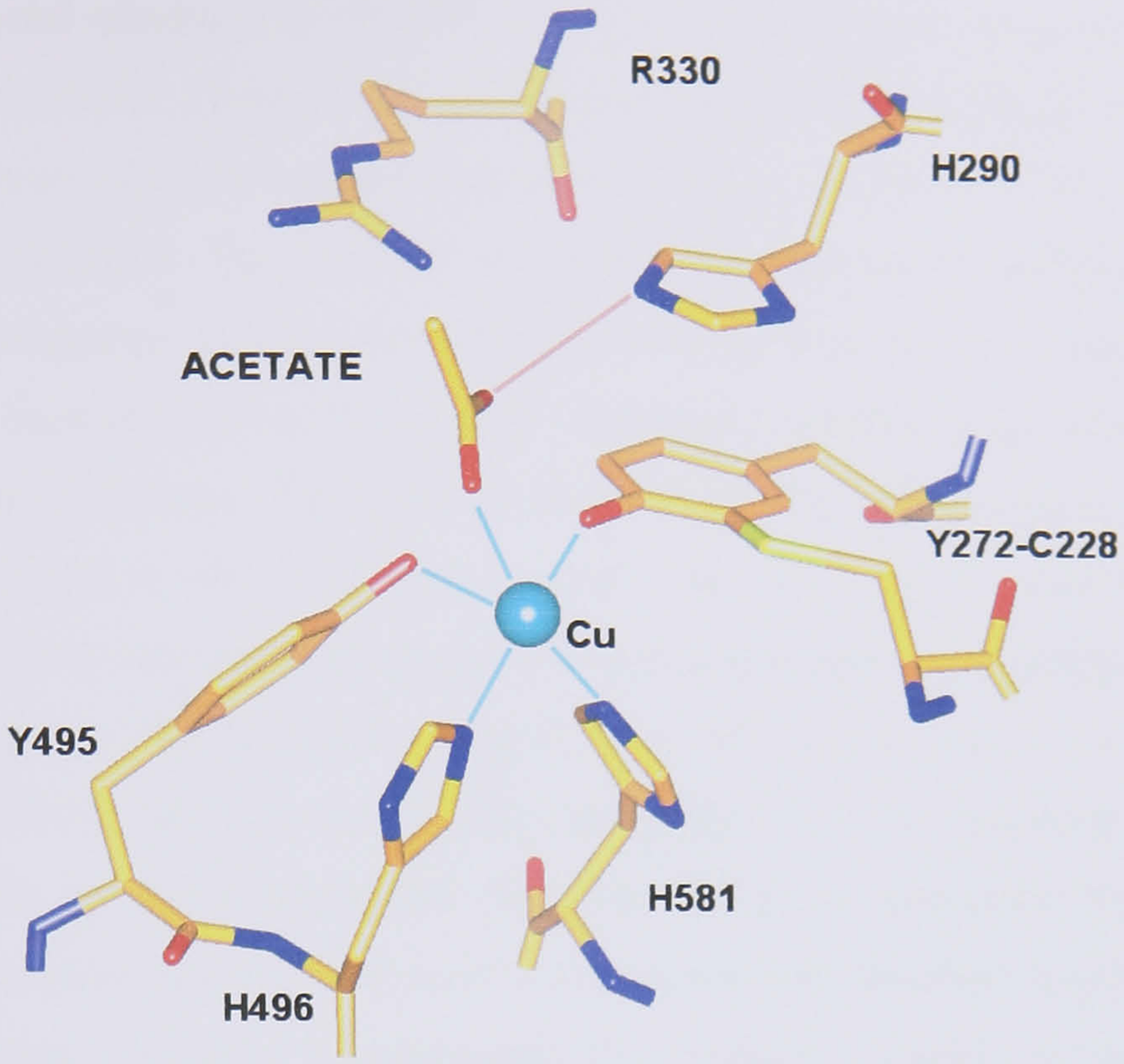


Figure 4.7: Active site of the W290H crystal structure (PDB=2EIB).

His 290 sits directly over the ring of Tyr 272 and forms a hydrogen bond with an acetate ion. Acetate occupies the fourth equatorial copper ligand position. Cyan lines represent copper coordination, red line indicate the hydrogen bond.

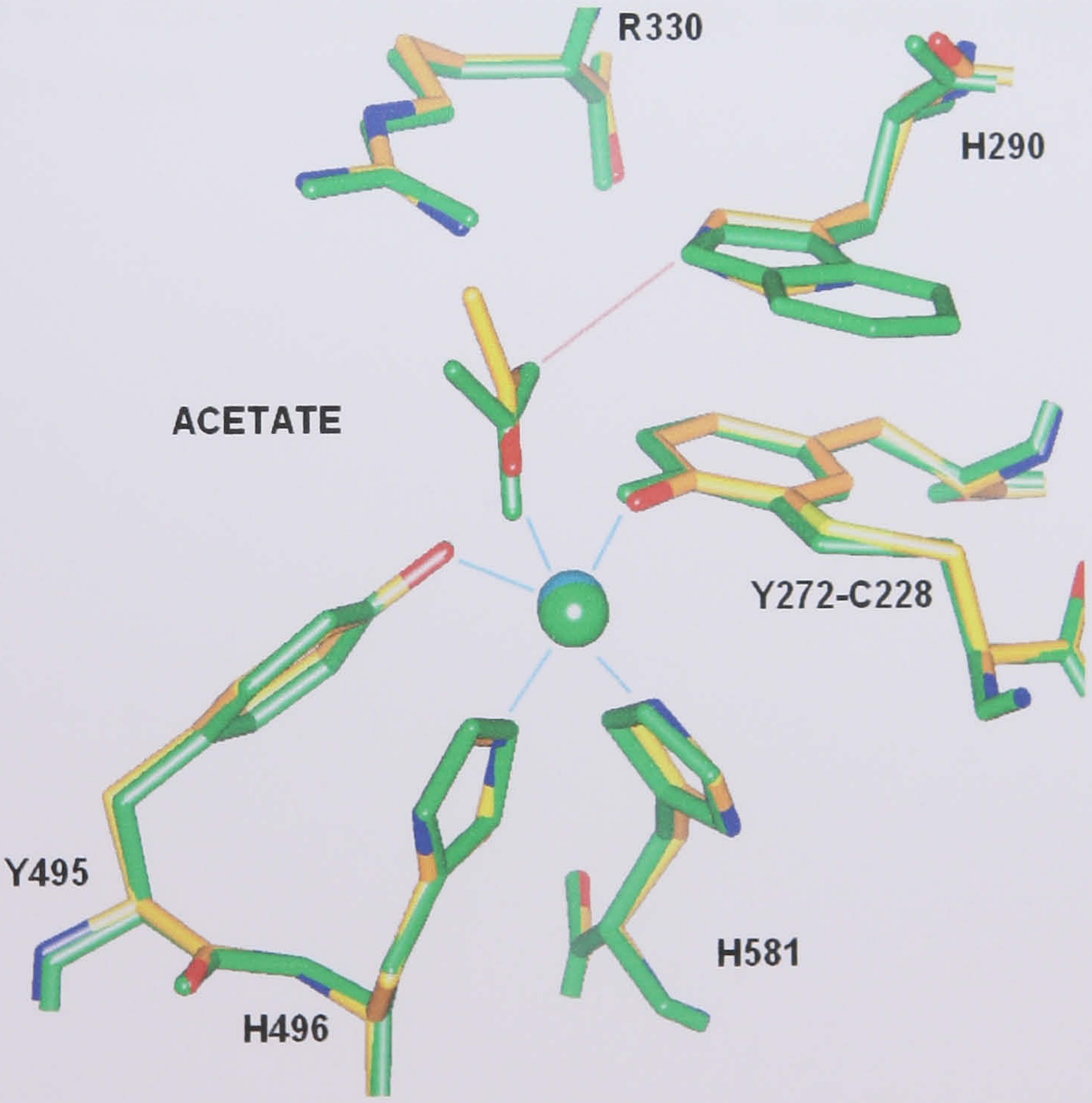


Figure 4.8: Alignment of the W290H with native, wild type galactose oxidase structure.

Atoms from the wild type structure are coloured green. The acetate ligand is not as close to the copper in W290H (PDB=2EIB) as it is in wild type (PDB=1GOF), but it is still able to maintain the same hydrogen bond interactions.

The 2.8 Å crystal structure of W290F was previously solved using 25% glycerol as a cryoprotectant, however, glycerol was not present at the active site (Figure 4.9) [Rogers *et al.*, 2007]. Furthermore, glycerol was not reported to be in the domain I or III locations identified in the W290G structure. The inability to observe glycerol in the W290F crystal structure could be a consequence of the lower resolution of the structure determination, where it is likely to have been difficult to distinguish confidently between a glycerol molecule and a cluster of water molecules. The benzene ring of the Phe 290 occupies the space where tryptophan is found in the wild type enzyme, the α, β and γ carbons of the mutant phenylalanine refine onto the same atomic position as the equivalent carbons of the wild type tryptophan (Figure 4.10). It is interesting to note that acetate was not a copper ligand in W290F and W290G, the variants showing the highest K_M for galactose. This observation provides support for early predictions that Trp 290 is an important residue in substrate binding. Other residues predicted to serve a role in substrate binding, based on the modelling studies, have been confirmed as important. For example mutation of Arg 330 results in substantially elevated K_M values (K_M for Galactose: R330A, 2240 ± 328 mM; R330K, 895 ± 86 mM), as does mutation of Phe 464 (K_M for Galactose: F464A, 937 ± 124 mM) [Deacon *et al.*, 2004]. Similarly, mutation of the active site residues R330K/Q406T/W290F was shown to result in 1000 fold reduced galactose oxidation activity, but glucose oxidase activity is detected [Sun *et al.*, 2002].

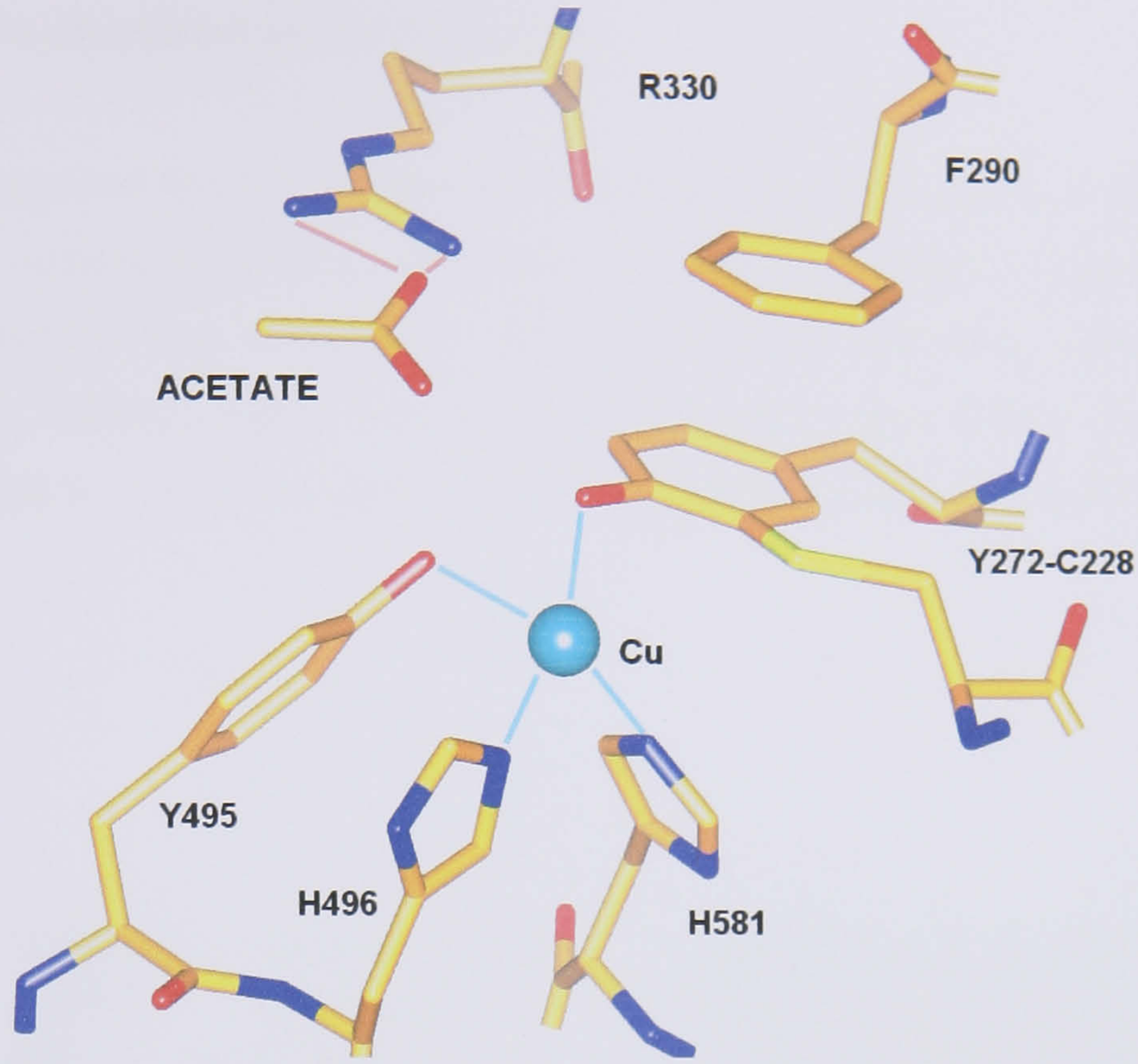


Figure 4.9: Active site of W290F crystal structure (PDB=1EIC).

The cyan lines represent copper coordination, and the red lines are potential hydrogen bond interactions. Phe 290 stacks over Tyr 272, but the hydrophobic nature of the benzene ring repels the acetate ion. There was no fourth equatorial ligand bound to the copper.

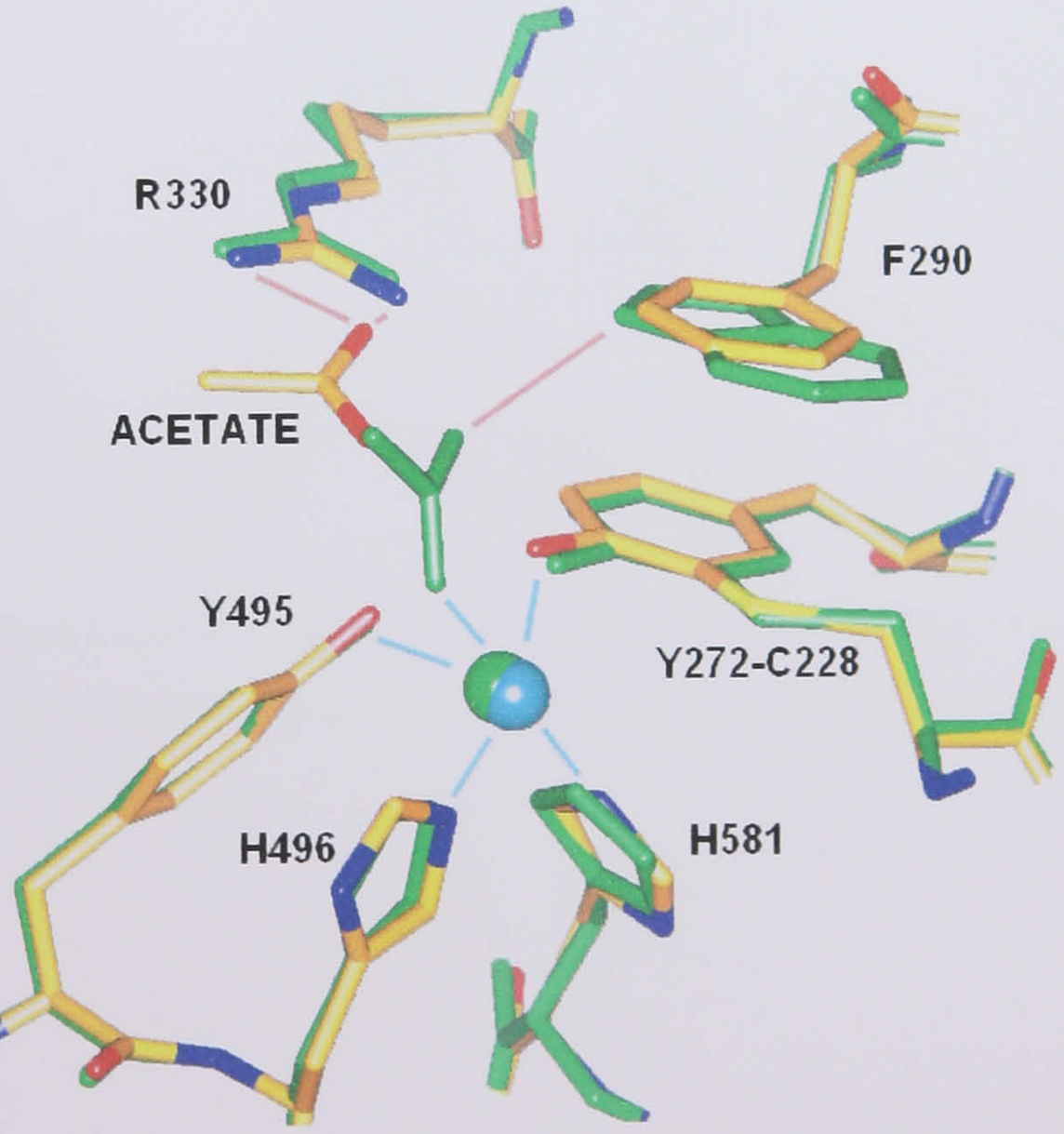


Figure 4.10: Alignment of the W290F active with native wild type galactose oxidase structure.

Atoms from the wild type structure (PDB=1GOF) are coloured green. There is no solvent derived copper ligand found at the active site of W290F (PDB= 2EIC). An acetate ion is near copper, but is interacting with Arg 330 and is over 5 Å away from the copper.

4.4.2 Active site solvent accessible surface

It has been suggested that Trp 290 may have an important role in substrate specificity [Saysell *et al.*, 1997], stabilising the free radical site via the stacking interaction of with the Tyr 272 ring and protection from solvent [Ito, 1991]. The solvent accessible surfaces of wild type (Figure 4.11), W290G (Figure 4.6), W290H (Figure 4.12) and W290F (Figure 4.13) have been examined to assess the difference in the thioether bond accessibility of the variant enzymes.

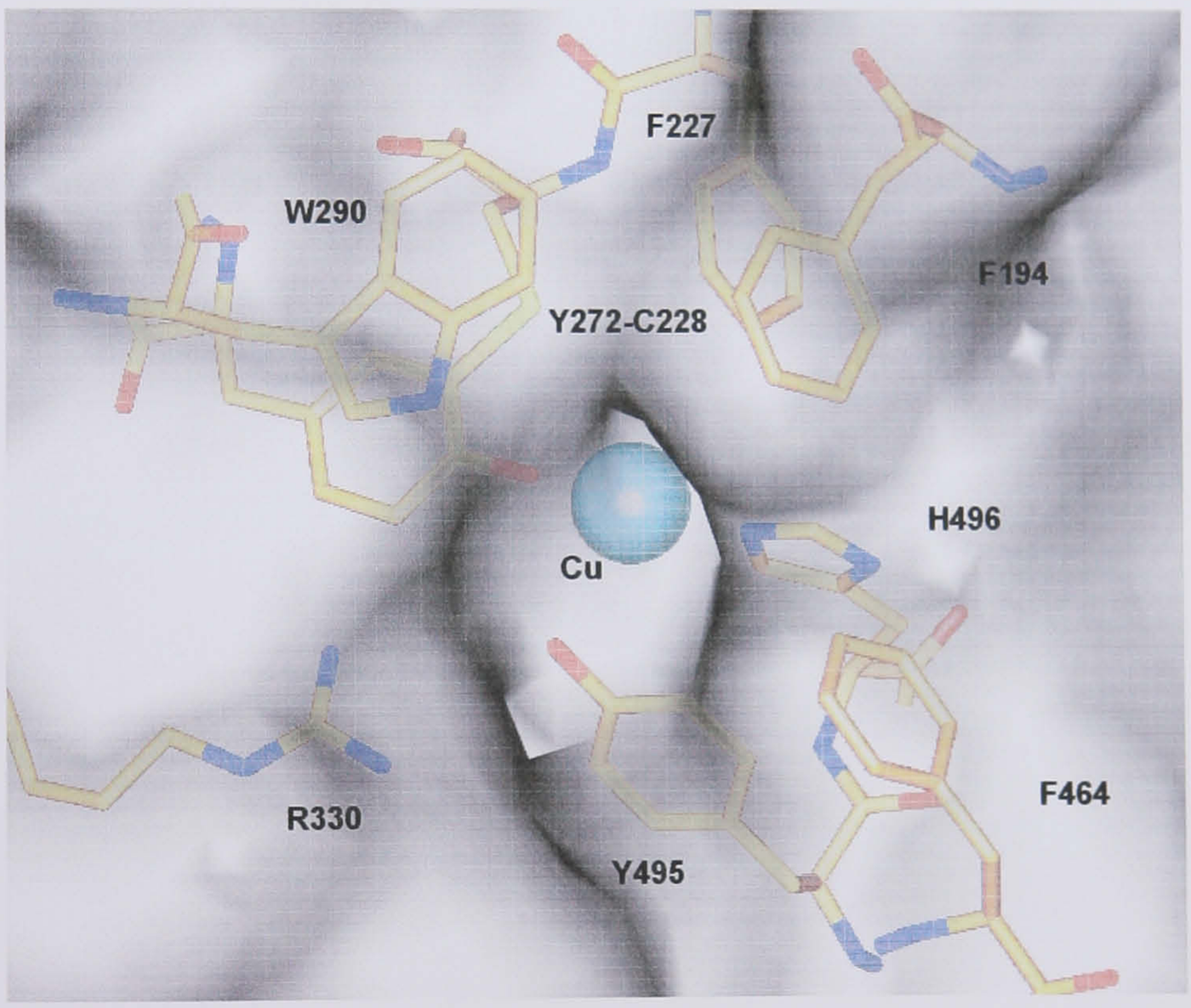


Figure 4.11: The solvent accessible active site surface of wild type galactose oxidase.

The tryptophan stacks over the Tyr 272-Cys 228 cofactor, which in inaccessible to solvent.

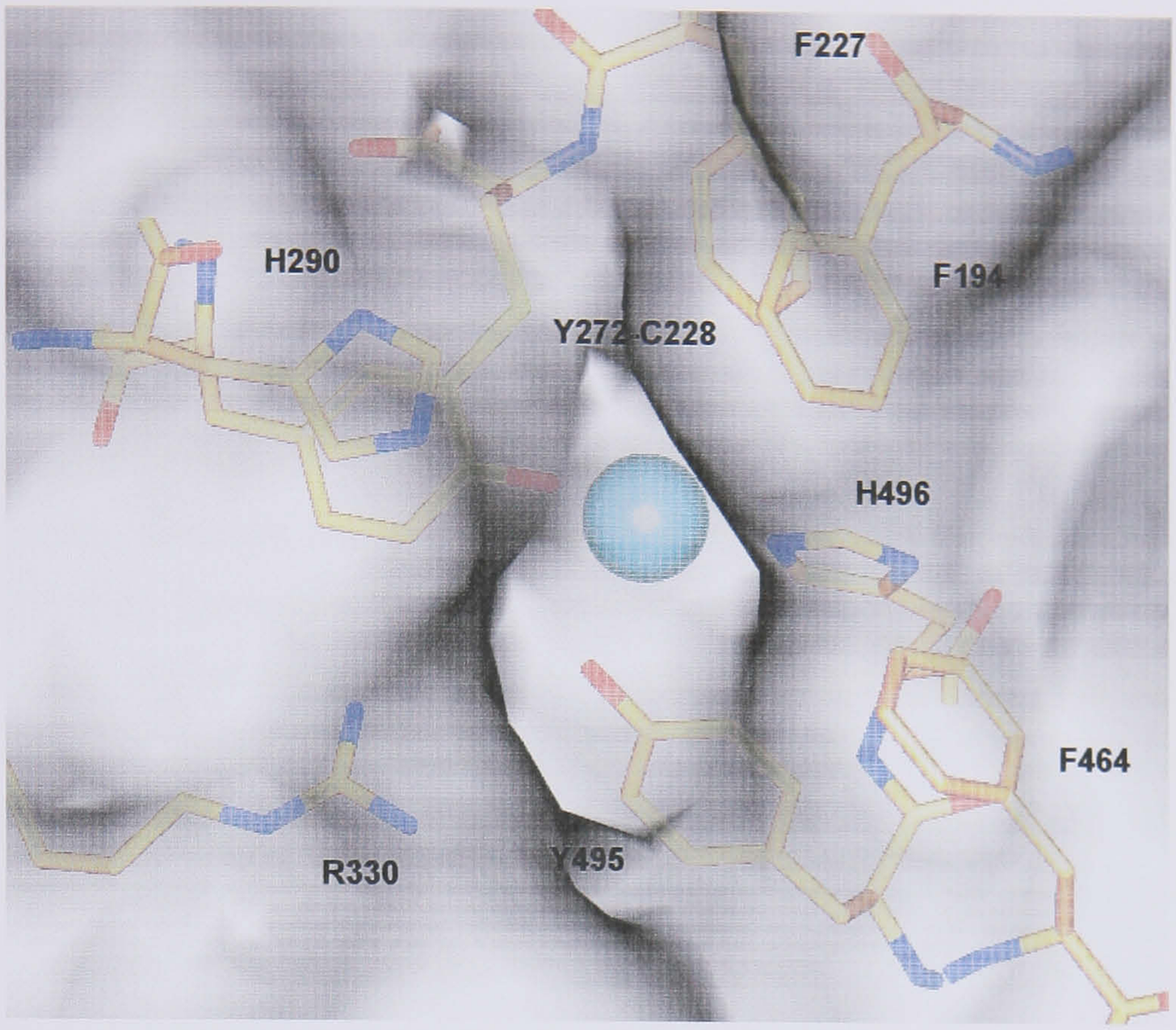


Figure 4.12: The solvent accessible active site surface of W290H.

The histidine stacks over Tyr 272, but the thioether bond is solvent accessible

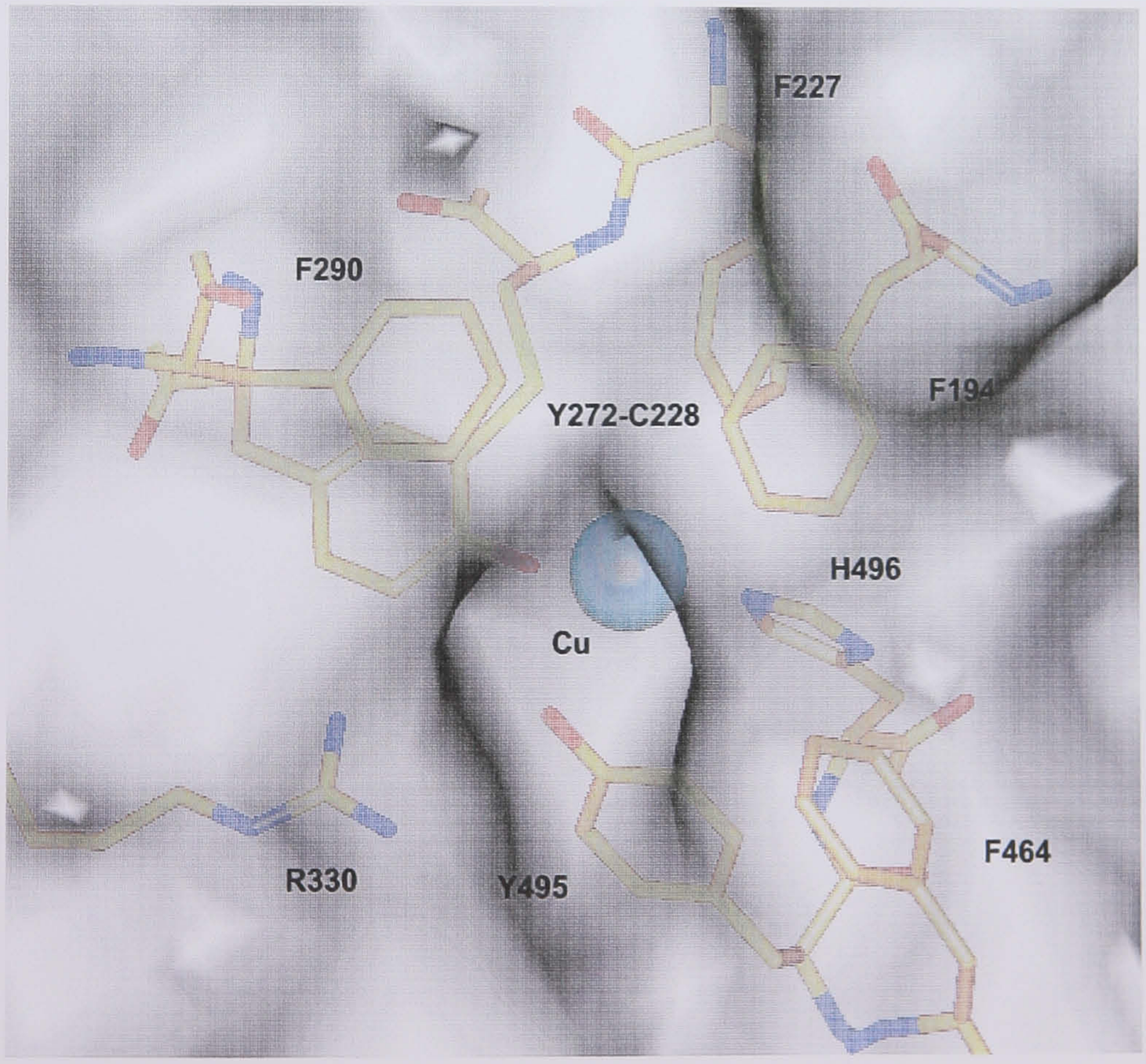


Figure 4.13: The solvent accessible active site surface of W290F.

The phenylalanine does not stack directly over Tyr 272, it occupies a space that appears to offer the thioether bond more protection

The active site pocket in W290G is more open to solvent compared to the wild type enzyme and the other Trp 290 mutants. The effect of the mutations on radical stability, however, is varied. Wild type galactose oxidase and W290F can be oxidised to the radical form of the enzyme using the mild oxidant potassium ferricyanide, while oxidation of W290H and W290G requires a stronger oxidant, caesium octacyanomolybdate(V) [($Cs_3[Mo(CN)_8]\cdot 2H_2O$: Rogers *et al.*, 2007]. Furthermore, the decay of the W290H radical is more rapid than observed in the other mutants. The hydrophilic nature of histidine seems to be favourable for substrate binding, however, the radical is quickly lost, thus explaining the poor catalytic activity for W290H. The reason for radical instability is not immediately clear, although it is possible that the size of the histidine is not sufficient to protect the thioether bond from solvent exposure or the hydrophilic nature of the residue adversely affects radical stability. The latter theory is supported by results showing the W290F radical is as stable as the wild type enzyme [Rogers *et al.*, 2007]. The current evidence indicates that the hydrophobic nature of Phe 290 contributes to the high K_M for galactose, but it may also act as a repellent to solvent attack and thus protect the radical site from rapid decay.

The structural and kinetic investigation of the current W290 mutants has effectively enabled the exploration and dissemination of the amphoteric nature of Trp 290 in galactose oxidase. It appears His 290 is capable of forming hydrogen bonds to anions such as acetate, an ability that appears to be necessary for binding of galactose to wild type galactose oxidase. On the other hand, Phe 290 is better at maintaining radical stability and so the W290F variant has a higher catalytic turn-over rate compared to the W290H variant. In the wild type enzyme Trp 290 efficiently combines both substrate binding and the radical stabilising effects.

5 Structural studies of galactose oxidase cofactor processing

5.1 Introduction

Galactose oxidase contains a free radical stabilising cofactor, derived from a covalent cross-link between Cys 228 and Tyr 272. When the presence of the protein derived cofactor was identified in the first crystal structure of the protein [Ito *et al.*, 1991], such post-translational modifications were extremely rare, although now there are many reports of similarly modified redox active amino acids [Okeley and van der Donk, 2000]. The processes by which they are generated, however, are not fully understood in most cases and are a subject of great interest.

This chapter investigates Cu^{2+}-dependent galactose oxidase cofactor biogenesis, using premature galactose oxidase. Thioether bond formation has been investigated in the past by soaking copper into premature galactose oxidase crystals under anaerobic conditions for 4 minutes or 24 hours (see Chapter 1, section 1.5.8). The resulting observation of Cys 228 interaction with copper prior to thioether bond formation led to a new mechanistic proposal (Figure 1.25). The focus of this investigation is to trap cofactor processing under anaerobic and aerobic conditions within two minutes after exposure to copper, with the hope of identifying further snap-shots of the mechanistic intermediates. The time frame was selected after solution studies suggested cofactor biogenesis was complete, in the majority of protein, two minutes after aerobic exposure to copper (as judged by SDS PAGE analysis, see Figure 1.22).

5.2 Results

5.2.1 Introduction

Galactose oxidase was expressed in *Aspergillus nidulans* and purified under strict metal-free conditions by Dr M. Rogers, Montana State University, using the procedures described in chapter two. Prior to crystallisation all plastic-ware were soaked in EDTA (0.1 M), glass-ware were acid washed using a mixture of nitric acid/Sulphuric acid (1M), solutions were treated with Chelex-100 resin and the protein crystallised from PIPES buffer containing 50 mM of the copper chelator sodium diethyldithiocarbamate in order to maintain the protein in the pro-GO form. Crystal screens were set up around the standard copper-free crystallisation conditions described in chapter two. A few grains of CHELEX resin were placed into the mother liquor solution in the bottom of each well of the crystal trays and left to mix on a

rocker for 20 minutes. These protein and crystal preparation conditions have previously been successful in the production of the crystallised pro-GO enzyme [Firbank *et al.*, 2001].

5.2.2 Structure of galactose oxidase expressed in copper-free conditions

The optimum crystallisation conditions were 14% PEG 8000, MES pH 5.7 and 200 mM calcium acetate at 18°C using the sitting drop method of vapour diffusion. Small plate crystals were obtained after four weeks with triangular tipped morphology. The crystals had grown to 0.2 – 0.5 mm in length and 0.1 – 0.2 mm in width (Figure 5.1). After 12 weeks the largest crystals were approximately 0.5 – 0.7 mm in length and 0.3 – 0.5 mm in width.

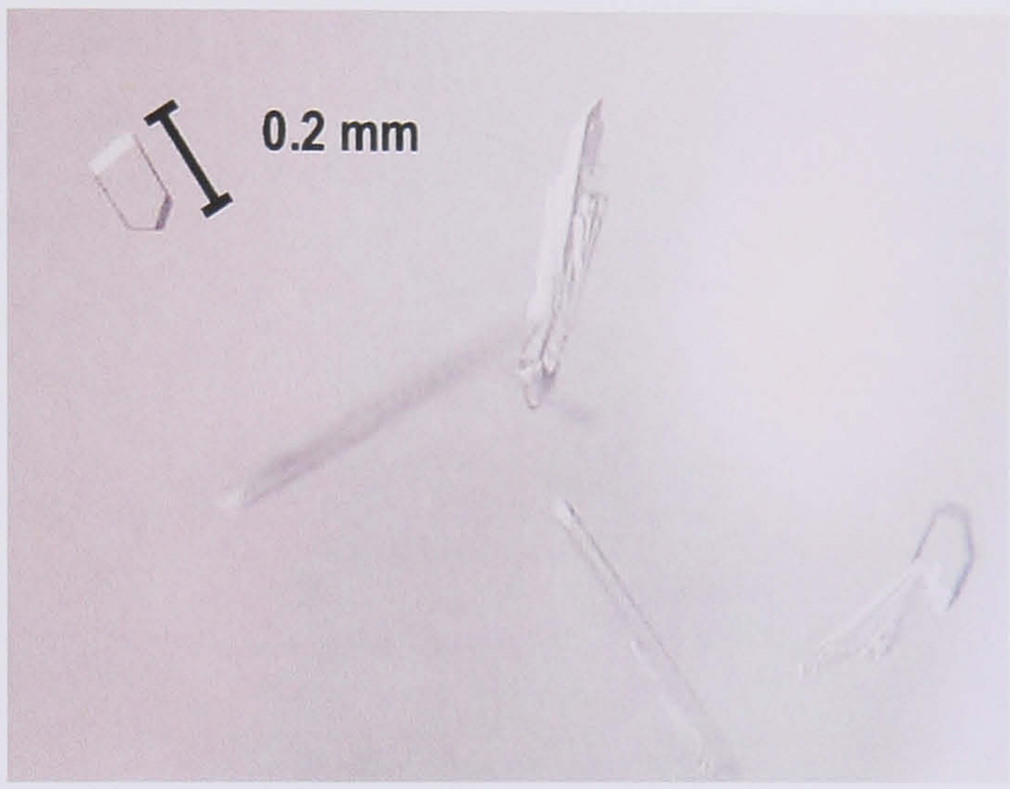


Figure 5.1: Crystals of galactose oxidase from copper free *Aspergillus nidulans* expression.

The triangular tipped crystals had similar morphology compared to the crystals obtained from the C383S mutant protein. Rectangular rod shaped crystals were also identified but not used for structure determination.

A crystal was briefly placed in cryoprotectant comprising mother liquor and 20% PEG 400, and then flash frozen in liquid nitrogen. Diffraction data from a single crystal were collected at 100 K using beamline 14.1, Daresbury SRS, with a wavelength 0.98 Å and equipped with an ADSC quantum IV CCD detector. The data were processed using MOSFLM [Leslie, 1992] and scaled with SCALA [CCP4, 1994], using the standard procedures described in chapter two.

The crystals obtained physically resembled the C2 crystals reported for the original native structure [Ito *et al.*, 1991; Ito *et al.*, 1994] and C383S (chapter 3). Pro-GO crystals grown in the past, however, were reported to form rods with hexagonal tips, belonging to the space group $P2_12_12_1$ [Firbank *et al.*, 2001]. A crystal diffracted to a resolution of 1.6 Å with 93% completeness (Table 5.1). Data indexing confirmed the crystals to be in the C2 space group, and isomorphous with the original native crystal structure determined by Ito [1991]. Initial

phases for structure determination were obtained from the native structure with all non-protein atoms and residues in the copper coordination sphere removed.

Table 5.1: Data collection parameters and statistics for copper-free premature galactose oxidase expressed in *Aspergillus nidulans*.	
Crystallisation conditions	14% PEG 8000, 100mM MES pH 5.7, 200mM calcium acetate.
X-ray facility	Station 14.1, Daresbury SRS, UK
Wavelength (Å)	0.98
Space group	C2
Cell dimensions	**a**=97.5 Å, **b**=89.0 Å, **c**=85.9 Å, **β**=117.6°
Resolution range (Å)	76.6 – 1.6
No. of observed reflections	239596
No. of unique reflections	78156
Mean I/σI	8.9
Completeness (%)	93.0 (93.0)
R_{sym} (%)	4.6 (27.7)
Multiplicity	2.8 (2.0)
Refinement statistics	
Refinement resolution (Å)	76.6 – 1.7
R_{cryst} (%)	17.3
R_{free} (%)	20.6
Number of atoms	5308
No. of protein atoms	4829
No. of water molecules	473
Average overall B-factor ($Å^2$)	16.9
Rms deviation bonds (Å)	0.019
Rms deviation angles (°)	1.7

Figures in parentheses denote highest resolution shell

The Fo-Fc difference and 2Fo-Fc electron density maps revealed copper was not present at the active site and the thioether bond had not formed. Examination of the N-terminal region of the structure, however, revealed no evidence of the pro-sequence (Figure 5.2). Therefore, the protein appeared to be the premature form of the enzyme (no pro-sequence, no thioether bond and no copper), and shall be referred to as such for the remainder of this chapter.

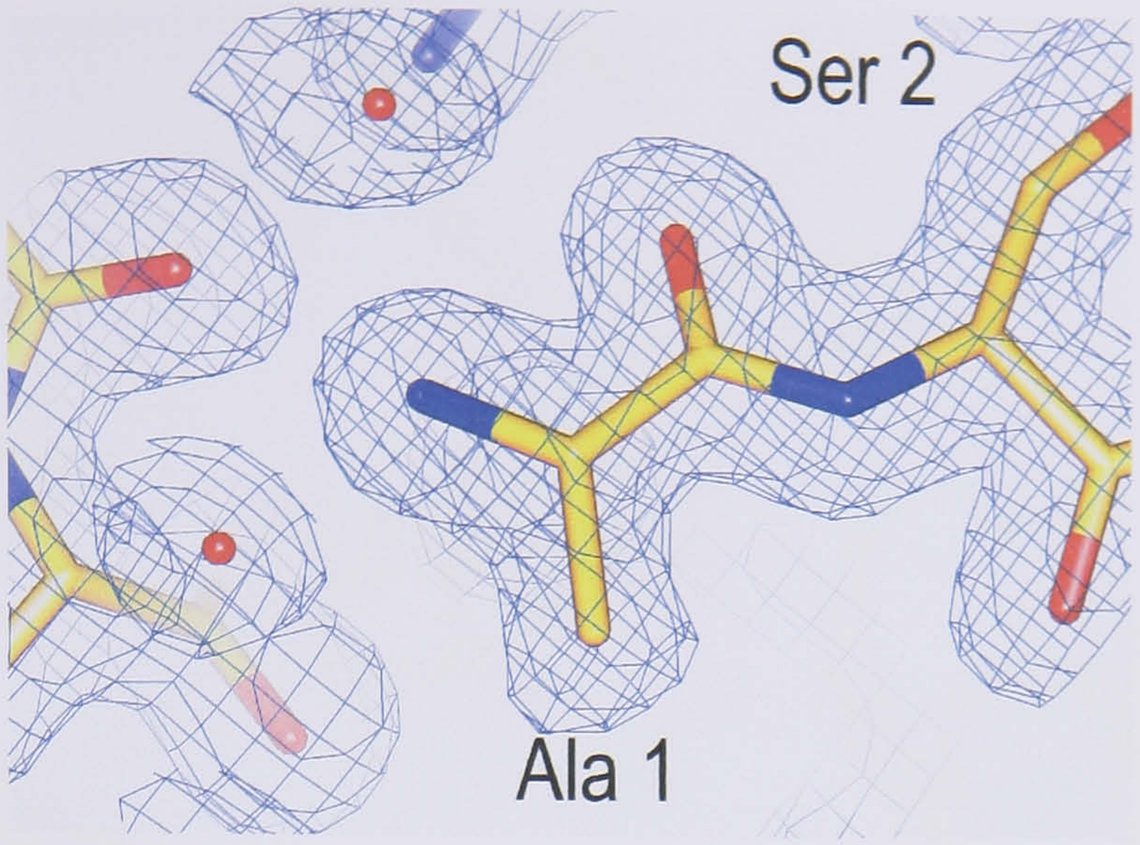


Figure 5.2: 2Fo-Fc map of the N-terminus of copper-free premature galactose oxidase.

The 2Fo-Fc electron density map of copper-free premature galactose oxidase. There is no indication of the presence of the pro-sequence at the N-terminus of this protein, in fact the density around Ala 1 is very well defined.

Examination of the active site revealed Cys 228 is in a conformation similar to that observed in the pro-GO structure (Figure 1.20). There is no evidence of oxidative modification of the residue as reported in the pro-GO structure [Firbank *et al.*, 2001]. Residue Trp 290 was particularly difficult to model as it appeared to be in a mixture of conformations, including the mature and the pro-GO conformations. The electron density for residues 290-295 is not very strong, particularly around Gly 292 and Gly 293. The inherent flexibility of the adjacent glycine residues allows the polypeptide loop comprising Trp 290 to be mobile, hence the disorder observed in the electron density. Trp 290 was built into the mature conformation, however, modelling of a second conformation for Trp 290 was unsuccessful due to the close proximity of the 290-295 loop to a symmetry related molecule next to this region. Steric hindrance caused by interaction between the flexible loop regions appears to be the cause of the difficulties encountered when attempting to model an alternate conformation for Trp 290. Similarly, modelling of Tyr 272 into the mature conformation resulted in poor Fo-Fc electron density, suggesting this residue was not well ordered (Figure 5.3). Two conformations of the Tyr 272 side chain were subsequently modelled at 50% occupancy, while Trp 290 was modelled at 50% occupancy into the mature conformation (Figure 5.4). It is possible that Trp 290 is being forced into the mature conformation by the crystal packing in the C2 space group and as a result Tyr 272 is forced to adopt two conformations. The position of the remaining active site residues have not been significantly altered by the lack of copper or cofactor. The model refined well, with a final R-factor of 17% and a free-R of 21%.

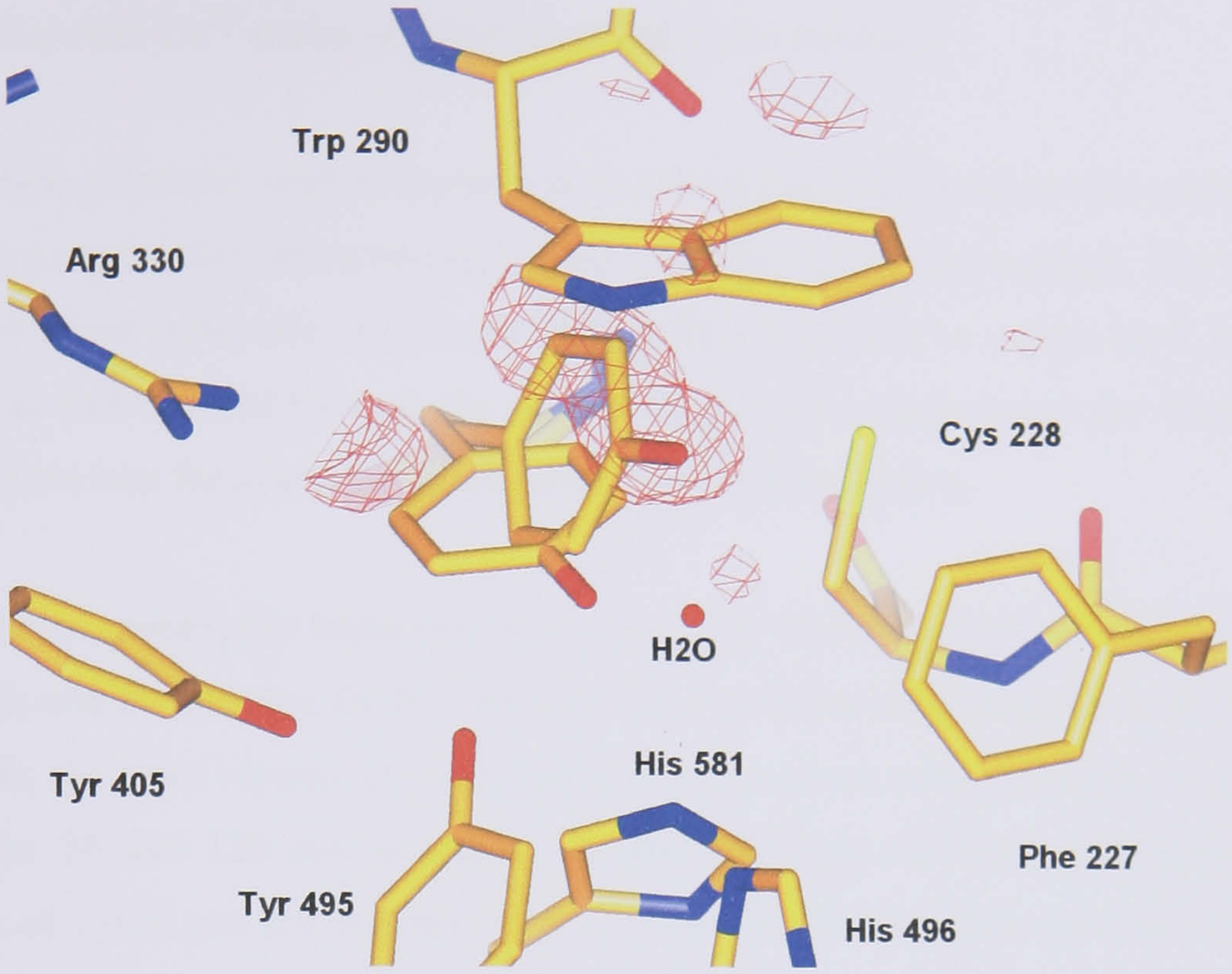


Figure 5.3: Fo-Fc difference map at the active site of copper-free premature galactose oxidase.

The Fo-Fc difference density map at the active site of copper-free premature galactose oxidase contoured at 3 x rms density. The positive density is coloured red. An alternative conformation for Tyr 272 was modelled into the structure to accommodate the positive density.

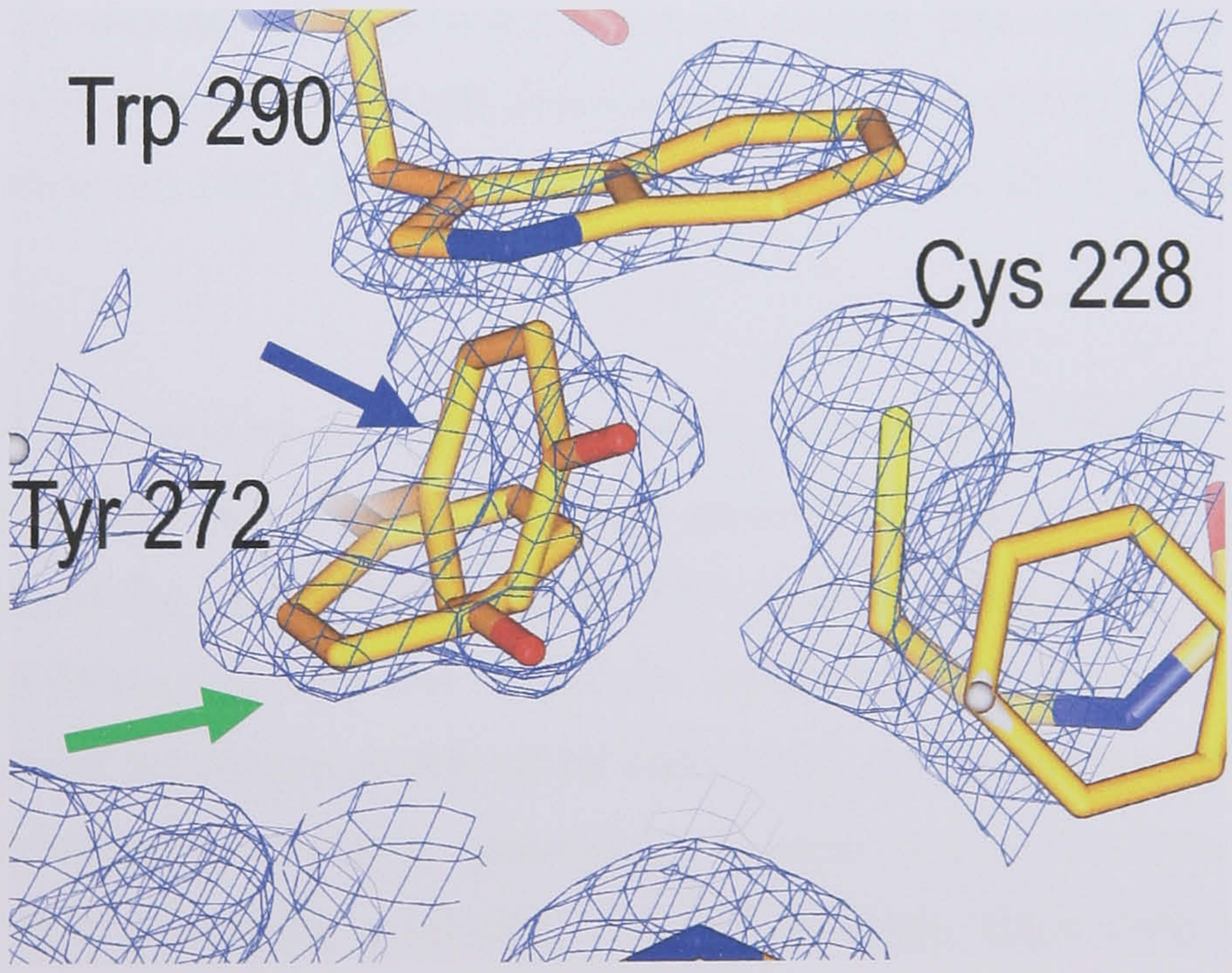


Figure 5.4: Refined active site 2Fo-Fc map of copper-free premature galactose oxidase.

The 2Fo-Fc electron density map at the active site of copper-free premature galactose oxidase after refinement and contoured at 1 x rms density of the map. Tyr 272 shows very clear density for two conformations, one similar to the mature position (green arrow) and the second closer to the pro-GO position (blue arrow). The density for Cys 228 is also extremely well defined, and is clearly in the premature conformation, while that of Trp 290 is weak.

5.2.3 Anaerobic Cu^{2+} soaks of premature galactose oxidase

The premature galactose oxidase crystals were introduced into an anaerobic environment, by transferring a crystal tray into a nitrogen filled glove box (Belle technology). The crystal trays were placed under anaerobic conditions at least 24 hours prior to copper soak experiments. Nitrogen gas was bubbled through the mother liquor and cryoprotectant for 10 minutes and left in the glove box for at least 24 hours prior to crystal incubation.

Crystals were anaerobically incubated with copper for 20 seconds, 47 seconds, 120 seconds, 270 seconds and 20 minutes. Diffraction data were collected from crystals soaked for 20 and 120 seconds, the other crystals unfortunately did not produce diffraction data of good quality. Data for the 20 and 120 second copper incubations were collected at Daresbury SRS to resolutions of 2.0 Å and 2.6 Å respectively (Table 5.2). The crystals used for the anaerobic copper soaks and the premature structure determination (described earlier in this chapter) were taken from the same crystallisation trays. Despite their identical crystal morphologies, initial X-ray analysis and indexing (space group P222 was initially chosen, pending full identification) revealed the unit cell dimensions of the copper-soaked crystals (**a**=78 Å, **b**=89 Å and **c**=107 Å) differed from previous galactose oxidase unit cells. The new unit cell dimensions were similar to the unit cell dimensions of pro-GO crystals (**a**=69 Å, **b**=90 Å and **c**=94 Å) [Firbank *et al.*, 2001]. It was therefore necessary to solve the structures by molecular replacement.

The structures were solved by molecular replacement in the program AMoRe [Navaza, 1994]. The data were initially processed in the space group P222 but analysis of the reflection intensities indicated the crystals belong to the $P2_12_12_1$ space group, since the alternate axial reflections were missing or very low. Molecular replacement was carried out on the space group $P2_12_12_1$ using the mature protein (PDB code 1GOF, with all exogenous atoms and the copper coordination sphere residues removed) as a search model. A solution was obtained which gave a high correlation coefficient and low R-factor. Rigid body refinement was applied to the best solution from the translation search and resulted in a solution with an R-factor of 36% and a correlation coefficient of 63% in the resolution range 9 – 4 Å. The same process was applied to the data obtained from the 120 second copper soaked crystals, resulting in a solution with an R-factor of 29% and a correlation coefficient of 76% in the resolution 9 – 4 Å (Table 5.3).

Table 5.2: Data collection and refinement statistics from anaerobic incubation with 20 mM copper acetate

Crystal incubation time	20 seconds	120 seconds
Crystallisation conditions	16% PEG 8000, 14% PEG 400, 100 mM MES pH 6.3, 200 mM calcium acetate.	16% PEG 8000, 14% PEG 400, 100 mM MES pH 6.3, 200 mM calcium acetate
X-ray facility	Station 14.1, Daresbury SRS, UK	Station 14.2, Daresbury SRS, UK
Wavelength (Å)	1.49	0.98
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions	**a**=77.6 Å, **b**=89.6 Å, **c**=107.6 Å, **α**=**β**=**γ**=90°	**a**=77.1 Å, **b**=89.1 Å, **c**=107.9 Å, **α**=**β**=**γ**=90°
Resolution range (Å)	27.0 – 2.0	63.2 – 2.6
No. of observed reflections	481239	193184
No. of unique reflections	50869	25540
Mean I/σI	3.6 (0.7)	7.1 (2.1)
Completeness (%)	98.7 (98.7)	100 (100)
R_{sym} (%)	12.7 (58.3)	10.6 (36.8)
Multiplicity	4.9 (4.6)	5.9 (5.9)
Refinement statistics		
Refinement resolution (Å)	27.0 – 2.1	63.2 – 2.6
R_{cryst} (%)	20.5	21.4
R_{free} (%)	23.1	21.6
Number of atoms	5337	5078
No. of protein atoms	4829	4929
No. of water molecules	453	240
Average B factor	25.6	27.9
Rms deviation bonds	0.01	0.005
Rms deviation angles	1.8	1.6

Figures in parentheses represent the highest resolution shell.

Table 5.3: Molecular Replacement rotation, translation and rigid body refinement statistics

Space group	Rotation Solution			Translation solution			R-factor (%)	Correlation coefficient (%)
	α^o	β^o	γ^o	Tx	Ty	Tz		
20 second anaerobic copper soak								
$P2_12_12_1$	125.5	70.9	345.5	0.07	0.35	0.06	38.2	60.1
Rigid body refinement solution	125.0	70.8	346.4	0.07	0.35	0.06	36.3	65.8
120 second anaerobic copper soak								
$P2_12_12_1$	122.4	70.1	344.0	0.07	0.35	0.07	31.5	74.1
Rigid body refinement solution	122.3	70.2	344.7	0.07	0.35	0.07	29.6	76.6

5.2.3.1 20 second anaerobic Cu^{2+} soak of premature galactose oxidase

The crystal soaked for 20 seconds in copper acetate diffracted to a maximum resolution of 2 Å. The I/σI at the highest resolution shell, however, was only 0.7 (Table 5.2) and so the resolution was cut back to 2.1 Å (I/σI for the 2.1 resolution shell was 3.4). Fo-Fc difference electron density contained a strong positive peak at the active site region and indicated that copper was coordinated by the Nε of His 581 and Nε of His 496 (Figure 5.5). Copper was modelled into the density, but high temperature factors and difference density after refinement indicated that the copper was approximately 50% occupied. Examination of the 2Fo-Fc electron density map indicated the Sγ of Cys 228 was orientated towards the metal and could be partially copper bound, while another portion of density was consistent with Cys 228 occupying the premature conformation. Cys 228 was built into both positions at 50% occupancy (Figure 5.6). The density for Tyr 272 and Trp 290 was very poor, especially for the tryptophan. There were no indications of any alternative orientations in the electron density, this is indicative of disorder in the crystal. Surprisingly the Cζ of Phe 227 was found only 2.7 Å from the copper site. This residue cannot be a copper ligand and its presence in this position is likely to be a result of the partial occupancy of copper in the active site.

Examination of the N-terminal region of the protein revealed a large Fo-Fc positive peak next to the peptide bond between Ala 1 and Ser 2 (Figure 5.7). This peak was too large to be a water molecule and attempts to model water at the site still left a large positive peak. A copper ion was successfully modelled into the density at 50% occupancy and the model refined well. The N-terminus of the native galactose oxidase structure was orientated so the Cβ of Ala 1 was pointing towards the copper density. Consequently rotation about the main chain ψ was required to enable sensible copper coordination in the model.

The coordination geometry of the N-terminal copper comprises one water molecule and three protein atoms, contributed by the N-terminal nitrogen (Ala 1), the de-protonated nitrogen involved in the peptide bond between Ala 1 – Ser 2, and the carbonyl oxygen of Ser 2 (Figure 5.8). This unusual galactose oxidase N-terminal copper association has been reported earlier in this thesis, but has not been previously observed, even after similar copper incubation experiments [Firbank, 2002]. It is not certain whether the presence of copper at the N-terminal is an artefact of incubation using high copper concentration, or if this phenomenon has possible catalytic significance with regards to pro-sequence cleavage.

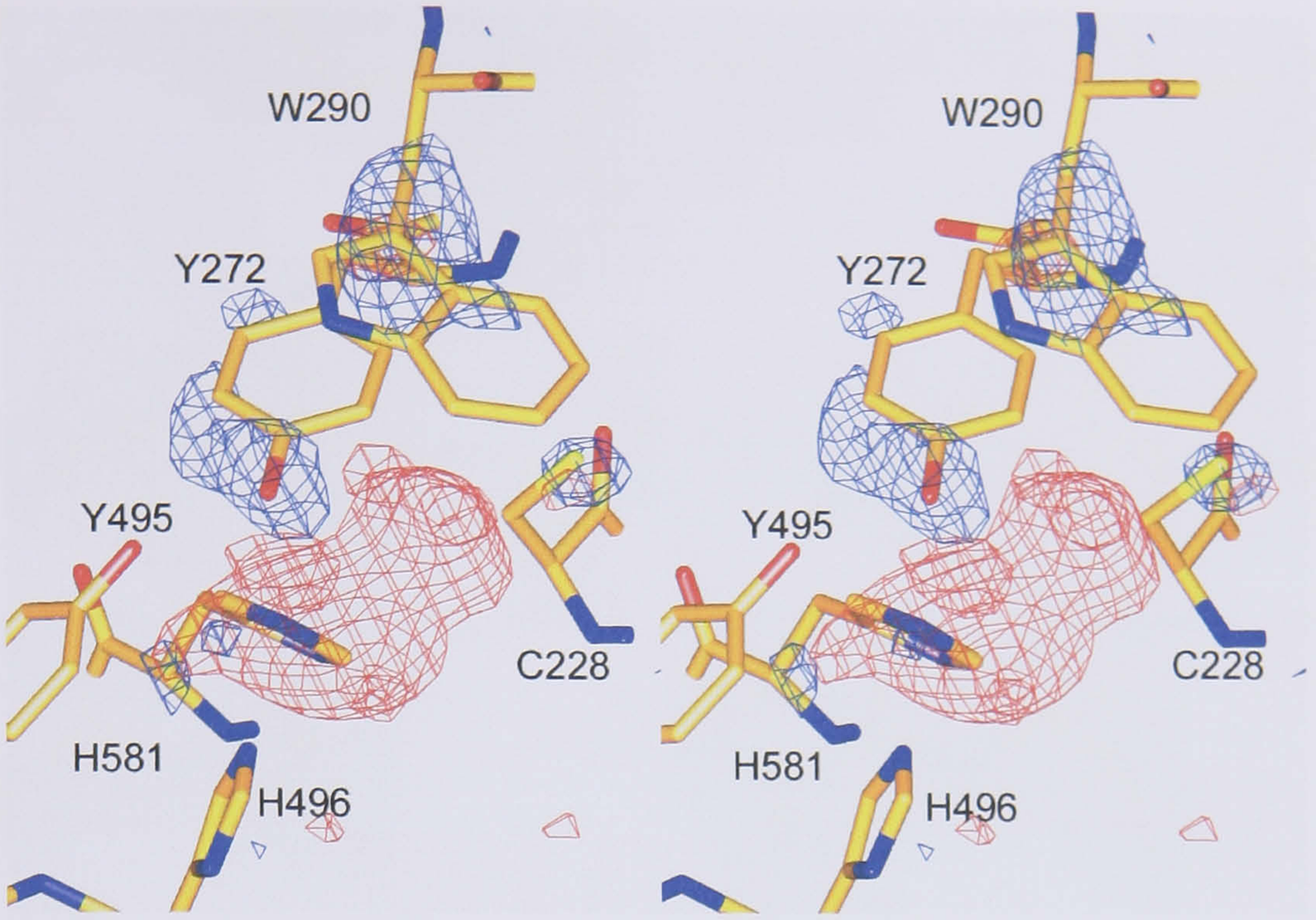


Figure 5.5: Stereoview Fo-Fc active site difference map of 20 second anaerobic Cu^{2+} soak.

The map is contoured at 4 x rms density of the map. Negative density is coloured blue and positive density is coloured in red. The negative density suggests the sulphur of Cys 228 is not in the premature position. Negative density is also found around Trp 290 and the hydroxyl of Tyr 272, suggesting that these residues may also be in alternative conformations. Positive density adjacent to the copper binding site could represent an alternate conformation for Cys 228.

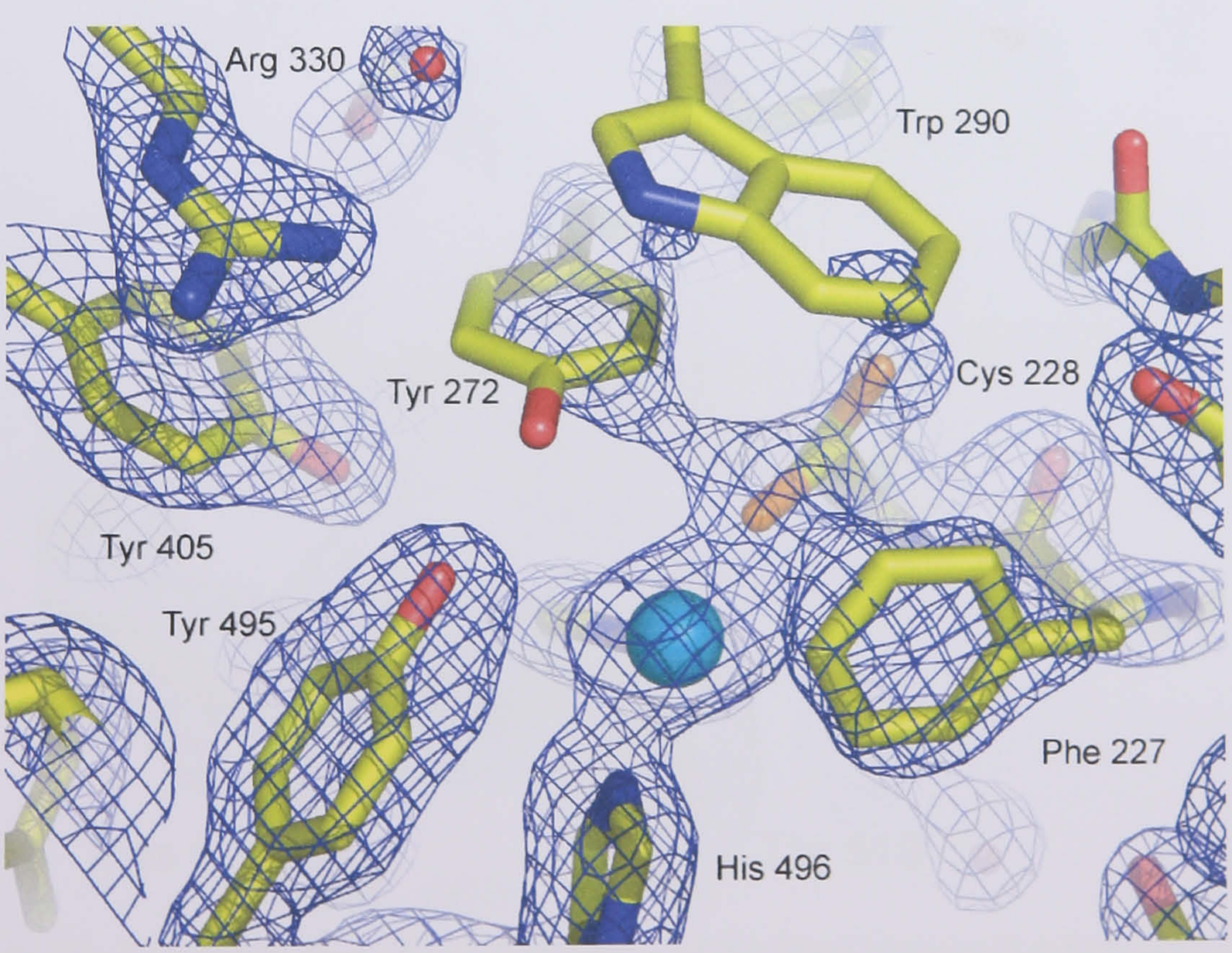


Figure 5.6: 2Fo-Fc active site density map of the final 20 second anaerobic Cu^{2+} soak model.

Cys 228 shows good density for two conformations, one similar to the premature position and another that appears to be interacting with the active site copper. The map is contoured at 1 rms density.

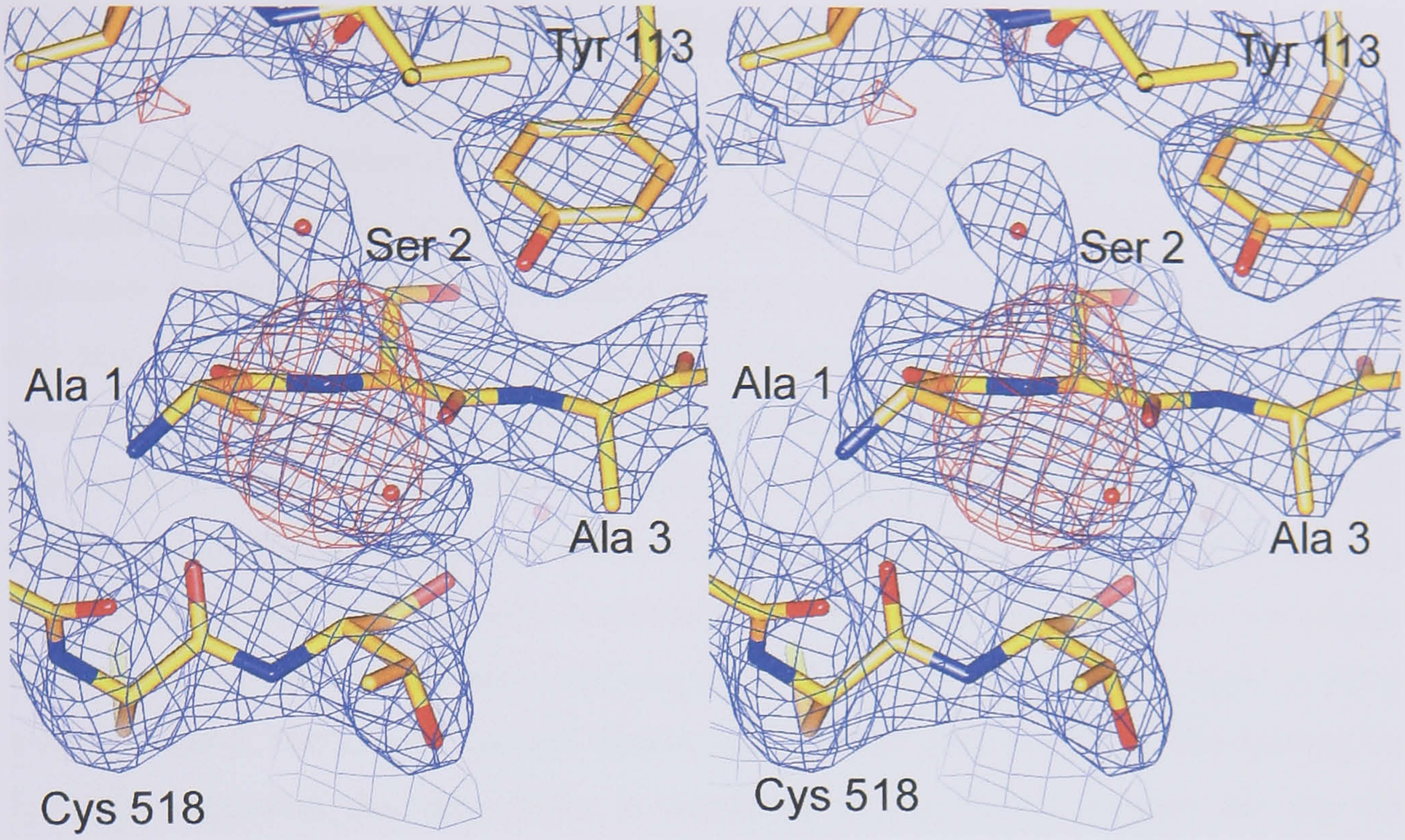


Figure 5.7: Initial density map overlay of the N-terminus of 20 second anaerobic Cu^{2+} soak.

Stereoview of the 2Fo-Fc electron density map of the N-terminal region of 20 second copper soaked premature galactose oxidase contoured at 1 x rms density. The red density represents the positive Fo-Fc electron density contoured at 3 x rms density. The positive density is located at the N-terminal, near the Cys 518-Cys 515 disulphide bridge.

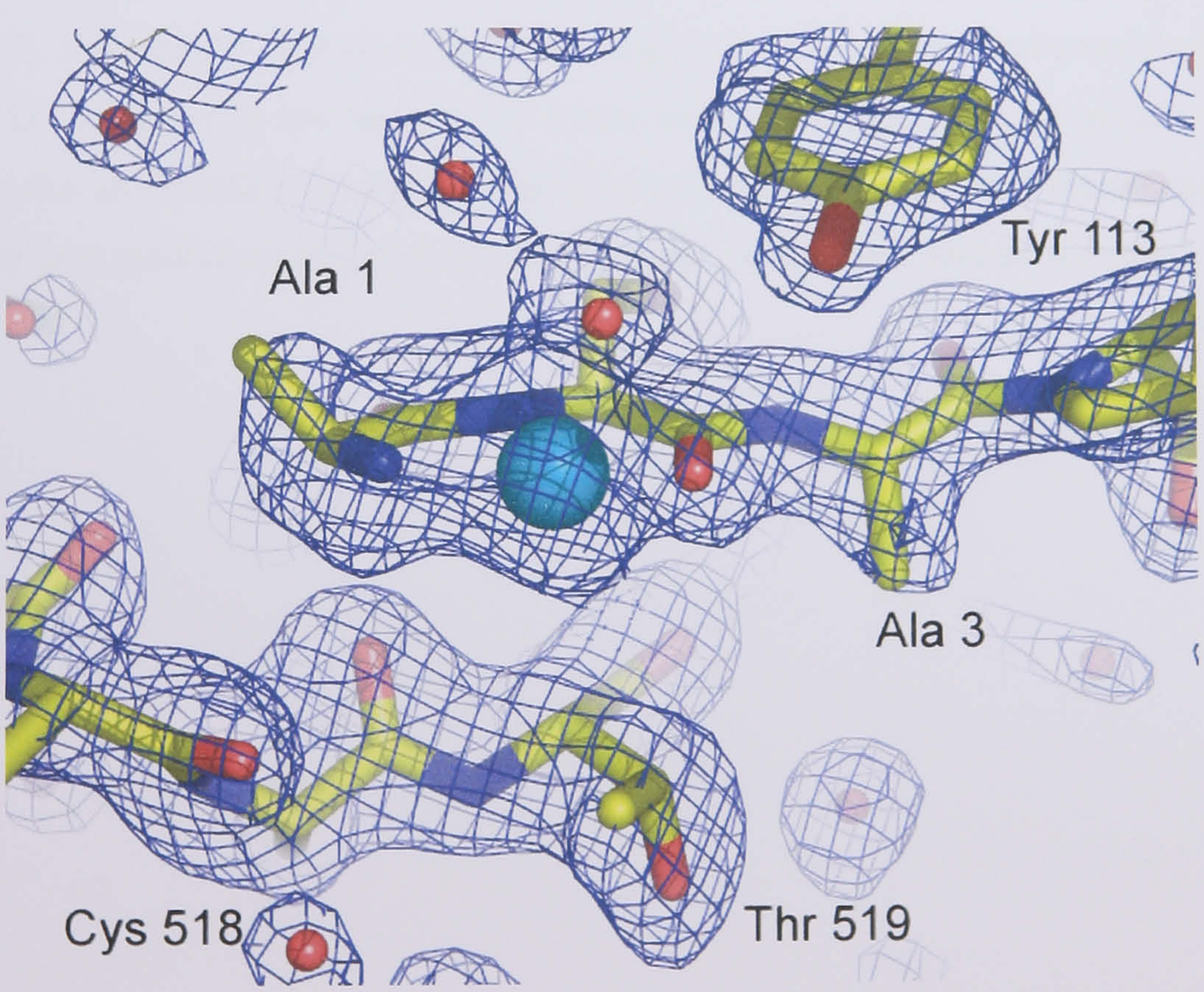


Figure 5.8: 2Fo-Fc density map of the N-terminus of the final 20 second anaerobic Cu^{2+} soak model.

View of the 2Fo-Fc electron density map of the N-terminal region of 20 second soaked premature galactose oxidase, contoured 1 x rms density. Copper is represented by the cyan sphere and water by red spheres. Copper was modelled at the N-terminus and refined reasonably well with two nearby water molecules.

5.2.3.2 120 second anaerobic Cu^{2+} soak of a premature galactose oxidase

A second crystal, incubated anaerobically with 20 mM copper acetate for 120 seconds, diffracted to 2.6 Å resolution on beamline 14.2, Daresbury SRS (Table 5.2). The initial Fo-Fc difference electron density maps contained a strong positive peak at the active site, similar to that seen after the 20 second copper soak. The peak suggested that copper had been incorporated into the active site and was coordinated by His 581 and His 496. A copper ion was modelled into this density, however, high temperature factors and difference density after refinement indicated the copper occupancy was approximately 80%. The density for Cys 228 was very good and indicated one conformation. Cys 228 refined well, with the sulphur residing 3.2 Å away from the active site copper. At this distance it is unlikely that Cys 228 is a copper ligand. The 2Fo-Fc electron density map showed poor density for the hydroxyl of Tyr 272, suggesting that the residue is disordered in the crystal. Density for Trp 290 confirmed the residue was occupying the mature conformation (Figure 5.9). The hydroxyl group of Tyr 495 (2.9 Å) and Tyr 272 (2.8 Å) are found close to copper and could tentatively be assigned as copper ligands, along with His 496 (2.1 Å) and His 581 (2.2 Å).

An unusual elongated electron density was identified between the hydroxyl groups of Tyr 272 and Tyr 495, this density appears to be interacting with both residues and possibly the copper ion too. The density was too large to be water and did not resemble an acetate ion. After refinement the unidentified density was still present and could represent disordered water or acetate, but at a resolution of 2.6 Å the true identity is uncertain and so the density was left empty.

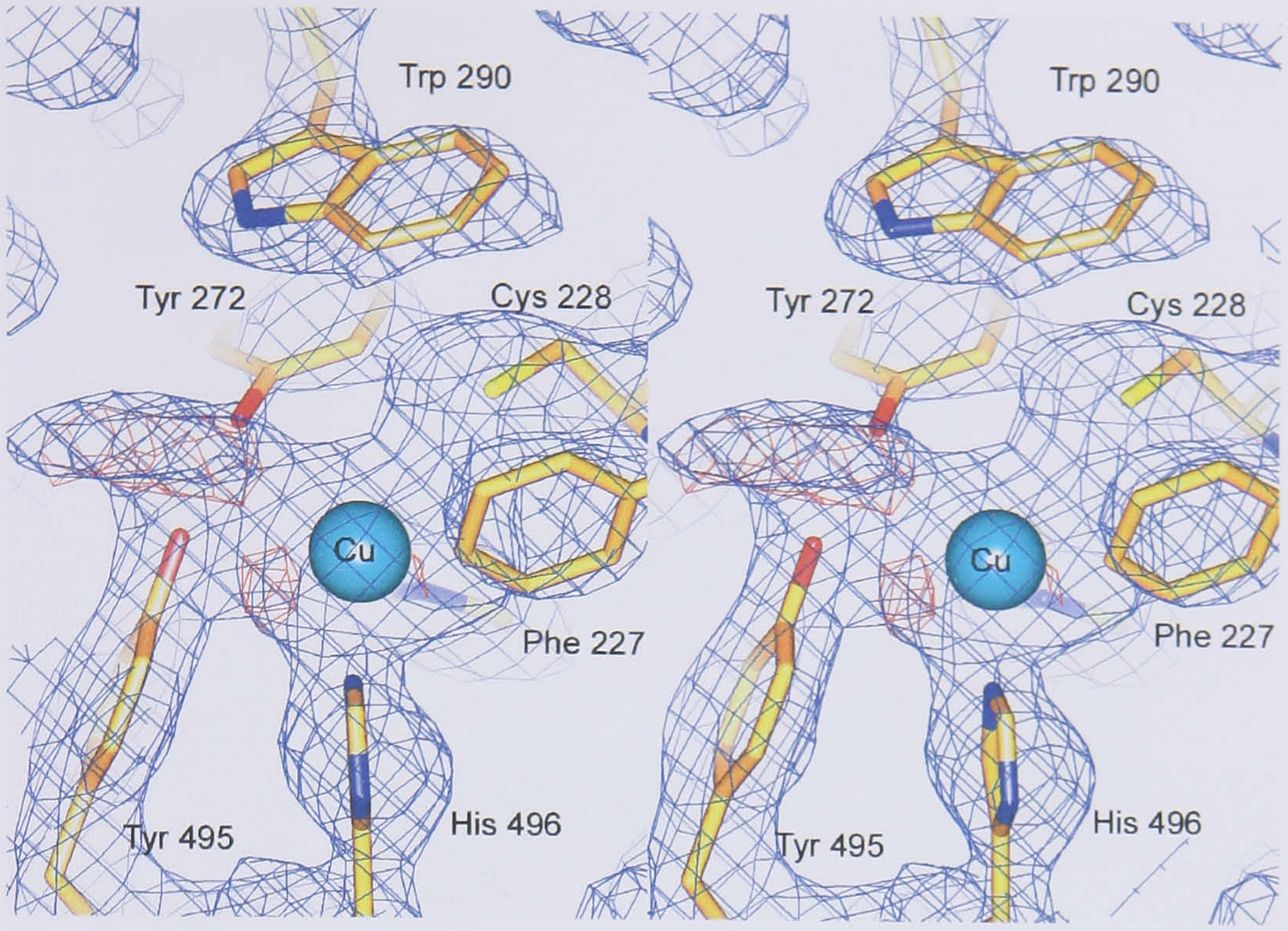


Figure 5.9: 2Fo-Fc active site density map of the 120 second anaerobic Cu^{2+} soak.

Stereoview of the 2Fo-Fc electron density map of the N-terminal region of 20 second soaked premature galactose oxidase, contoured 1 x rms density. The red density is the Fo-Fc positive density contoured at 3 x rms density. Positive density exists between the two tyrosines and the copper, this density was too large to be water and did not resemble acetate. Copper appears to be interacting with His 496, 581, Cys 228 and possibly Tyr 272 and 495.

Two peaks of electron density were found adjacent to Cys 383 (Figure 5.10). The peaks lie close to the active site and strongly resemble metal ions, being only 2.0-2.1 Å from the sulphur of the cysteine. The peaks were modelled in as two copper ions, with one found between cysteine and the Nδ of His 334, while the other has only Cys 383 and a water molecule as ligands. Both putative copper ions appear to have only two ligands each, although Met 382 is close by, it is not within coordination distance. One peak is weaker than the other, but both refined well with an occupancy of 60% (Histidine, Cysteine coordinated copper) and 40% (Cysteine, water coordinated copper). It is not known whether both copper positions are occupied simultaneously or if their occupancies represent the percentage of premature galactose oxidase molecules containing just one of the ions. Binding to this site was identified during anaerobic Cu^{1+} soak experiments presented earlier in this thesis (see section 3.5.3) and during previous copper soak experiments [Firbank, 2002]. A fully occupied copper ion was also identified coordinated by the N-terminus main chain, with geometry similar to the binding described for the 20 second copper soak. An extra water molecule is occupying an equatorial position in the square pyramidal coordination geometry (Figure 5.11). The N-terminal binding here is reminiscent of the binding described following copper incubation of apo-galactose oxidase (Section 3.4.2 and section 3.5.4).

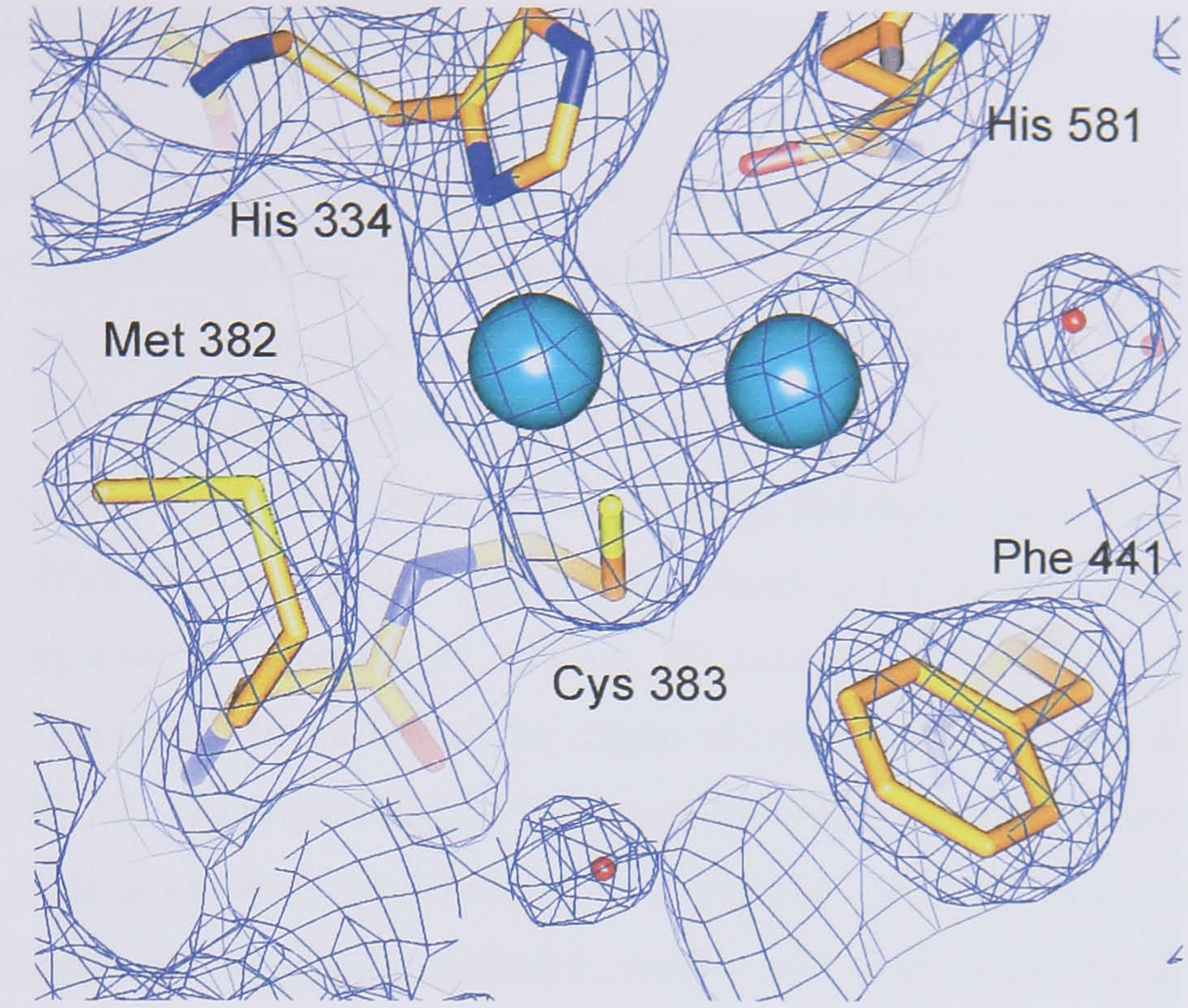


Figure 5.10: 2Fo-Fc density map of the cysteine 383 region after 120 second Cu^{2+} soak.

Refined 2Fo-Fc electron density map of the region around cysteine 383. Two copper atoms were modelled into the density next to residue 383. One copper is coordinated to the cysteine and the Nδ of Histidine 334, while the other copper is interacting with cysteine and a nearby water molecule.

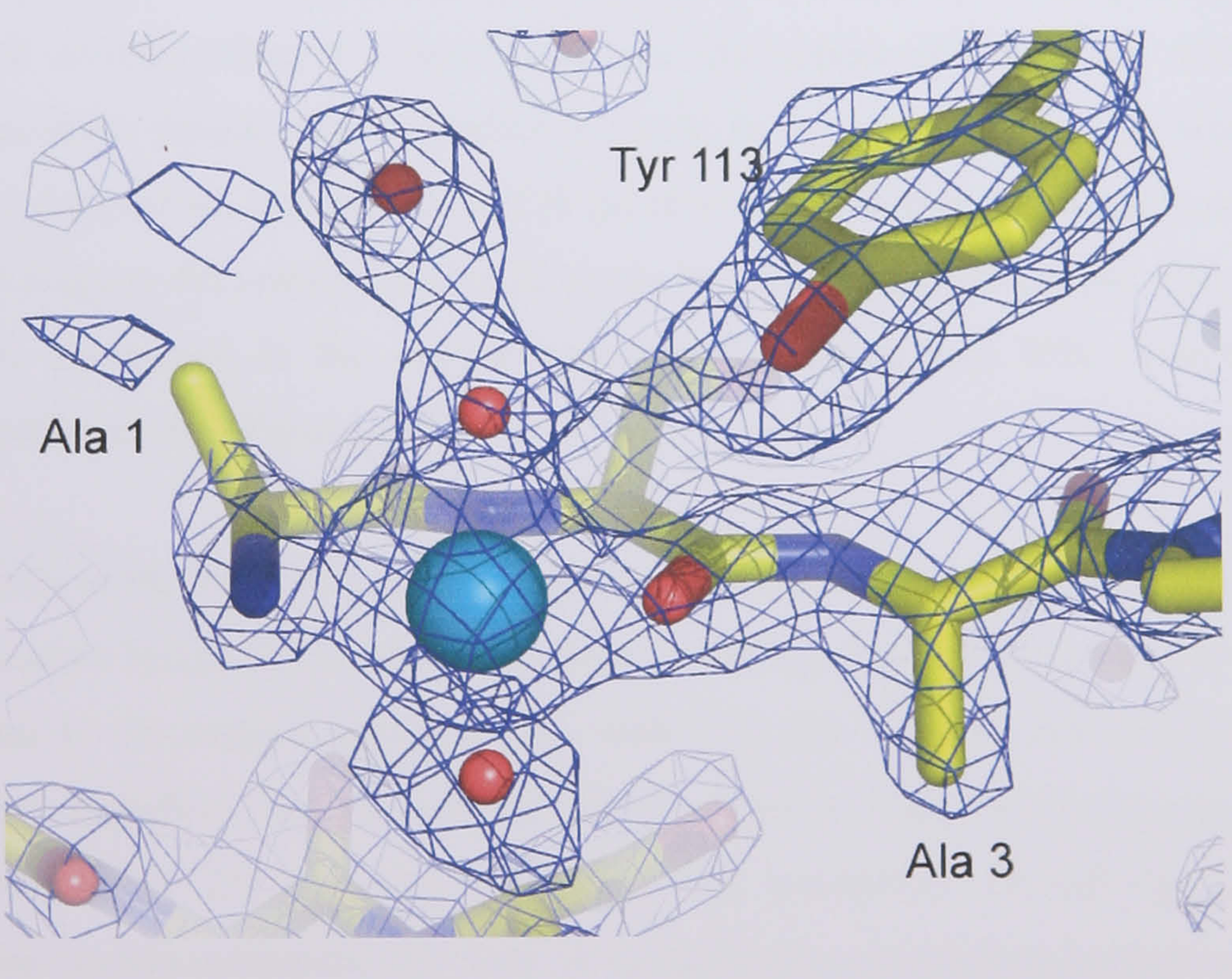


Figure 5.11: N-terminal 2Fo-Fc density from 120 second anaerobic Cu^{2+} incubation

N-terminal 2Fo-Fc density from 120 second anaerobic copper incubation, contoured at 1 x rms density. Copper was bound by three main chain atoms and two water molecules occupying an axial and equatorial position in the square pyramidal coordination geometry.

5.2.4 Aerobic addition of Cu^{2+} to the premature galactose oxidase

Aerobic addition of copper to premature galactose oxidase is reported to result in the formation of enzyme displaying the characteristics of the mature, radical containing enzyme [Rogers, *et al.*,2001]. Several crystals were taken to Daresbury SRS, each had been incubated aerobically in 20 mM copper acetate for differing lengths of time and stored in liquid nitrogen. To aid in the high throughput crystal screening and data collection at station 14.2, an ACTOR automated sample changer and crystal mounting robot was utilised. The crystal incubation times were 47 seconds, 52 seconds, 90 seconds, 120 seconds, 180 seconds, 10 minutes and 20 minutes. The success of the copper incubation was variable. Initial diffraction from the 47, 180, 52, 120 second and 20 minute soaked crystals revealed data judged to be of poor quality. Data from these crystals were not collected. Data were successfully collected from the 90 second and 10 minute incubated crystals with both diffracting to a resolution of 1.6 Å (Table 5.4).

5.2.4.1 90 second aerobic Cu^{2+} soak of premature galactose oxidase crystal

The crystal soaked aerobically in 20 mM copper acetate for 90 seconds, diffracted to 1.6 Å resolution at station 14.2 Daresbury SRS. The Fo-Fc difference density map revealed a positive peak corresponding to copper, however, the copper site was only 50% occupied. Additional positive density was identified adjacent to copper and Cys 228, while negative density was observed on the Sγ of Cys 228 positioned in the premature conformation. Both observations suggest the sulphur of Cys 228 may be interacting with copper. Cys 228 refined well at 50% occupancy in the copper coordinated position and 50% occupancy of the premature conformation (Figure 5.12).

Positive density at the Tyr 272 position suggests the residue also exists in two conformations. One conformation brings the residue very close to the region occupied by Trp 290, while the other position is the mature conformation, with Trp 290 stacking over the residue. Both residues were modelled into the mature conformation at 70% occupancy, an alternative conformation for Tyr 272 was also present at 30% occupancy. Overall the two Tyr 272 conformations closely resembled that seen in the premature structure presented earlier in this chapter (Figure 5.4). An alternative position for Trp 290 could not be found (Figure 5.12).

Table 5.4: Data collection and refinement statistics from aerobic incubation with 20 mM copper acetate		
Crystal incubation time	90 seconds	10 minutes
Crystallisation conditions	16% PEG 8000, 14% PEG 400, 100mM MES pH 6.3, 200mM calcium acetate.	16% PEG 8000, 14% PEG 400, 100mM MES pH 6.3, 200mM calcium acetate
X-ray facility	Station 14.2, Daresbury SRS, UK	Station 14.2, Daresbury SRS, UK
Wavelength (Å)	0.98	0.98
Space group	C2	C2
Unit ell dimensions	**a**=97.1 Å, **b**=89.0 Å, **c**=85.7 Å, **β**=117.5 °	**a**=97.2 Å, **b**=89.0 Å, **c**=85.3 Å, **β**=117.3 °
Resolution range (Å)	41.5 – 1.6	62.0 – 1.6
No. of observed reflections	582927	310795
No. of unique reflections	79644	81664
Mean I/σI	6.27	7.1
Completeness (%)	91.2 (91.2)	96.1 (96.1)
R_{sym} (%)	4.7 (14.1)	5.4 (40.1)
Multiplicity	2.7 (2.6)	2.8 (2.2)
Refinement statistics		
Refinement resolution (Å)	41.5 – 1.7	62.0 – 1.7
R_{cryst} (%)	23.1	21.6
R_{free} (%)	18.3	21.7
Number of atoms	5432	5295
No. of protein atoms	4829	4829
No. water molecules	596	455
Average B-factor	23.7	29.2
Rms deviation bonds (Å)	0.007	0.007
Rms deviation angles (Å)	1.5	1.7

Figures in Parentheses correspond to highest resolution shell

Copper appears to be coordinated by a three ligand geometry comprising His 496, His 581 and Cys 228, although a water molecule and Tyr 272 are close (2.8 Å and 2.4 Å away respectively) and could be ligands. The partial occupancy and disorder of Cys 228 and Tyr 272 in the active site, however, make it difficult to unequivocally assign these two residues as copper ligands. There is no indication of Tyr 495 binding to copper at this stage of biogenesis. Electron density for the side chain of Phe 227 is disordered and appears close to the copper density (2.4 Å distance). This apparent proximity is unlikely to occur when both molecules are present in the active site at full occupancy. Copper was again identified at the N-terminus, no additional copper sites were found. Several of the surface regions of the protein had high B-factors and poor density, indicative of increased side chain disorder at these regions. For this reason the model refined to an R-factor of 18% and R-free of 23%.

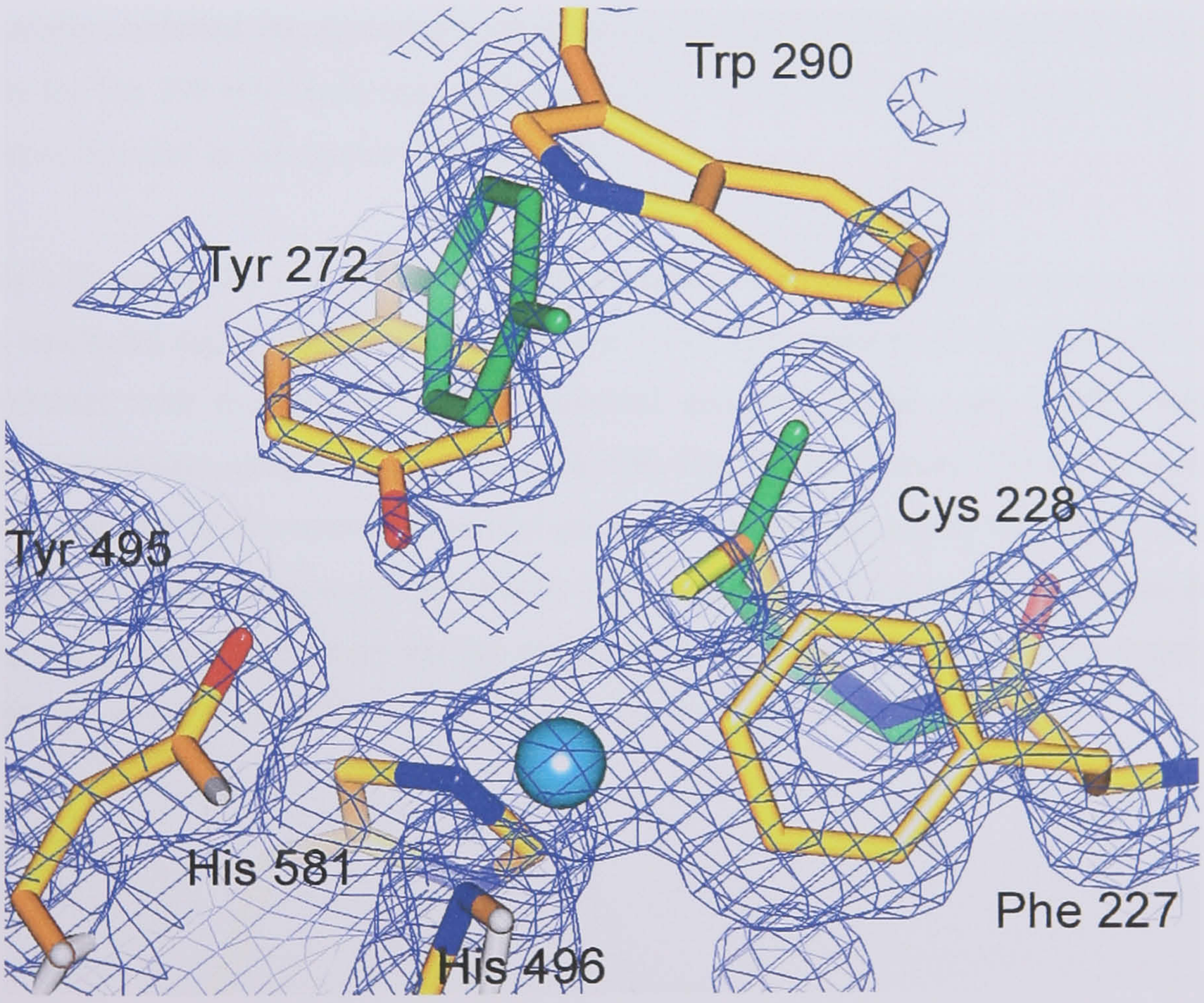


Figure 5.12: Active site 2Fo-Fc map from 90 second aerobic Cu^{2+} incubation.

2Fo-Fc electron density map of premature galactose oxidase after 90 seconds aerobic copper incubation. The map is contoured at 1 x rms density. Tyr 272 occupies two conformations. One conformation (green) is prevented by steric hindrance between Trp 290 in the mature position. Cys 228 refined into a premature conformation (green) and an alternate position in which the Sγ points towards copper. Copper is coordinated by His 581, His 496 and possibly Cys 228. Phe 227 is close to the copper, but Tyr 272 and 495 do not appear to be copper ligands.

5.2.4.2 10 minute aerobic Cu^{2+} soak of a premature galactose oxidase

A second crystal aerobically incubated in 20 mM copper acetate for 10 minutes diffracted to a maximum resolution of 1.6 Å at Daresbury SRS (Table 5.4). 2Fo-Fc electron density contained a strong peak at the active site region, indicating that copper was present, and was coordinated by Tyr 272, Tyr 495, His 581 and His 496 (Figure 5.13). Density between Tyr 272 and Cys 228 indicated the presence of the thioether bond, although partial coordination of the Sγ of Cys 228 to copper cannot be ruled out. More density was present around the copper and radical cofactor than is usual for the galactose oxidase active site structure at high resolution. Perhaps indicative of residual disorder in the active site, suggesting that cofactor processing has not been completed throughout the whole crystal. There was reasonable density for Tyr 272 and the loop region comprising residue Trp 290, hence Tyr 272 was successfully modelled into the mature conformation, with Trp 290 stacking over the residue. Density for Trp 290 side chain was slightly distorted, which suggests the residue has a certain amount of disorder in the crystal (Figure 5.14).

Density adjacent to the copper indicated the presence of an acetate molecule occupying the fourth equatorial copper ligand position (Figure 5.14). The geometry of the active site copper is consistent with the mature square pyramidal geometry (PDB code 1GOF), with the equatorial positions occupied by Tyr 272, His 496, His 581 and acetate. Tyr 495 is located in the axial position of the square pyramidal geometry. N-terminal copper was again identified and modelled into the region coordinated by three main chain atoms and two water molecules in a square pyramidal geometry similar to the active site coordination. No other copper ions were present in the structure.

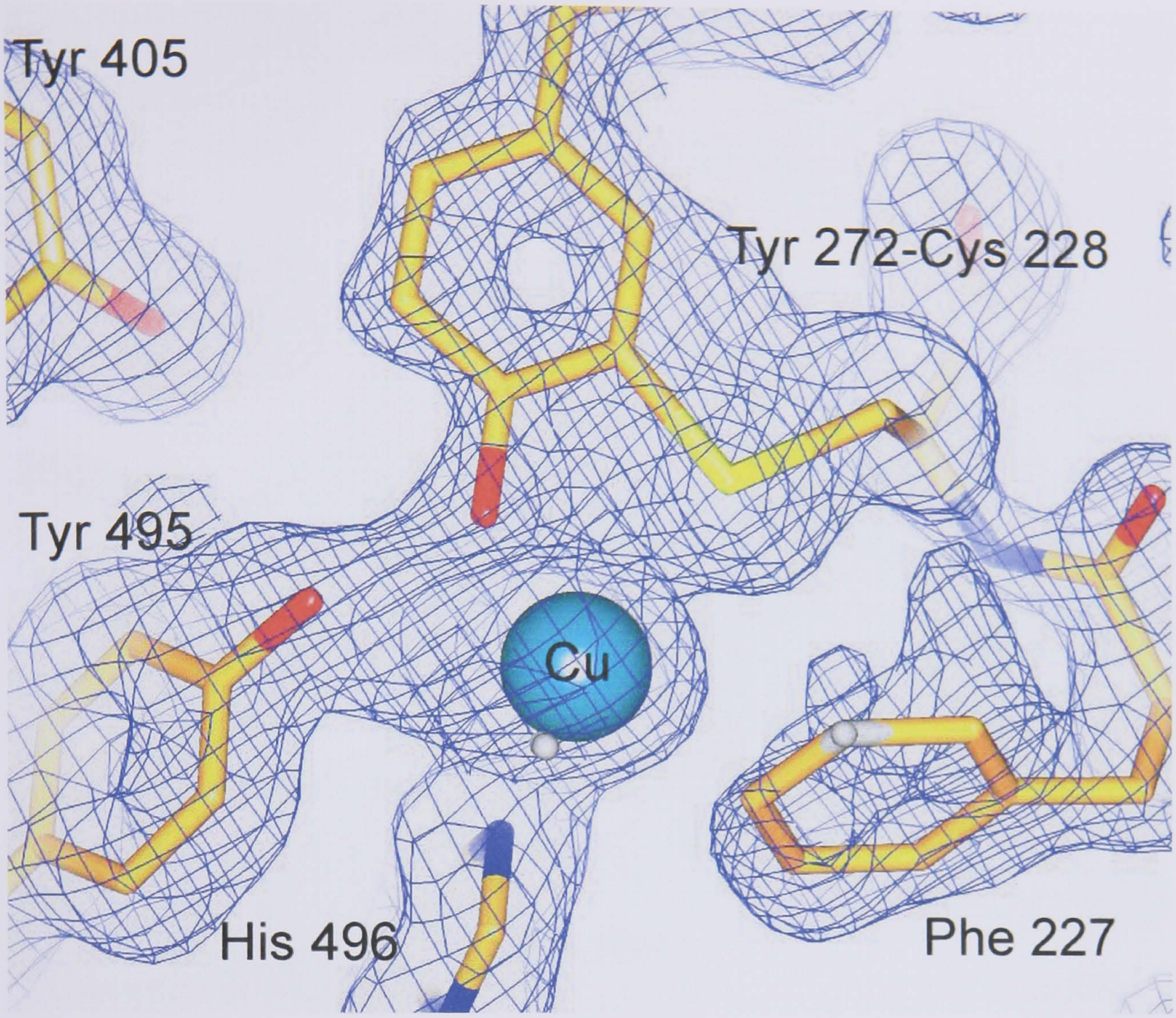


Figure 5.13: 2Fo-Fc map of cofactor after 10 minute aerobic Cu^{2+} incubation.

2Fo-Fc electron density map of the active site of premature galactose oxidase incubated aerobically for 10 minutes in copper. The map was contoured at 1 x rms density. Density between Cys 228 and Tyr 272 is indicative of thioether bond formation. Density for Phe 227 is distorted and appears to be close to Cys 228. There is a large peak for the copper density, Tyr 272 and Tyr 495 are adjacent to the metal and are within coordination distance.

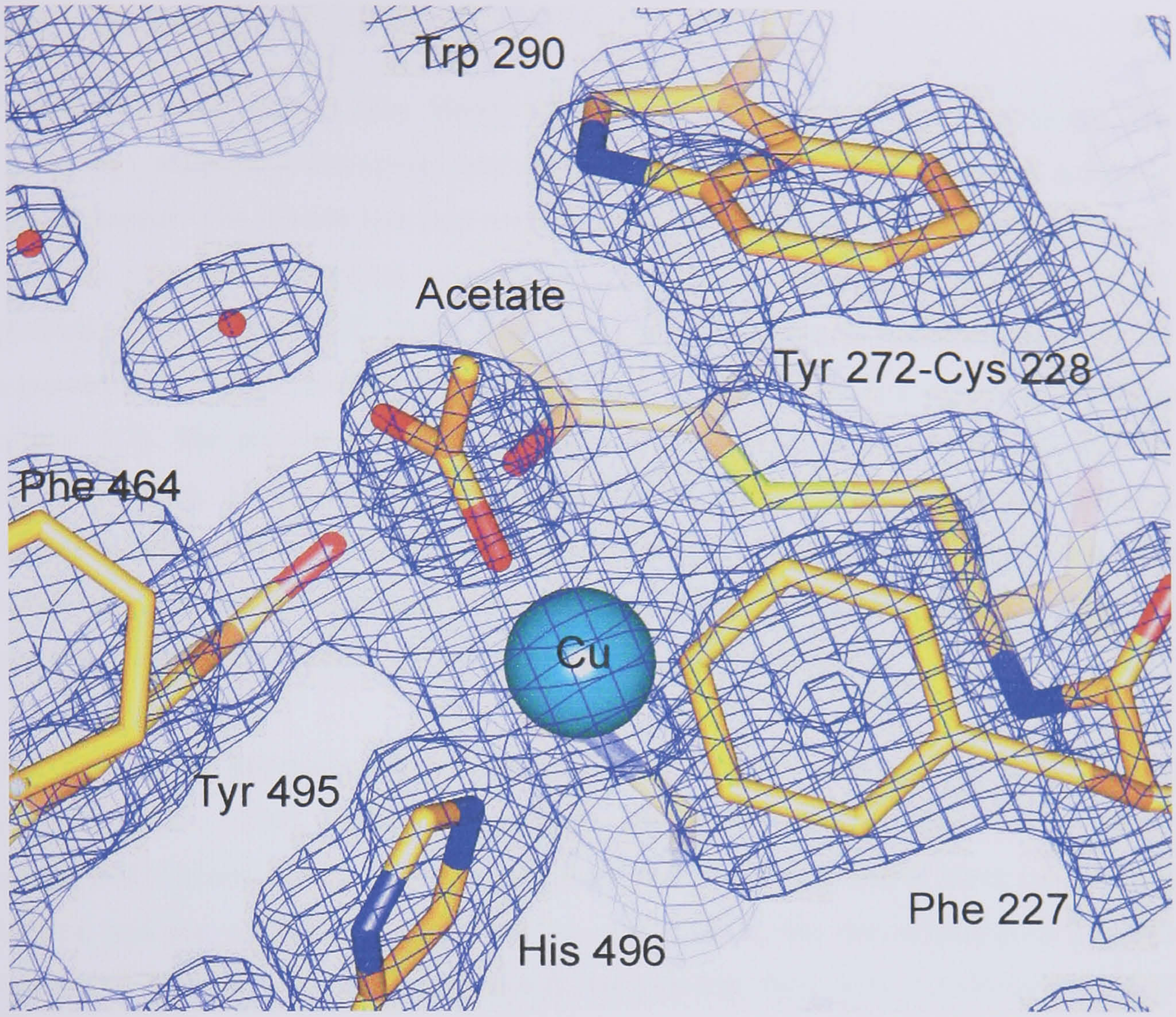


Figure 5.14: 2Fo-Fc map of square pyramidal geometry after 10 minute Cu^{2+} soak

2Fo-Fc of the active site of premature galactose oxidase following 10 minutes incubation with copper, contoured at 1 x rms density. Acetate was modelled into a peak at the fourth equatorial copper ligand position. Density for Trp 290 and Phe 227 was slightly distorted.

5.3 Discussion

Galactose oxidase protein was kindly supplied by Dr Melanie Rogers (Montana State University), the protein was crystallised and resembled the original native crystals despite the lack of copper. The protein was expected to contain the pro-sequence and lack the thioether bond, as observed in previous copper-free *A.nidulans* preparations of galactose oxidase [Firbank, 2002; Rogers *et al.*, 2000]. The crystal structure, however, revealed no evidence of a pro-sequence, instead revealing surprisingly good density for the N-terminus of the protein (Figure 5.2). The protein structure contained no indication of copper binding or thioether bond formation, so it appears the pro-sequence has been cleaved before crystallisation, leaving the protein in the premature form. As the thioether bond was not present either, it can not be ruled out that the supplied protein was originally the premature galactose oxidase form expressed from the BD1construct (lacking n-terminal pro-sequence).

The structure of galactose oxidase expressed from the BD1 construct has been published [Deacon *et al.*, 2004], while the premature copper-free structure has been previously determined [Firbank, 2002], but never published. The overall structure of premature galactose oxidase and the original mature protein are very similar, the rms deviation of the alpha carbons of the two structures is only 0.4 Å. Despite this, there were significant differences identified at the active site. The thioether bond residues Tyr 272 and Cys 228, are found occupying pro-GO conformations. The Sγ of Cys 228 is found pointing away from the position it would occupy after thioether bond formation, while Tyr 272 seems to be influenced by the disorder of Trp 290 and thus adopts two conformations. One orientation is similar to Tyr 272 in pro-GO while the other conformation closely resembles what is expected in the mature structure. Tyr 272 and Trp 290 are likely to preferentially occupy their mature conformations at the same time, as when Tyr 272 is in the pr-GO conformation there is a steric clash with Trp 290 occupying the mature conformation.

Anaerobic addition of copper to premature galactose oxidase for 20 seconds allowed the visualisation of copper bound at the active site. Copper appeared to be coordinated by His 581, His 496 and Cys 228 (Table 5.5). The density maps suggested Cys 228 was occupying two conformations. It appears as though the presence of the metal has caused Cys 228 to change conformation, thus allowing Sγ to interact with copper. The copper site is approximately 50% occupied, consequently, it is difficult to know which Cys 228 conformation is present at the same time as copper. Observation of a proportion of the Sγ of Cys 228 rotated around χ2 to point towards the copper site following the addition of copper, suggests the addition of copper could be responsible for the conformational change.

After a 120 second anaerobic incubation with copper the structure was very similar to the 20 second copper soak. The sulphur of Cys 228 was no longer occupying the premature conformation, while density adjacent to copper indicated Cys 228 could be associated with the metal. After refinement, however, the sulphur atom was at a distance of 3.2 Å from copper ion, therefore too far from the metal to be a coordinated ligand (Table 5.5). The active site copper occupancy was higher compared to the shorter soak, but after the longer copper exposure, the Cys 228-copper distance was greater. Furthermore, Tyr 272 and Tyr 495 appear to be closer to copper compared to the shorter soak, suggesting the phenolate coordination and Cys 228 dissociation from the metal. Although, it is noted that the low resolution of this structure (2.6 Å) does not allow very reliable bond distances to be measured.

Elongated electron density was identified adjacent to the hydroxyl of Tyr 272, Tyr 495 and the copper ion. The density did not resemble acetate and was too large to be water molecule. The position is very close to the fourth equatorial ligand position where acetate [Ito *et al.*, 1991], water [Ito *et al.*, 1994] and azide [Firbank, 2002; Vinecombe, 1999] have been observed in previous crystal structures. The fourth equatorial position is predicted to be the substrate (alcohol and oxygen) binding site, but as the crystal was incubated under anaerobic conditions and in the absence of substrate, no oxygen binding was expected. The possibility of low concentrations of molecular oxygen being present during the anaerobic experiment would raise several possibilities for the identity of the elongated density.

All biogenesis theories suggest an oxygen-copper complex is formed during cofactor formation. In solution studies under anaerobic conditions, Cu^{2+}-dependent biogenesis was reported to result in slow development of the 406 nm spectral feature, which develops the mature spectrum upon exposure to oxygen [Whittaker and Whittaker, 2003]. In light of these observations all proposed mechanisms of cofactor biogenesis suggest a copper-enzyme complex is formed prior to the introduction of oxygen. Therefore, if oxygen has invaded the anaerobic copper soak experiment, the elongated density observed after the 120 second incubation could be an oxygenated enzyme complex. Alternatively, if the reaction conditions were truly anaerobic, then it is possible the peak represents disordered water or acetate ion. It is not clear what species is responsible for the elongated density as the maps show disorder in the active site. In addition to a fairly low resolution structure, it is not possible to be certain of the data interpretation.

Table 5.5: Active site distances for premature galactose oxidase incubated with copper, compared with the native mature galactose oxidase structure. The distances for Cys 228-Sγ and Tyr 272 marked (*), are taken from the residues in their non-premature conformations.

Table 5.5: Active site distances for premature galactose oxidase copper soaked crystals (Å)

Atom distance measured	**20 seconds anaerobic soak**	**120 seconds anaerobic soak**	**90 seconds aerobic soak**	**600 seconds aerobic soak**	**Mature galactose oxidase (1GOF)**
Tyr 272-OH to Cu	3.9	2.8	2.5*	2.3	1.9
Try 495-OH to Cu	3.6	2.9	3.4	3.0	2.7
His 581-Nε to Cu	1.9	2.2	2.2	2.0	2.1
His 496-Nε to Cu	2.3	2.1	2.2	2.0	2.1
Cys 228-Sγ to Cu	2.2*	3.2	2.5*	2.7	3.6
Cys 228-Sγ to Tyr 272-Cε	2.9*	2.8	2.6*	1.8	1.8
Phe 227-Cε2 to Cys 228- Sγ	3.1*	3.1	2.8*	3.3	3.3
Phe 227-Cζ to Cu	2.7	2.7	2.4	2.8	3.6
Tyr 272 – OH to Tyr 495 OH	2.7	2.2	2.1*	2.1	2.9
Exogenous ligand to Cu	-	-	2.8[a]	2.4[b]	2.3[b]

[a] denotes water ligand, [b] denotes an acetate ligand.

The protein ligand to copper distances in each soak structure, appear to indicate an active site in transition from the premature conformation through to the mature geometry, via the processing intermediates described earlier. There were similarities between the three structures showing copper exposure at 20, 90 and 120 seconds, despite 90 second soak experiment having been carried out in aerobic conditions. All three structures suggest that the first event to occur following exposure to copper, is a Cys 228 conformational change. This step occurs under aerobic and anaerobic conditions prior to cofactor formation. The structures presented in this chapter support the copper coordinated Cys 228 structures reported by Firbank and consequently the thioether bond formation mechanism proposed as a result of the findings (Figure 1.25).

Aerobic incubation of premature galactose oxidase for 10 minutes produced a structure containing the thioether bond. The active site copper is coordinated by the four equatorial ligands Tyr 272, His 496, His 581 and acetate. Tyr 495 occupies the axial position of a square pyramidal geometry. The active site structure looks remarkably similar to the native mature galactose oxidase active site (PDB code 1GOF). The alpha carbon atoms of the two structures were aligned, revealing an rms deviation of 0.23 Å, indicating that there is very little difference between them. A minor difference is noted for the position of copper, His 496 and Tyr 495 (Figure 5.15). The difference may be due to remaining disorder in the crystal, caused by partially processed enzymes.

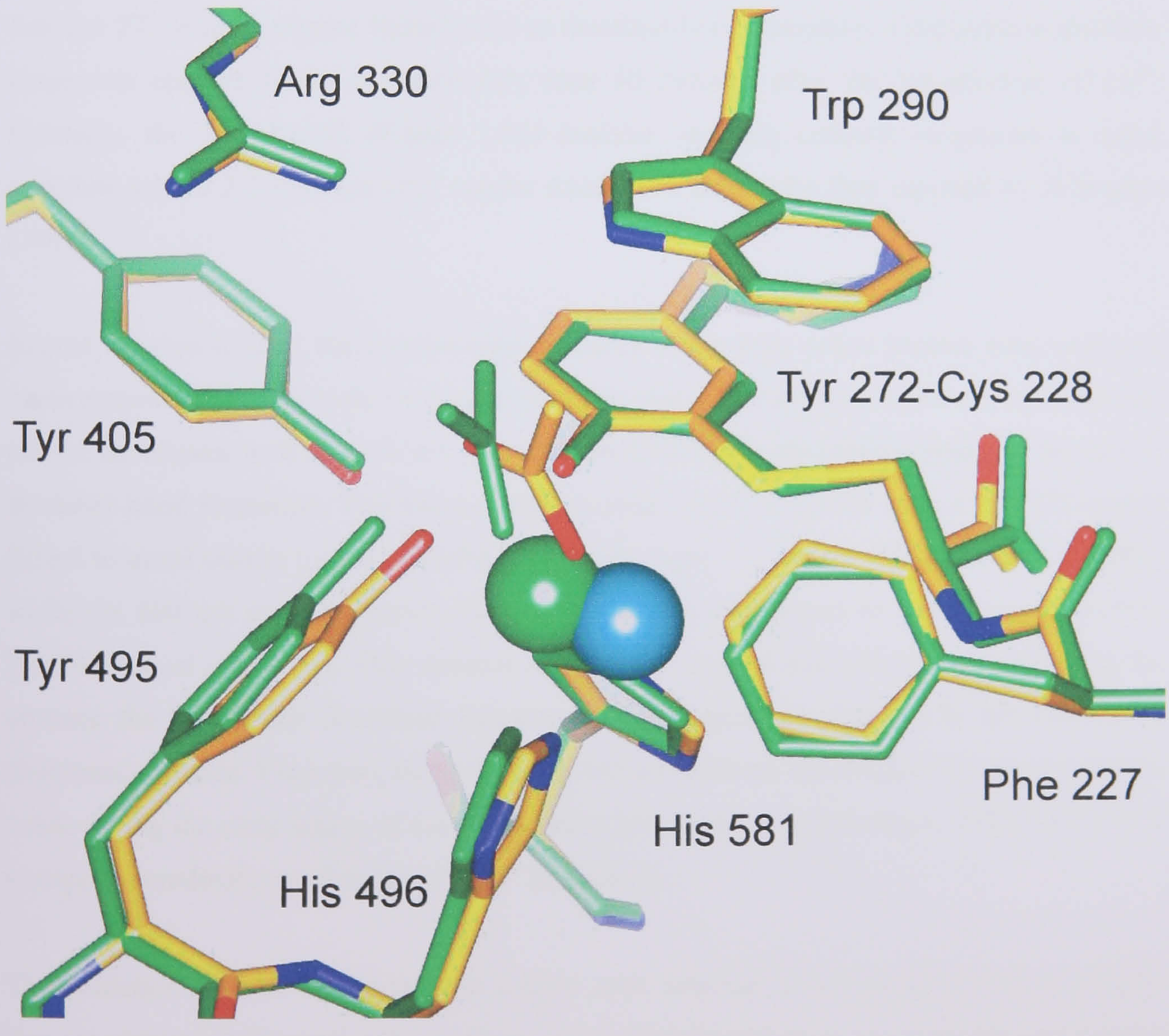


Figure 5.15: Alignment of native enzyme with 10 minute aerobic Cu^{2+} incubated structure.

Overlay of native galactose oxidase (PDB code 1GOF) onto premature galactose oxidase aerobically incubated with copper for 10 minutes. The atoms from the native structure are coloured in green. The movement of His 496 alters the position of the active site copper, thus similar movements occur with Tyr 495 and the copper bound acetate.

The 10 minute aerobic structure is reminiscent of the original native structure, with Tyr 272, His 581, His 496 and acetate all present as ligands. The copper-Tyr 272 and copper-Tyr 495 distances are longer than observed in the original mature structure (Table 5.5). The significance of this observation is uncertain, but dissociation or partial dissociation of Tyr 495 and Tyr 272 from the copper could explain the lack of the 800 nm absorbance peak reported during Cu^{2+}-dependent cofactor biogenesis in solution studies [Rogers *et al.*, 2001, Whittaker and Whittaker, 2003].

It has been suggested that Cu^{2+}-dependent cofactor biogenesis has a half-life of several hours, and in fact is so slow that reduction of the copper to the Cu^{1+} oxidation state may be responsible for the biogenesis reaction observed when Cu^{2+} is incubated with premature galactose oxidase [Whittaker and Whittaker, 2003]. The structural data, however, indicates that Tyr 272 is not a copper ligand prior to thioether bond formation. Furthermore, thioether biogenesis appears to be complete less than 10 minutes after the introduction of Cu^{2+}. Similarly, the SDS-PAGE (Figure 1.22) analysis suggests cofactor biogenesis is rapid, complete within 2-3 minutes after copper addition, a faster rate than reported by Whittaker [2003].

In past solution studies, the spectroscopic features of the fully active protein were used as a measurement of the rate cofactor biogenesis. This method of analysis, however, interprets the rate of development of the 445 nm and 800 nm UV-visible absorbance peaks as the rate of thioether bond formation. The 445 nm peak is proposed to originate from a Tyr 272-copper ligand to metal charge transfer and the radical transition $Tyr^{\bullet}$ $\pi \rightarrow \pi^*$ [Whittaker *et al.*, 1989], while the 800 nm peak represents the Tyr 495 to Tyr 272 ligand to ligand charge transfer [McGleshen *et al.*, 1995]. This method of analysis implies that cofactor biogenesis in the absence the previously described interactions will display no features in the UV-visible absorbance spectra. Therefore, the lack of the 445 and 800 nm absorbance features reported to occur during the early stages of Cu^{2+}-dependent biogenesis could be due to a lack of Tyr 272 to copper coordination rather than "slow" biogenesis.

The oxidation state of the 10 minutes copper soak structure is not known, but it is possible that the enzyme is the semi-reduced form. The initial formation of the semi-reduced enzyme would explain the initial lack of strong absorbance at 445 and 800 nm during Cu^{2+}-dependent biogenesis. Prolonged exposure to oxygen would then result in the formation of the oxidised enzyme via the reported auto-oxidation reaction [Wright and Sykes, 2001a]. This does not explain, however, why Cu^{1+}-dependent biogenesis results in the apparently rapid attainment of the oxidised enzyme.

Additional copper binding sites were identified during the soaking experiments. Copper was identified coordinated by Cys 383, which lies close to the active site (see Figure 1.14). This was observed in the 120 second anaerobic soak, but also occurred in the 24 hour anaerobic soak mentioned earlier [Firbank, 2002] and in copper reconstituted orthorhombic galactose oxidase crystals (Chapter three). This binding is not thought to play a role in cofactor biogenesis.

A copper binding site was also found at the N-terminus of each of the copper incubated protein structures. The metal is coordinated by the terminal nitrogen of Ala 1, the de-protonated peptide nitrogen between Ala 1 and Ser 2, the carbonyl oxygen of Ser 2. This three ligand geometry was also identified following a 25 minute Cu^{1+} incubation of apo-galactose oxidase. The three coordinate geometry, however, was associated with partially occupied copper sites, while higher occupancy copper sites revealed two additional water molecules coordinated at the axial and equatorial positions of a square pyramidal geometry (Figure 5.11).

Although unusual in proteins, copper binding to peptides has been observed in small molecules, such as copper-dipeptide complexes [Facchin *et al.*, 2002]. The crystal structures of copper coordinated with alanine-isoleucine, alanine-threonine or alanine-tyrosine revealed that each dipeptide complex possessed equatorial N_2O_2 ligand geometry, with an axial oxygen ligand (Figure 5.17). In all three complexes the axial and one of the equatorial ligands are provided by oxygen atoms from symmetry related molecules. The complexes were formed by dissolving the di-peptide in hot water, adding solid $CuCO_3$, followed by crystallisation at 80°C. Although the conditions for complex formation were harsh, the only major difference between the geometry of these structures and the galactose oxidase observation is the increased length the water equatorial ligand. Manual alignment of the Ala-Thr structure onto the N-terminal copper bound structure reveals almost identical atom position for the main chain positions (Figure 5.18).

After observation of copper binding at the N-terminus of galactose oxidase, it is tempting to suggest copper may also bind at the cleavage site in pro-GO, thus initiating cleavage of the pro-peptide. Examination of the cleavage site in pro-GO indicates metal binding of the sort seen at the mature N-terminus is not possible without a change in conformation (Figure 5.19). The nitrogen of Ala 1 is pointing in a different direction in the pro-GO structure and so could not contribute to metal coordination. The N-terminal binding of copper presented in this

thesis may be an artefact of the 20 mM copper concentration. It remains, however, the first observation of metal binding at the pro-sequence cleavage site.

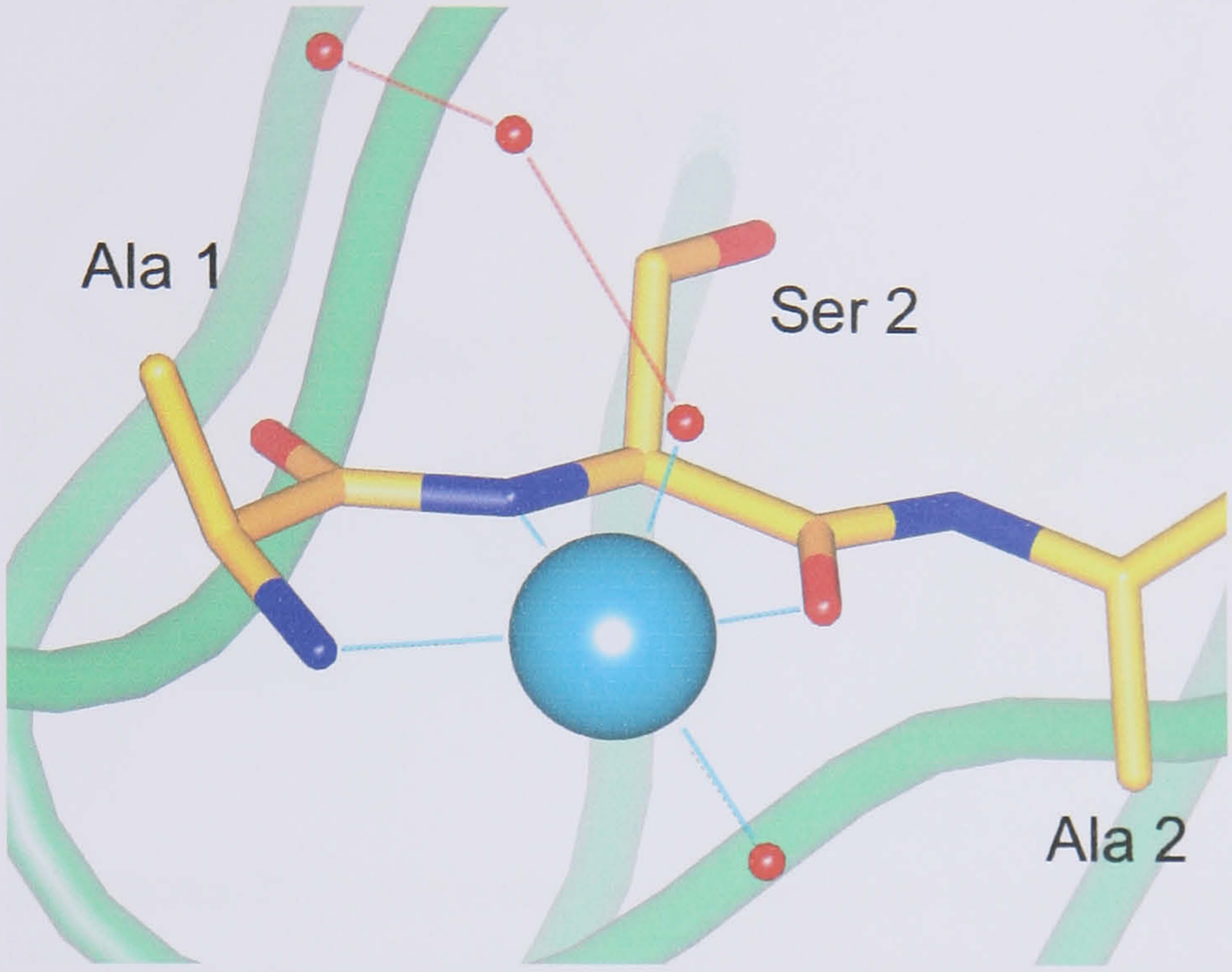


Figure 5.16: The refined N-terminal Cu^{2+} binding geometry in galactose oxidase.

Cyan lines are copper bond distances, red lines are hydrogen bonds. The copper is found coordinating three polypeptide main chain atoms; the terminal NH_2 of Ala 1, the deprotonated amide, and a carbonyl oxygen. Water molecules are found occupying the fourth equatorial and the axial ligand positions. The overall geometry can be described as elongated square pyramidal.

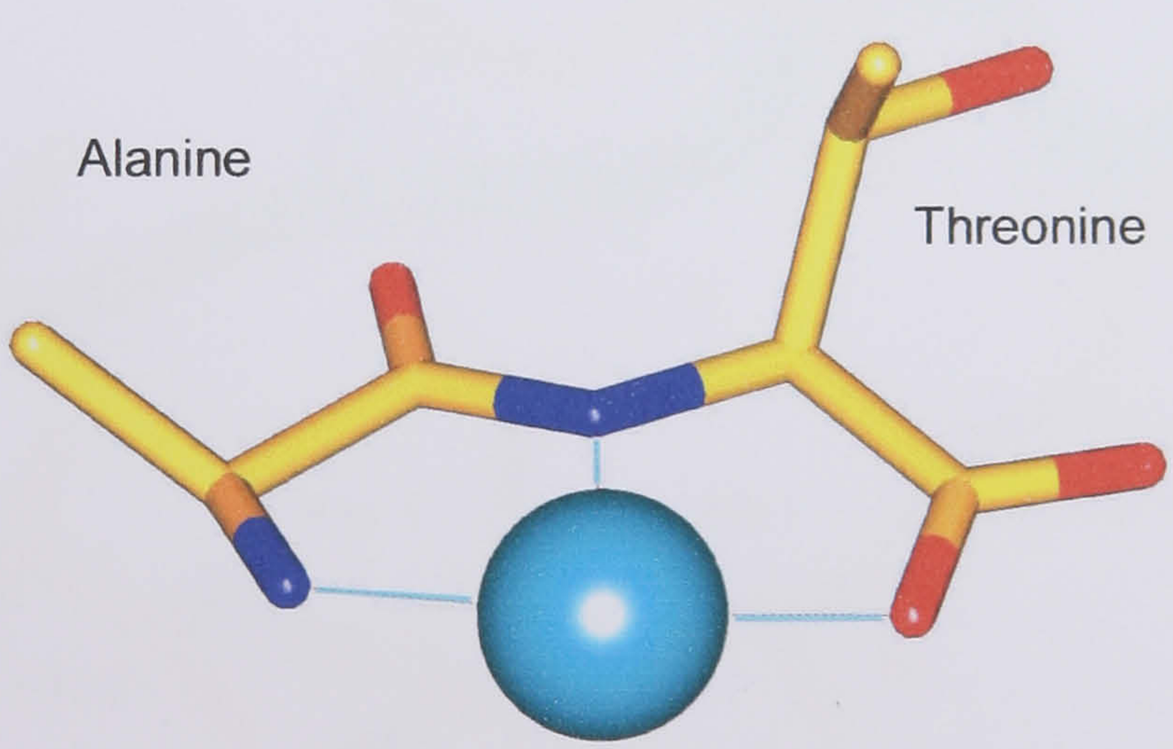


Figure 5.17: The structure of a Cu^{2+} di-peptide complex.

The copper is seen coordinated by an N-terminal NH_2, a de-protonated amide and carbonyl oxygen. The fourth equatorial and the axial positions are occupied by oxygen atoms from symmetry related molecules (not shown) [Facchin *et al.*, 2002].

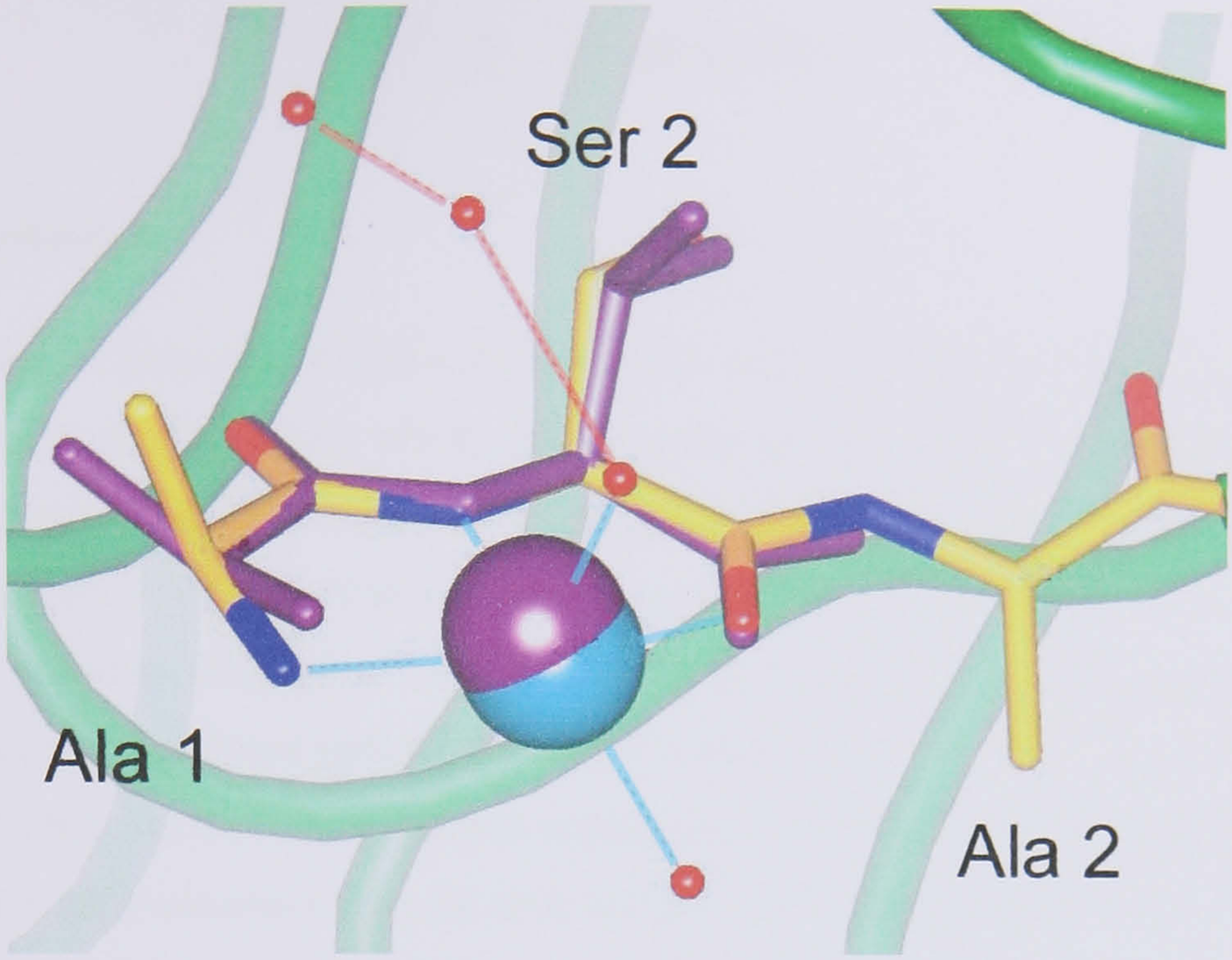


Figure 5.18: Synthetic Ala-Thr-copper complex aligned with the N-terminus copper site.

Synthetic Ala-Thr-copper complex (purple) manually overlaid onto the N-terminal copper binding site in galactose oxidase. The position of the metals and the copper ligands are remarkably similar.

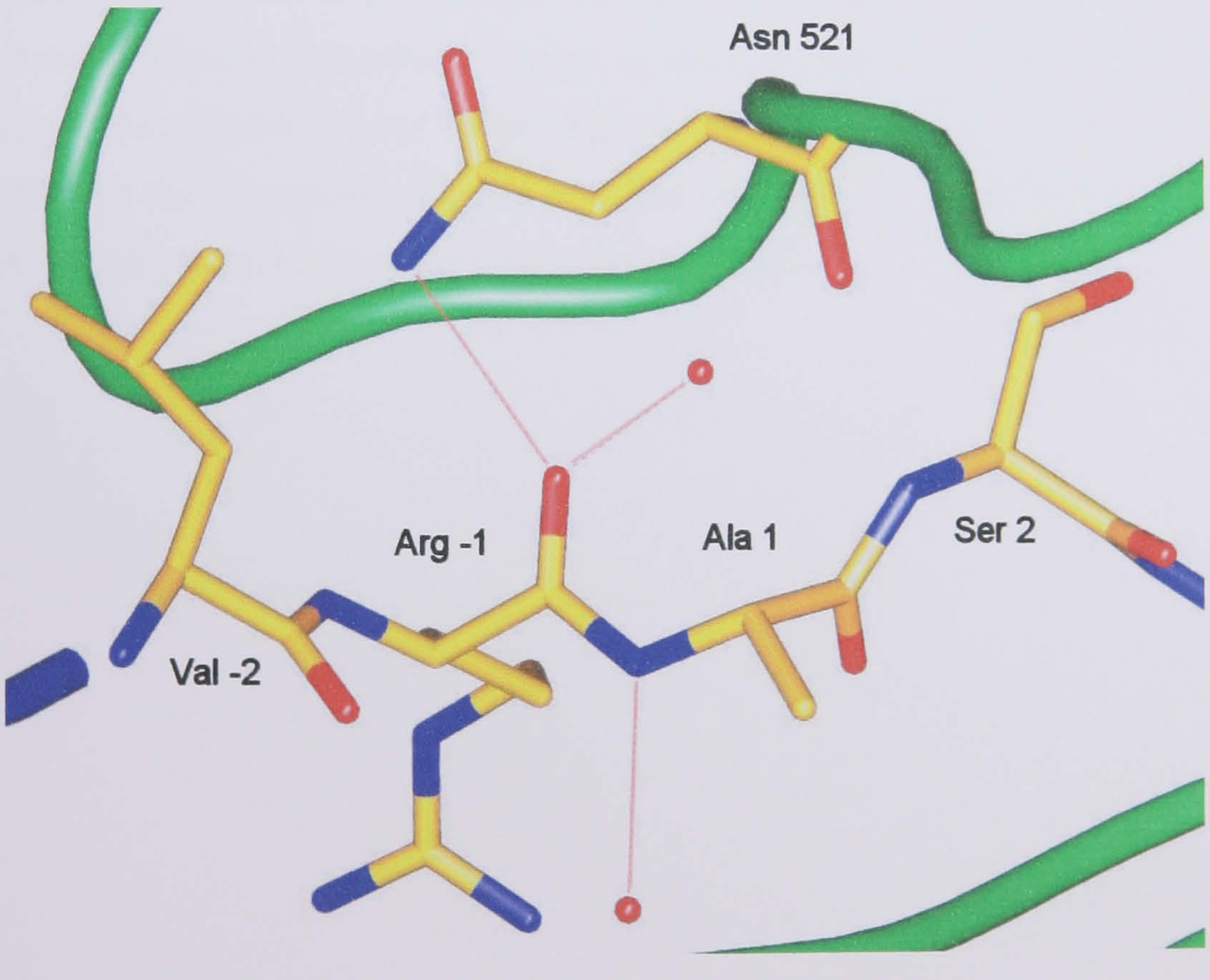


Figure 5.19: The cleavage site of the pro-GO N-terminus.

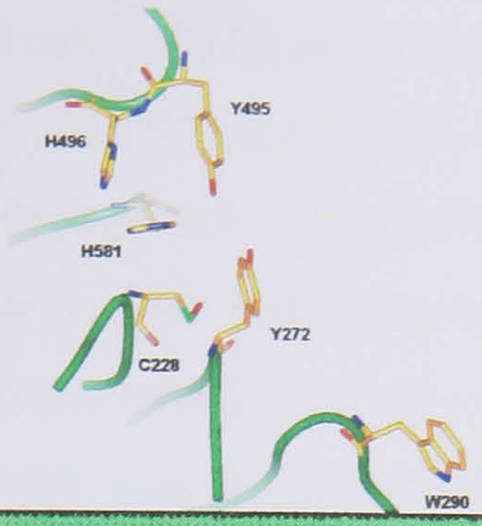
The pro-sequence region of galactose oxidase (PDB code, 1KI3). Cleavage occurs between Arg -1 and Ala 1. Copper soaks have established that the metal can bind to the mature N-terminal, associating with Ser 2 and Ala 1.

6 Discussion

6.1 Introduction

The processing of galactose oxidase from the pro-sequence form to the catalytically active enzyme requires several steps which are currently poorly understood. Upon the aerobic addition of copper, the 17 amino acid pro-sequence is cleaved and the active site Tyr-Cys cofactor is formed. The order of processing appears to commence with pro-sequence cleavage followed by cofactor biogenesis. This is implied by previous results during which limited exposure to copper produced galactose oxidase containing no cofactor and no pro-sequence, but as yet no form has been detected that contained a cofactor and a pro-sequence [Rogers *et al.*, 2000]. It is still uncertain how the enzyme is processed from pro-GO to premature GO, and from there to the mature form (Figure 6.1, steps 1 and 2 respectively).

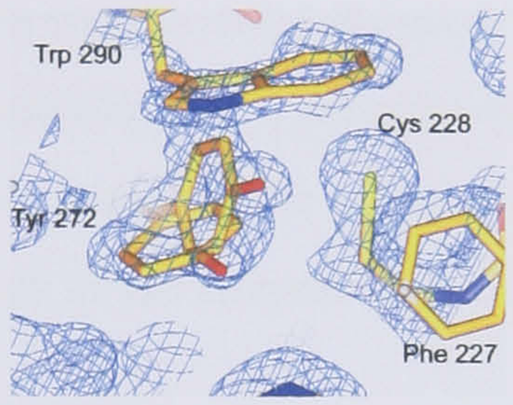
I have presented four structures of premature galactose oxidase determined after crystals were exposed to copper under aerobic and anaerobic conditions. These structures provide a view of initial copper binding to the enzyme active site, followed by apparent cysteine coordination to the bound metal. Finally, following a 10 minute aerobic exposure to copper, the active site structure looks remarkably similar to the original native structure. A copper binding site is also identified at the N-terminus, the pro-sequence cleavage site, thus raising the possibility direct involvement of the metal in the cleavage mechanism.



Pro-Galactose oxidase [Firbank *et al.*, 20011

C228 Y272

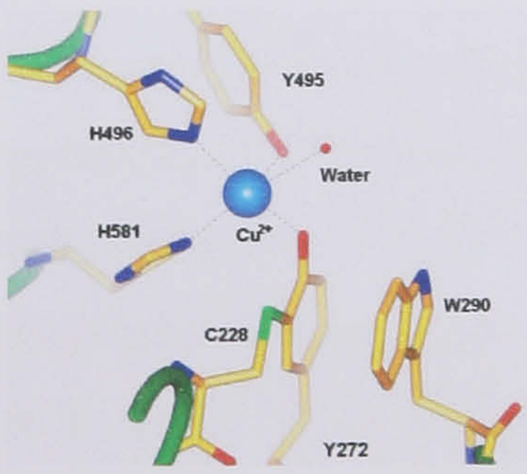
1.Pro-sequence cleavage



Premature galactose oxidase

C228 Y272

2. Cofactor processing



Mature galactose oxidase [Ito *et al.*, 1991]

C228 Y272 Cu^{2+}

Figure 6.1: Galactose oxidase structures, representing different stages of processing.

Structures of pro-GO, premature GO and mature GO have been previously described. Step 1) No structures have been presented which depict pro-sequence cleavage intermediates. Step 2) I have determined structures of cofactor processing intermediates.

After processing and activation, the catalytically active form of the enzyme contains Cu^{2+} and a free radical situated at the cofactor site. The stability of the radical appears to be influenced by the presence of Trp 290, which is proposed to shield the radical site from solvent. Trp 290 also appears to have a catalytic role, via the stabilisation of substrates during catalysis. The W290G variant greatly increased the accessibility of solvent to the cofactor site, in contrast to previously solved structures of W290F and W290H, which offer more protection [Rogers *et al.*, 2007]. Overall these results suggest the hydrophobic portion of Trp 290 protects the radical site from solvent attack, while the hydrophilic part is able to stabilise substrate binding at the active site by providing a hydrogen bonding partner via the Nε1 amide nitrogen.

Structures of a galactose oxidase C383S variant in complex with a possible substrate in one instance and a diatomic molecule in another were shown (Chapter 3). Furthermore, the crystal structure of galactose oxidase containing Cu^{1+}, presents the first observation of the enzyme with a lower oxidation state copper bound at the active site. The active site configuration is consistent with spectroscopic studies of the enzyme in the catalytically fully reduced state.

6.2 Galactose oxidase pro-sequence cleavage

The galactose oxidase cleavage site immediately preceding the mature sequence does not correspond to any known protease site and it was revealed that cleavage occurs spontaneously upon the aerobic addition of copper even in the presence of protease inhibitors [McPherson *et al.*, 1992]. How copper is associated with the pro-sequence prior to cleavage is not known. Galactose oxidase processing has been suggested to be Cu^{1+}-dependent, partly because of the preference of metallochaperones to transport copper in the Cu^{1+} oxidation state [Whittaker *et al.*, 2003]. Such preferences, however, are not true of all chaperones. For example the CopC protein, found in tomato plants, is proposed to transport copper in the oxidising environment of the periplasmic space. The protein features two distinct copper binding sites, one specific for Cu^{1+} and another for Cu^{2+} coordination [Koay *et al.*, 2005]. Information about copper chaperones is increasing, but there is currently no information about how or what form of copper is delivered to galactose oxidase. Furthermore, galactose oxidase is a secreted enzyme, so it is possible that processing occurs outside the cell, where copper is more likely to be freely available.

The role of the galactose oxidase pro-sequence is not certain. In general pro-sequences are known to direct the targeting and secretion of proteins, or act as intramolecular chaperones, activators, or inhibitors [Baardsnes *et al.*, 1998; Eder and Fersht, 1995; Sellman and Tweten,

1997; Tang *et al.*, 2003]. Galactose oxidase does not appear to display any of these activities. The nature of the pro-sequence cleavage event is also unusual, as it is copper-dependent and mutation of the residues at the cleavage site (Arg -1 and Ala 1) to phenylalanines did not prevent cleavage [Ogel, 1993]. This suggests the cleavage event is not sequence dependent.

Currently there are no proposed mechanisms for GO pro-sequence cleavage, as the four known protease classes comprise; serine proteases, cysteine proteases, aspartic acid proteases and zinc proteases. The Cys 515 - Cys 518 disulphide bridge is approximately 6 Å from the pro-sequence cleavage site and a serine is also nearby, however, there are no histidine residues close enough to either of these residues to constitute the catalytic triad of a serine/cysteine protease. Cysteine residues readily bind copper, thus the close proximity of two cysteine residues to the cleavage site raises the possibility of copper binding to one of the residues prior to disulphide formation. This possibility requires further investigation, where the study of the double mutant C515A/C518A processing would be informative. It is interesting to note that two aspartate residues, Asp 517 and Asp 192, are near the cleavage site in the mature enzyme (Figure 6.2), suggesting the possibility of an aspartate protease cleavage mechanism. In the previously solved pro-GO structure [Firbank *et al.*, 2001], however, Asp 517 is 6.6 Å away from the Arg -1 carbonyl, while Asp 192 is displaced approximately 13 Å from its final position in the mature (Figure 6.3).

Copper-dependent proteolysis was reported to be more efficient than the native zinc reaction in the astacin-like metalloendopeptidase serralysin [Park and Ming, 2002]. Serralysin is a Zn endopeptidase excreted by the opportunistic pathogen *Serratia marcescens*. The enzyme active site consists of Zn coordinated by three histidines, tyrosine and a water molecule in a trigonal bipyramidal geometry. The water molecule is within hydrogen bonding distance to the tyrosine and glutamate. The zinc protease motif has not been identified in galactose oxidase, although a tyrosine and glutamate do reside close to the cleavage site in pro-GO (Figure 6.2). Previous studies in which galactose oxidase was incubated with zinc did not induce pro-sequence cleavage, effectively dismissing the possibility of zinc protease activity [Rogers *et al.*, 2000].

I have determined the structures of apo-galactose oxidase and premature galactose oxidase, incubated with copper. Although these forms do not possess the pro-sequence, site specific copper binding has been observed at the N-terminus. The copper is coordinated in a square pyramidal geometry (Figure 6.3).

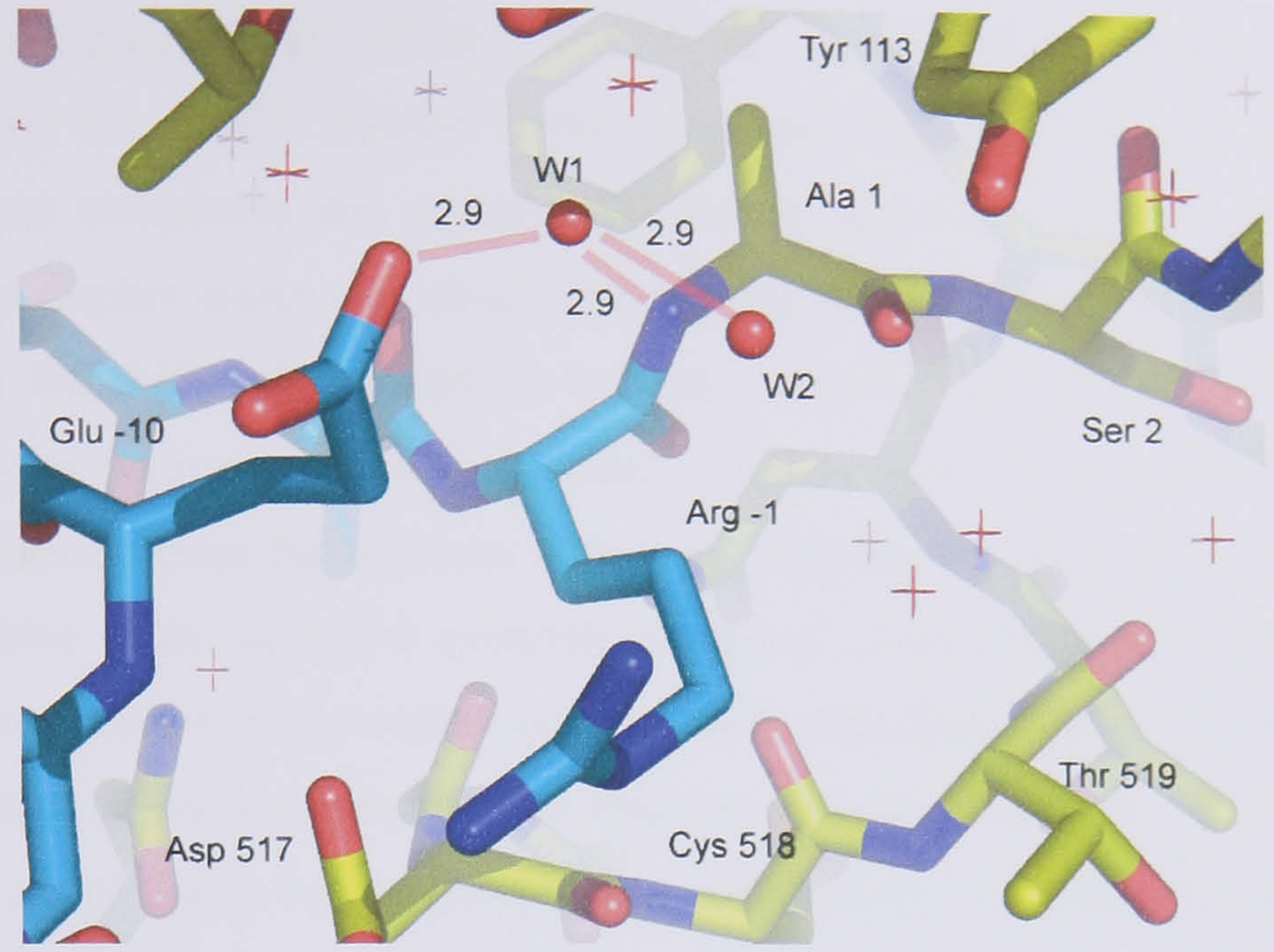


Figure 6.2: The pro-sequence region of galactose oxidase.

The galactose oxidase pro-sequence cleavage site in the pro-GO structure (PDB 1KI3, Firbank *et al.*2001). A glutamate residue (Glu -10) is close to the cleavage site, along with several water molecules. Water is represented by red spheres and crosses, and red lines represent hydrogen bond distances (Å).

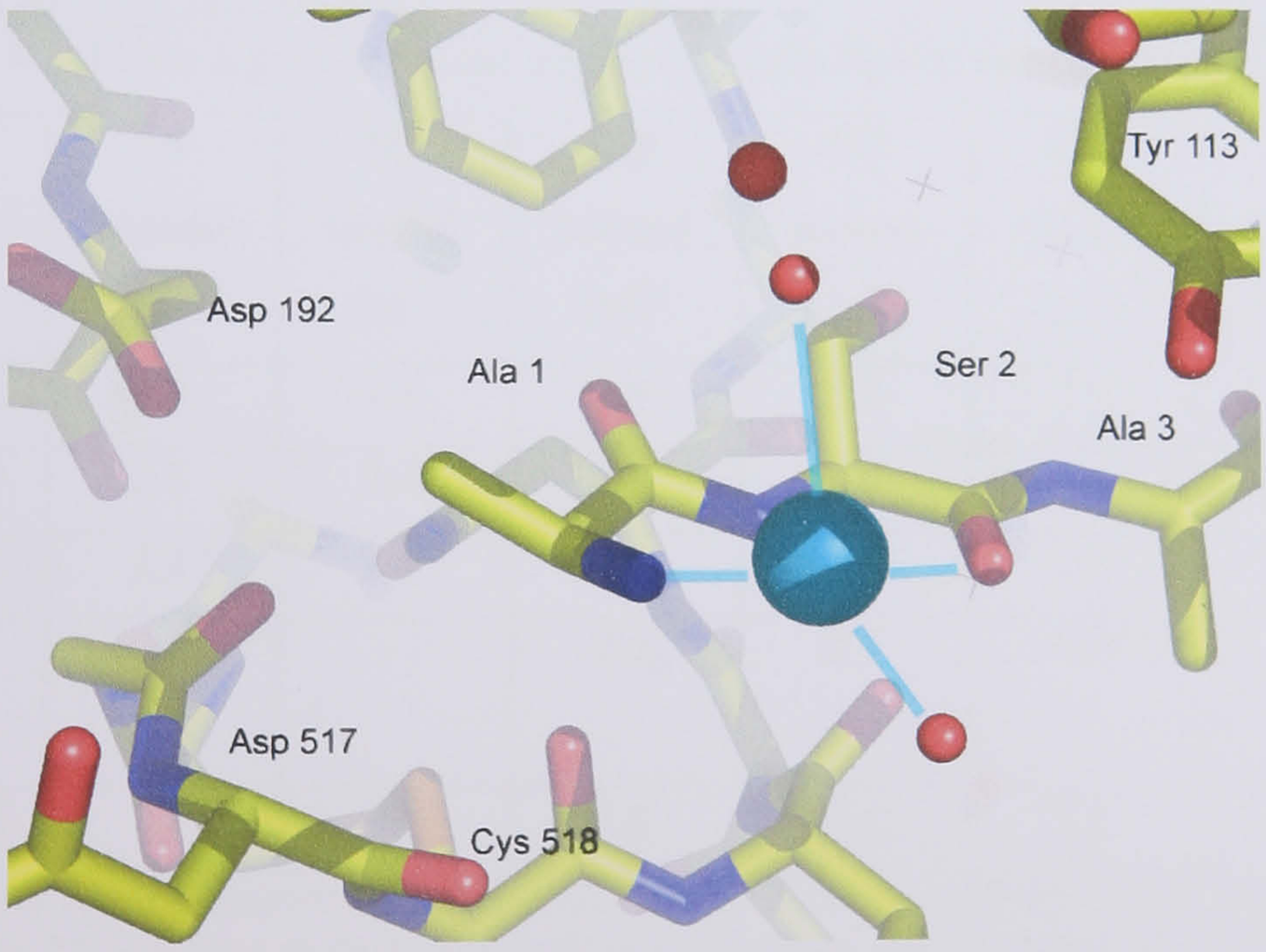


Figure 6.3: N-terminal copper binding in mature galactose oxidase.

Two aspartate residues are located close to the cleavage site. The disulphide bridge between Cys 515 - Cys 518 is also nearby. Water molecules are represented by red spheres, copper ligand bonds are shown in cyan.

Copper binds to the N-terminal region with a five, four or three coordinate geometry (Table 6.1). There is no precedence in the literature for this manner of copper binding to the N-terminal regions of proteins, although peptides mixed with Cu^{2+} have been shown to form very similar five coordinate complexes [Facchin *et al.*, 2002] (Figure 5.18).

Dipeptides comprising Ala and either Val, Phe, Ile, Thr or Tyr each form complexes with copper that look remarkably similar to the galactose oxidase N-terminal complex. These dipeptide-copper complexes are reported to display anti-inflammatory activity and it is suggested they display superoxide-dismutase (SOD)-like properties. The SOD-like activity was judged by the ability of the peptides to mimic the inhibitory effects of SOD in a redox reaction involving superoxide. SOD inhibition of the reaction was derived from its ability to sequester superoxide from the reaction, thus suggesting the dipeptide may possess the same property. Another study utilised a synthetic complex, 2,6-bis-(benzimidazol-2-yl) pyridine, with copper. Introduction of the copper complex to Bovine Serum Albumin (BSA) was found to initiate oxygen-dependent site specific cleavage [Shrivastava *et al.*, 2002]. The authors suggest metal coordination by the polypeptide, followed by activation of molecular oxygen by the metal, induced polypeptide cleavage. The metal binding site, however, was not reported.

Table 6.1: The N-terminal copper complexes observed after copper incubation with galactose oxidase. The 20 second, 90 second, 120 second and 600 second data were observed following copper incubation with premature galactose oxidase. The 180 second and 25 minutes data were observed after incubation with apo-galactose oxidase.

Table 6.1: N-terminal copper complex ligand distances (Å)

Ligands	20 second Cu^{2+}	90 second Cu^{2+}	120 second Cu^{2+}	180 second Cu^{2+}	600 second Cu^{2+}	25 minute Cu^{1+}	Synthetic Cu^{2+} complex[a]
Cu –Ala 1 (N1)	1.9	2.4	1.9	2.3	1.9	2.6	2.0
Cu – Ser 2 (N1)	2.1	2.4	1.8	2.0	2.1	2.4	1.9
Cu – Ser 2 (O)	2.1	2.1	2.2	1.9	2.0	2.4	2.0
Cu – H_2O (equatorial)	-	-	2.2	2.7	2.1	-	2.0*
Cu – H_2O (axial)	2.3	2.4	3.0	2.2	2.5	-	2.3**

[a]distances refer to Ala-Thr peptide [Facchin *et al.*, 2002]. * alanine carbonyl oxygen from a symmetry related molecule. **Threonine methyl oxygen from a symmetry related molecule

The discovery of a new N-terminal copper binding site represents an opportunity to examine the possibility of copper binding directly to the cleavage site in pro-GO, although, it is noted that the orientation of the polypeptide chain at Ala 1 of the mature sequence is quite different from the same region in pro-GO (Figure 6.2 and Figure 6.3 respectively). In the mature protein the carbonyl oxygen of Ala 1 is pointing in the opposite direction compared to pro-GO, therefore copper would not be able to coordinate the amide nitrogen of Ser 2 as observed

in the mature protein without a conformational change. Considerably more water molecules are present at the pro-GO cleavage site region than observed in the mature enzyme. One of these molecules (W1) is located 2.9 Å from the peptide nitrogen between Ala 1 and Arg -1, and it is tempting to speculate that W1 could be occupying a potential copper binding site. In pro-GO it is not known if copper could be coordinated by the carbonyl oxygen of Arg -1, or if the metal would replace water in the W1/W2 position. Copper soak experiments with pro-GO crystals are needed in order to investigate this possibility. Previous studies, however, have shown that incubation of pro-GO crystals with copper cause the dissolution of the said crystals [Firbank, 2002], so copper soak experiments with pro-GO protein crystals are difficult. Further investigation of the new copper binding site and the cleavage reaction is required.

6.3 Thioether bond formation

Premature galactose oxidase can be expressed in *Pichia pastoris* from a gene lacking the pro-sequence. This form of the enzyme processes in the presence of copper and oxygen to form the fully active enzyme [Rogers and Dooley, 2001], so it is useful for studying thioether bond formation in isolation from the pro-sequence cleavage event. SDS-PAGE analysis of premature galactose oxidase processing has shown that the mature enzyme can be detected after aerobic incubation with Cu^{2+} for 2 minutes (Figure 1.22) [Firbank, 2002]. Past copper soak experiments, however, were carried out under anaerobic conditions, with a copper incubation time of four minutes. The structure revealed no thioether bond formation and the active site metal coordinated to the two active site histidines and the Sγ of Cys 228 [Firbank, 2002]. This thesis presents an extension of this work, in which the investigation of processing intermediates is observed following copper incubation (anaerobic and aerobic) for less than two minutes and a 10 minute soak.

The structures I have determined suggest Cys 228 begins to change conformation from the premature galactose oxidase position to one that is closer to the copper atom, prior to the final mature conformation. Furthermore, the structural evidence confirms the presence of the thioether bond within 10 minutes of the introduction of copper. This contradicts claims by Whittaker that Cu^{2+}-dependent cofactor biogenesis is an extremely slow process with a half-life of several hours [Whittaker and Whittaker, 2003]. The Cu^{1+}-dependent processing reaction was proposed to be faster, as assessed by the rapid development of the characteristic absorbance spectra of the oxidised enzyme. The galactose oxidase UV-visible spectrum is used to detect the presence of the radical, via the absorbance peak at 445 nm. Consequently,

previous Cu^{2+}-dependent biogenesis experiments have interpreted a slow development of the free radical as a slow cofactor biogenesis reaction.

Generally, cofactor biogenesis mechanisms suggest initial copper coordination by Tyr 272, His 581 and His 496, thus forming a pre-cofactor complex. Molecular oxygen is then proposed to bind to copper and induce radical attack of Cε of Tyr 272 by Cys 228-Sγ$^{\bullet}$ [Rogers and Dooley, 2001; Schwartz and Klinman, 2001; Whittaker and Whittaker, 2003]. There were no indications, however, that copper is initially coordinated by Tyr 272. Taking the previous structural and solution studies into account, the structures I have determined support the Firbank cofactor biogenesis mechanism (Figure 1.25), in which copper associates with Cys 228 prior to thioether bond formation (Figure 6.4).

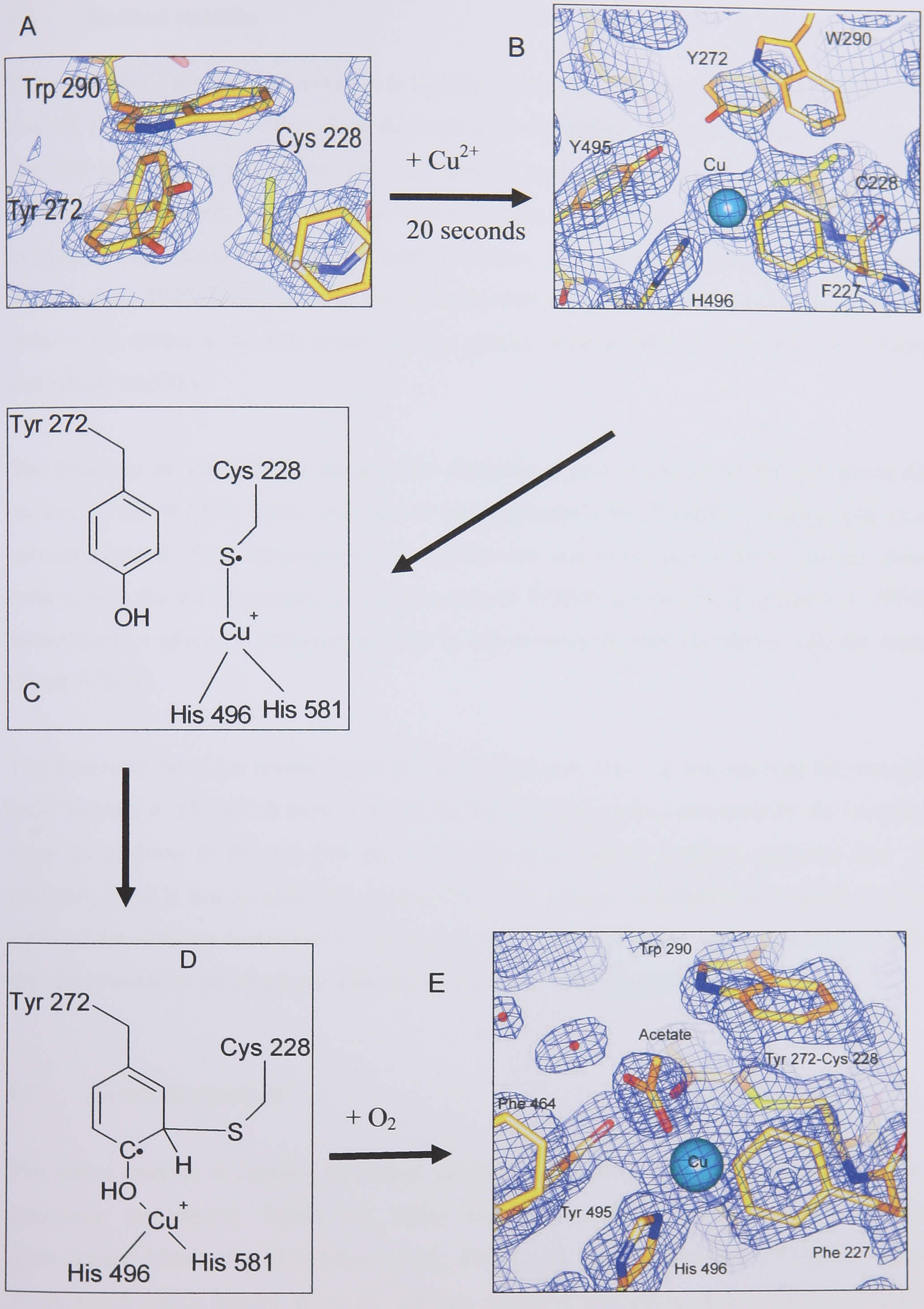


Figure 6.4: Galactose oxidase cofactor biogenesis mechanism.

A: premature galactose oxidase. **B**: copper binding. **C**: Cu^{2+}/Cu^{1+} resonance leads to formation of sulphur radical. **D**: thioether bond formed via radial attack. **E**: Oxygen oxidises copper producing the semi-reduced enzyme. 10 minute aerobic copper soak revealed an active site similar to the native enzyme.

6.4 Radical stability

The presence of the Tyr-Cys cofactor is thought to increase the ease with which the radical is formed, however, the stability of the free radical is dependent on the active site environment. Trp 290 stacks over the radical site and appears to make a major contribution to radical stability. Furthermore, the residue also seems to assist substrate binding. Mutation of Trp 290 to histidine maintains the ability to bind galactose, but reduces the stability of the radical [Baron *et al.*, 1993]. Alternatively, the phenylalanine mutation maintains radical stability but reduces the ability to bind to galactose. The glycine mutant showed poor galactose binding and radical stability.

The structure of the W290G mutant was determined and revealed the solvent accessible surface of the W290G active site pocket and particularly the thioether cofactor was more solvent exposed. As a consequence the radical site was more accessible to solvent attack compared to the wild type enzyme and the mutants W290H and W290F [Rogers *et al.*,2007]. Interestingly a glycerol molecule, present as cryoprotectant, was associated with the active site of W290G.

The structural data also revealed that the tryptophan side chain is not essential for thioether bond formation. The 6.3 Å movement of the Trp 290 side chain (measured by the Cα atom) from its position in the pro-GO protein to the final mature location, stacking over the thioether bond is not an absolute requirement with regards to cofactor processing. Further study of the cofactor formation in this mutant form would be necessary to establish whether the rate of cofactor processing is affected by the lack of the tryptophan.

6.5 Substrate catalysis

The active enzyme is capable of radical catalysis of primary alcohols to aldehydes, via a ping-pong mechanism. While the initial enzyme reductive half-cycle has been well characterised kinetically [Whittaker *et al.*, 1998] and spectroscopically [Whittaker *et al.*, 2000], much less is known about the enzyme oxidative half-cycle. Azide inhibits galactose oxidase [Whittaker and Whittaker, 1988] by binding to the active site copper at the substrate-binding site. Several structures of azide bound galactose oxidase have been determined [Ito *et al.*, 1995; Vinecombe, 1999; Rogers *et al* 2007], all confirming azide replaces the acetate or water molecule in the mature enzyme structure (Figure 6.5) [Ito *et al.*, 1995]. Crystals soaked with sugar substrates, however, failed to produce an enzyme-substrate structure. This thesis presents the first clear observation of a true substrate binding to the active site (chapter 3).

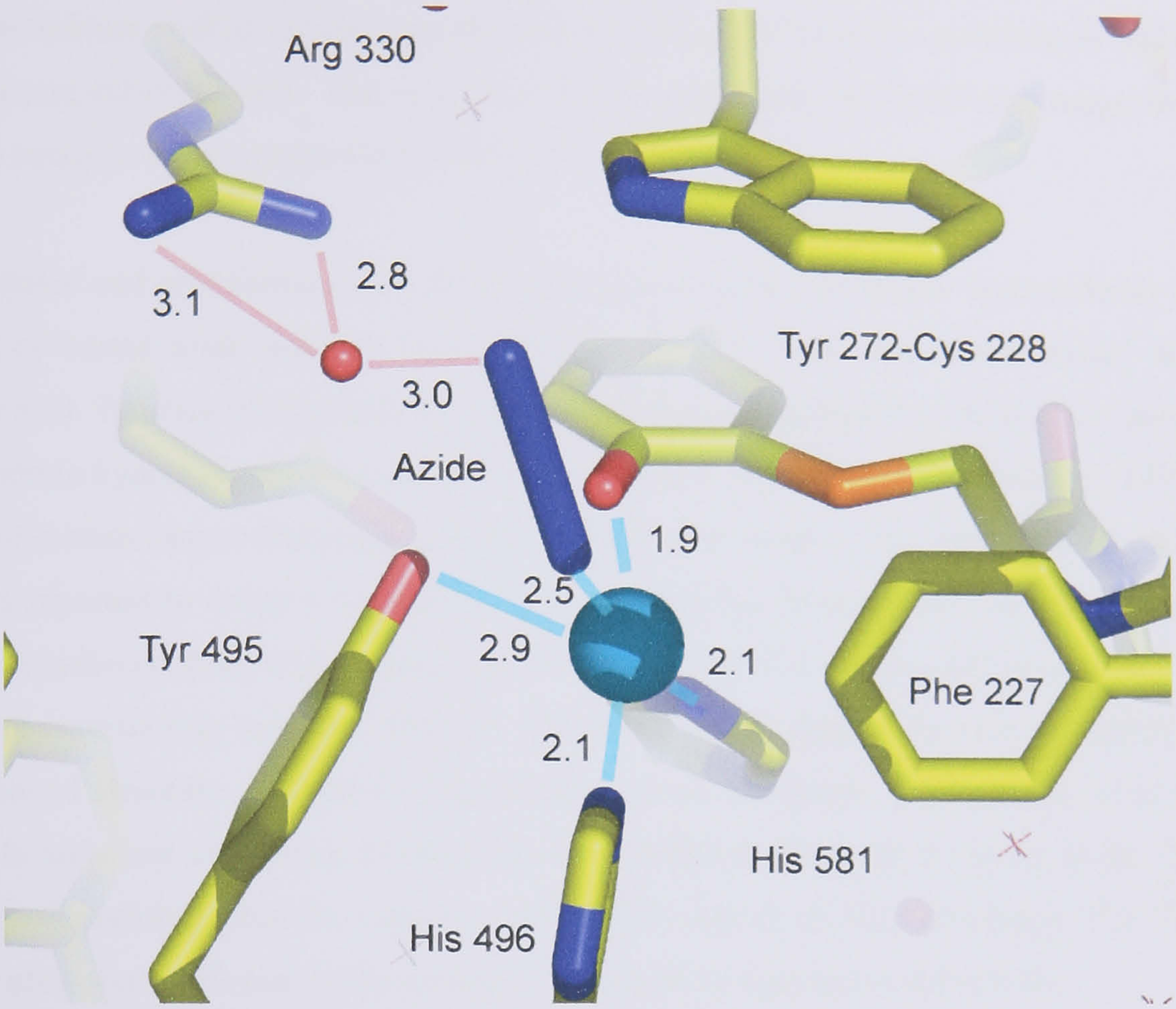


Figure 6.5: The active site structure of azide bound wild type galactose oxidase.

1.8 Å crystal structure of azide bound wild type galactose oxidase [Rogers *et al.*,2007]. Tyr 495 is dissociated from copper upon azide binding in solution and appears to do the same in the crystal structure, thus the copper appears to adopt a four ligand coordination.

Crystals of the galactose oxidase mutant C383S were used for substrate binding studies. The crystal incubated with 2MPD for 60 seconds revealed a copper bound "substrate", whereas the three minute soaked crystal revealed the presence of a diatomic molecule in addition to more distant substrate peak. The molecules bind to the fourth equatorial site, suggesting they may be catalytic substrates/products (Section 3.3.2).

The position and orientation of the diatomic molecule in the active site is remarkably similar to that of bound azide and the bound putative 2MPD in the short soak crystal structure (Figure 6.6). Binding of substrate to galactose oxidase is proposed to initiate the transfer of the substrate hydroxyl proton to the phenolate Tyr 495 [Whittaker and Whittaker, 1993]. Tyr 495 protonation causes dissociation of the residue from copper. The binding of the inhibitor azide is reported to cause a similar Tyr 495 dissociation from copper. In the absence of a substrate hydroxyl from which to acquire the proton, Tyr 495 is proposed to acquire a proton from the surrounding solution. The Tyr 495 bond length appears to have increased in the azide bound structure compared to the substrate soak structures, although the error is such that it is only just significant (Table 6.2). It is difficult, however, to know if the Tyr 495 ligand is dissociated from the copper in the 2MPD soaked crystals structures. The Tyr 495-copper distances in the non-azide structures do not show significant differences.

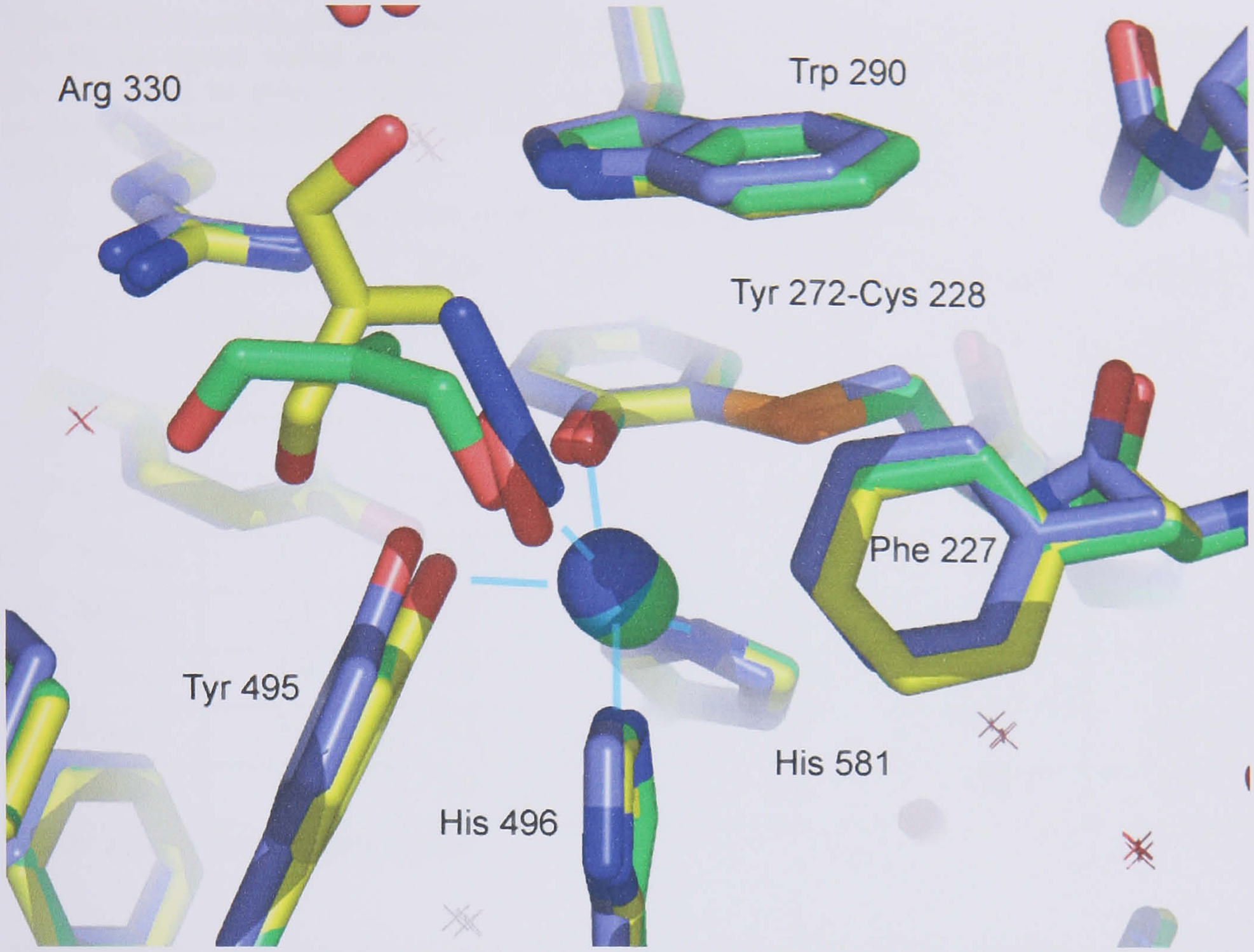


Figure 6.6: Alignment of the substrate/diatomic/azide bound active sites structures.

Active site overlay of azide bound wild type galactose oxidase (blue carbon atoms), C383S 60 second substrate soak (green carbon atoms), and C383S 3 minute substrate soak (yellow carbon atoms). The binding orientation of the putative peroxide, the copper-associated substrate hydroxyl and azide are very similar. Tyr 495 in the azide structure has the longest hydroxyl to copper distance.

Table 6.2: Comparison of bond distances for copper and its ligands in the original pH 4.5 structure (1GOF), the crystal soaked into non-acetate buffer at pH 7.0 (1GOG), the azide soaked structure [Firbank, 2002], the galactose soaked crystal, and the 60 second and 3 minute C383S 2-methylene-1,3-propandiol soaked structures. L is the fourth equatorial copper ligand, while * indicates L in each structure.

Table 6.2: Galactose oxidase active site copper ligand Distances (Å)

	1GOF (acetate*)	1GOG (H_2O*)	Azide* [Firbank, 2002]	MPD* 60 second soak	MPD soak (diatomic*)	Galactose soak (H_2O*)
Cu^{2+} –L*	2.3	2.8	2.5	3.0	2.6	3.1
Cu^{2+} –$O_{(495)}$	2.7	2.6	2.9	2.7	2.6	2.8
Cu^{2+} –$N\varepsilon2_{(496)}$	2.1	2.2	2.1	2.0	2.1	2.0
Cu^{2+} –$N\varepsilon2_{(581)}$	2.1	2.2	2.1	2.1	2.2	2.2
Cu^{2+} –$O_{(272)}$	1.9	1.9	1.9	2.0	1.9	2.0
L*–$O_{(495)}$	3.4	3.7	3.6	3.4	3.6	4.2
L*–$O_{(272)}$	2.8	3.0	2.6	2.7	2.9	2.7

MPD= 2-methylene-1,3-propandiol.

The C383S crystals were not treated with oxidant to ensure the presence of the radical and thus catalytic activity, so the exact identity of the copper bound molecules is uncertain. If the majority of enzyme molecules were in the semi-reduced state, the enzyme would be catalytically inactive. Introduction of substrate would therefore be expected to produce a structure representing an average of enzymes in the initial substrate to copper binding step and the Tyr 495 protonation step of substrate oxidation (Step A in Figure 1.16). The putative 2MPD peak from the shorter soak was fully occupied, suggesting the vast majority of enzyme active sites in the crystal have some kind of bound substrate. Alternatively, if the crystals were catalytically active, the structure would be expected to represent the state preceding the rate determining step of the mechanism (Figure 6.7). This is proposed to be the C-H bond cleavage by Tyr 272$^{\bullet}$, which occurs during substrate oxidation (Step B in Figure 1.16) [Whittaker *et al.*, 1998].

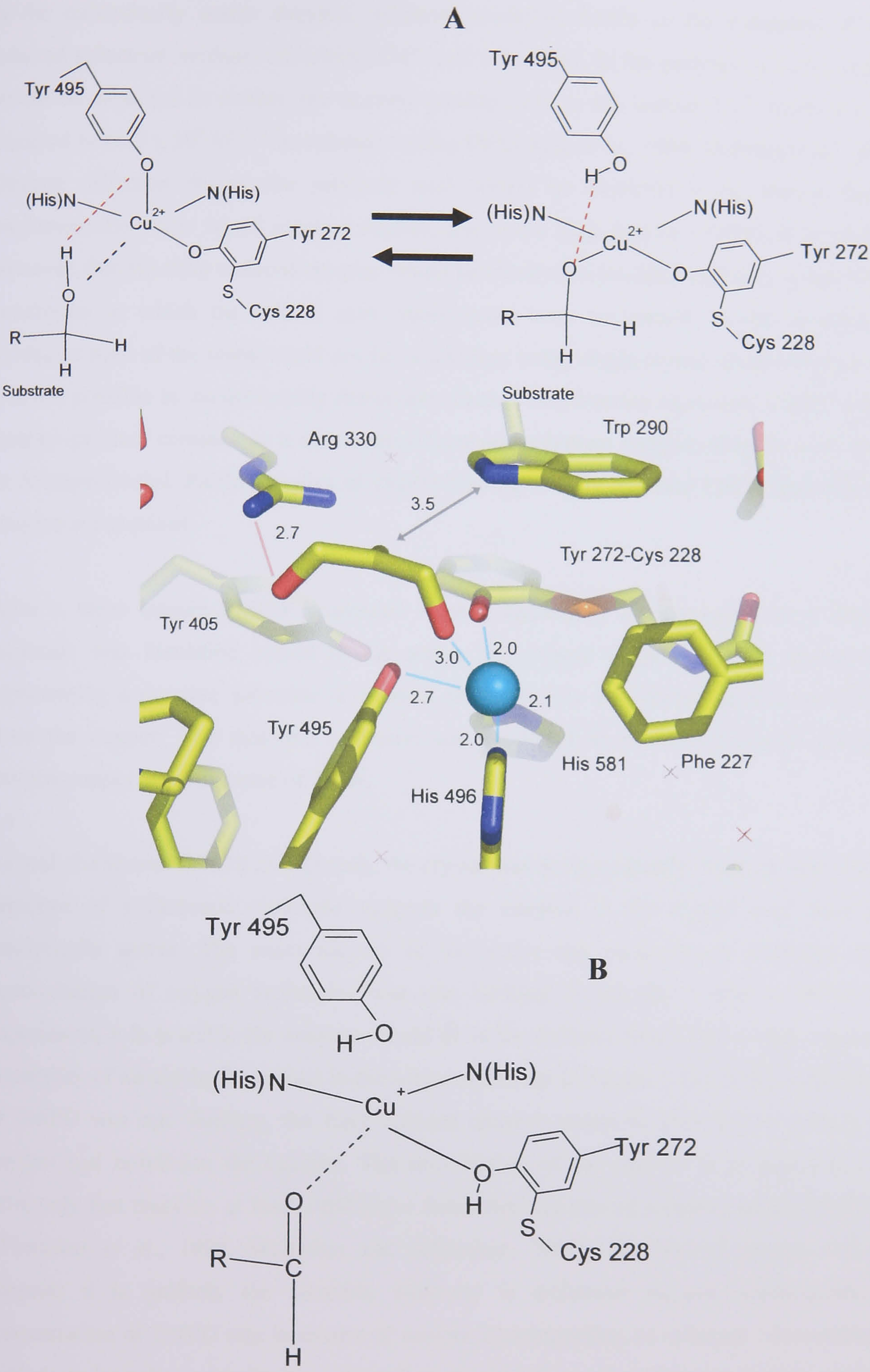


Figure 6.7: The possible identity of the copper ligand after the 60 second substrate soak.

The "substrate" bound active site structure which could represent an initial substrate deprotonation complex (**A**) or a product bound enzyme (**B**). Red lines represent hydrogen bonding.

In the catalytically active enzyme, substrate oxidation results in the formation of fully reduced galactose oxidase, containing Cu^{1+} and no radical. In the presence of Cu^{1+}, oxygen would be expected to oxidise the enzyme quickly back to the radical, Cu^{2+} form, at a rate reported to be 8×10^6 $M^{-1}s^{-1}$ in solution studies [Whittaker *et al.*, 1998: Borman *et al.*, 2001]. Oxygen diffusion during the substrate soak would be expected to be slower than in predominantly water based solution studies, due to the high density of PEG. It is unlikely, however, that the fully reduced enzyme would be observed crystallographically in the aerobic conditions in which the C383S soak experiments were performed. In this instance the oxidation state of the metal could not be established using single crystal spectroscopy and so it is not possible to unequivocally determine whether the structure represents a fully reduced enzyme-product complex or a semi-reduced enzyme-substrate complex (Figure 1.16, step C or A respectively). Further studies of 2MPD binding to fully oxidised C383S and wild type enzyme are required.

After a three minute 2MPD incubation a peak apparently corresponding to a diatomic molecule was identified bound to the active site copper (Figure 3.31), a second peak representing a putative substrate or product molecule was also identified at some distance from the copper. The diatomic molecule was modelled as putative peroxide, giving an oxygen-copper bond distance of 2.6 Å.

As with the shorter C383S 2MPD soak, the crystal was not specifically oxidised, however, the presence of a diatomic molecule suggests the enzyme in the crystal may have been catalytically active. The exact identity of the active site molecules is uncertain. If the concentration of oxygen molecules was rate limiting during the 3 minute 2MPD soak experiment, it is possible the enzyme would be in the reduced form. This would suggest the possibility of an aldehyde product in the active site (Step D, Figure 1.17). If the concentration of 2MPD was rate limiting, the fully reduced enzyme would be expected to quickly bind oxygen and re-oxidise the enzyme. The re-oxidation of the enzyme is proposed to be an extremely fast reaction, at least 1000 times faster than the rate of substrate induced reduction [Whittaker *et al.*, 1998; Whittaker and Whittaker, 1988]. The rate of oxygen reduction suggests it is unlikely the diatomic molecule is molecular oxygen. Furthermore, the concentration of 2MPD was in excess of enzyme concentration, so substrate concentration is unlikely to have been rate limiting. Therefore, the crystal structure is more likely to represent hydrogen peroxide molecules associated to the copper prior to product release (Figure 6.8).

A hydrogen peroxide bound structure could be in resonance with a hydroxyl peroxide (HOO^-) bound structure, prior to deprotonation of Tyr 495. The identity of the second exogenous

molecule nearby could be a substrate or product molecule (Figure 1.17 step D to E). The structure could represent an enzyme in which the rate determining step is product release. In this scenario, strong hydrogen bond interactions between the active site residues, the product and the high concentration of surrounding substrate/product would delay release.

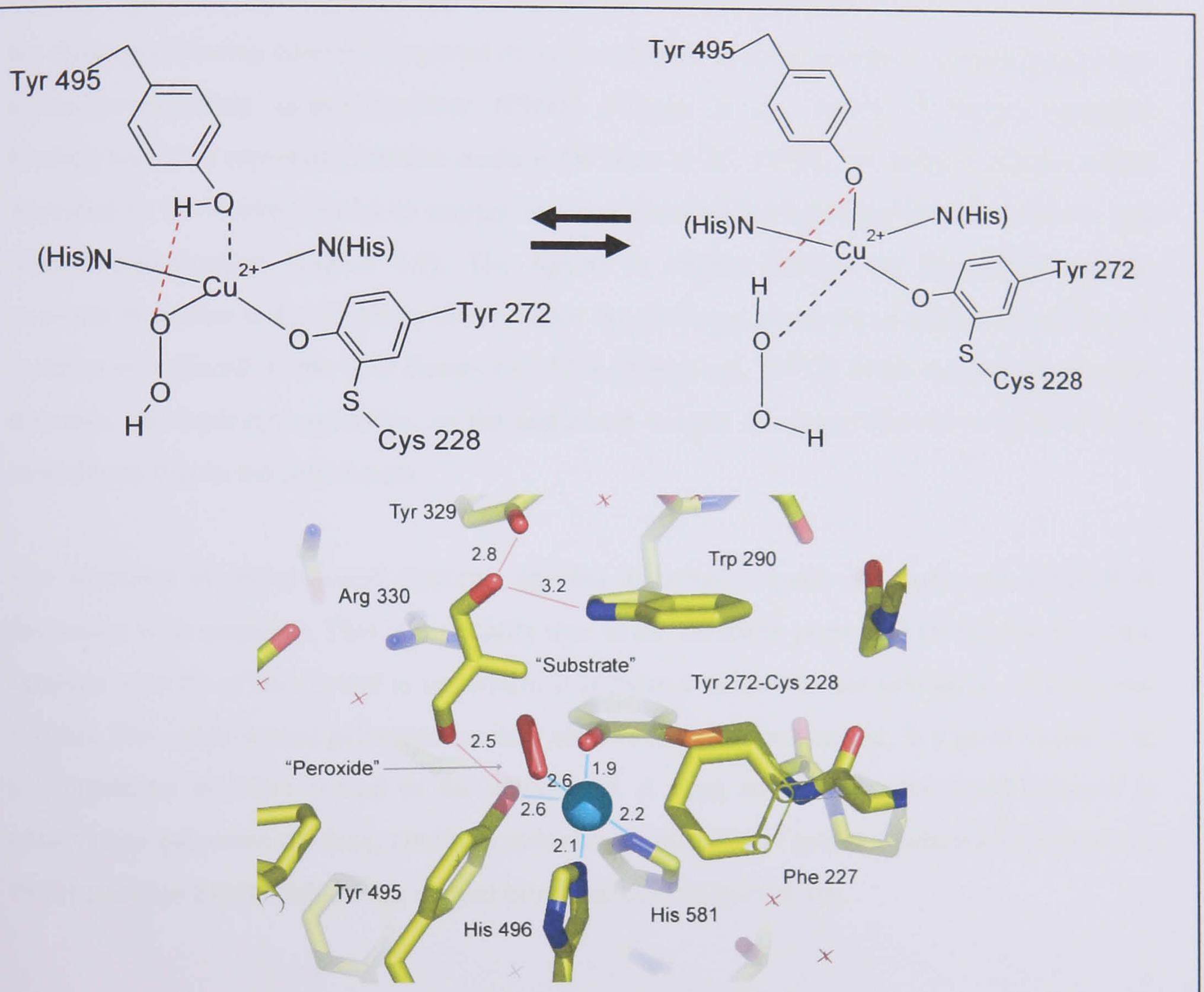


Figure 6.8: Possible identity of the copper ligand following the 3 minute substrate soak.

The 3 minute substrate soak revealed a diatomic bound copper at the active site. The structure resembles the proposed peroxide binding predicted to occur during catalysis. The high concentration of substrate/product could be obstructing the release of hydrogen peroxide and thus allowing observation of an enzyme-product/substrate species. Red lines represent hydrogen bonding.

Synthetic chemistry has provided models for mononuclear copper coordination complexes containing dioxygen bound to the metal in two distinct coordination modes: side-on and end-on [Aboelella *et al.*, 2002]. Side-on oxygen-binding requires the availability of two adjacent coordination sites on the metal ion, therefore, the end-on Cu^{2+}-superoxide adduct would be expected to be the favoured configuration for galactose oxidase. The "peroxide" ligand in the C383S mutant structure approaches copper in an end-on configuration. This is similar to the hydrogen peroxide binding observed with ascorbate oxidase [Messerschmidt *et al.*, 1993] and the dioxygen binding recently observed in a pre-catalytic enzyme complex of peptidylglycine-alpha-hydroxylating monooxygenase (PHM) [Prigge *et al.*, 2004]. Hydrogen peroxide binding was also reported in amine oxidase [Wilmot *et al.*, 1999], revealing a species which appeared to be weakly bound to copper in an orientation intermediate between end-on and side-on coordination (Figure 6.9). The ligand to copper distance in the amine oxidase structure is similar to the distance observed for the diatomic molecule in galactose oxidase. A restraint was placed on the O-O distance (1.47 Å [Kleywegt, 1998]) in the assumption that the diatomic was hydrogen peroxide, so the angle and oxygen to copper distance will have been sensitive to this in the refinement

The identities of metal bound diatomic species in protein crystal structures are difficult to determine with certainty. This is especially true in the structure presented in this thesis, as the catalytic activity of the crystal is uncertain. It is more common to use inhibitors, or dioxygen mimics. The azide-bound galactose oxidase structure, mentioned earlier, is a good example of a competitive inhibitor bound to the active site. A long azide to copper bond distance is observed in galactose oxidase, similarly azide is an inhibitor of amine oxidase [Wilmot *et al.*, 1999] and also binds with a long copper bond distance (Figure 6.10).

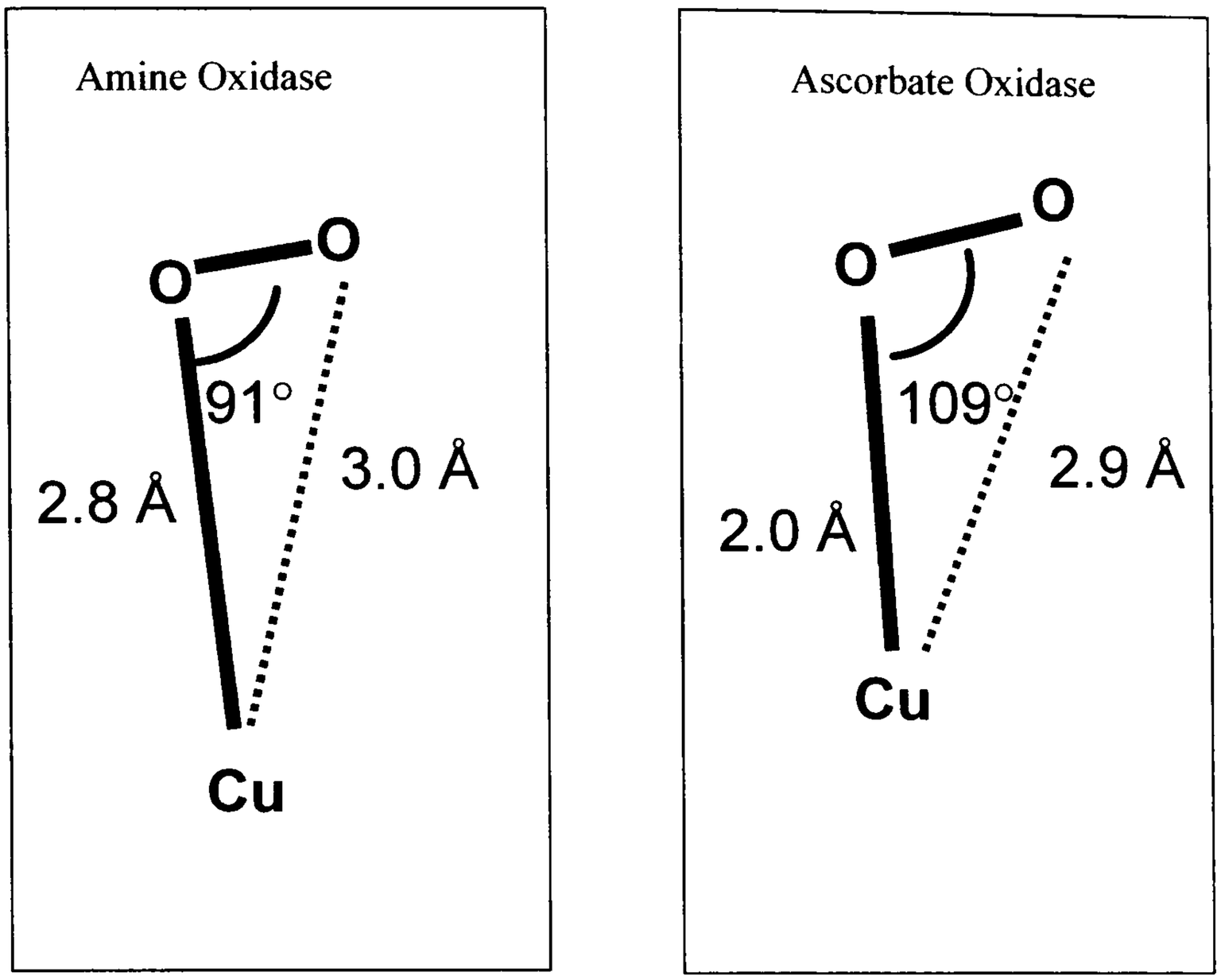
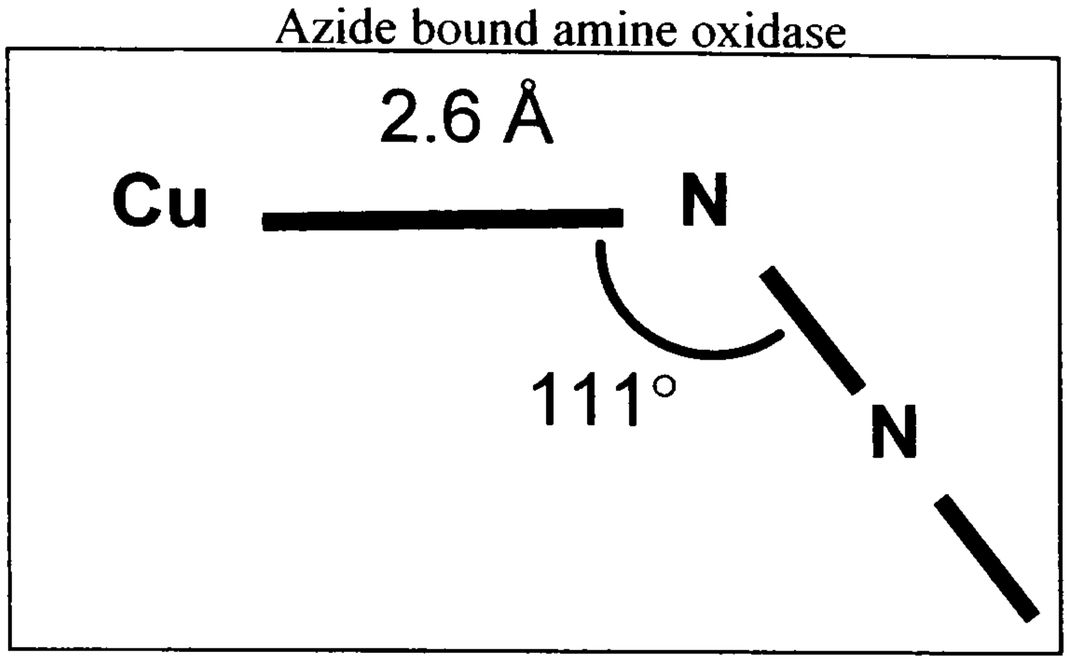


Figure 6.9: Copper oxidases diatomic ligand geometry.

Crystal structures have described the diatomic ligand in amine oxidase [Wilmot *et al.*, 1999] and ascorbate oxidase [Messerschmidt *et al.*, 1993] as hydrogen peroxide adducts. While PHM is described as a dioxygen bound copper [Prigge *et al.*, 2004]. The ligand to copper distance of galactose oxidase is similar to that in amine oxidase, however, the end-on geometry is closer to that of dioxygen coordination in PHM.

Azide bound amine oxidase



Azide bound galactose oxidase

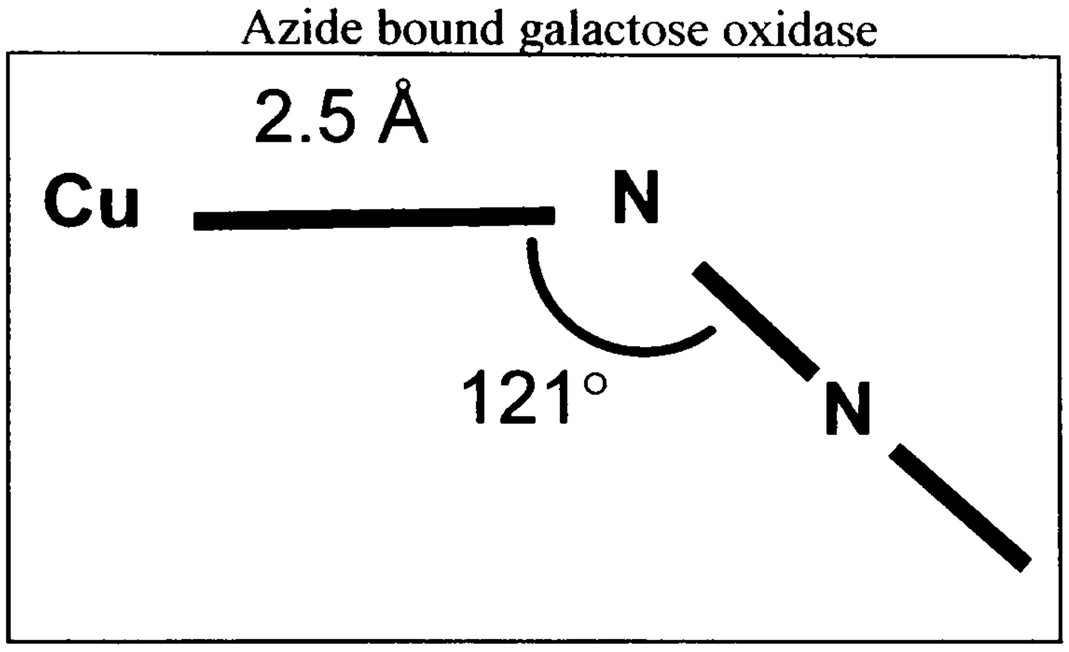


Figure 6.10: The Azide to copper geometry in amine oxidase and galactose oxidase.

Azide binds closer to the galactose oxidase copper and with a more obvious end-on configuration.

Although there is some uncertainty about the identity of the exogenous copper bound ligands observed following the 2MPD soak. A comparison of the galactose manual docking model [Ito *et al.*, 1995] with the experimentally derived 2MPD soak structure reveals some intriguing similarities (Figure 6.11). The O5 position of galactose aligns reasonably well with methylene groups of 2-methylene-1,3-propandiol, but is approximately 3.5 Å away from the Nε of Trp 290. This was unexpected, as the galactose O5 would be expected to be closer to Arg 330. This observation suggests manual adjustment of the galactose model docking model may be required to reflect the 2MPD data. Similarly the O4 position is close to the hydroxyl of the copper binding 2-methylene-1,3-propandiol (green in Figure 6.11), but is 3.0 Å from Nη2 of Arg 330.

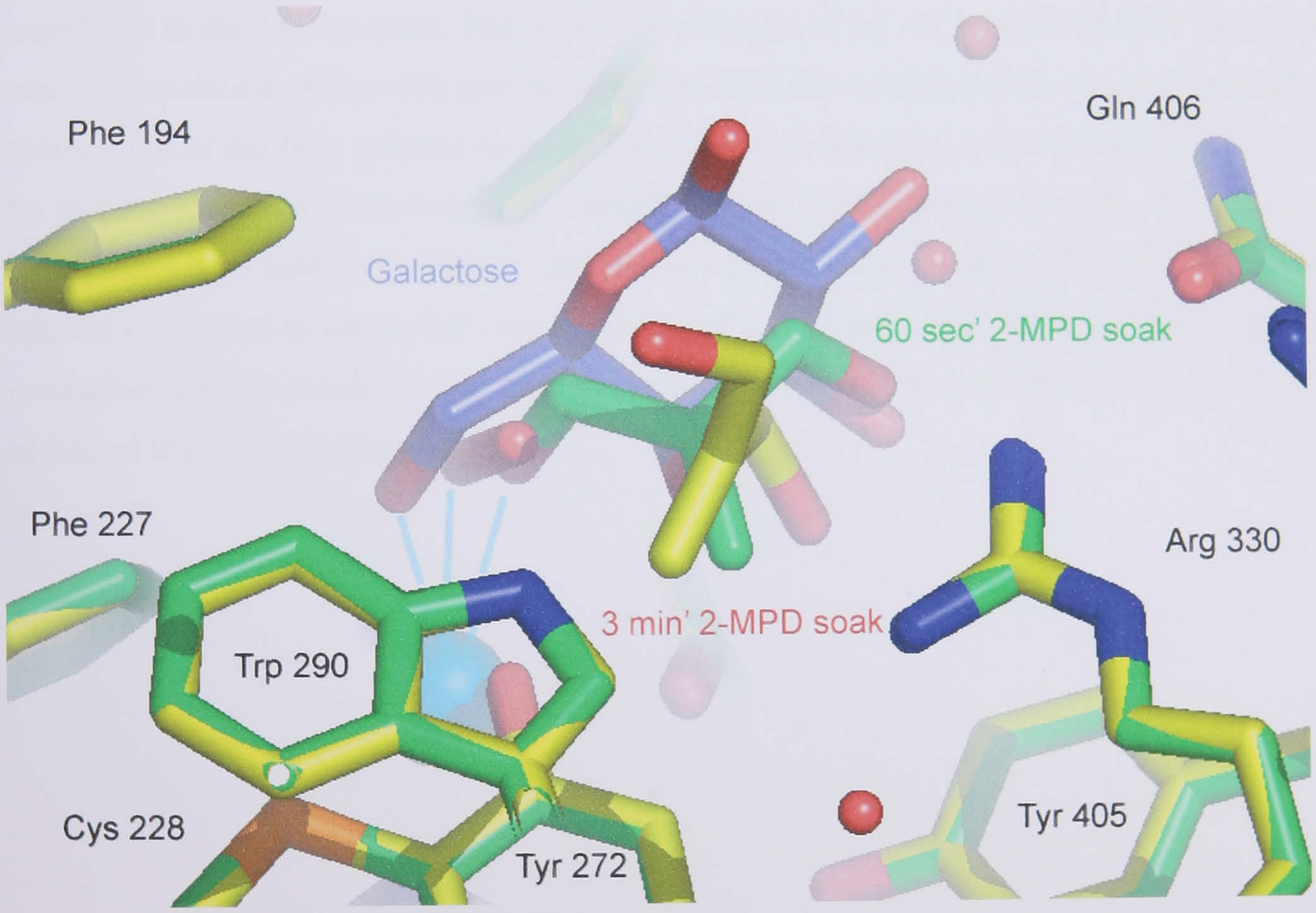


Figure 6.11: Galactose oxidase substrate complexes.

Overlay of the manually docked galactose (carbons coloured purple) with the 60 second 2-methylene-1,3-propandiol (2MPD) soak structure (carbons coloured green) and the 3 minute 2-MPD soak structure (carbons coloured yellow). The methylene group and the non-copper coordinated hydroxyl of 2MPD from the 60 second soak (green), closely matches the O4 and O5 positions of galactose. The methylene of the non-copper binding 2MPD (yellow) is in a similar position to the O5 of galactose.

The active site of the reduced enzyme is proposed to contain a three coordinate Cu^{1+} ion and no radical. The crystal structure of fully reduced galactose oxidase, however, has previously proved to be elusive. In the work presented in this thesis, Cu^{1+} was introduced into apo-galactose oxidase crystals under anaerobic conditions, revealing the first active site structure of the Cu^{1+} containing enzyme (Figure 6.12). As predicted, Tyr 272, His 581 and His 496 appear to be copper ligands, although the Tyr 272 bond distance is longer than the 2.0 Å predicted by EXAFS [Knowles *et al.*, 1995]. The optical spectrum of reduced galactose oxidase has been reported to lack the strong absorbance at 445 and 800 nm associated with the oxidised protein [Whittaker and Whittaker, 1993]. These absorbance peaks are partially attributed to Tyr 272 and Tyr 495 interaction with copper. Dissociation of Tyr 495 from copper is reported to abolish the 800 nm peak, while a Tyr 272 to copper charge transfer contributes to the 445 nm peak. The active site structure of the Cu^{1+} soaked enzyme revealed copper dissociation of Tyr 495 and possibly Tyr 272, this corresponds to the spectroscopic observations of the fully reduced enzyme. The long bond distances, however, could be due to the partial occupation of the copper site, which was not fully occupied even after 25 minutes incubation with 5 mM Cu^{1+}. This in contrast to Cu^{2+} soak experiments, in which only 5-10 minutes is required to obtain full copper loading of apo-crystals. It is not certain why copper incubation did not result in full copper loading, however, this problem will have to be addressed in future experiments.

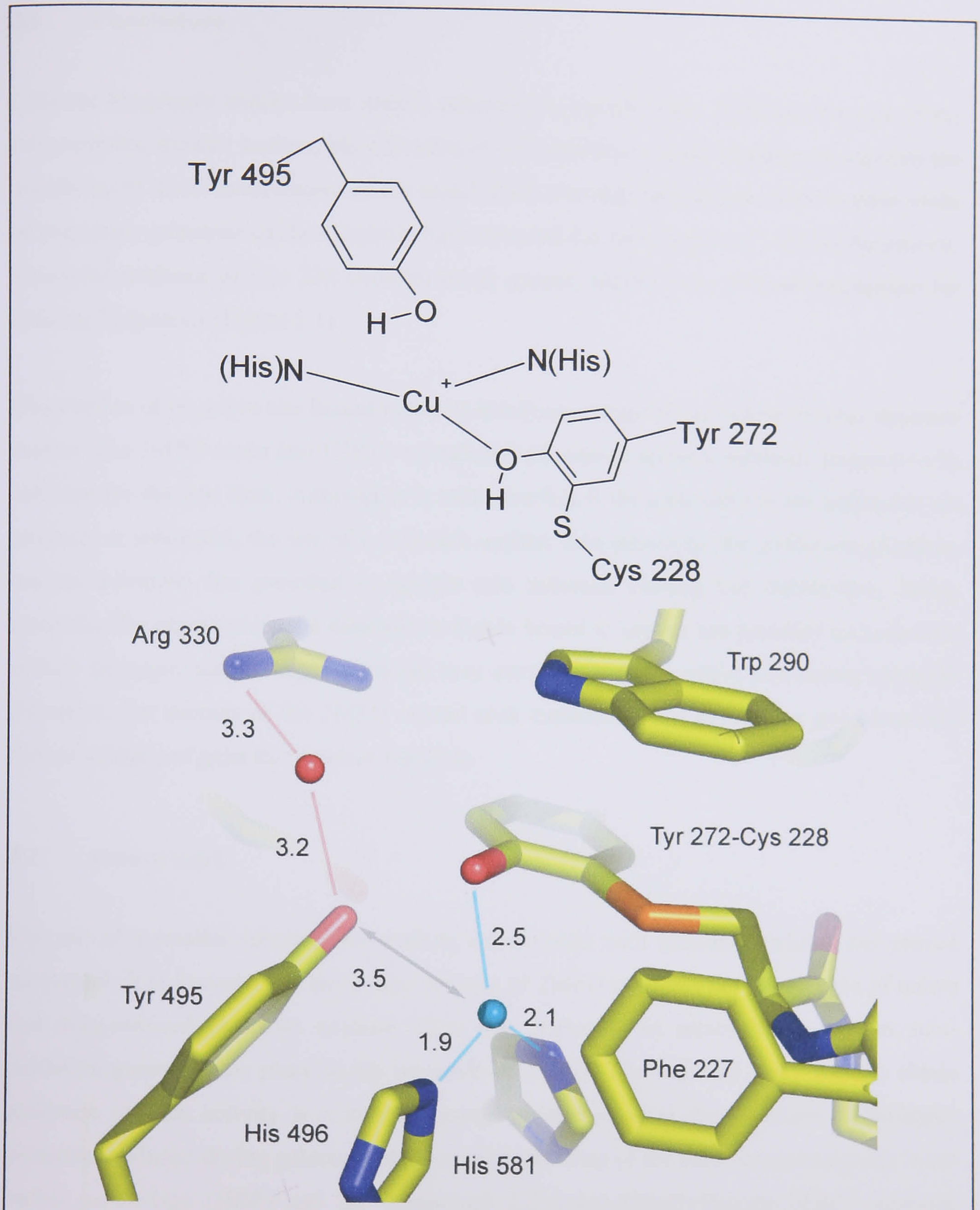


Figure 6.12: Cu^{1+} soak structure compared to the predicted fully reduced enzyme geometry.

Incubation of apo-galactose oxidase with Cu^{1+} revealed an active site structure reminiscent of the fully reduced active site, proposed to exist during the catalytic cycle (Figure 1.17) [Whittaker, 2003].

6.6 Conclusions

Cofactor biogenesis studies have greatly advanced in recent years. Although the role of the pro-sequence remains unclear, identification of an N-terminal copper binding site suggests the possibility of direct involvement of the metal in the cleavage mechanism. Short copper soaks of premature galactose oxidase crystals have revealed the early stages of cofactor biogenesis. Structural evidence of Cys 228 coordinated to copper, supports the Firbank mechanism for cofactor biogenesis (Figure 5.1).

Observation of an active site bound substrate has been a target for galactose oxidase structure studies. The 2MPD soaks into C383S crystals has allowed an enzyme-substrate complex to be obtained for the first time. Although it is still uncertain if the molecules in the active site are products or substrates, the use of a substrate smaller than galactose, the traditional galactose oxidase substrate, has provided an insight into substrate binding and stabilisation during catalysis. The observation of a diatomic molecule bound to copper has provided an indication of how dioxygen and hydrogen peroxide may associate with the active site during catalysis. Therefore, the success of the 2MPD crystal soak experiment has provided a good basis for further analysis of galactose oxidase catalysis.

6.7 Future work

The use of a smaller substrate for soaking experiments with galactose oxidase has proved successful. It is important to determine the rate of 2MPD catalysis in the presence of native and semi-reduced galactose oxidase. It is also important to establish whether an auto-oxidation process takes place in the presence of 2MPD. The standard assay used to obtain galactose oxidase activity is a coupled assay which measures the presence of hydrogen peroxide produced during galactose oxidase catalysis. Two of the assay components are horse radish peroxidase (HRP) and the compound 2,2'-azino-bis(ethylbenzthizoline-6-suphonic acid) (ABTS). ABTS turns green in the presence of hydrogen peroxide, while HRP is present to facilitate this reaction [Baron *et al.*, 1993]. Measurement of the change of colour and thus the reaction rate can be achieved spectrophotometrically. The standard activity assays could not be used for 2MPD, however, due to its interaction with components of the assay in the absence of galactose oxidase. Using an oxygen electrode, however, will allow the measurement of the rate of oxygen disappearance during 2MPD catalysis, thus allowing the direct measurement of the enzyme re-oxidation rate during catalysis [Rogers *et al.*, 2007]. Solution studies could be combined with UV-visible absorbance spectra to follow the formation and disappearance of the radical during catalysis. The same technique can be

applied to crystals, thus ensuring the solution experiments are in keeping with the structural observations. For example, the anaerobic incubation of crystals with substrate would be expected to produce the fully reduced enzyme structure, possibly with substrate or product present in the active site. The presence of Cu^{1+} containing enzyme could be verified by single crystal spectrophotometry.

Introduction of Cu^{1+} into the active site of galactose oxidase was only partially successful. There are several different ways to attain the fully reduced enzyme. One method is to repeat the Cu^{1+} incubations with the apo-enzyme. 5 mM copper soaks for up to 25 minutes were not successful, so perhaps a larger copper concentration is necessary. A second method, mentioned earlier, is by substrate reduction of enzyme under anaerobic conditions. A third method is to chemically reduce the native enzyme using ferrocyanide [Whittaker and Whittaker, 1988], or ruthenium [Wright and Sykes, 2001b]. Unfortunately, attempts so far to obtain fully reduced enzyme via chemical or substrate reduction of crystallised galactose oxidase have proved unsuccessful. Once the fully reduced enzyme is attained, the oxygen binding reaction can be probed using an oxygen mimic, such as nitrous oxide.

More research on the pro-sequence cleavage process is needed. It is still uncertain how the reaction occurs and what parameters affect the rate. The crystal structure of pro-GO was crystallised by Firbank, but was never observed to undergo full processing. Further investigation of this form of the enzyme is required in order to understand the mechanism of the auto-catalytic cleavage event. As mentioned earlier, UV-visible spectra of crystal structures can be monitored, thus ensuring the biogenesis and transient intermediates during processing are fully characterised. Confirmation of the nature of the 410 nm absorbing species observed during Cu^{2+}-dependent biogenesis could be obtained by measuring the absorbance spectra of premature galactose oxidase crystals incubated with copper under anaerobic and aerobic conditions.

References

Abergel C, Bouveret E, Claverie JM, Brown K, Rigal A, Lazdunski C, Benedetti H (1999): Structure of the Escherichia coli TolB protein determined by MAD methods at 1.95 A resolution. *Structure Fold Des* **7**:1291-300.

Aboelella NW, Lewis EA, Reynolds AM, Brennessel WW, Cramer CJ, Tolman WB (2002): Snapshots of dioxygen activation by copper: the structure of a 1:1 Cu/O(2) adduct and its use in syntheses of asymmetric Bis(mu-oxo) complexes. *J Am Chem Soc* **124**:10660-1.

Ahmad KF, Engel CK, Prive GG (1998): Crystal structure of the BTB domain from PLZF. *Proc Natl Acad Sci U S A* **95**:12123-8.

Avigad G (1985): Oxidation rates of some desialylated glycoproteins by galactose oxidase. *Arch Biochem Biophys* **239**:531-7.

Avigad G, Amaral D, Asensio C, Horecker BL (1962): The D-galactose oxidase of Polyporus circinatus. *J Biol Chem* **237**:2736-43.

Baardsnes J, Sidhu S, MacLeod A, Elliott J, Morden D, Watson J, Borgford T (1998): Streptomyces griseus protease B: secretion correlates with the length of the propeptide. *J Bacteriol* **180**:3241-4.

Babcock GT, El-Deeb MK, Sandusky PO, Whittaker MM, Whittaker JW (1992): Electron-paramagnetic resonance and electron nuclear double-resonance spectroscopies of the radical site in galactose oxidase and of thioether-substituted phenol model compounds. *J Am Chem Soc* **114**:3727-3734.

Babcock GT, Espe M, Hoganson C, Lydakis-Simantiris N, McCracken J, Shi W, Styring S, Tommos C, Warncke K (1997): Tyrosyl radicals in enzyme catalysis: some properties and a focus on photosynthetic water oxidation. *Acta Chem Scand* **51**:533-40.

Banci L, Bertini I, Ciofi-Baffoni S, Su XC, Borrelly GP, Robinson NJ (2004): Solution structures of a cyanobacterial metallochaperone: insight into an atypical copper-binding motif. *J Biol Chem* **279**:27502-10.

Baron AJ, Stevens C, Wilmot C, Seneviratne KD, Blakeley V, Dooley DM, Phillips SEV, Knowles PF, McPherson MJ (1994): Structure and mechanism of galactose oxidase. The free radical site. *J Biol Chem* **269**:25095-105.

Baron AJ, Stevens C, Wilmot CM, Knowles PF, Phillips SEV, McPherson MJ (1993): Preliminary studies of two active site mutants of galactose oxidase. *Biochem Soc Trans* **21 (Pt 3)**:319S.

Beers, M. H. and R. Berkow (1999). "The Merck manual of Diagnosis and Therapy." *The Merck Manual* **17**: Section 1, Chapter 4.

Beisel HG, Kawabata S, Iwanaga S, Huber R, Bode W (1999): Tachylectin-2: crystal structure of a specific GlcNAc/GalNAc-binding lectin involved in the innate immunity host defense of the Japanese horseshoe crab Tachypleus tridentatus. *Embo J* **18**:2313-22.

Betts M.J., Russell R.B (2003) Amino Acid properties and consequences of substitutions. *Bioinformations forGenetisists*, Barnes M.R., Grays I.C., eds, *Wiley*.

Blaschko H (1958): Amine oxidase in mammalian blood. *Farmaco.***13**:521-33.

Borman C.D., Saysell C.G., and Sykes A.G (1997). Kenetic Studies on the reactios of *Fusarium* galactose oxidase with five different substrate in the presence of dioxygen . *J. Biol. Inorg. Chem.* **2**, 480-487.

Bragg WH, Brag WL (1913): The structure of a diamond. *Nature*:557.

Branchaud BP, Turner BE (2002): Galactose oxidase: probing radical mechanism with ultrafast radical probe. *Methods Enzymol* **354**:415-25.

Brandstetter H, Kim JS, Groll M, Huber R (2001): Crystal structure of the tricorn protease reveals a protein disassembly line. *Nature* **414**:466-70.

Brown K, Djinovic-Carugo K, Haltia T, Cabrito I, Saraste M, Moura JJ, Moura I, Tegoni M, Cambillau C (2000a): Revisiting the catalytic CuZ cluster of nitrous oxide (N2O) reductase. Evidence of a bridging inorganic sulfur. *J Biol Chem* **275**:41133-6.

Brown K, Tegoni M, Prudencio M, Pereira AS, Besson S, Moura JJ, Moura I, Cambillau C (2000b): A novel type of catalytic copper cluster in nitrous oxide reductase. *Nat Struct Biol* **7**:191-5.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998): Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 (Pt 5)**:905-21.

Brunger AT, Karplus M, Petsko GA (1989): Crystallographic refinement by simmulated annealing - application to crambin. *Acta Crystallogr* **A**:50-61.

Brush EJ, Lipsett KA, Kozarich JW (1988): Inactivation of Escherichia coli pyruvate formate-lyase by hypophosphite: evidence for a rate-limiting phosphorus-hydrogen bond cleavage. *Biochemistry* **27**:2217-22.

Cai D, Klinman JP (1994): Evidence of a self-catalytic mechanism of 2,4,5-trihydroxyphenylalanine quinone biogenesis in yeast copper amine oxidase. *J Biol Chem* **269**:32039-42.

Cai D, Williams NK, Klinman JP (1997): Effect of metal on 2,4,5-trihydroxyphenylalanine (topa) quinone biogenesis in the Hansenula polymorpha copper amine oxidase. *J Biol Chem* **272**:19277-81.

Carr S, Penfold CN, Bamford V, James R, Hemmings AM (2000): The structure of TolB, an essential component of the tol-dependent translocation system, and its protein-protein interaction with the translocation domain of colicin E9. *Structure Fold Des* **8**:57-66.

Carter P, Wells JA (1988): Dissecting the catalytic triad of a serine protease. *Nature* **332**:564-8.

CCP4 (1994): Collaborative Computational Project, Number 4. *Acta Crystallogr.* **D**:760-763.

Chang MC, Yee CS, Stubbe J, Nocera DG (2004): Turning on ribonucleotide reductase by light-initiated amino acid radical generation. *Proc Natl Acad Sci U S A* **101**:6882-7.

Chang TK, Iverson SA, Rodrigues CG, Kiser CN, Lew AY, Germanas JP, Richards JH (1991): Gene synthesis, expression, and mutagenesis of the blue copper proteins azurin and plastocyanin. *Proc Natl Acad Sci U S A* **88**:1325-9.

Chirpich TP, Zappia V, Costilow RN, Barker HA (1970): Lysine 2,3-aminomutase. Purification and properties of a pyridoxal phosphate and S-adenosylmethionine-activated enzyme. *J Biol Chem* **245**:1778-89.

Christopher JA (1998): SPOCK: The Structural Properties Observation and Calculation Kit (program manual).

Clark K, Penner-Hahn JE, Whittaker M, Whittaker JW (1994): Structural characterization of the copper site in galactose oxidase using X-ray absorption spectroscopy. *Biochemistry* **33**:12553-7.

Cooper JA, Smith W, Bacila M, Medina H (1959): Galactose oxidase from Polyporus circinatus, Fr. *J Biol Chem* **234**:445-8.

Crennell SJ, Garman EF, Laver WG, Vimr ER, Taylor GL (1993): Crystal structure of a bacterial sialidase (from Salmonella typhimurium LT2) shows the same fold as an influenza virus neuraminidase. *Proc Natl Acad Sci U S A* **90**:9852-6.

Crowther RA (1972): in "The Molecular Replacement Method" (M.G. Rossmann, ed) New York: Gordonand Breach, pp 173-178.

Crowther RA, Blow DM (1967): A method of positioning a known molecule in an unknown crystal structure. *Acta Crystallogr.*:544-548.

Darwin C (1859): "On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life." London:

Datta S, Mori Y, Takagi K, Kawaguchi K, Chen ZW, Okajima T, Kuroda S, Ikeda T, Kano K, Tanizawa K, Mathews FS (2001): Structure of a quinohemoprotein amine dehydrogenase with an uncommon redox cofactor and highly unusual crosslinking. *Proc Natl Acad Sci U S A* **98**:14268-73.

Deacon SE, Mahmoud K, Spooner RK, Firbank SJ, Knowles PF, Phillips SEV, McPherson MJ (2004): Enhanced fructose oxidase activity in a galactose oxidase variant. *Chembiochem* **5**:972-9.

Delagrave S, Murphy DJ, Pruss JL, Maffia AM, 3rd, Marrs BL, Bylina EJ, Coleman WJ, Grek CL, Dilworth MR, Yang MM, Youvan DC (2001): Application of a very high-throughput digital imaging screen to evolve the enzyme galactose oxidase. *Protein Eng* **14**:261-7.

DeLano WL (2004): The PyMOL Molecular Graphics System San Carlos, CA, USA: DeLano Scientific LLC.

Dove JE, Schwartz B, Williams NK, Klinman JP (2000): Investigation of spectroscopic intermediates during copper-binding and TPQ formation in wild-type and active-site mutants of a copper-containing amine oxidase from yeast. *Biochemistry* **39**:3690-8.

Eder J, Fersht AR (1995): Pro-sequence-assisted protein folding. *Mol Microbiol* **16**:609-14.

Engstrom M, Himo F, Agren H (2000): *Ab initio* g-tensor calculations of the thioether bond substituted tyrosyl radical in galactose oxidase. *Chem. Phys. Lett.* **319**:191-196.

Faber HR, Groom CR, Baker HM, Morgan WT, Smith A, Baker EN (1995): 1.8 A crystal structure of the C-terminal domain of rabbit serum haemopexin. *Structure* **3**:551-9.

Facchin G, Torre MH, Kremer E, Piro OE, Castellano EE, Baran EJ (2002): Synthesis and characterization of three new Cu(II)-dipeptide complexes. *J Inorg Biochem* **89**:174-80.

Firbank SJ (2002): Structural Investigation of Processing and Catalysis in Galactose Oxidase. *PhD Thesis, University of Leeds*:1-216.

Firbank SJ, Rogers MS, Wilmot CM, Dooley DM, Halcrow MA, Knowles PF, McPherson MJ, Phillips SEV (2001): Crystal structure of the precursor of galactose oxidase: an unusual self-processing enzyme. *Proc Natl Acad Sci U S A* **98**:12932-7.

Frey PA (2001): Radical mechanisms of enzymatic catalysis. *Annu Rev Biochem* **70**:121-48.

Fulop V, Moir JW, Ferguson SJ, Hajdu J (1995): The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd1. *Cell* **81**:369-77.

Fulop V, Bocskei Z, Polgar L (1998): Prolyl Oligopeptidase: An unusual ß-Propeller Domain Regulates Proteolysis.*Cell*, **94**: 161-170.

Garman EF, Schneider TR (1997): Macromolecular Cryocrystallography. *J. Appl Cryst* **30**:211-237.

Gerfen GA, Bellew B, Griffin R, Singel D, Ekberg CA, Whittaker JW (1996): High-frequency electron paramagnetic resonance spectroscopy of the apogalactose oxidase radical. *J. Phys. Chem.* **100**:16739-16748.

Ghosh M, Anthony C, Harlos K, Goodwin MG, Blake C (1995): The refined structure of the quinoprotein methanol dehydrogenase from Methylobacterium extorquens at 1.94 A. *Structure* **3**:177-87.

Gohlke U, Gomis-Ruth FX, Crabbe T, Murphy G, Docherty AJ, Bode W (1996): The C-terminal (haemopexin-like) domain structure of human gelatinase A (MMP2): structural implications for its function. *FEBS Lett* **378**:126-30.

Ha NC, Oh BC, Shin S, Kim HJ, Oh TK, Kim YO, Choi KY, Oh BH (2000): Crystal structures of a novel, thermostable phytase in partially and fully calcium-loaded states. *Nat Struct Biol* **7**:147-53.

Hamilton GA, Adolf PK, de Jersey J, DuBois GC, Dyrkacz GR, Libby RD (1978): Trivalent copper, superoxide, and galactose oxidase. *J Am Chem Soc* **100**:1899-1912.

Hamilton GA, Dyrkacz GR, Libby RD (1976): The involvement of superoxide and trivalent copper in the galactose oxidase reaction. *Adv Exp Med Biol* **74**:489-504.

Hamilton GA, Libby RD, Hartzell CR (1973): The valence of copper and the role of superoxide in the D-galactose oxidase catalyzed reaction. *Biochem Biophys Res Commun* **55**:333-40.

Harriman A (1987): Further comments on the redox potentials of tryptophan and tyrosine. *J. Phys. Chem.* **91**:6102-6104.

Hart PJ, Nersissian AM, Herrmann RG, Nalbandyan RM, Valentine JS, Eisenberg D (1996): A missing link in cupredoxins: crystal structure of cucumber stellacyanin at 1.6 A resolution. *Protein Sci* **5**:2175-83.

Ito N (1991): X-ray Crystallographic studies on Galactose Oxidase. *PhD Thesis, University of Leeds*:163.

Ito N, Keen JN, Knowles PF, McPherson MJ, Phillips SEV, Stevens C, Yadav KD (1990): Structural analysis of galactose oxidase. *Biochem Soc Trans* **18**:931-2.

Ito N, Knowles PF, Phillips SEV (1995): X-ray crystallographic studies of cofactors in galactose oxidase. *Methods Enzymol* **258**:235-62.

Ito N, Phillips SEV, Stevens C, Ogel ZB, McPherson MJ, Keen JN, Yadav KD, Knowles PF (1991): Novel thioether bond revealed by a 1.7 A crystal structure of galactose oxidase. *Nature* **350**:87-90.

Ito N, Phillips SEV, Stevens C, Ogel ZB, McPherson MJ, Keen JN, Yadav KD, Knowles PF (1992): Three-dimensional structure of galactose oxidase: an enzyme with a built-in secondary cofactor. *Faraday Discuss*:75-84.

Ito N, Phillips SEV, Yadav KD, Knowles PF (1994): Crystal structure of a free radical enzyme, galactose oxidase. *J Mol Biol* **238**:794-814.

Jack A, Levitt M (1978): Refinement of large structures by simultaneous minimisation of energy and R factor. *Acta Crystallogr* **A**:931-935.

Janes SM, Mu D, Wemmer D, Smith AJ, Kaur S, Maltby D, Burlingame AL, Klinman JP (1990): A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site of bovine serum amine oxidase. *Science* **248**:981-7.

Jawad Z, Paoli M (2002): Novel sequences propel familiar folds. *Structure (Camb)* **10**:447-54.

Jeon H, Meng W, Takagi J, Eck MJ, Springer TA, Blacklow SC (2001): Implications for familial hypercholesterolemia from the structure of the LDL receptor YWTD-EGF domain pair. *Nat Struct Biol* **8**:499-504.

Jian F, Zhao P, Xiao H, Wang H (2004): Crystal structure of tetranuclear copper (II) cluster compound [$Cu_4OCl_6(BBIm)_4$] (BBIm = 3-Benzyl-Benzimidazole). *The Japan Society for Analytical Chemistry* **20**:23-24.

Jiang W, Bond JS (1992): Families of metalloendopeptidases and their relationships. *FEBS Lett* **312**:110-4.

Jones TA, Zou JY, Cowan SW, Kjeldgaard (1991): Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**:110-9.

Keilin D, Hatree EF (1948): Properties of glucose oxidase (notatin): Addenudum. Sedimentation and diffusion of glucose oxidase (notatin). *Biochem J.* **42 (2)**: 221-9.

Keitel T, Diehl A, Knaute T, Stezowski JJ, Hohne W, Gorisch H (2000): X-ray structure of the quinoprotein ethanol dehydrogenase from Pseudomonas aeruginosa: basis of substrate specificity. *J Mol Biol* **297**:961-74.

Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC (1958.): A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*:662-666.

Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR (1960): Structure of Myoglobin. A Three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature*:422-427.

Kersten PJ, Cullen D (1993): Cloning and characterization of cDNA encoding glyoxal oxidase, a H2O2-producing enzyme from the lignin-degrading basidiomycete Phanerochaete chrysosporium. *Proc Natl Acad Sci U S A* **90**:7411-3.

Kim M, Okajima T, Kishishita S, Yoshimura M, Kawamori A, Tanizawa K, Yamaguchi H (2002): X-ray snapshots of quinone cofactor biogenesis in bacterial copper amine oxidase. *Nat Struct Biol* **9**:591-6.

Klabunde T, Eicken C, Sacchettini JC, Krebs B (1998): Crystal structure of a plant catechol oxidase containing a dicopper center. *Nat Struct Biol* **5**:1084-90.

Kleywegt GJ, Brunger AT (1996): Checking your imagination: Applications of free R value. *Structure* **4**:897-904.

Kleywegt GJ, Henrick K, Dodson EJ, Van Aalten DMF (2003): Pound-wise but penny-foolish - How well do micromolecules fare in macromolecular refinement? *Structure* **11**:1051-1059.

Kleywegt GJ, Jones TA (1997): Detecting folding motifs and similarities in protein structures. *Methods Enzymol* **277**:525-545.

Kleywegt GJaJ, T.a. (1998): Databases in crystallography. *Acta Crystallogr* **D**:1119-1131.

Knappe J, Sawers G (1990): A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of Escherichia coli. *FEMS Microbiol Rev* **6**:383-98.

Knowles PF, Brown RDI, Koenig SH, Wang S, Scott RA, McGuirl MA, Brown DE, Dooley DM (1995): Spectroscopic studies of the active-site of galactose oxidase. *Inorg Chem* **34**:3895-3902.

Koay M, Zhang L, Yang B, Maher MJ, Xiao Z, Wedd AG (2005): CopC protein from Pseudomonas syringae: intermolecular transfer of copper from both the copper(I) and copper(II) sites. *Inorg Chem* **44**:5203-5.

Kosman DJ (1984): "Copper proteins and copper enzymes." Boca Raton: Boca Raton Publishers.

Kosman DJ, Ettinger MJ, Weiner RE, Massaro EJ (1974): The molecular properties of the copper enzyme galactose oxidase. *Arch Biochem Biophys* **165**:456-67.

Kosman DJ, Peisach J, Mims WB (1980): Pulsed electron paramagnetic resonance studies of the copper(II) site in galactose oxidase. *Biochemistry* **19**:1304-8.

Kumar V, Dooley DM, Freeman HC, Guss JM, Harvey I, McGuirl MA, Wilce MC, Zubak VM (1996): Crystal structure of a eukaryotic (pea seedling) copper-containing amine oxidase at 2.2 A resolution. *Structure* **4**:943-55.

Laskowski RA, Moss DS, Thornton JM (1993): Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**:1049-67.

Lerch K (1982): Primary structure of tyrosinase from Neurospora crassa. II. Complete amino acid sequence and chemical structure of a tripeptide containing an unusual thioether. *J Biol Chem* **257**:6414-9.

Leslie AGW (1992): Recent changes to MOSFLM packages for processing film and image plate data.: "*Joint CCP4 and ESRF EACBM Newsletter on protein Crystallography.*"

Leuthner B, Leutwein C, Schulz H, Horth P, Haehnel W, Schiltz E, Schagger H, Heider J (1998): Biochemical and genetic characterization of benzylsuccinate synthase from Thauera aromatica: a new glycyl radical enzyme catalysing the first step in anaerobic toluene metabolism. *Mol Microbiol* **28**:615-28.

Li J, Brick P, O'Hare MC, Skarzynski T, Lloyd LF, Curry VA, Clark IM, Bigg HF, Hazleman BL, Cawston TE, et al. (1995): Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed beta-propeller. *Structure* **3**:541-9.

Li X, Zhang D, Hannink M, Beamer LJ (2004): Crystal structure of the Kelch domain of human Keap1. *J Biol Chem* **279**:54750-8.

Lim D, Park HU, De Castro L, Kang SG, Lee HS, Jensen S, Lee KJ, Strynadka NC (2001): Crystal structure and kinetic analysis of beta-lactamase inhibitor protein-II in complex with TEM-1 beta-lactamase. *Nat Struct Biol* **8**:848-52.

Liu X, Kilner CA, Halcrow MA (2002): 3,4-(4-Methoxybenzo):8,9-benzobicyclo[4.4.1]undeca-3,8-dien-11-one ethylene acetal. *Acta Crystallogr C* **58**:O218-9.

Lonsdale K (1929): The structure of the Benzene ring in Hexamethylbenzene. *Proceedings of the Royal Society* **124A**:494.

Lonsdale K, Astbury WT (1924): Tabulated Data for the Examination of the 230 Space groups by Homogeneous X-Rays. *Philosophical Transactions of the Royal Society* **224A**:221.

Lu Y, Roe JA, Bender CJ, Peisach J, Banci L, Bertini I, Gralla EB, Valentine JS (1996): New Type 2 Copper-Cysteinate Proteins. Copper Site Histidine-to-Cysteine Mutants of Yeast Copper-Zinc Superoxide Dismutase. *Inorg Chem* **35**:1692-1700.

Mahmoud K (2001): Protein engineering and heterologous expression studies of galactose oxidase. *PhD Thesis, University of Leeds*.

McCall KA, Huang C, Fierke CA (2000): Function and mechanism of zinc metalloenzymes. *J Nutr* **130**:1437S-46S.

McGleshen ML, Eads DD, Spiro TG, Whittaker JW (1995): Resonance Raman spectroscopy of galactose-oxidase - a new interpretation based on model-compound free radical spectra. *J. Phys. Chem.* **99**:4918-4922.

McIntire WS, Wemmer DE, Chistoserdov A, Lidstrom ME (1991): A new cofactor in a prokaryotic enzyme: tryptophan tryptophylquinone as the redox prosthetic group in methylamine dehydrogenase. *Science* **252**:817-24.

McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, Sekhon M, Wylie K, Mardis ER, Wilson RK, Fulton R, Kucaba TA, Wagner-McPherson C, Barbazuk WB, Gregory SG, Humphray SJ, French L, Evans RS, Bethel G, Whittaker A, Holden JL, McCann OT, Dunham A, Soderlund C, Scott CE, Bentley DR, Schuler G, Chen HC, Jang W, Green ED, Idol JR, Maduro VV, Montgomery KT, Lee E, Miller A, Emerling S, Kucherlapati, Gibbs R, Scherer S, Gorrell JH, Sodergren E, Clerc-Blankenburg K, Tabor P, Naylor S, Garcia D, de Jong PJ, Catanese JJ, Nowak N, Osoegawa K, Qin S, Rowen L, Madan A, Dors M, Hood L, Trask B, Friedman C, Massa H, Cheung VG, Kirsch IR, Reid T, Yonescu R, Weissenbach J, Bruls T, Heilig R, Branscomb E, Olsen A, Doggett N, Cheng JF, Hawkins T, Myers RM, Shang J, Ramirez L, Schmutz J, Velasquez O, Dixon K, Stone NE, Cox DR, Haussler D, Kent WJ, Furey T, Rogic S, Kennedy S, Jones S, Rosenthal A, Wen G, Schilhabel M, Gloeckner G, Nyakatura G, Siebert R, Schlegelberger B, Korenberg J, Chen XN, Fujiyama A, Hattori M, Toyoda A, Yada T, Park HS, Sakaki Y, Shimizu N, Asakawa S, Kawasaki K, Sasaki T, Shintani A, Shimizu A, Shibuya K, Kudoh J, Minoshima S, Ramser J, Seranski P, Hoff C, Poustka A, Reinhardt R, Lehrach H (2001): A physical map of the human genome. *Nature* **409**:934-41.

McPherson MJ, Ogel ZB, Stevens C, Yadav KD, Keen JN, Knowles PF (1992): Galactose oxidase of Dactylium dendroides. Gene cloning and sequence analysis. *J Biol Chem* **267**:8146-52.

McPherson MJ, Stevens C, Baron AJ, Ogel ZB, Seneviratne K, Wilmot C, Ito N, Brocklebank I, Phillips SEV, Knowles PF (1993): Galactose oxidase: molecular analysis and mutagenesis studies. *Biochem Soc Trans* **21 (Pt 3)**:752-6.

Mendonca MH, Zancan GT (1987): Purification and characterization of intracellular galactose oxidase from Dactylium dendroides. *Arch Biochem Biophys* **252**:507-14.

Mendonca MH, Zancan GT (1988): Role of carbohydrate content on the properties of galactose oxidase from Dactylium dendroides. *Arch Biochem Biophys* **266**:427-34.

Messerschmidt A, Luecke H, Huber R (1993): X-ray structures and mechanistic implications of three functional derivatives of ascorbate oxidase from zucchini. Reduced, peroxide and azide forms. *J Mol Biol* **230**:997-1014.

Minasian SG, Whittaker MM, Whittaker JW (2004): Stereoselective hydrogen abstraction by galactose oxidase. *Biochemistry* **43**:13683-93.

Muirhead H, Perutz MF (1963): Structure of haemoglobin. A three-dimensional Fourier synthesis of reduced human haemoglobin at 5.5 Angstrom resolution. *Nature*:633-636.

Multhaup G, Ruppert T, Schlicksupp A, Hesse L, Bill E, Pipkorn R, Masters CL, Beyreuther K (1998): Copper-binding amyloid precursor protein undergoes a site-specific fragmentation in the reduction of hydrogen peroxide. *Biochemistry* **37**:7224-30.

Murzin AG (1992): Structural principles for the propeller assembly of beta-sheets: the preference for seven-fold symmetry. *Proteins* **14**:191-201.

Navaza J (1993): On the computation of the fast rotation function. *Acta Crystallogr D Biol Crystallogr* **49**:588-91.

Navaza J (1994): AMoRe - an automated package for molecular replacement. *Acta Crystallogr* **A**:157.

Navaza J (2001): Implementation of molecular replacement in AMoRe. *Acta Crystallogr D Biol Crystallogr* **57**:1367-72.

Neer EJ, Smith TF (1996): G protein heterodimers: new structures propel new questions. *Cell* **84**:175-8.

Nersissian AM, Immoos C, Hill MG, Hart PJ, Williams G, Herrmann RG, Valentine JS (1998): Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: plant-specific mononuclear blue copper proteins. *Protein Sci* 7:1915-29.

Nishihira J, Tachikawa H (1999): Theoretical evaluation of a model of the catalytic triads of serine and cysteine proteases by ab initio molecular orbital calculation. *J Theor Biol* **196**:513-9.

Ogel ZB (1993): Molecular analysis of a fungal galactose oxidase gene. *PhD Thesis, University of Leeds*.

Okeley NM, van der Donk WA (2000): Novel cofactors via post-translational modifications of enzyme active sites. *Chem Biol* 7:R159-71.

Otwinowski Z, Minor W (1997): Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**:307-325.

Oubrie A, Rozeboom HJ, Kalk KH, Duine JA, Dijkstra BW (1999): The 1.7 A crystal structure of the apo form of the soluble quinoprotein glucose dehydrogenase from Acinetobacter calcoaceticus reveals a novel internal conserved sequence repeat. *J Mol Biol* **289**:319-33.

Oubrie A, Rozeboom HJ, Kalk KH, Huizinga EG, Dijkstra BW (2002): Crystal structure of quinohemoprotein alcohol dehydrogenase from Comamonas testosteroni: structural basis for substrate oxidation and electron transfer. *J Biol Chem* **277**:3727-32.

Pannu NS, Read RJ (1996): Improved structure refinement through maximum likelihood. *acta Crystallogr.* **A**:659-668.

Paoli M (2001): Protein folds propelled by diversity. *Prog Biophys Mol Biol* **76**:103-30.

Park HI, Ming LJ (2002): Mechanistic studies of the astacin-like Serratia metalloendopeptidase serralysin: highly active (>2000%) Co(II) and Cu(II) derivatives for further corroboration of a "metallotriad" mechanism. *J Biol Inorg Chem* **7**:600-10.

Parsons MR, Convery MA, Wilmot CM, Yadav KD, Blakeley V, Corner AS, Phillips SEV, McPherson MJ, Knowles PF (1995): Crystal structure of a quinoenzyme: copper amine oxidase of Escherichia coli at 2 A resolution. *Structure* **3**:1171-84.

Pelmenschikov V, Blomberg MR, Siegbahn PE (2002): A theoretical study of the mechanism for peptide hydrolysis by thermolysin. *J Biol Inorg Chem* **7**:284-98.

Ponting CP, Pallen MJ (1999): A beta-propeller domain within TolB. *Mol Microbiol* **31**:739-40.

Prag S, Adams JC (2003): Molecular phylogeny of the kelch-repeat superfamily reveals an expansion of BTB/kelch proteins in animals. *BMC Bioinformatics* **4**:42.

Prasad BV, Suguna K (2002): Role of water molecules in the structure and function of aspartic proteinases. *Acta Crystallogr D Biol Crystallogr* **58**:250-9.

Prigge ST, Eipper BA, Mains RE, Amzel LM (2004): Dioxygen binds end-on to mononuclear copper in a precatalytic enzyme complex. *Science* **304**:864-7.

Read RJ (1986): Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta. Crystallogr.* **A**:140-149.

Reichard P (1993): The anaerobic ribonucleotide reductase from Escherichia coli. *J Biol Chem* **268**:8383-6.

Reichard P, Ehrenberg A (1983): Ribonucleotide reductase--a radical enzyme. *Science* **221**:514-9.

Renault L, Nassar N, Vetter I, Becker J, Klebe C, Roth M, Wittinghofer A (1998): The 1.7 A crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature* **392**:97-101.

Reynolds MP, Baron AJ, Wilmot CM, Phillips SEV, Knowles PF, McPherson MJ (1995): Tyrosine 495 is a key residue in the active site of galactose oxidase. *Biochem Soc Trans* **23**:510S.

Robinson RC, Turbedsky K, Kaiser DA, Marchand JB, Higgs HN, Choe S, Pollard TD (2001): Crystal structure of Arp2/3 complex. *Science* **294**:1679-84.

Rogers MS, Baron AJ, McPherson MJ, Knowles PF, Dooley DM (2000): Galactose oxidase pro-sequence cleavage and cofactor assembly are self-processing reactions. *J Am Chem Soc* **122**:990-991.

Rogers MS, Dooley DM (2001): Posttranslationally modified tyrosines from galactose oxidase and cytochrome c oxidase. *Adv Protein Chem* **58**:387-436.

Rogers MS, Tyler EM, Akumanyi N (miss-typed as Akyumani), Kurtis CR, Spooner RK, Deacon SE, Tamber S, Firbank SJ, Mahmoud, K, Knowles PF, Phillips SEV, McPherson MJ, Dooley DM (2007): The stacking tryptophan of galactose Oxidase: A second-Coordination sphere residue that has profound effects on tyrosyl radical behaviour and enzyme catalysis. *Biochemistry, **46**, 4606-4618*

Rosenzweig AC (2002): Metallochaperones: bind and deliver. *Chem Biol* **9**:673-7.

Rossmann MG, Blow DM (1962): the detection of sub-units with the crystallographic asymmetric unit. *Acta crystallogr.*:24-31.

Saysell CG, Barna T, Borman CD, Baron AJ, McPherson MJ, Sykes AG (1997): Properties of the Trp290His variant of *Fusarium* NRRL 2903 galactose oxidase: interactions of the $GOase_{semi}$ state with different buffers, its redox activity and ability to bind azide. *Journal of Biological Inrganic Chemistry* **2**:702-709.

Scharff EI, Koepke J, Fritzsch G, Lucke C, Ruterjans H (2001): Crystal structure of diisopropylfluorophosphatase from Loligo vulgaris. *Structure (Camb)* **9**:493-502.

Schlegel RA, Gerbeck CL, Montgomery R (1968): Substrate specifity of D-galactose oxidase. *Carbohydr Res* **7**:193-199.

Schwartz B, Dove JE, Klinman JP (2000): Kinetic analysis of oxygen utilization during cofactor biogenesis in a copper-containing amine oxidase from yeast. *Biochemistry* **39**:3699-707.

Schwartz B, Klinman JP (2001): Mechanisms of biosynthesis of protein-derived redox cofactors. *Vitam Horm* **61**:219-39.

Schwartz B, Olgin AK, Klinman JP (2001): The role of copper in topa quinone biogenesis and catalysis, as probed by azide inhibition of a copper amine oxidase from yeast. *Biochemistry* **40**:2954-63.

Sellman BR, Tweten RK (1997): The propeptide of Clostridium septicum alpha toxin functions as an intramolecular chaperone and is a potent inhibitor of alpha toxin-dependent cytolysis. *Mol Microbiol* **25**:429-40.

Shrivastava HY, Kanthimathi M, Nair BU (2002): Copper II complex of a tridentate ligand: an artificial metalloprotease for bovine serum albumin. *Biochim Biophys Acta* **1573**:149-55.

Solomon EI, Szilagyi RK, DeBeer George S, Basumallick L (2004): Electronic structures of metal sites in proteins and models: contributions to function in blue copper proteins. *Chem Rev* **104**:419-58.

Sondek J, Bohm A, Lambright DG, Hamm HE, Sigler PB (1996): Crystal structure of a G-protein beta gamma dimer at 2.1A resolution. *Nature* **379**:369-74.

Sprague ER, Redd MJ, Johnson AD, Wolberger C (2000): Structure of the C-terminal domain of Tup1, a corepressor of transcription in yeast. *Embo J* **19**:3016-27.

Springer TA (1998): An extracellular beta-propeller module predicted in lipoprotein and scavenger receptors, tyrosine kinases, epidermal growth factor precursor, and extracellular matrix components. *J Mol Biol* **283**:837-62.

Sun L, Bulter T, Alcalde M, Petrounia IP, Arnold FH (2002): Modification of galactose oxidase to introduce glucose 6-oxidase activity. *Chembiochem* **3**:781-3.

Sun L, Petrounia IP, Yagasaki M, Bandara G, Arnold FH (2001): Expression and stabilization of galactose oxidase in Escherichia coli by directed evolution. *Protein Eng* **14**:699-704.

Sun X, Ollagnier S, Schmidt PP, Atta M, Mulliez E, Lepape L, Eliasson R, Graslund A, Fontecave M, Reichard P, Sjoberg BM (1996): The free radical of the anaerobic ribonucleotide reductase from Escherichia coli is at glycine 681. *J Biol Chem* **271**:6827-31.

Szilagyi RK, Solomon EI (2002): Electronic structure and its relation to function in copper proteins. *Curr Opin Chem Biol* **6**:250-8.

Tang B, Nirasawa S, Kitaoka M, Marie-Claire C, Hayashi K (2003): General function of N-terminal propeptide on assisting protein folding and inhibiting catalytic activity based on observations with a chimeric thermolysin-like protease. *Biochem Biophys Res Commun* **301**:1093-8.

ter Haar E, Musacchio A, Harrison SC, Kirchhausen T (1998): Atomic structure of clathrin: a beta propeller terminal domain joins an alpha zigzag linker. *Cell* **95**:563-73.

Thomson AJ (1991): Radical copper in oxidases. *Nature* **350**:22-3.

Troger W, Ctortecka B, Faller P, Decker H, collaboration. I (2002): Coordination studies of the Metal Center of Hemocyanin by ^{199m}Hg Nuclear Quadroupole Interaction.: "XVIth International Symposim on Nuclear Quadrupole Interactions." Hiroshima, Japan., pp 1-4.

Trylska J, Grochowski P, McCammon JA (2004): The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease. *Protein Sci* **13**:513-28.

Turner BE, Branchaud BP (1999): Probing the radical mechanism of galactose oxidase using an ultrafast radical probe. *Bioorg Med Chem Lett* **9**:3341-6.

van Amsterdam IM, Ubbink M, van den Bosch M, Rotsaert F, Sanders-Loehr J, Canters GW (2002): A new type 2 copper cysteinate azurin. Involvement of an engineered exposed cysteine in copper binding through internal rearrangement. *J Biol Chem* **277**:44121-30.

van der Meer RA, Jongejan JA, Duine JA (1989): Pyrroloquinoline quinone as cofactor in galactose oxidase (EC 1.1.3.9). *J Biol Chem* **264**:7792-4.

Varghese JN, Laver WG, Colman PM (1983): Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution. *Nature* **303**:35-40.

Vellieux FM, Huitema F, Groendijk H, Kalk KH, Jzn JF, Jongejan JA, Duine JA, Petratos K, Drenth J, Hol WG (1989): Structure of quinoprotein methylamine dehydrogenase at 2.25 A resolution. *Embo J* **8**:2171-8.

Vinecombe E (1999): X-Ray crystallographic studies on copper-containing oxidases.: *PhD Thesis, University of Leeds*.

Wachter RM, Branchaud BP (1996a): Molecular modeling studies on oxidation of hexapyranoses by galactose oxidase. An active site topology aparently designed to catalyze radical reactions, either concerted or stepwise. *J Am Chem Soc*:2782-2789.

Wachter RM, Branchaud BP (1996b): Thiols as mechanistic probes for catalysis by the free radical enzyme galactose oxidase. *Biochemistry* **35**:14425-35.

Wang SX, Mure M, Medzihradszky KF, Burlingame AL, Brown DE, Dooley DM, Smith AJ, Kagan HM, Klinman JP (1996): A crosslinked cofactor in lysyl oxidase: redox function for amino acid side chains. *Science* **273**:1078-84.

Watson J, Crick F (1953): Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**:737-738.

Whittaker JW (1995): Spectroscopic studies of galactose oxidase. *Methods Enzymol* **258**:262-77.

Whittaker JW (2002): Galactose oxidase. *Adv Protein Chem* **60**:1-49.

Whittaker JW (2003): Free radical catalysis by galactose oxidase. *Chem Rev* **103**:2347-63.

Whittaker JW (2005): The radical chemistry of galactose oxidase. *Arch Biochem Biophys* **433**:227-39.

Whittaker MM, Ballou DP, Whittaker JW (1998): Kinetic isotope effects as probes of the mechanism of galactose oxidase. *Biochemistry* **37**:8426-36.

Whittaker MM, DeVito VL, Asher SA, Whittaker JW (1989): Resonance Raman evidence for tyrosine involvement in the radical site of galactose oxidase. *J Biol Chem* **264**:7104-6.

Whittaker MM, Duncan WR, Whittaker JW (1996a): Synthesis, Structure, and Properties of a Model for Galactose Oxidase. *Inorg Chem* **35**:382-386.

Whittaker MM, Ekberg CA, Peterson J, Sendova MS, Day EP, Whittaker JW (2000): Spectroscopic and magnetochemical studies on the active site copper complex in galactose oxidase. *Journal of Molecular Catalysis B: Enzymatic* **8**:3-15.

Whittaker MM, Kersten PJ, Cullen D, Whittaker JW (1999): Identification of catalytic residues in glyoxal oxidase by targeted mutagenesis. *J Biol Chem* **274**:36226-32.

Whittaker MM, Kersten PJ, Nakamura N, Sanders-Loehr J, Schweizer ES, Whittaker JW (1996b): Glyoxal oxidase from Phanerochaete chrysosporium is a new radical-copper oxidase. *J Biol Chem* **271**:681-7.

Whittaker MM, Whittaker JW (1988): The active site of galactose oxidase. *J Biol Chem* **263**:6074-80.

Whittaker MM, Whittaker JW (1993): Ligand interactions with galactose oxidase: mechanistic insights. *Biophys J* **64**:762-72.

Whittaker MM, Whittaker JW (2001): Catalytic reaction profile for alcohol oxidation by galactose oxidase. *Biochemistry* **40**:7140-8.

Whittaker MM, Whittaker JW (2003): Cu(I)-dependent biogenesis of the galactose oxidase redox cofactor. *J Biol Chem* **278**:22090-101.

Wilkinson D, Akumanyi N, Hurtado-Guerrero R, Dawkes H, Knowles PF, Phillips SEV, McPherson MJ (2004): Structural and kinetic studies of a series of mutants of galactose oxidase identified by directed evolution. *Protein Eng Des Sel* **17**:141-8.

Wilmot CM, Hajdu J, McPherson MJ, Knowles PF, Phillips SEV (1999): Visualization of dioxygen bound to copper during enzyme catalysis. *Science* **286**:1724-8.

Wright C, Sykes AG (2001a): Autoredox interconversion of two galactose oxidase forms GOase(ox) and GOase(semi) with and without dioxygen. *Inorg Chem* **40**:2528-33.

Wright C, Sykes AG (2001b): Interconversion of Cu(I) and Cu(II) forms of galactose oxidase: comparison of reduction potentials. *J Inorg Biochem* **85**:237-43.

Wunsch P, Herb M, Wieland H, Schiek UM, Zumft WG (2003): Requirements for Cu(A) and Cu-S center assembly of nitrous oxide reductase deduced from complete periplasmic enzyme maturation in the nondenitrifier Pseudomonas putida. *J Bacteriol* **185**:887-96.

Xia ZX, Dai WW, Xiong JP, Hao ZP, Davidson VL, White S, Mathews FS (1992): The three-dimensional structures of methanol dehydrogenase from two methylotrophic bacteria at 2.6-A resolution. *J Biol Chem* **267**:22289-97.

Xiong JP, Stehle T, Diefenbach B, Zhang R, Dunker R, Scott DL, Joachimiak A, Goodman SL, Arnaout MA (2001): Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **294**:339-45.

Xue F, Cooley L (1993): kelch encodes a component of intercellular bridges in Drosophila egg chambers. *Cell* **72**:681-93.

Yoshikawa S, Shinzawa-Itoh K, Tsukihara T (1998): Crystal structure of bovine heart cytochrome c oxidase at 2.8 A resolution. *J Bioenerg Biomembr* **30**:7-14.

Appendix

Protein Engineering, Design & Selection vol. 17 no. 2 pp. 141–148, 2004
Published online January 12, 2004 DOI: 10.1093/protein/gzh018

Structural and kinetic studies of a series of mutants of galactose oxidase identified by directed evolution

D.Wilkinson[1], N.Akumanyi, R.Hurtado-Guerrero, H.Dawkes[2], P.F.Knowles, S.E.V.Phillips and M.J.McPherson[3]

Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK

[1]Present address: Delta Biotechnology Ltd, Castle Court, 59 Castle Boulevard, Nottingham NG7 1FD, UK

[2]Present address: Biologics SRC, Pfizer Ltd, Sittingbourne, Kent ME9 8AG, UK

[3]To whom correspondence should be addressed.
E-mail: m.j.mcpherson@leeds.ac.uk

Galactose oxidase (GO; E.C. 1.1.3.9) is a copper-containing enzyme that oxidizes a range of primary alcohols to aldehydes. This broad substrate specificity is reflected in a high K_M for substrates. Directed evolution has previously been used to select variants of GO that exhibit enhanced expression and kinetic properties. In assays using unpurified enzyme samples, the variant C383S displayed a 5-fold lower K_M than wild-type GO. In the present study, we have constructed, expressed, purified and characterized a number of single, double and triple mutants at residues Cys383, Tyr436 and Val494, identified in one of the directed evolution studies, to examine their relative contributions to improved catalytic activity of GO. We report kinetic studies on the various mutant enzymes. In addition, we have determined the three-dimensional structure of the C383S variant. As with many mutations identified in directed evolution experiments, the availability of structural information does not provide a definitive answer to the reason for the improved K_M in the C383S variant protein.

Keywords: copper oxidase/directed evolution/galactose oxidase/guar gum/substrate binding

Introduction

Galactose oxidase (GO; E.C. 1.1.3.9), a 68 kDa mononuclear copper-containing enzyme from *Fusarium graminearum*, oxidizes primary alcohols to the corresponding aldehyde with coupled reduction of molecular oxygen to hydrogen peroxide (Kosman *et al.*, 1974) according to the reaction scheme:

$$RCH_2OH + O_2 \rightarrow RCHO + H_2O_2$$

The biologically relevant substrate of GO is not known as the enzyme exhibits broad substrate specificity from small alcohols through sugars to oligo- and polysaccaharides (Avigad *et al.*, 1962; Maradufu *et al.*, 1971; Maradufu and Perlin, 1974; Mendonca and Zancan, 1987).

The crystal structure of GO has been determined to 1.7 Å resolution, and the protein is predominantly β-structure with three domains (Ito *et al.*, 1991, 1994). The copper at the active site is coordinated by His496, His581, Tyr272 and a solvent molecule as equatorial ligands, with Tyr495 as an axial ligand at a distance of 2.69 Å from the copper. One of the copper ligands, Tyr272, is covalently bonded through Cε to the sulphur of Cys228 and is the site of the free radical. The side chain of Trp290 stacks over this Tyr–Cys feature (Figure 1A). The semi-reduced form of the enzyme [Cu(II)-Tyr] can undergo a one electron oxidation to the oxidized state [Cu(II)-Tyr$^{\cdot}$] which catalyses the two electron oxidation of the substrate.

There is interest in the use of GO for industrial processes such as derivatization of guar gum and related polymers. Guar gum is a galactomannan isolated from *Cyamopsis tetragonoloba*. It comprises a 1→4-linked β-D-mannopyranose backbone with 1→6-linked α-D-galactopyranose residues in a ratio of 1 D-galactose to ~1.5–2 D-mannose residues. Guar gum is a complex polymer with molecules that comprise ~10 000 monosaccharide residues. It is widely used as a food thickener and emulsifier, and is also used in a variety of other industries including textiles, printing, cosmetics, pharmaceuticals and paper manufacture (Hall and Yalpani, 1980; Andreana *et al.*, 2002).

In order to improve the activity of GO towards appropriate substrates, there have been recent reports of directed evolution based on error-prone PCR. In the first example (Delagrave *et al.*, 2001), a digital imaging system was used to screen microcolonies for activity against 1-methyl-α-D-galactoside as an analogue for guar gum, and also against guar gum itself. Using an assay system based on crude lysates of bacterial cultures, these authors determined K_M and V_{max} values. The results identified mutations at Cys383, Tyr436 and Val494 that were subsequently combined by subcloning and assayed for additive contributions to enhanced activity. The triple mutant C383S/Y436H/V494A showed a reduced K_M and increased V_{max} contributing to a 15.9-fold greater V_{max}/K_M with 1-methyl-α-D-galactoside as substrate, relative to wild-type. This variant also showed 10.4-fold higher activity against guar gum than the wild-type. This study was designed to identify improved enzymatic activity in crude extracts and so there was no measurement of relative expression levels of the variant proteins. Enhanced activity towards substrate in such an assay system may therefore be due to contributions from increased levels of expressed protein and/or enhanced catalytic activity.

The importance of enhanced expression is demonstrated in a second directed evolution study on GO (Sun *et al.*, 2001). In this case, individual bacterial colonies were cultured, lysed and assayed in microtitre plates for increased activity. Further rounds of mutagenesis led to increased activity and variants were purified for further analysis. The most active variant (A3E7) displayed a 30-fold increase in activity resulting from an 18-fold higher expression and 1.7-fold greater catalytic efficiency. This variant contained a silent mutation in the codon for P136 and five amino acid substitutions: S10P, M70V, G195E, V494A and N535D. Two of these mutations were

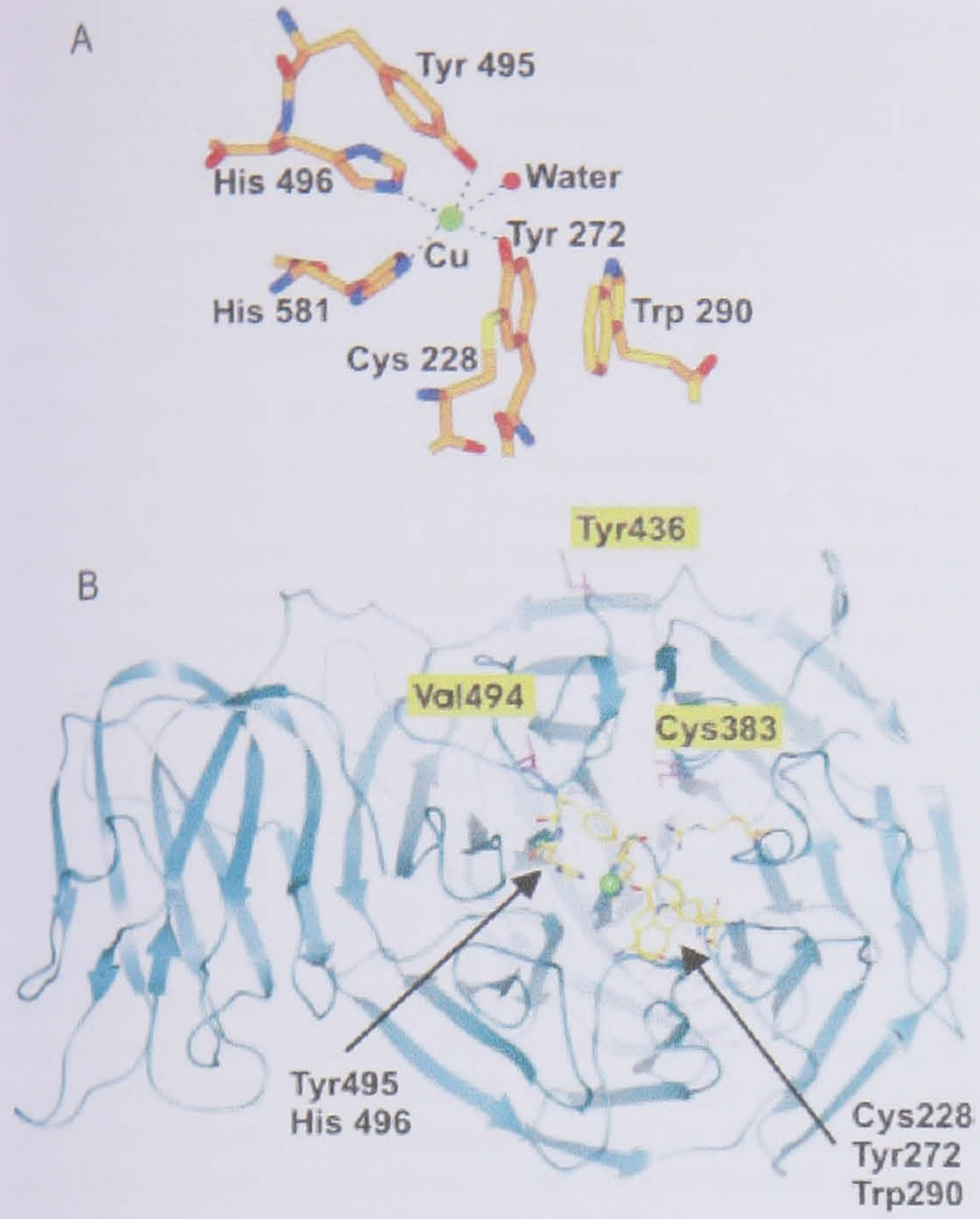


Fig. 1. Structure of GO. (**A**) The active site showing the copper ligands, the thioether bond between Tyr272 and Cys228, and the stacking interaction with Trp290. (**B**) An overview of the GO monomer showing domains 1 and 2 and showing the locations of the residues Cys383, Tyr436 and Val 494 (shown in mauve and highlighted by yellow labels), which form the basis of the present study. The active site residues are shown in atom colouring and are labelled by arrows. Non-covalent bond interactions are shown as dotted lines.

common to both studies where V494A was identified as enhancing V_{max}, while N535D was shown to contribute a modest increase in activity.

We have investigated in detail the variants identified by Delagrave *et al.* (2001) by generating a series of single and double mutants comprising V494A, C383S and Y436H for comparison with the triple mutant clone. We also studied the variants C383A and Y436A. The locations of the mutated residues are shown in Figure 1B. The most interesting effect is reported to be due to the C383S mutation (Delagrave *et al.*, 2001) which leads to a significant reduction in K_M. Cys383 lies in a pocket at the back of the active site behind the copper atom and therefore is not able to make direct interactions with the alcohol substrate. It is one of five cysteines that can be titrated in the denatured, reduced enzyme, but no role has yet been assigned to this residue. To explore the structural basis for improvement in K_M we have determined the three-dimensional structure of the C383S variant by X-ray crystallography.

Materials and methods

Generation of GO mutants

GO from *Fusarium* spp. (NRRL 2903) was expressed using the pPICZα vector (Invitrogen) in *Pichia pastoris* X-33. The GO gene encodes a 41 amino acid N-terminal leader sequence comprising a secretion signal and a 17 amino acid pro-sequence that remains associated with the protein when it is purified under copper-free conditions from the heterologous host *Aspergillus nidulans*.

The constructs used in this work contained the coding sequence for the 17 amino acid pro-form of the enzyme linked in-frame to the α-mating factor secretion signal of the vector, and are therefore referred to as Pro-GO. The single mutant C383S and the triple mutant C383S/Y436H/V494A, referred to as variant 2-1 and 8-1, respectively (Delagrave *et al.*, 2001), were kindly provided by Hercules European Research Center (Barnveld, The Netherlands). Other mutants were generated using the plasmid pPICZαproGO as template for the QuikChange Multi Site-Directed mutagenesis kit (Stratagene). The mutagenic oligonucleotides 5′-CTGATGCCATG**T**C**T**GGAAACGCTGTC-3′ (C383S), 5′-CTGATGCCATG**GC**T**T**GGAAACGCTGTC-3′ (C383A), 5′-CCATTGTTCGC**G**C**TT**ACCATAGCATTTCC-3′ (V494A), 5′-GCAATGGGTTG**GCT**TTTGCCCGAACG-3′ (Y436A) and 5′-GCAATGGGTTGC**A**TTTTGCCCGAACG-3′ (Y436H) were purchased from MWG Biotech; the altered codon is underscored, while nucleotide substitutions are shown in bold. The presence of the expected mutation and the absence of further mutations were confirmed by DNA sequencing of the complete GO-encoding gene.

Pichia pastoris transformation and characterization

Plasmid DNA was linearized with *Sac*I (New England Biolabs) and transformed into Mut$^+$ *P.pastoris* X-33 cells using the EasyComp transformation kit (Invitrogen). Colonies were then screened for activity by plating onto BMMY plates (buffered complex medium containing methanol; Invitrogen manual) containing the components for a chromogenic assay {D-galactose (2%), 2,2′-azino-bis(3-ethylbenz-thiazoline-6-sulphonic acid) (ABTS) (1.2 mg/ml) and horseradish peroxidase [0.04 units (0.4 μg)/ml]}. Transformants that expressed active GO resulted in formation of a green coloration following incubation for between 5 min and several hours at 30°C.

To confirm the identity of the protein being produced by *Pichia* transformants, genomic DNA was extracted. Briefly, *P.pastoris* transformants were grown in 10 ml of YPD medium for 16 h. The cells were harvested (1500 *g*, 10 min) and resuspended in 1 ml of SCED buffer (1 M sorbitol, 10 mM Na citrate, pH 7.5, 10 mM EDTA, 10 mM DTT) containing 100 units/ml of lyticase (Sigma). After incubating at 37°C for 50 min the resulting spheroplasts were harvested (10 000 *g*, 10 min). The pellet was used for genomic DNA preparation using the Qiagen DNeasy plant kit (Qiagen) by following the manufacturer's instructions. The genomic DNA was used as template for PCR using the pPICαZ vector-specific primers AOX1 (gactggttccaattgacaagc) or ALPHAF (actattgccagcattgctgc) as forward primers, and 3′AOXB (GTCGACGGCGCTATTCAGATC) as the reverse primer, and the following conditions: 94°C for 2 min, then 25 cycles of 94°C for 1 min; 55°C for 1 min; 72°C for 1 min, with a final 72°C for 7 min incubation. Each reaction contained 1× buffer (Amersham Pharmacia), 1 μg of genomic DNA, 50 μM each dNTP, 50 pmol of each primer and 2.5 units of *Taq* DNA polymerase (Amersham Pharmacia). The PCR products were purified through a Qiagen PCR clean-up column and were used as templates for DNA sequencing using the following primers, labelled at their 5′-ends with an appropriate IRD700 or IRD800 dye. Forward primers (IRD700): AOX1, 5′(IRD700)GACTGGTTCCAATTGACAAGC3′; ALPHAF 5′(IRD700)ACT-

ATTGCCAGCATTGCTGC3′; GOFOR, 5′(IRD700)GGCAGCCCTGTTGCGTCAG3′; GOFORB, 5′(IRD700)CCAGTCTAACCGTGGTGTAG3′; reverse primers (IRD800): GOREV, 5′(IRD800)GATCTCAGGTGTAAATACC3′; 3′AOX1, 5′(IRD800)GCAAATGGCATTCTGACATCC3′; 3′AOXB, 5′(IRD800)GTCGACGGCGCTATTCAGATC3′.

The reaction products were analysed on a Li-Cor IR^2 automated DNA sequencing instrument.

Protein expression and purification

Cultures (1 litre) of *P.pastoris* transformants were grown overnight at 30°C in BMGY (Invitrogen manual). When the culture OD_{600} reached 2–6, cells were harvested by centrifugation and resuspended in BMMY containing 0.5 mM copper sulphate and 0.5% methanol to induce expression of GO and were grown for a further 48 h at 25°C with a further addition of methanol to 0.5% at 24 h (Whittaker and Whittaker, 2000). Secreted GO was purified from the culture supernatant using a modification of the affinity/size exclusion procedure described by Hatton and Rodgoeczi (1982) and Sun *et al.* (2001). Proteins were concentrated from the 50 mM Tris–HCl, pH 8, buffered culture supernatant by adding ammonium sulphate to 65% saturation. Protein was redissolved in and dialysed against 100 mM ammonium acetate buffer, pH 7, prior to loading onto a pre-equilibrated Sepharose 6B column. The eluted enzyme was concentrated by ultrafiltration using an Amicon UF unit and YM30 membranes before being copper-loaded according to the dialysis procedure described by Baron *et al.* (1994).

Kinetic measurements

Protein concentration was determined by absorbance at 280 nm (ε_{280} 105 400 mM^{-1} cm^{-1}). Kinetic measurements were made using the coupled assay system to measure hydrogen peroxide production (McPherson *et al.*, 2001). A sample of diluted enzyme (50 µl) was added to 1 ml of the assay buffer containing horseradish peroxidase (0.165 mg/ml), 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid), diammonium salt, ABTS (1 mg/ml), D-galactose or 1-methyl-α-D-galactose (0–700 mM) or guar [1% (w/v)], sodium phosphate (100 mM, pH 7.0) and the absorbance change at 405 nm was recorded. Assays were performed at 25°C. The chromophore ε_{405} used was 31 300 mM^{-1} cm^{-1}. To ensure that copper was not limiting in the reactions, we also performed kinetic measurements following pre-incubation of the enzyme in the presence of 50 µM copper sulphate and conducted the assay in the presence of this concentration of copper sulphate.

For experiments with guar gum, 0.5 g of supercol guar gum (food grade; kindly provided by Hercules Inc.) was suspended in 200 ml of sterile water to give a 0.25% solution. This was incubated at 50°C for 1 h and then centrifuged at 14 000 *g* for 30 min, and the supernatant was used in kinetic measurements. Assays were performed by mixing 500 µl of 2× assay buffer, 450 µl of 0.25% guar and 50 µl of enzyme. Three independent measurements were made to determine the specific activity for each enzyme being investigated, where the concentration of protein varied between 1 and 0.1625 µg/ml in the assay.

Mass spectrometry

Samples were analysed on a Q-Tof orthogonal acceleration quadrupole time-of-flight mass spectrometer equipped with nano-electrospray ionization (Micromass UK Ltd, Manchester, UK). Samples were dissolved in 1:1 (v/v) aqueous methanol containing 1% formic acid. Data were acquired over the appropriate *m/z* range and spectra processed using the MassLynx software supplied with the mass spectrometer. The spectra were calibrated with a separate introduction of horse heart myoglobin (MW 16 951.49 Da) and a mass accuracy of 0.01% is expected.

Enzyme oxidation and reduction

Protein samples (2–15 µM) in 0.1 M sodium phosphate buffer pH 7 were chemically oxidized by incubating at room temperature for 5–10 min in the presence of 50 mM potassium ferricyanide. Oxidant was subsequently removed by size exclusion chromatography using a Micro Bio-Spin 6 column (Bio-Rad) equilibrated in 50 mM sodium acetate, pH 4.5. This pH jump procedure for simultaneous oxidant removal and buffer exchange was designed to stabilize the radical form of GO (Rogers *et al.*, 1998). Samples were then used for UV–Vis spectral studies in a Shimadzu UV2401-PC spectrophotometer over the range 300–900 nm.

Crystallography

The C383S mutant protein was dialysed into 20 mM HEPES (pH 7.0) overnight and then concentrated to 2 mg/ml using an Amicon Concentrator. The protein was crystallized at 18°C using the sitting drop method where the well solution was 0.2–1.0 M ammonium sulphate, 200 mM sodium acetate (pH 4.6) and 20–30% PEG 4000. Four to 7 µl of well solution were added to 4–7 µl of C383S protein solution in the drop and mixed thoroughly by pipetting. Crystals appeared after 2 weeks and reached ~400 µm in the largest dimension after 8 weeks. Prior to X-ray diffraction data collection, a crystal was placed in cryoprotectant comprising 0.3 M ammonium sulphate, 20% PEG 8000, 25% PEG 400. The crystal was soaked for 2–3 min and then flash frozen in liquid nitrogen. The data were collected on station 14.2 at Daresbury Laboratory, UK.

Data processing and structure determination

The data set was processed using the HKL suite comprising DENZO and SCALEPACK (Otwinowski and Minor, 1997). A single frame was used for indexing. Data were scaled and merged in SCALEPACK (Otwinowski and Minor, 1997). The output from SCALEPACK was converted to an MTZ format using the CCP4 program MTZ2VARIOUS (Bailey, 1994) and then to structure factor amplitudes using the CCP4 program TRUNCATE (Bailey, 1994). Phase information was obtained by using the original GO structure (PDB code 1GOG) since the space group and unit cell were isomorphous to the C383S variant. Rigid body refinement was carried out using CNS (Brunger *et al.*, 1998) with data to a resolution of 3 Å until the *R*-values converged (final R = 23.97, free R = 24.42). Water molecules were added using WATER_PICK.

Results

Pichia pastoris transformation and protein purification

DNA sequence analysis confirmed the presence of the desired mutations in the pPICZαproGO constructs and these were then transformed into *P.pastoris* strain X-33. Transformants expressing active GO formed a green colour on BMMY plates containing the chromogenic assay components. All GO mutants tested in this qualitative assay displayed some level of enzyme activity. Cultures (1 l) were induced by the addition of methanol to 0.5% for production of GO protein that was purified from the culture medium by ammonium sulphate precipitation followed by dialysis and then affinity purification using a Sepharose 6B column essentially according to the procedure of Sun *et al.* (2001). The resulting proteins were pure

as judged by SDS–PAGE. The GO was isolated as the inactive blue form of the enzyme as reported by Whittaker and Whittaker (2000), probably due to the presence of methanol in the culture medium leading to the semi-reduced form (Cu^{2+}/Tyr).

The wild-type and variant proteins were analysed by electrospray mass spectrometry (ES-MS). The PRO-GO protein produced in *P.pastoris* displayed a mixture of N-terminal extended species rather than the expected 17 amino acid N-terminal pro-sequence observed when produced in *A.nidulans* (Rogers *et al.*, 2000). The majority of the protein has the additional residues Ser3, Leu2 and Arg1 at the N-terminal end. The expected and observed molecular masses of the mature proteins are shown in Table I. The C383S/V494A variant exhibited a low solubility that prevented ES-MS analysis. For the majority of the samples there is good agreement between the expected and observed molecular masses. The V494A variant shows a substantially higher mass than expected of +42 mass units, which may be due to the presence of other metal ions, such as a calcium or two sodium ions. The variants Y436H/V494A, C383S/Y436H/V494A and the wild-type control all differ from the expected value by 10 or more mass units, although the reason for this is unclear. To confirm that the proteins being characterized contained only the expected mutations, the GO coding region was PCR amplified from *P.pastoris* genomic DNA isolated from each transformant, and was subjected to DNA sequence analysis. These data confirmed the correct gene sequence with no undesired mutations.

Table I. Molecular masses of proteins determined by ES-MS

Protein	Molecular weight		Difference
	Theoretical	Actual	
Wild-type proGO	68 516	68 527	+11
Wild-type matGO	68 516	68 520	+4
C383A	68 484	68 485	+1
C383S	68 500	68 507	+7
Y436A	68 424	68 426	+2
Y436H	68 490	68 492	+2
V494A	68 488	68 530	+42
C383S/Y436A	68 408	68 399	–9
C383S/Y436H	68 474	68 468	–6
C383S/V494A	68 472	ND	ND
Y436A/V494A	68 396	68 399	+3
Y436H/V494A	68 462	68 452	–10
C383S/Y436H/V494A	68 446	68 460	+14

ND, not determined due to the low solubility of this mutant.

Enzyme kinetics

The initial selection process that gave rise to these variants (Delagrave *et al.*, 2001) was designed to identify variants that would function efficiently using guar gum as substrate. However, difficulties in preparing working solutions of guar gum led initially to the use of 1-methyl-α-D-galactose as a guar substrate mimic. It was interesting, therefore, to compare the kinetic parameters of the various mutants against this substrate and unmodified D-galactose, the commonly used substrate for GO. Crude cell extracts were previously used to make kinetic measurements (Delagrave *et al.*, 2001), so the use of purified protein in the present work should provide more accurate information about the influence of the various mutations on enzyme activity. Initial enzyme activity was measured for the wild-type and each variant across a range of substrate concentrations (10–700 mM) for both D-galactose and 1-methyl-α-D-galactose, and K_M and k_{cat} values were determined. The inclusion of additional copper sulphate at 50 μM did not alter the assay results, indicating that protein samples were not deficient in copper. Table II provides a summary of the kinetic data for both substrates and also reports the catalytic efficiency (k_{cat}/K_M) as well as the relative catalytic efficiency compared with wild-type GO. In general, the kinetic parameters for the wild-type GO and the variants were similar with both substrates. We have confirmed the observation that the K_M of GO variants containing the C383S substitution is reduced by 3.5- to 4.5-fold compared with wild-type GO (Delagrave *et al.*, 2001). This difference is not simply due to the loss of the sulphur atom, as this effect is not observed for the C383A variant, implying that enhancement of substrate binding is due to the serine side chain. The C383S variant shows a minor reduction in k_{cat}, but in terms of catalytic efficiency the dominant effect of K_M results in an enzyme with a 3- to 4-fold enhanced k_{cat}/K_M. In contrast, the C383A variant represents a slightly worse enzyme with a k_{cat}/K_M of 0.5- to 0.6-fold that of the wild-type due to a reduction in k_{cat}.

Combination of the Y436A or H and V494A variants reveals an interesting feature, with the Y436H/V494A variant showing a catalytic efficiency essentially the same as wild-type, indicating that the positive effect of the V494A variant on

Table II. Kinetic properties of GO mutants with D-galactose and 1-methyl-α-D-galactose as substrates

Mutant	D-Galactose				1-Methyl-α-D-galactopyranoside			
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})	k_{cat}/K_M WT/variant	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat}/K_M WT/variant
Wild-type proGO	68 ± 5	1090 ± 40	16 000 ± 1300	1	56 ± 5	995 ± 34	17 800 ± 1700	1
C383S	15 ± 1	834 ± 18	55 600 ± 3900	3.5	16 ± 3	839 ± 63	52 400 ± 11000	2.9
C383A	62 ± 6	639 ± 30	10 300 ± 1100	0.6	47 ± 6	457 ± 34	9700 ± 1400	0.5
V494A	53 ± 8	1119 ± 66	21 100 ± 3400	1.3	46 ± 4	1444 ± 86	31 400 ± 3300	1.8
Y436A	61 ± 9	1100 ± 71	18 000 ± 2900	1.1	46 ± 5	1039 ± 103	22 600 ± 3300	1.3
Y436H	53 ± 6	756 ± 36	14 300 ± 1800	0.9	55 ± 5	1267 ± 97	23 000 ± 2700	1.3
C383S/V494A	16 ± 2	915 ± 33	57 200 ± 7400	3.6	17 ± 3	1495 ± 114	87 900 ± 17 000	4.9
C383S/Y436A	17 ± 2	1266 ± 63	74 500 ± 9500	4.7	25 ± 3	811 ± 63	32 400 ± 4600	1.8
C383S/Y436H	22 ± 3	1871 ± 96	85 000 ± 12 400	5.3	17 ± 3	1091 ± 69	64 200 ± 12 000	3.6
Y436A/V494A	30 ± 8	1398 ± 148	46 600 ± 13 400	2.9	46 ± 3	1548 ± 95	33 600 ± 3000	1.9
Y436H/V494A	68 ± 4	1181 ± 36	17 400 ± 1200	1.1	45 ± 7	851 ± 91	18 900 ± 3600	1.1
C383S/Y436H/V494A	19 ± 2	1301 ± 46	68 500 ± 7600	4.3	18 ± 2	1067 ± 40	59 300 ± 7000	3.3

k_{cat} is not evident in this double mutant. In contrast, the Y436A/V494A variant does reveal the enhanced k_{cat}, thus proving to be a more efficient enzyme with a 2- to 3-fold enhanced catalytic efficiency. The triple mutant C383S/Y436H/V494A displays an enhanced k_{cat}/K_M predominantly due to the K_M effect of C383S but with a small k_{cat} enhancement against both substrates tested.

There are some interesting differences in the data for the two substrates, for example, the C383S/Y436A variant displays a 2-fold lower catalytic efficiency against 1-methyl-α-D-galactose than against D-galactose, largely due to an effect on k_{cat}. The reason for such differences remain unclear at present. Figure 2 provides an overview of the comparative catalytic efficiencies of the variant enzymes with wild-type and for the substrates D-galactose and 1-methyl-α-D-galactose. The improvement afforded by proteins carrying the C383S variant is clearly demonstrated.

In order to assess the relative efficiencies of the variants against guar gum we conducted coupled assays using a 1% (w/v) guar gum solution. Since it was not possible to measure K_M data for guar gum, Table III shows a comparison of k_{cat} data for 1-methyl-α-D-galactose and guar gum as substrates. It is clear that activity against guar gum is poor, being only 3 to 6% of the activity with 1-methyl-α-D-galactoside as substrate. Even for variants such as C383S/V494A and the triple mutant C383S/Y436H/V494A, which display a 4.9- and 3.3-fold enhancement in catalytic efficiency with the monosaccharide substrate, the activity with guar gum remains low.

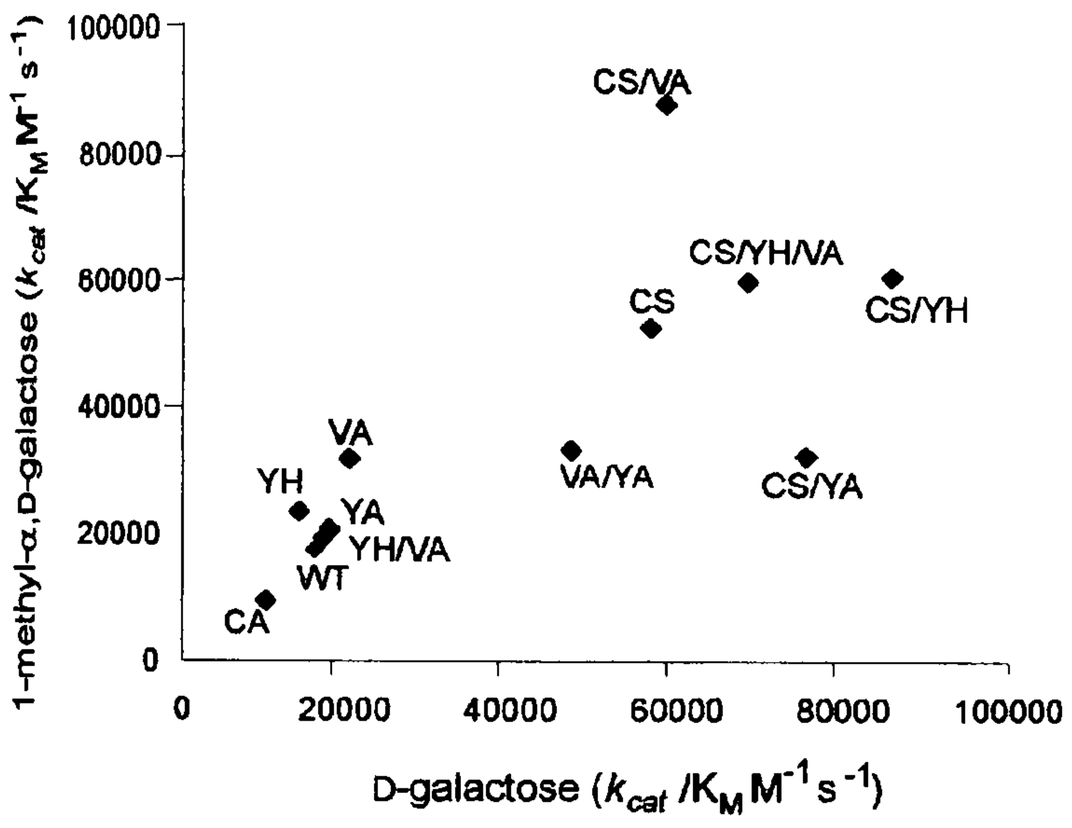


Fig. 2. Catalytic efficiencies of the GO variants relative to wild-type. The k_{cat}/K_M values for the variants are plotted for the two monosaccharides, D-galactose and 1-methyl-α-D-galactose, used in this study. The plot shows the enhanced catalytic efficiencies of variants that include the C383S variant.

Table III. Comparison of k_{cat} for GO and variant proteins with 1-methyl-α-D-galactoside and guar

Mutant	k_{cat} (s^{-1}) 1-methyl-α-D-galactose	k_{cat} (s^{-1}) guar	% Activity of wild-type and variants with guar compared with 1-methyl-α-D-galactoside
Wild-type proGO	995 ± 34	42.8 ± 2.3	4.3
C383S	839 ± 63	41.1 ± 1.4	4.1
C383A	457 ± 34	13.7 ± 2.3	1.4
V494A	1444 ± 86	32 ± 2.3	3.2
Y436A	1039 ± 103	26.3 ± 6.9	2.7
Y436H	1267 ± 97	30.4 ± 3.4	3.1
C383S/V494A	1495 ± 114	58 ± 4.6	5.8
C383S/Y436A	811 ± 63	41.1 ± 3.4	4.2
C383S/Y436H	1091 ± 69	37.7 ± 2.3	3.8
Y436A/V494A	1548 ± 95	29.7 ± 2.3	3.0
Y436H/V494A	851 ± 91	26.3 ± 1.1	2.6
C383S/Y436H/V494A	1067 ± 40	44.5 ± 3.4	4.5

UV–Vis spectroscopy

Samples of the proteins were oxidized by incubation with ferricyanide at pH 7 to generate the oxidized (radical-containing) form of the enzyme (Cu^{2+}/$Tyr^{\cdot}$) and this species was then trapped by a pH jump desalting step into 50 mM sodium acetate buffer at pH 4.5 (Rogers *et al.*, 1998). The UV–Vis spectra showed the 445 nm band characteristic of the oxidized spectrum. The broad feature centred around 800 nm is not seen, as this arises from a ligand-through-metal-ligand charge transfer and at low pH Tyr495 becomes protonated and dissociates from the metal, so reducing this spectral feature (Whittaker and Whittaker, 1993; McGlashen *et al.*, 1995). All the variants showed essentially similar spectra for their oxidized states.

X-ray crystallography

We determined the crystal structure of the C383S variant in order to explore structural differences between this variant and wild-type GO that might explain the reduction in K_M on substitution of the sulphur atom with an oxygen atom. C383S was crystallized in space group *C*2 and the three-dimensional structure was solved to a resolution of 2.4 Å. A summary of the crystallographic data collection and refinement statistics is provided in Table IV. Apart from the alteration of Cys383 to a serine residue (Figure 3A and B), there are no major structural changes. Figure 3B shows the presence of a mixed occupancy site with acetate (~40%) and water (~60%) close to the acetate binding site in the original crystal structure (Ito *et al.*, 1991, 1994). It seems likely this mixed occupancy has arisen during the transfer of crystals from the crystallization liquor (acetate, pH 4.5) to cryoprotectant (no acetate, pH 6.3). The position of the acetate provides an indication of the site of substrate binding, consistent with modelling studies (Ito *et al.*, 1994). The Ser383 side chain forms a H-bond interaction to a water that is part of a H-bond network to other waters and the

Table IV. Crystallographic data collection and refinement statistics for C383S

Space group	*C*2
Unit cell dimensions	a = 97.54 Å, b = 89.02 Å, c = 85.79 Å, α = 90°, β = 117.61°, γ = 90°
Total number of reflections	142 517
Number of unique reflections	28 426 (836)
R_{sym} (%)	8.5 (25.4)
Completeness (%)	97.1 (71.1)
$I/\sigma I$	10.0
Resolution limits (Å)	80–2.30
R_{cryst} (%)	21.2
R_{free} (%)	22.8
Number of atoms	5244
Wilson *B*-factor (Å^2)	12.02
Average overall *B*-factor (Å^2)	32.10
R.m.s. bond length (Å)	0.01
R.m.s. bond angles (°)	1.95

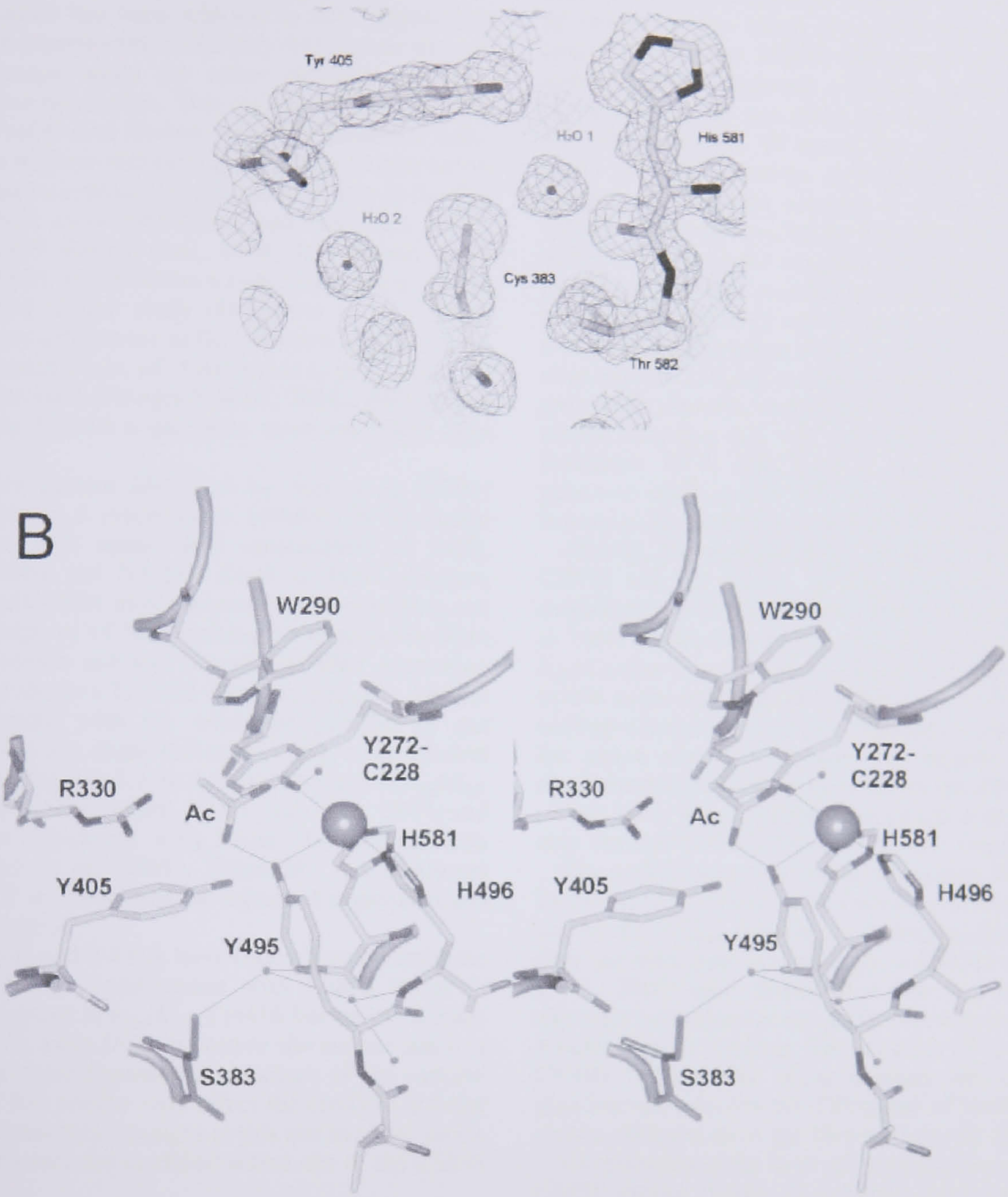


Fig. 3. (**A**) Electron density map of wild-type GO showing Cys383. (**B**) Stereo view of the active site of the C383S variant. Copper is shown as the large green sphere and water molecules as small red spheres, acetate is labelled Ac and indicate the site at which substrate would bind. The structure shows the H-bond interaction between Ser383 and a water that leads to a further water-mediated interaction with the backbone nitrogen between residues Tyr495 and His496.

backbone nitrogen between residues Tyr495 and His496 (Figure 3B).

Discussion

The construct used for GO expression in these experiments contains the 17 amino acid N-terminal pro-sequence (Rogers *et al.*, 2000; Firbank *et al.*, 2001) and mass spectrometry data revealed that the major form of GO contains an additional three residues at the N-terminus (Ser3, Leu2 and Arg1). This finding contrasts with the production of recombinant enzyme from a pro-sequence construct in the filamentous fungus *A.nidulans* or of the native enzyme from *F.graminearum* where in both cases there is a single mature enzyme species starting at Ala +1 (McPherson *et al.*, 1992; Baron *et al.*, 1994). This observation is consistent with a previous report of mixed N-terminal sequences from constructs expressing GO containing a pro-sequence in *P.pastoris* (Whittaker and Whittaker, 2000). The reason for the differences in N-terminal processing by these two hosts is unclear; however, the recombinant enzymes prepared from *P.pastoris* and *A.nidulans* exhibit similar properties including kinetics, UV–Vis spectra and three-dimensional structure (Deacon *et al.*, 2004).

The wild-type GO and variants display Michaelis–Menten kinetics. The K_M for wild-type GO is high at ~60–70 mM. This seems to be a consequence of the broad substrate specificity of

the enzyme resulting in an active site capable of binding a range of different alcohol substrates, but therefore being relatively weak at binding any particular substrate.

The V494A variant has been selected in two independent directed evolution experiments, indicating that it must provide a selective advantage under the rather different selection criteria used in these two studies. This residue is adjacent to the copper ligands Tyr495 and His496, and may therefore influence the positions of these residues. In particular, it may affect Tyr495, which plays a critical role as a general base in the first step of the catalytic cycle (Whittaker and Whittaker, 1993; Reynolds *et al.*, 1997; Rogers *et al.*, 1998). It is perhaps more surprising that C383S, which shows a substantial effect on K_M, was only identified in one study (Delagrave *et al.*, 2001). Perhaps this reflects differences in the selection regimes used. In one case, concentrations of 1-methyl-α-D-galactoside of 0.72–200 mM were used (Delagrave *et al.*, 2001), while in the other, 50 and then 25 mM D-galactose were employed (Sun *et al.*, 2001).

The most active mutant identified by Sun *et al.* (2001) contains six mutations, a synonymous mutation in the codon for P136 together with amino acid substitutions of S10P, M70V,G195E,V494A and N535D. Some of these changes, such as N553D and G195E, may enhance protein solubility and reduce the very high p*I* of GO, making its expression more efficient in *Escherichia coli* and potentially other expression hosts. This variant shows a 1.7-fold increase in k_{cat}/K_M against D-galactose compared with the wild-type protein. In our V494A variant there is a slight reduction in K_M and a modest increase in k_{cat}, leading to a 1.3- to 1.8-fold increase in k_{cat}/K_M. This is quite similar to the variant A3.E7 (Sun *et al.*, 2001) and is also in general agreement with initial observations with V494A (Delagrave *et al.*, 2001). Therefore, the additional mutations in A3.E7 are important for enhanced expression and solubility of the protein.

Mutations Y436H and Y436A have little effect on catalytic efficiency alone but in combination with C383S result in substantial enhancement in k_{cat}/K_M. Tyr436 lies on the surface of the protein ~27 Å away from the active site copper and it is not clear from the three-dimensional structure of the enzyme how alterations at this residue may affect the catalytic activity of the enzyme. Presumably, changes at this site mediate subtle alterations that influence the modified active site of the C383S variant.

Single and double mutants exhibit general agreement in the kinetics measurement trends observed in crude assays (Delagrave *et al.*, 2001) and the present study. Some of the previously reported relative V_{max}/K_M values are somewhat higher than the relative k_{cat}/K_M values reported here. For example, the reported increase in V_{max}/K_M for the triple mutant (C383S/Y436H/V494A) was 15.9 (Delagrave *et al.*, 2001) compared with a level of 3.4–4.4 in the present study. This is probably because the original assays were performed in crude extracts compared with the use of purified proteins here.

Since one aspect of enhancing the activity of GO is to improve activity against polymeric substrates, we assayed the variants with guar as substrate. It should be noted, however, that activity is expressed in terms of k_{cat} as it was not possible to determine K_M values for guar gum as substrate. For variants that showed improved activity against 1-methyl-α-D-galactose, a comparable pattern of increases in relative activity against guar gum was observed (Delagrave *et al.*, 2001). This trend is also seen in our data where, for example, C383S/V494A is the most efficient variant against 1-methyl-α-D-galactose and also shows the highest activity against guar. However, as shown in Table III, the wild-type and variants, including C383S/V494A, all show very low levels of activity against guar in a coupled assay system. Thus, despite differences in catalytic efficiencies against a monosaccharide substrate, these enhanced activities of only 1–2% do not make the enzymes substantially more efficient with guar. Of course the assay system we used was based on initial activity and therefore may not be a good predictor of residual activity in a longer incubation assay system. The major reason for low activity against guar gum is likely to relate to the complex structure of this polymer. The guar gum preparation used was cationic and cold extracted, which should enhance solubility, but could only be used in the assays at concentrations of 1%, resulting in a low concentration of galactose (<20 mM). Since the galactomannan isolated from guar seeds comprises some 38 ± 2% galactose with a complex tertiary structure that will limit enzyme access (Hoffman and Svensson, 1978; McCleary *et al.*, 1985), many of the D-galactose residues will be inaccessible to the enzyme, thereby lowering the apparent concentration further.

Clearly, the most important single mutation in this study is C383S and the effects of this mutation are maintained in combination with all the other mutations tested. As illustrated in Table II, this mutation results in a 3- to 3.6-fold increase in k_{cat}/K_M due to a reduction in K_M to ~17 mM. The substitution results in the exchange of a sulphur atom for a smaller oxygen, leading to more space in a pocket that lies behind the copper in the active site. Cys383 does not directly interact with the alcohol substrate, but the presence of the -OH group on Ser383 clearly leads to an enhancement of substrate binding, presumably through indirect interactions with substrate.

The importance of the serine at position 383 is demonstrated by the Ala383 variant, which shows a similar K_M to wild-type and a reduced k_{cat}, resulting in an enzyme that is slightly poorer than the wild-type. Interestingly, a C383G variant (Delagrave *et al.*, 2001) also displayed a low K_M similar to C383S, although the reason for this difference between the Ala and Gly substitutions is unclear. The absence of a side chain in the C383G variant may allow a water molecule to occupy a position equivalent to the -OH group of Ser383, thus mediating similar differences in the H-bond network in the active site.

Examination of the three-dimensional crystal structure of the C383S variant (Figure 3) provides little understanding of the reason for the beneficial effect of this substitution. Cys383 lies within a pocket at the back of the active site. Substitution by Ser replaces a weak H-bond group with a strong H-bonding group and would lead to a stronger interaction with neighbouring water molecules. Such a change could be transmitted through the H-bond network to the substrate binding site, subtly altering its structure or flexibility. Perhaps more obvious differences within the active site would become apparent upon substrate binding. Unfortunately, there is no crystal structure for GO with any bound substrate, therefore preventing direct testing of this possibility at present.

A feature of directed evolution experiments as a tool for the identification of beneficial mutations within proteins is that unexpected substitutions can be observed that lie some distance from the active site. Rationalization of the mechanism(s) by which such changes mediate their beneficial effects is often difficult, presumably as a consequence of the subtle manner in which they lead to concerted effects on protein structure and H-bond networks. The detailed characterization of these

mutational variants of GO that display improvements in catalytic properties provides an excellent starting point for further selection strategies to optimize the enzyme to act on new polysaccharide substrates.

Acknowledgements

We are grateful to Denise Ashworth for DNA sequence analysis and Dr Alison Ashcroft for mass spectrometry analysis. This work was supported by funding from Hercules Inc. and BBSRC 24/B11662.

References

Andreana,P.R., Xie,W., Cheng,H.N., Qiao L, Murphy,D.J., Qu-Ming G and Wang,P.G. (2002) *Org. Lett.*, **4**, 1863–1866.

Avigad,G., Amaral,D., Asensio,C. and Horecker,B.L. (1962) *J. Biol. Chem.*, **237**, 2736–2743.

Bailey,S. (1994) *Acta Crystallogr. D*, **50**, 760–763.

Baron,A.J., Stevens,C., Wilmot,C., Seneviratne,K.D., Blakeley,V., Dooley,D.M., Phillips,S.E., Knowles,P.F. and McPherson,M.J. (1994) *J. Biol. Chem.*, **269**, 25095–25105.

Brunger,A.T *et al.* (1998) *Acta Crystallogr. D*, **54**, 905–921.

Deacon,S., Mahmoud,K., Spooner,R.K., Firbank,S.J., Knowles,P.F., Phillips,S.E.V. and McPherson,M.J. (2004) *Chembiochem*, in press.

Delagrave,S. *et al.* (2001) *Protein Eng.*, **14**, 261–267.

Firbank,S.J., Rogers,M.S., Wilmot,C.M., Dooley,D.M., Halcrow,M.A., Knowles,P.F., McPherson,M.J. and Phillips,S.E.V. (2001) *Proc. Natl Acad. Sci. USA*, **98**, 12932–12937.

Hall,L.D. and Yalpani,M.A. (1980) *Carbohydr. Res.*, **81**, C10–C12.

Hatton,M.W.C. and Regoeczi,E. (1982) *Methods Enzymol.*, **89**, 172–176.

Hoffman,J. and Svensson,S. (1978) *Carbohydr. Res.*, **65**, 65–71.

Ito,N., Phillips,S.E., Stevens,C., Ogel,Z.B., McPherson,M.J., Keen,J.N., Yadav,K.D. and Knowles,P.F. (1991) *Nature*, **350**, 87–90.

Ito,N., Phillips,S.E.V., Yadav,K.D.S. and Knowles,P.F. (1994) *J. Mol. Biol.*, **238**, 794–814.

Kosman,D.J., Ettinger,M.J., Weiner,R.E. and Massaro,E.J. (1974) *Arch. Biochem. Biophys.*, **165**, 456–467.

Maradufu,A. and Perlin,A.S. (1974) *Carbohydr. Res.*, **32**, 93–99.

Maradufu,A., Cree,G.M. and Perlin,A.S. (1971) *Can. J. Chem.*, **49**, 3429–3437.

McCleary,B.V., Clark,A.H., Dea,I.C.M. and Rees,D.A. (1985) *Carbohydr. Res.*, **139**, 237–260.

McGlashen,M.L., Eads,D.D., Spiro,T.G. and Whittaker,J.W. (1995) *J. Phys. Chem.*, **99**, 4918–4922.

McPherson,M.J., Ogel,Z.B., Stevens,C., Yadav,K.D.S., Keen,J.N. and Knowles,P.F. (1992) *J. Biol. Chem.*, **267**, 8146–8152.

McPherson,M.J., Parsons,M.R., Spooner,R.K. and Wilmot,C.M. (2001) In Messerschmidt,A., Huber,R., Poulos,T. and Wieghardt,K. (eds), *Handbook of Metalloproteins*. John Wiley & Sons, Ltd, Chichester, Vol. 2, pp. 1272–1283.

Mendonca,M.H. and Zancan,G.T. (1987) *Arch. Biochem. Biophys.*, **252**, 507–514.

Otwinowski,Z. and Minor,W. (1997) *Methods Enzymol.*, **276**, 307–326.

Reynolds,M.P., Baron,A.J., Wilmot,C.M., Vincombe,E., Stevens,C., Phillips,S.E.V., Knowles,P.F. and McPherson,M.J. (1997) *J. Biol. Inorg. Chem.*, **2**, 327–335.

Rogers,M.S., Knowles,P.F., Baron,A.J., McPherson,M.J. and Dooley,D.M. (1998) *Inorg. Chim. Acta*, **276**, 175–181.

Rogers,M.S., Baron,A.J., McPherson,M.J., Knowles,P.F. and Dooley,D.M. (2000) *J. Am. Chem. Soc.*, **122**, 990–991.

Sun,L.H., Petrounia,I.P., Yagasaki,M., Bandara,G. and Arnold,F.H. (2001) *Protein Eng.*, **14**, 699–704.

Whittaker,M.M. and Whittaker,J.W. (1993) *Biophys. J.*, **64**, 762–772.

Whittaker,M.M. and Whittaker,J.W. (2000) *Protein Expr. Purif.*, **20**, 105–111.

Received September 19, 2003; revised December 19, 2003; accepted December 24, 2003

Edited by Alan Fersht

The Stacking Tryptophan of Galactose Oxidase: A Second-Coordination Sphere Residue that Has Profound Effects on Tyrosyl Radical Behavior and Enzyme Catalysis[†,‡]

Melanie S. Rogers,[§] Ejan M. Tyler,[§] Nana Akyumani,[∥] Christian R. Kurtis,[∥] R. Kate Spooner,[∥] Sarah E. Deacon,[∥] Sarita Tamber,[∥] Susan J. Firbank,[∥] Khaled Mahmoud,[∥] Peter F. Knowles,[∥] Simon E. V. Phillips,[∥] Michael J. McPherson,*[,∥] and David M. Dooley*[,§]

Department of Chemistry and Biochemistry, Montana State University, Bozeman, Montana 59717, and Astbury Centre for Structural Molecular Biology and Institute of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, U.K.

Received October 13, 2006; Revised Manuscript Received January 19, 2007

ABSTRACT: The function of the stacking tryptophan, W290, a second-coordination sphere residue in galactose oxidase, has been investigated via steady-state kinetics measurements, absorption, CD and EPR spectroscopy, and X-ray crystallography of the W290F, W290G, and W290H variants. Enzymatic turnover is significantly slower in the W290 variants. The K_m for D-galactose for W290H is similar to that of the wild type, whereas the K_m is greatly elevated in W290G and W290F, suggesting a role for W290 in substrate binding and/or positioning via the NH group of the indole ring. Hydrogen bonding between W290 and azide in the wild type−azide crystal structure are consistent with this function. W290 modulates the properties and reactivity of the redox-active tyrosine radical; the Y272 tyrosyl radicals in both the W290G and W290H variants have elevated redox potentials and are highly unstable compared to the radical in W290F, which has properties similar to those of the wild-type tyrosyl radical. W290 restricts the accessibility of the Y272 radical site to solvent. Crystal structures show that Y272 is significantly more solvent exposed in the W290G variant but that W290F limits solvent access comparable to the wild-type indole side chain. Spectroscopic studies indicate that the Cu(II) ground states in the semireduced W290 variants are very similar to that of the wild-type protein. In addition, the electronic structures of W290X−azide complexes are also closely similar to the wild-type electronic structure. Azide binding and azide-mediated proton uptake by Y495 are perturbed in the variants, indicating that tryptophan also modulates the function of the catalytic base (Y495) in the wild-type enzyme. Thus, W290 plays multiple critical roles in enzyme catalysis, affecting substrate binding, the tyrosyl radical redox potential and stability, and the axial tyrosine function.

Over the past 20 years, there has been a growing appreciation for the catalytic utility of protein-derived free radical cofactors in enzymes (*1*−*3*). Free radical chemistry is harnessed to catalyze bond activation and molecular rearrangements in a wide variety of enzymes, including ribonucleotide reductase (*4*−*7*), DNA photolyase (*8*), cytochrome *c* peroxidase (*9*), pyruvate-formate lyase (*10*), lysine-2,3-aminomutase (*11*), prostaglandin H synthase (*12*), glyoxal oxidase (*13*), and galactose oxidase (*14*).

It is becoming increasingly clear that the environment of a protein-based radical can play a significant role in modulating its properties and reactivity (*15*−*18*). The case of *Escherichia coli* ribonucleotide reductase, where alteration of the hydrophobic environment of the tyrosyl radical by mutation of second-coordination sphere residues dramatically lowers radical stability, is especially pertinent (*19*, *20*). Additionally, substantial evidence now indicates that second-coordination shell effects in metalloproteins may profoundly influence the reactivity and electronic structure of active site metal ions, while weak interactions (such as hydrogen bonds) have been strategically manipulated to control reactivity in proteins and model complexes (*21*−*28*).

We have initiated a systematic investigation of the influence of the active site microenvironment on the structure and reactivity of the tyrosyl radical (Y•) in the copper radical enzyme, galactose oxidase (*29*, *30*). A notable feature of the Cu(II)-Y• active site is the presence of the so-called "stacking tryptophan", W290 (Figure 1). As presented herein, our results indicate that the second-coordination sphere residue W290 has significant effects on the generation and reactivity of the tyrosyl radical (Y272) in galactose oxidase and, consequently, the catalytic activity of the enzyme.

Galactose oxidase (secreted by *Fusarium graminearum*) contains a mononuclear copper ion coordinated to a post-translationally modified, cross-linked cysteine−tyrosine radical cofactor (Figure 1) (*31*).

[†] This work was supported by a grant from the National Institutes of Health (GM27659 to D.M.D.) and from the Biotechnology and Biological Sciences Research Council (M.J.M.). N.A. and S.J.F. thank the BBSRC for studentship awards. K.M. thanks the Egyptian Government for studentship funding.

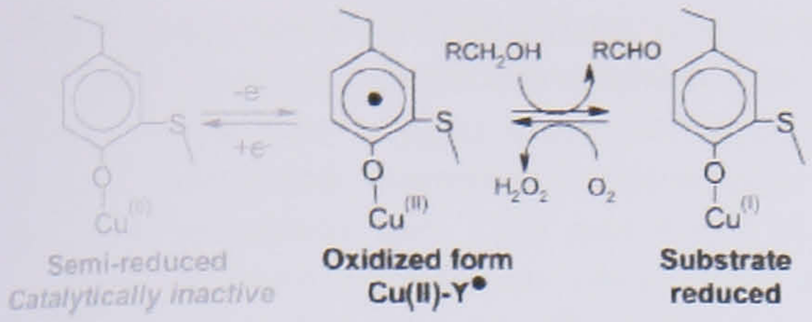
[‡] The atomic coordinates and structure factors for W290G, W290F, and W290H have been deposited in the Protein Data Bank as entries 2EID, 2EIC, and 2EIB, respectively.

* To whom correspondence should be addressed. D.M.D.: e-mail, dmdooley@montana.edu; telephone, (406) 994-4373; fax, (406) 994-7989. M.J.M.: e-mail, m.j.mcpherson@leeds.ac.uk; telephone, +44 113 233-2595; fax, +44 113 233-3167.

[§] Montana State University.

[∥] University of Leeds.

10.1021/bi062139d CCC: $37.00
Published on Web 03/27/2007



(1)

A growing number of enzymes have been found to have cofactors derived from post-translationally modified amino acids (*32*). Galactose oxidase catalyzes the stereospecific (*33*) two-electron oxidation of primary alcohol to the corresponding aldehyde, reducing O_2 to hydrogen peroxide (eq 1). The enzyme has K_M values in the millimolar range (*30*, *34*) and a broad substrate range (*35*, *36*). The precise biological roles of galactose oxidase and related enzymes are not known but may be associated with peroxide formation (*37*, *38*).

The structure of galactose oxidase has been determined by X-ray crystallography to 1.7 Å resolution (*31*, *39*) and the chemistry of its active site studied extensively. Y272, the site of the free radical in the oxidized form of galactose oxidase (*40*), is modified by a thioether cross-link to C228, and W290 "stacks" over this unit (Figure 1). In model systems, thioether substitution of the tyrosyl radical stabilizes the one-electron oxidized state by ~540 mV relative to an unsubstituted tyrosyl radical (*2*, *41*). The coordinated H_2O can be displaced by acetate and azide or by alcohol substrates, suggesting that this is the site of substrate binding (*42*). The axial ligand Y495 may act as a base to accept a proton from the bound alcohol substrate, activating the alcohol for oxidation and dissociating from the copper ion (*43*, *44*).

Previous studies had suggested that W290 influences the redox properties and reactivity of the tyrosyl radical and may play other roles in catalysis. The W290H[1] tyrosyl radical redox potential is 730 mV (*45*) compared to 430 mV in the wild-type enzyme (*41*), and the radical of W290H appears unstable at pH 7.0 (*29*, *30*). Modeling studies suggested that W290 may be involved in binding D-galactose and perhaps in stabilizing the transition state for oxidation of this substrate (*39*, *46*). W290 might also modulate the reactivity of the catalytic base, Y495 (*47*).

This paper seeks to define the role(s) of W290 in galactose oxidase by comparing the properties of the wild-type enzyme to those of three W290 mutational variants (W290F, W290H, and W290G). The crystal structures of these variant proteins have been determined, and the reactivity and electronic structure have been defined through spectroscopic and kinetic methodologies. The roles of W290 are discussed in light of these new findings.

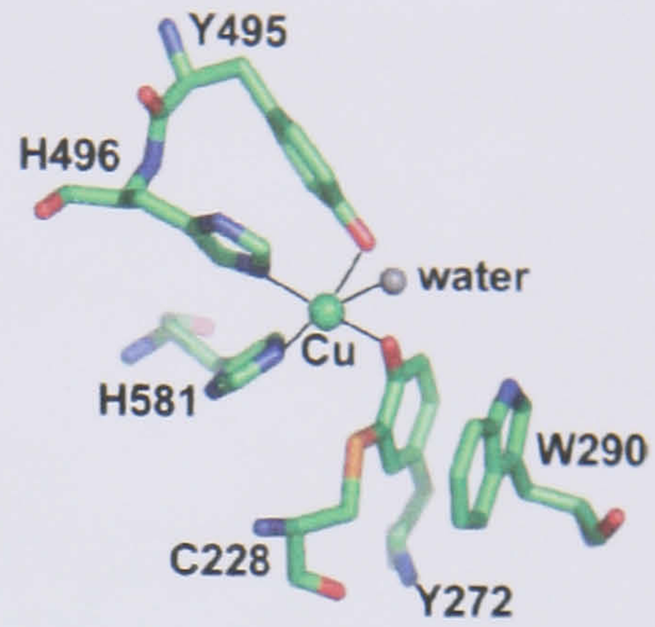


FIGURE 1: Active site of wild-type galactose oxidase at pH 7.0 (1GOG). The copper is shown as a green sphere, and the ligand bonds to the copper are shown as solid lines. Figure produced using Insight (Accelrys).

EXPERIMENTAL PROCEDURES

Enzyme Purification. Wild-type, W290H, and W290G *F. graminearum* galactose oxidase were purified from the *Aspergillus nidulans* overexpression system (*30*). Purity and molecular weight were confirmed by SDS−PAGE. Wild-type and W290F galactose oxidase were alternatively isolated via a *Pichia pastoris* expression system (*48*, *49*). Galactose oxidase purified from the yeast expression system had kinetic and structural properties essentially indistinguishable from those of the enzyme isolated from the fungal expression system. Protein concentrations were calculated using an ϵ_{280} of 104 900 M^{-1} cm^{-1} (*50*). The concentrations of W290 variants were corrected by multiplying by $^{16}/_{15}$ to account for the lowered tryptophan content. Thioether bond formation was assessed by SDS−PAGE analysis, taking advantage of the anomalous migration characteristic of the protein with the tyrosine−cysteine cross-link (*30*).

Copper Analysis by Atomic Absorption Spectrometry. The copper content of each galactose oxidase protein was analyzed using a Buck Scientific model 210VGP atomic absorption spectrophotometer. Protein samples were prepared at a concentration of 1 mg/mL using 50 mM potassium phosphate (pH 7.0) as a background control.

Enzyme Kinetics. The specific activity was determined as described previously via detection of H_2O_2 using a coupled assay (*30*). Suitably diluted enzyme was added to 1 mL of assay buffer containing 600 mM D-galactose, 0.8 mM ABTS, and 20 units/mL horseradish peroxidase (HRP) in 100 mM sodium phosphate (pH 7.0). The absorbance change at 415 nm was recorded on an HP8453 spectrophotometer at 25 °C. Specific activity was calculated from the rate of production of the ABTS radical cation (ϵ_{415} = 31 300 M^{-1} cm^{-1}) assuming two molecules of the ABTS radical cation are produced from every molecule of hydrogen peroxide (*48*). One unit of specific activity of galactose oxidase corresponds to 1 μmol of H_2O_2 produced min^{-1} mg^{-1}.

k_{cat} and K_m Determination. k_{cat} and K_m values for D-galactose were determined as previously described (*30*, *49*) via coupled specific activity assays containing HRP and ABTS, performed over a D-galactose concentration range of 2.5−907.5 mM. A substrate-depleted assay mixture was supplemented with D-galactose to the required concentration. The amount of protein present in each assay was as follows: 1.16×10^{-7} μmol for the wild type, 9.46×10^{-6} μmol for W290F, 6.8×10^{-5} μmol for W290G, and 2.88×10^{-4} μmol for W290H. Nonlinear curve fitting to the Michaelis−Menten equation was used to determine K_m and k_{cat} (OriginLab).

[1] Abbreviations: wild type, wild-type galactose oxidase; W290F, W290F galactose oxidase; W290G, W290G galactose oxidase; W290H, W290H galactose oxidase; 2MP, 2-methylene-1,3-propanediol; 8CNMO, cesium octacyanomolybdate; LMCT, ligand-to-metal charge transfer; LLCT, ligand-to-ligand charge transfer; ABTS, 2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid), diammonium salt; HRP, horseradish peroxidase.

Kinetic parameters for 2-methylene-1,3-propanediol (2MP) were determined via oxygen consumption assays (0.680 mL) using a Clark-type oxygen electrode (Instech). Reaction buffer [15 units/mL horseradish peroxidase in 100 mM potassium phosphate (pH 7.0)] was added to the reaction chamber, followed by substrate (from 2 to 500 mM) and finally enzyme. The amounts of protein present in these assays were as follows: 7.6×10^{-12} mol for the wild type, 6.9×10^{-12} mol for W290F, 5.13×10^{-9} mol for W290G, and 9.44×10^{-9} mol for W290H. k_{cat} and K_m were determined as described above.

X-ray Crystallographic Data Collection and Structure Determination. Crystals of W290F and W290G were grown from protein as isolated, using hanging drop vapor diffusion from 1 to 2 M $(NH_4)_2SO_4$ and 0.1 M sodium acetate at pH 4.0−5.0. To prepare the azide complex, wild-type galactose oxidase was dialyzed into 50 mM PIPES (pH 7.0) for 2 h, before addition of 20 mM sodium diethyl dithiocarbamate (DDC). The copper chelator DDC promotes the formation of orthorhombic crystals, which have more favorable crystal packing for substrate soaking experiments. The copper-free protein was crystallized in 13.5% PEG 8000, 200 mM calcium acetate, and 100 mM MES (pH 6.3). The crystals grew to full size (approximately 400 μm × 75 μm) in 2−3 weeks and were loaded with copper via addition of 250 mM copper acetate to the crystallization mother liquor for 4 h. After the sample had been washed with mother liquor to remove excess copper acetate, a suitable crystal was transferred to mother liquor containing 10 mM azide, before the transfer to a cryoprotectant comprising mother liquor, 10 mM azide, and 20% PEG 400 followed by flash-cooling in liquid nitrogen. Crystals of W290F and W290G were briefly transferred to cryoprotectant, comprising mother liquor and 25% glycerol, before being flash-cooled in liquid nitrogen. Diffraction data for all crystals were collected at 100 K on station 9.6 at the Daresbury Synchrotron Radiation Source (SRS), using an ADSC quantum IV CCD detector.

Diffraction image data for W290F and the azide complex were integrated using MOSFLM (*51*) and processed with programs from the CCP4 suite (*52*). W290F crystals diffracted to only 2.8 Å but were isomorphous with the original galactose oxidase structure (*39*), and phase information was obtained from the wild-type galactose oxidase model (1GOG) with F290 modeled as alanine. Following rigid body, positional, and *B*-factor refinement with CNS, water molecules were added, F290 was built into the model, and refinement continued until R equaled 17.7% and R_{free} equaled 23.3%. The wild-type azide complex crystal diffracted to 1.8 Å and was isomorphous with a previously determined orthorhombic wild-type structure (*53*). An azide ligand was added to this model at the copper site, and only minor changes to the protein and water molecules were needed before refinement to an R of 19.2% and an R_{free} of 20.9%. Coordination distances to copper were subjected to weak restraints to standard values in all structures.

W290G crystallized in space group $P6_522$, with a very long c axis, requiring the detector to be moved further from the crystal, limiting the resolution of the data to 2.2 Å. The diffraction images were integrated using the HKL suite (*54*), scaled and merged in SCALEPACK, and subsequently processed with CCP4. The structure of another variant, Y495K, that had been determined in the same crystal form (*55*) was used as a starting model, and refinement proceeded in a manner similar to that of W290F to an R of 19.1% and an R_{free} of 22.7%. Data collection, processing, and refinement statistics are listed in Table 1.

Table 1: Crystallographic Statistics for W290F and W290G Galactose Oxidase Variants and the Wild Type−Azide Complex

	W290G	W290F	wild type−N_3 complex
space group	$P6_522$	$C2$	$P2_12_12_1$
unit cell dimensions	$a = 89.88$ Å	$a = 97.77$ Å	$a = 59.31$ Å
	$b = 89.88$ Å	$b = 88.89$ Å	$b = 89.25$ Å
	$c = 415.33$ Å	$c = 86.19$ Å	$c = 134.37$ Å
		$\beta = 117.9°$	
resolution (Å)	10−2.2	30−2.8	39−1.8
wavelength (Å)	0.864	0.864	0.864
no. of observed reflections	33161	42956	682200
no. of unique reflections	12672	14432	65987
completeness (%)	64.6	89.3	97.0
R_{merge} (%)	7.0	8.4	4.7
R_{cryst} (%)	19.1	17.7	19.2
R_{free} (%)	22.7	23.3	20.9
no. of atoms	5341	5096	5373
rms deviation			
bond lengths (Å)	0.013	0.010	0.011
bond angles (deg)	1.60	1.55	1.62

All structures were superimposed using LSQMAN (*56*) to fit C_α atoms of the refined models for W290G, W290F, W290H, and the azide complex to the wild-type structure (1GOF). Solvent contact surface areas (in square angstroms on the van der Waals surface of protein atoms contacted by a sphere with a radius of 1.4 Å) for galactose oxidase proteins were calculated using AREAIMOL (*52*), utilizing the Lee and Richards method (*57*). Hydrogens, waters, and nonprotein atoms, but not the copper ion, were excluded from the calculation. Solvent−protein contact surface representations were produced using PyMOL (DeLano Scientific, San Carlos, CA). This is the surface traced out by the surface of a water atom (van der Waals radius of 1.4 Å) in contact with the protein molecule represented by atomic spheres with van der Waals radii.

Generation of Oxidized and Semireduced Galactose Oxidase. As isolated, galactose oxidase is a mixture of the oxidized and semireduced forms (*14*), so treatment with redox agents is required to obtain homogeneous redox states of the enzyme prior to spectroscopic experiments. Fully activated enzyme samples [Cu(II)-Y•] were prepared by oxidizing galactose oxidase (2 mg) with either a 500-fold excess of potassium ferricyanide ($E^{\circ\prime} = 424$ mV) (*58*) for the wild type and W290F or a 12-fold excess of cesium octacyanomolybdate (8CNMO) ($E^{\circ\prime} = 892$ mV) (*59*) for W290G and W290H variants. 8CNMO was synthesized as previously described (*29*). The concentration of 8CNMO solutions, prepared immediately before use in double-deionized water, was determined using an ϵ_{390} of 1339 M^{-1} cm^{-1}. Enzymatically inactive semireduced galactose oxidase [Cu(II)-Y] was prepared by reduction with a 500-fold excess of potassium ferrocyanide. Enzyme samples were in 100 mM potassium phosphate (pH 7.0). The oxidant or reductant was rapidly removed by gel filtration using a 10 mL Bio-Gel P-6DG column equilibrated with either 100 mM sodium phosphate (pH 7.0) or 50 mM sodium acetate (pH 4.5). Typically, the concentration of eluted protein was 15−25 μM. The time from the addition of oxidant or reductant to data collection was consistently between 2 and 3 min.

Tyrosyl Radical Decay. Oxidized galactose oxidase was prepared as described above. Following the removal of excess oxidant by gel filtration, the stability of the tyrosyl radical was measured by the decrease in absorbance at 445 nm (pH 7.0) or 455 nm (pH 4.5). The rate of radical decay was determined using nonlinear curve fitting methods (OriginLab) with the following equations: $y = A_0 \exp(-kt) + A_\infty$ for the wild type, W290H, and W290G (monophasic fit) and $y = A_1 \exp(-k_1 t) + A_2 \exp(-k_2 t) + A_\infty$ for W290F (biphasic fit).

Spectroscopy of Semireduced and Oxidized Galactose Oxidase. Optical absorption spectra were recorded using a HP8453 or Cary 6000I spectrophotometer. Circular dichroism (CD) spectra were recorded on a Jasco J-710 spectropolarimeter using either a cylindrical cell or a masked 1 mL rectangular quartz cell (path length of 1 cm). Baseline adjustments were made by subtracting the buffer CD spectrum. X-Band EPR spectra were recorded on a Bruker EMX EPR spectrometer at 10 K using an Oxford cryostat. Semireduced enzyme samples (200 μL) in 100 mM potassium phosphate buffer (pH 7.0) in quartz EPR tubes were flash-frozen in liquid nitrogen prior to insertion into the cryostat. Spectral baseline correction was applied using WinEPR, and the spectra were simulated using SimFonia (Bruker).

Interactions with Azide. Oxidized or semireduced galactose oxidase was prepared as described above. After the optical and CD spectra of the enzyme had been recorded, sodium azide was added to a final concentration of 5 mM and the optical and CD spectra were recorded immediately. Azide dissociation constants were determined by titration of semireduced proteins with aliquots of freshly prepared 50 mM sodium azide in 50 mM potassium phosphate buffer (pH 7.0). Dissociation constants were determined from plots of $A_{obs} - A_o$ versus azide concentration via nonlinear curve fitting using a hyperbolic function (OriginLab), where A is the absorbance corresponding to the $N_3^- \rightarrow$ Cu(II) ligand-to-metal charge-transfer band occurring in the range of 370–383 nm. Solvent proton uptake by galactose oxidase upon azide binding was assessed via thymol blue-mediated titrations of the "as-purified" proteins at pH 8.0 (*43*, *44*). All solutions were CO_2-free. Galactose oxidase (80 μM) and thymol blue (80 μM) in 50 mM NaCl (final volume of 1 mL) were transferred to a 3 mL gastight cuvette containing a microstir bar (25 °C). Small aliquots (5–10 μL) of 1 mM NaOH were added until the change in absorbance at 600 nm became linear. The rate of uptake of protons was then measured by the change in absorbance at 600 nm after the addition of two aliquots of sodium azide. The concentration of azide added was based on the K_D for azide previously determined for each protein ([N_3^-] added = $10K_D$).

RESULTS

Biochemical Properties of the Mutational Variant Proteins. W290F, W290G, and W290H galactose oxidase were isolated in a pure form as determined by SDS–PAGE. Each W290 variant migrated as a single band with an anomalous electrophoretic mobility (apparent $M_r \sim 65$ kDa) similar to that of the mature wild-type galactose oxidase, and in contrast to the mobility expected for a protein identical in size to the wild-type protein but lacking the C228–Y272 thioether bond ($M_r \sim 68$ kDa) (*30*). Therefore, the SDS–PAGE results indicate that the thioether bond is completely formed in the W290 variants. Copper analysis of the wild type and W290 variant confirmed full copper loading: 1.2 ± 0.08 mol of Cu/mol of protein for the wild type, 0.96 ± 0.09 mol of Cu/mol of protein for W290F, 1.0 ± 0.06 mol of Cu/mol of protein for W290G, and 1.3 ± 0.1 mol of Cu/mol of protein for W290H. Wild-type and W290F galactose oxidase were oxidized to Cu(II)-Y$^\bullet$ by potassium ferricyanide ($E^{\circ\prime} = 424$ mV), whereas oxidation to W290H-Y$^\bullet$ and W290G-Y$^\bullet$ required cesium octacyanomolybdate ($E^{\circ\prime} = 892$ mV). Potassium ferrocyanide treatment was sufficient to reduce all the W290 variants to the Cu(II)-Y form.

Table 2: Kinetic Parameters for the Wild Type and W290 Variant Forms of Galactose Oxidase

	specific activity (units/mg)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
		D-Galactose		
wild type	321 ± 6.7	82 ± 15.6	503 ± 16.2	6374 ± 709
W290F	96.6 ± 4.3	2950 ± 473.4	371 ± 43.0	129 ± 14
W290G	1.09 ± 0.04	1686 ± 457.3	1.66 ± 0.28	1.0 ± 0.2
W290H	0.445 ± 0.02	45 ± 4.5	0.24 ± 0.004	5.4 ± 0.3
		2-Methylene-1,3-propanediol		
wild type	221 ± 23	57 ± 4.1	283 ± 7.0	4965 ± 377
W290F	149 ± 6	29 ± 3.6	166 ± 6.4	5800 ± 763
W290G	0.005 ± 0.001	ND[a]	ND[a]	ND[a]
W290H	no measurable activity			

[a] Not determined.

Kinetic Properties. Using D-galactose as a substrate, the three W290 variants exhibited significantly lower specific activities (by factors of 3.3–720, Table 2) compared to that of wild-type galactose oxidase. The wild type and W290H exhibited saturation kinetics, with a k_{cat}/K_m for W290H 1000-fold lower than that of the wild type, whereas the activity of W290F and W290G exhibited an almost linear dependence on substrate concentration and could not be saturated with galactose. Values for k_{cat} and K_m could be reliably determined for only the wild type and W290H; approximate values were estimated for W290F and W290G obtained using the same data fitting method (Table 2).

Activities with 2-methylene-1,3-propanediol (2MP) were also examined. On the basis of modeling analysis, 2MP is unlikely to hydrogen bond to W290, as predicted for galactose binding (*39*, *46*), allowing the impact of this possible interaction to be assessed. Kinetic parameters are listed in Table 2. Note that turnover of 2MP by W290H could not be detected and that saturation kinetics were observed for only the wild-type enzyme and the W290F variant. It should be noted that the low activities of the W290G and W290H variants may be due, in part, to the difficulty of generating the tyrosyl radical under direct assay conditions, because of their elevated redox potentials.

Comparison of Crystal Structures of W290G, W290F, and the Wild Type–Azide Complex to Structures of the Wild Type and W290H. Crystallographic data for W290G, W290F, and the wild type–azide complex are listed in Table 1. All refined well, but the more limited resolution of W290F, and lower completeness for W290G, render these structures less accurate than the wild type–azide complex, W290H (*30*), and wild-type structures (*39*). Least-squares superposition of the

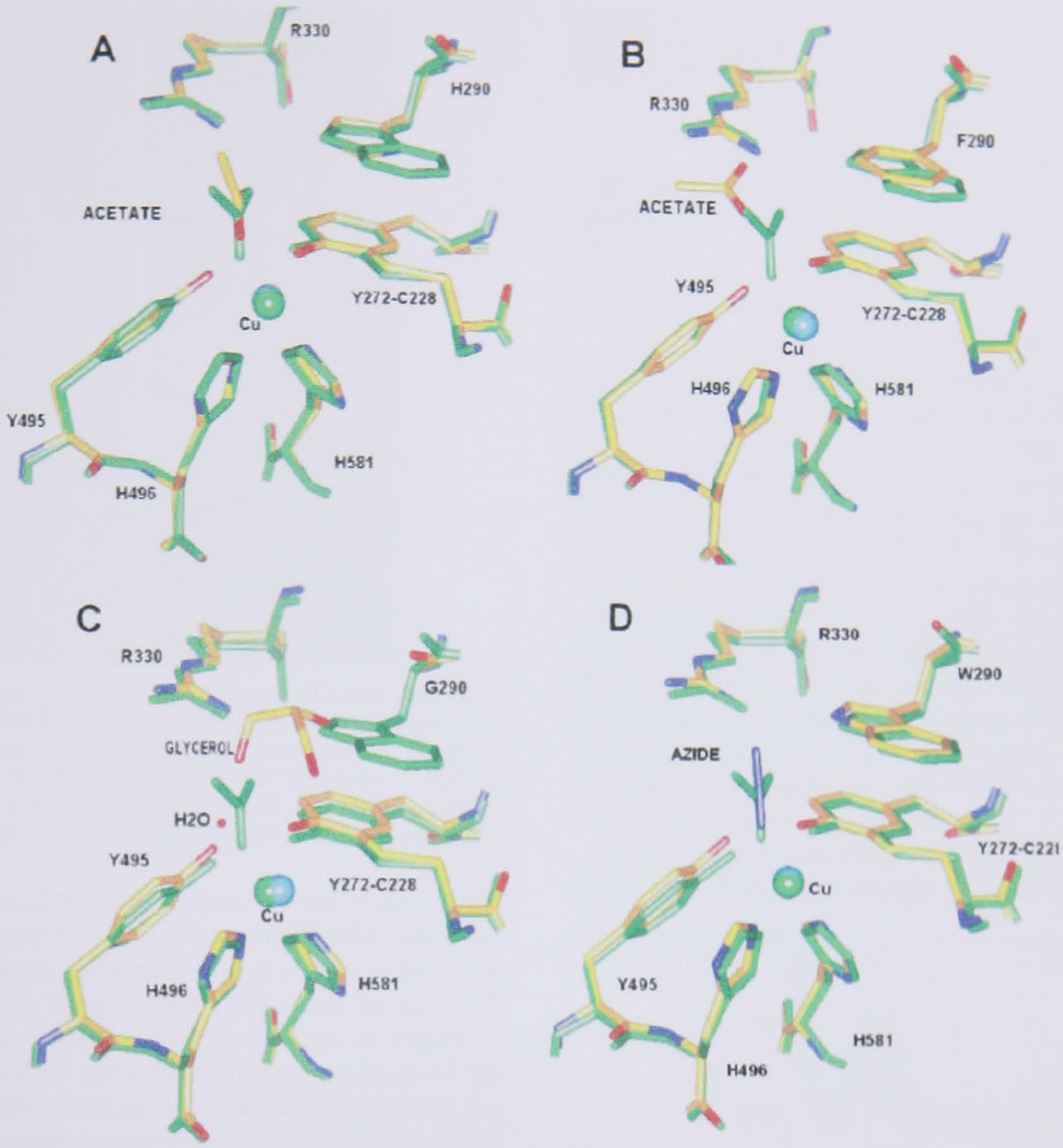


FIGURE 2: Active site of the W290 variants superimposed with wild-type galactose oxidase: (A) W290H, (B) W290F, (C) W290G, and (D) the wild type−azide complex. The structure of wild-type galactose oxidase is colored green, and that of the variant or azide structure is shown superimposed in atom-based color. The positions of the exogenous ligands in each structure are also shown. The figure was generated using SPOCK (*79*).

C_α atoms of W290G, W290F, and the azide complex with the wild-type structure (1GOF) give rms deviations of 0.37, 0.25, and 0.33 Å, respectively, showing the structures exhibit no significant overall changes. Notably, all the structures indicate that the thioether bond is fully formed and that copper is present at high occupancy, consistent with the SDS−PAGE and atomic absorption data reported above. As previously reported, the histidine ring in W290H superimposes well on the position of the five-membered ring of the tryptophan side chain (*30*). Additionally, the H290 ring nitrogen atom hydrogen bonds to the copper-bound acetate (2.92 Å, Figure 2A), suggesting the imidazole $N_\epsilon H$ group would be available for interaction with substrates, as proposed for the stacking tryptophan W290. H290 also occupies an alternative flipped (about χ_2) conformation in which N_ϵ points away and interacts with a water molecule. This result suggests that H290 displays greater mobility than W290. W290F lacks a ring NH group, which precludes hydrogen bonding to coordinated acetate. Acetate is observed in the W290F structure but interacts with R330 (Figure 2B). The phenyl ring of F290 shields the Y272−C228 radical site about as effectively as the indole ring of W290 and more completely than histidine. The absence of the side chain in the W290G variant permits a molecule of the cryoprotectant glycerol to bind, which partly occupies the position of the tryptophan side chain in the wild-type enzyme. The bound glycerol also encroaches on the site of the acetate ion, which has been replaced with a water molecule (Figure 2C). In the W290 variants, the axial tyrosine−copper distance is not significantly perturbed compared to the wild-type structure (Table 3) (*39*).

Table 3: Crystallographic Copper−Ligand Distances in Galactose Oxidase[a]

	temp (K)	Y272−Cu (Å)	Y495−Cu (Å)	exogenous ligand−Cu (Å, ligand)
wild type at pH 4.5	293	1.9	2.7	2.3, acetate
wild type at pH 7.0	293	1.9	2.6	2.8, water
W290H	293	1.8	2.5	2.7, acetate
W290F	100	2.1	2.7	4.1, acetate
W290G	100	2.1	2.8	3.2, water
wild type−azide complex	100	1.9	2.9	2.5, azide

[a] Errors in the coordination distances are difficult to estimate but are likely to be at least 0.1 Å for atoms in Y272 and Y495 and larger for water, azide, and acetate.

Replacement of W290 has implications for the solvent accessibility of the active site. In the wild-type enzyme, the

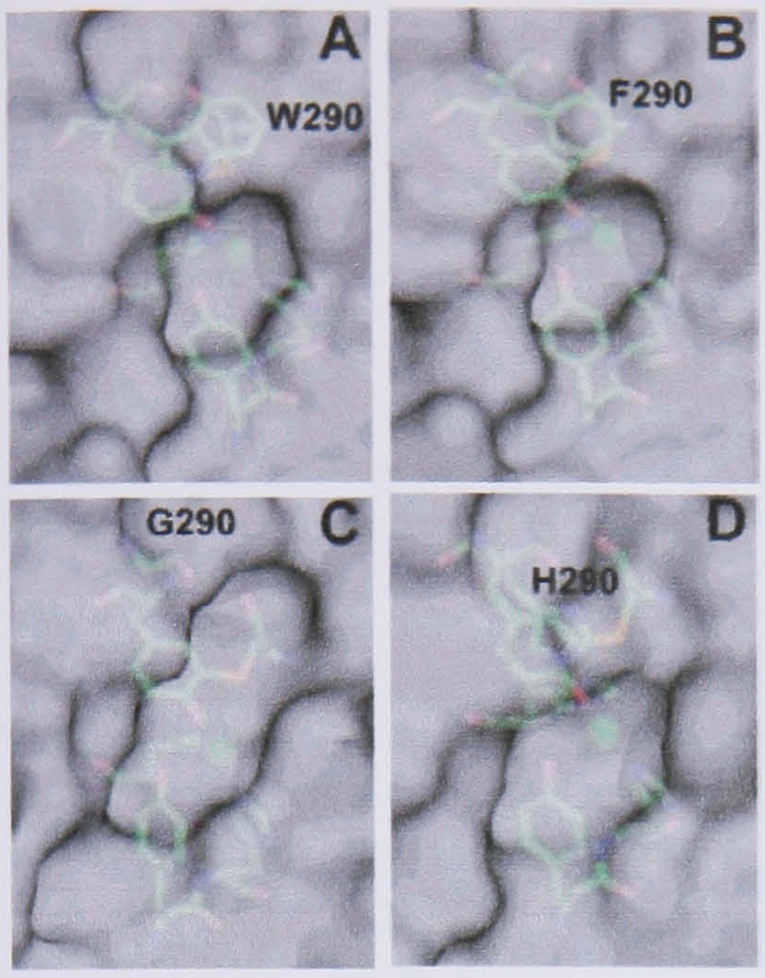


FIGURE 3: Solvent accessible contact surfaces of (A) the wild type, (B) W290F, (C) W290G, and (D) W290H. The copper ligands are shown with the solvent−protein contact surface represented in gray. Contact surface area for [C228−Y272] in the wild type (1GOF) is 1.69 Å^2, in W290F 2.02 Å^2, in W290G 9.28 Å^2, and in W290H 2.42 Å^2. The figure was produced using Pymol (DeLano Scientific).

radical site is protected from solvent by the hydrophobic indole side chain, and this is also true for W290F, since the phenyl ring is strongly hydrophobic and almost as large. In W290G, however, the absence of the side chain leaves Y272 exposed to solvent. This can be clearly seen in Figure 3, which shows a view of the solvent contact surfaces of the wild type and variants, looking into the active site from the solvent. While solvent accessibilities to the Y272−C228 cofactor in the wild type (Figure 3A) and W290F (Figure 3B) are broadly similar, W290G exhibits a dramatic opening of the cleft above the copper (Figure 3C), allowing solvent direct access to Y272, and this is where glycerol is observed to bind in the structure. H290 is calculated to shield the Y272−C228 cofactor (Figure 3D) more substantially than anticipated considering the instability of the W290H tyrosyl radical (vide infra).

In the wild type−azide complex (Figure 2D), N_3^- directly coordinates to the copper ion in bent, end-on geometry, with a copper-nitrogen distance of 2.5 Å and an angle of 121°, well within the range displayed by Cu(II)−N_3^- complexes of other copper-containing proteins. The terminal nitrogen atom of N_3^- hydrogen bonds (3.25 Å) to W290 $N_{\epsilon 1}$, which may influence binding of azide to the wild-type enzyme. Bond distances to the equatorial and axial tyrosines (Table 3) are consistent with an increase in the length of the Cu(II)−Y495 bond in the azide complex, as previously observed (*43*). Only small variations in the coordination geometry are evident in the various structures, and these are mostly below the significance level at the current resolution.

Stability of the Tyrosyl Radical. Rates for the Cu(II)-Y$^\bullet$ → Cu(II)-Y reaction in the W290 variant proteins varied widely (Table 4). At pH 7, W290F-Y$^\bullet$ had stability comparable ($k = 0.2$ h^{-1}) to that of the wild type ($k = 0.172$ h^{-1}), whereas W290G decayed more rapidly ($k = 2.72$ h^{-1}). The tyrosyl radical in W290H was the least stable, decaying rapidly at pH 7.0 ($k = 16.88$ h^{-1}).

Table 4: Tyrosyl Radical Decay Rates for Wild-Type Galactose Oxidase and W290 Variants

	decay rate (h^{-1})	
	phosphate buffer (pH 7.0)	acetate buffer (pH 4.5)
wild type	0.172 ± 0.0017	$3.95 \times 10^{-4} \pm 1.0 \times 10^{-4}$
W290F	0.20 ± 0.02	$1.06 \times 10^{-5} \pm 2.7 \times 10^{-4}$
W290G	2.72 ± 0.16	0.372 ± 0.13
W290H	16.88 ± 1.656	1.908 ± 0.216

While the kinetics of radical decay in the wild type, W290G, and W290H were monophasic (Figure 4 and Table 4), biphasic kinetics were observed for W290F (fast phase at 2.35 h^{-1}; slow phase at 0.2 h^{-1}). The decay rate of the second phase for W290F-Y$^\bullet$ is comparable to that of the wild type and represents 65% of the total absorbance change. The chemical nature of the initial fast phase has not yet been identified. Radical lifetimes were extended by rapid buffer exchange from pH 7.0 to 4.5. Although the tyrosyl radical in W290G and W290H is stabilized at pH 4.5, the decay is still faster than that of the wild type at pH 7.0. The W290F radical appears to be slightly more stable than wild-type galactose oxidase at pH 4.5. It should be noted that in addition to a pH shift, the acidic buffer also contains acetate that can bind to the copper ion.

Spectroscopy of Oxidized and Semireduced W290 Variants. As isolated, both the wild-type and W290F proteins display

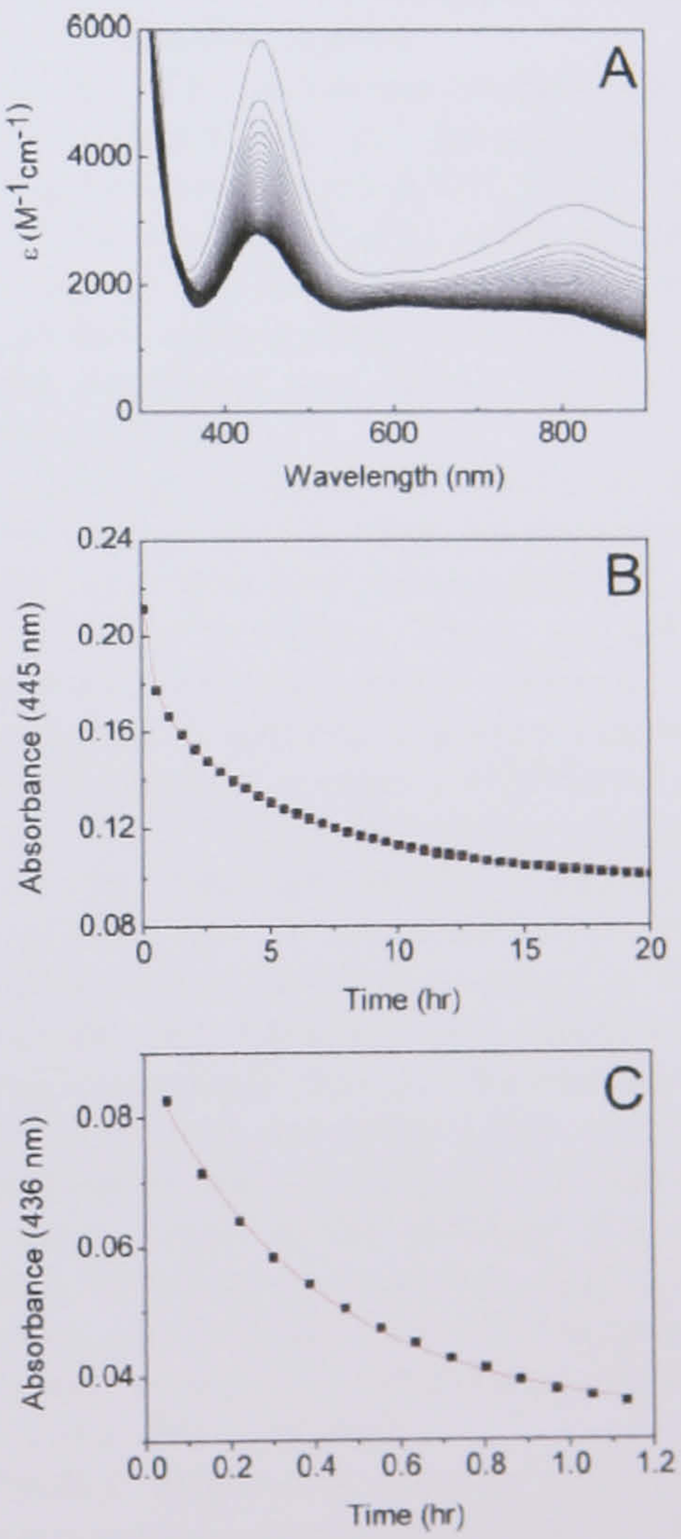


FIGURE 4: Optical spectra recorded at 30 min intervals showing (A) the decay of the tyrosyl radical in W290F in 100 mM potassium phosphate (pH 7.0) and (B) the change in A_{445} vs time. (C) Kinetics of tyrosyl radical decay in W290G in 100 mM potassium phosphate (pH 7.0).

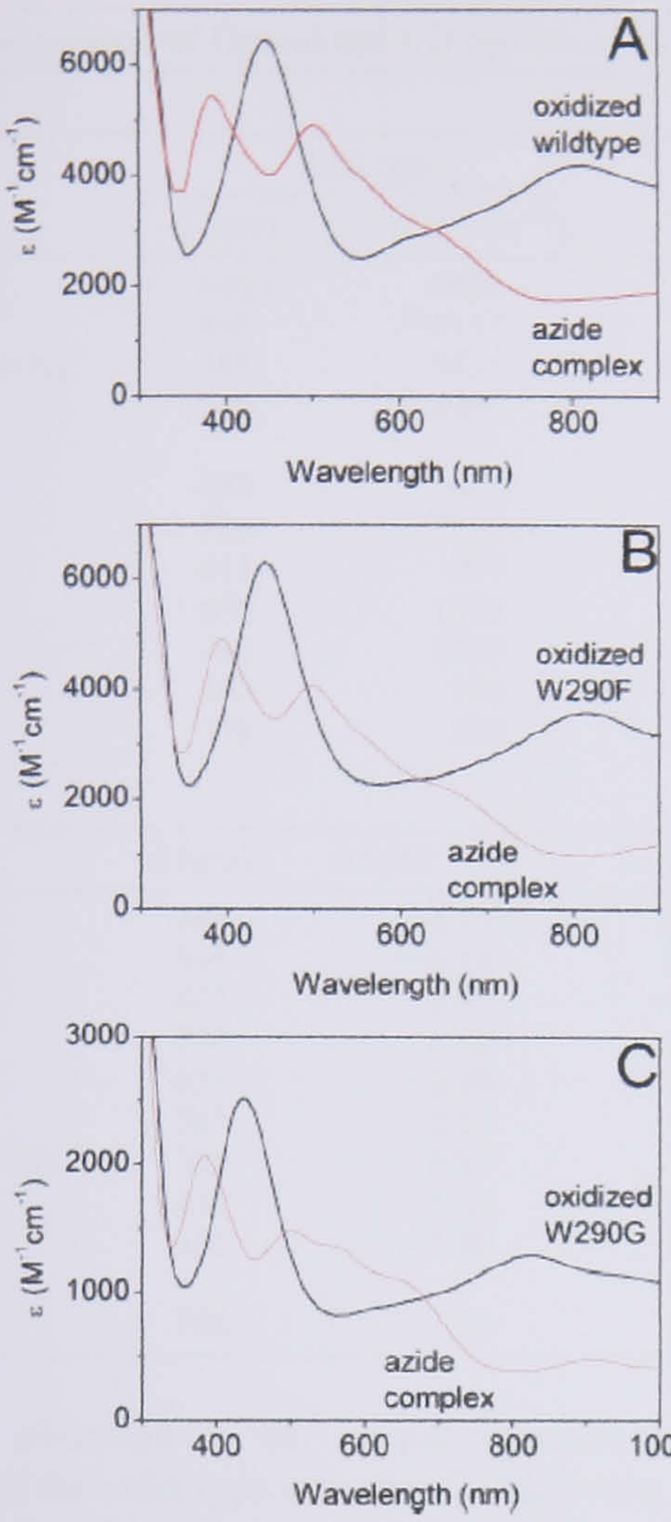


FIGURE 5: Optical spectra of oxidized galactose oxidase (black) and azide complexes (red) in 100 mM potassium phosphate (pH 7.0) for the wild type (A), W290F (B), and W290G (C).

thermochromic behavior (color changes from blue-green at 298 K to dusky pink at 193 K). In contrast, W290G and W290H were purple-pink and orange, respectively, at both temperatures. Note that as-isolated wild-type and W290F proteins are mixtures of the oxidized and semireduced states, whereas the W290G and W290H variants are probably in the semireduced state exclusively, because of their higher tyrosyl redox potentials.

Optical Spectra for the Oxidized Protein States. Wild-type Cu(II)-Y$^{\bullet}$ has a distinctive electronic absorption spectrum, with two intense bands in the visible-near-infrared region: a peak at 445 nm attributed to overlapping phenolate → Cu(II) ligand-to-metal charge transfer (LMCT) and Y$^{\bullet}$ $\pi \rightarrow \pi^*$ transitions and an intense low-energy transition at 810 nm, assigned as a ligand-to-ligand (LLCT) charge transfer transition (*60*), or a $\pi \rightarrow \pi^*$ transition of the radical (*61*). The optical spectra of the oxidized W290 variants are comparable to that of the wild-type enzyme (Figure 5 and Table 5). The spectra of W290F-Y$^{\bullet}$ are most similar to that of the wild-type protein, whereas the energies of the characteristic electronic transitions are perturbed to a greater extent in W290G and W290H. Note that the smaller extinction coefficients of the W290G bands at 437 and 825 nm may be attributable, at least in part, to the lower stability of this variant. Given the rapid decay of oxidized W290H, only the band at ~455 nm could be reliably detected; Sykes and co-workers observed the transition at 828 nm in a transient kinetic experiment (*45*).

Optical and CD Spectra for the Semireduced Protein States. The electronic transitions of the semireduced enzymes are generally less intense than those associated with the radical copper site. Two bands are observed in the absorption spectrum of the wild-type enzyme at 441 nm [assigned as a phenolate → Cu(II) LMCT transition] and 630 nm (mixed copper ligand-field transition and LMCT) (*40*, *43*, *62*). Spectra of the W290 variants are generally similar (Figure 6) to those of the wild-type enzyme, especially the spectra of W290F. The exception is W290G, with the most prominent ligand-field absorption band observed at ~500 nm, reminiscent of the low-temperature or anion-bound forms of semireduced galactose oxidase (*63*, *64*).

As is frequently the case, the different selection rules for CD spectroscopy permit additional electronic transitions to be resolved (*65*). Remarkably, the CD spectra are more similar among the proteins than the absorption spectra, indicating that the electronic structure of the Cu(II) center is not grossly perturbed by second-coordination shell variations at W290. More specifically, the most prominent CD band at 625 nm has been assigned as a $^2B_1 \rightarrow {}^2B_2$ ($d_{x^2-y^2} \rightarrow d_{xy}$) transition, assuming effective C_{4v} symmetry for the Cu(II) complex (*43*, *66*). This assignment predicts additional ligand-field transitions to both higher and lower energy, which is consistent with the CD spectra. Note the positive CD feature in the semireduced spectra at ~314−325 nm, which is assigned as an N(π) imidazole → Cu(II) LMCT arising from the histidine ligands.

Inspection of Table 5 shows that the three Cu(II) ligand-field transitions (at 510, 625, and 785 nm in the wild type) vary in energy by approximately 670−790 cm^{-1} (or 4−5%) among the W290 proteins. Clearly, variation of the second coordination sphere at W290 produces a modest but measurable effect on the Cu(II) electronic structure. The properties of the azide complexes (see below) provide additional evidence for this effect.

EPR Spectroscopy. X-Band EPR spectra of the semireduced W290 variants at 10 K (data not shown) are closely similar, which is evident from the simulated $g_\perp$, $g_{||}$, and $A_{||}$ parameters presented in Table 6. The $g_\perp$, $g_{||}$, and $A_{||}$ values correspond to an axial type II copper center, in a tetragonal coordination geometry with a $d_{x^2-y^2}$ ground state (*67*). Except for a somewhat decreased $A_{||}$ seen for W290G, the parameters for the three W290 variants are closely similar to those obtained for the wild-type protein, indicating that the electronic ground state is not significantly affected by variations in the second coordination sphere at W290 at 10 K. Collectively, the electronic and magnetic data are consistent with a common first coordination sphere for Cu(II) in the four proteins and indicate that variations in the second coordination shell at W290 have a measurable, but relatively minor, impact on the electronic structure of the Cu(II) center. Variability in equatorial copper coordination (water, acetate, or no ligand) seen in the crystal structures (Figure 2) likely results from the different crystallographic buffer conditions that were used.

Azide Binding. Binding of azide to either oxidized or semireduced galactose oxidase generates an absorbance band at ~380 nm attributed to an N_3^- → Cu(II) LMCT band (Figures 5 and 6, summarized in Table 5). The energy and intensities of the N_3^- → Cu(II) LMCT transitions are consistent with equatorial coordination of the anion in all

Table 5: Comparison of Optical and CD Spectra of the Wild Type and W290 Variants of Galactose Oxidase

	Optical							
	wild type		W290F		W290G		W290H	
	λ (nm)	ϵ (M^{-1} cm^{-1})	λ (nm)	ϵ (M^{-1} cm^{-1})	λ (nm)	ϵ (M^{-1} cm^{-1})	λ (nm)	ϵ (M^{-1} cm^{-1})
Cu(II)-Y$^{\bullet}$	445	6436	445	6311	437	1290		
	810	4133	815	3582	825	2515		
Cu(II)-Y$^{\bullet}$–N_3^-	385	5437	394	4920	384	2062		
	500	4894	500	4054	500	1486		
			570	3993	567	1348		
	650	2945	665	2121	660	1070		
	900	1782			910	474		
Cu(II)-Y	441	887	430	1257	493	2029	480	1395
	630	1112	610	1425			630	914
Cu(II)-Y–N_3^-	380	2049	385	2396	370	3593	370	2239
	565	736	552	1035	510	1940	535	951
	775	539	745	352			720	397

	CD							
	λ (nm)	$\Delta\epsilon$ (M^{-1} cm^{-1})	λ (nm)	$\Delta\epsilon$ (M^{-1} cm^{-1})	λ (nm)	$\Delta\epsilon$ (M^{-1} cm^{-1})	λ (nm)	$\Delta\epsilon$ (M^{-1} cm^{-1})
Cu(II)-Y	314	4.79	325	1.62	320	1.85	318	2.23
	365	−0.14	358	−0.9	359	0.008		
	422	0.97	420	1.27	399	1.47	400	1.16
	510	0.42	515	0.42	497	0.83	495	0.73
	625	−2.89	625	−4.40	606	−3.47	600	−3.03
	785	0.69			745	1.44	752	2.12
Cu(II)-Y–N_3^-	320	0.87			335	0.027	355	0.006
	412	−1.48	404	−2.06	407	−0.78	395	−1.30
	555	−3.32	560	−3.14	511	−1.82	530	−1.67
					597	−1.42		
	750	1.56	740	1.24	722	0.83	720	1.03

cases, as observed in the crystal structure of the azide complex of the wild-type enzyme. Collectively, the shifts in the energy of the N_3^- → Cu(II) LMCT transition induced by the substitution at W290 are relatively minor, spanning ~600 cm^{-1} (~2% shift compared to that of the wild-type enzyme) in the oxidized complexes and ~700 cm^{-1} (~3%) in the semireduced complexes. Loss of the ~800 nm band in the azide complexes of the oxidized proteins may reflect protonation (and dissociation from copper) of Y495 induced by azide binding (Figure 5), as detailed previously (*43*). The band at 500 nm may be assigned, in part, as a $\pi \rightarrow \pi^*$ radical transition, indicative of the persistence of the tyrosyl radical in the W290 variants upon anion binding (*68*).

At first inspection, the CD spectra for the semireduced azide complexes are quite similar, with perturbations of the Cu(II) ligand-field bands on azide binding being comparable in magnitude. However, there is some variation in the energy and intensity of the CD bands (Figure 6 and Table 5). As anticipated, the CD for W290F·N_3^- is the most similar to that of the wild-type complex. The persistence of a weak negative CD feature around 597 nm ($\Delta\epsilon = -1.42$ M^{-1} cm^{-1}) in the spectrum of W290G·N_3^- suggests the presence of uncomplexed enzyme, consistent with the lower affinity of the variant for azide (vide infra). The shift of the 625 nm band to a higher energy in the wild-type enzyme has been interpreted as a change in copper geometry from square pyramidal to tetragonal initiated by the protonation and/or dissociation of the axial tyrosine ligand by either anion complexation, low pH, or low temperature (*64*). This shift of the ligand-field band to a higher energy is also evident in the spectra of the variants.

Cupric azide dissociation constants (K_D) were determined using semireduced protein in 50 mM potassium phosphate buffer (pH 7.0). The K_D values are as follows: 0.130 ± 0.006 mM for W290H < 0.146 ± 0.007 mM for the wild type < 0.730 ± 0.006 mM for W290F < 1.69 ± 0.057 mM for W290G. Clearly, variations in the second coordination sphere significantly perturb binding of azide to Cu(II), with the largest perturbation (W290G) resulting in an order of magnitude decrease in the affinity for the azide anion. Azide binding induces dissociation of the axial tyrosinate (Y495) and proton uptake. The H^+ uptake stoichiometries were as follows: 0.86 ± 0.003 proton/mol for the wild-type protein, in agreement with previous reports (*43*, *44*); 0.57 ± 0.01 proton/mol for W290H > 0.41 ± 0.009 proton/mol for W290F > 0.24 ± 0.006 proton/mol for W290G. According to the proposed mechanism of proton uptake (*63*), these data suggest that the pK_a of the axial tyrosine (Y495) has been perturbed significantly by side chain variation in the W290 position. This finding concurs with the previous assignment of a pK_a of 6.9 to the protonation of Y495 in the W290H variant, compared to a pK_a of 7.9 in the wild-type enzyme (*45*).

DISCUSSION

As discussed here, our data demonstrate that mutation of W290 in the second coordination shell of the copper active site in galactose oxidase produces substantial changes in some enzyme characteristics (radical stability and kinetic parameters) while only slightly perturbing others (geometry and ground-state electronic structure). Structural, spectroscopic, and kinetic analyses establish that the predominant factors are likely to be the altered solvent accessibility of the [Cu(II)-Y$^{\bullet}$] site, although subtle second-shell electronic effects may also contribute to modulating reactivity. An additional effect is that loss of the W290 ring NH group apparently decreases the affinity for D-galactose.

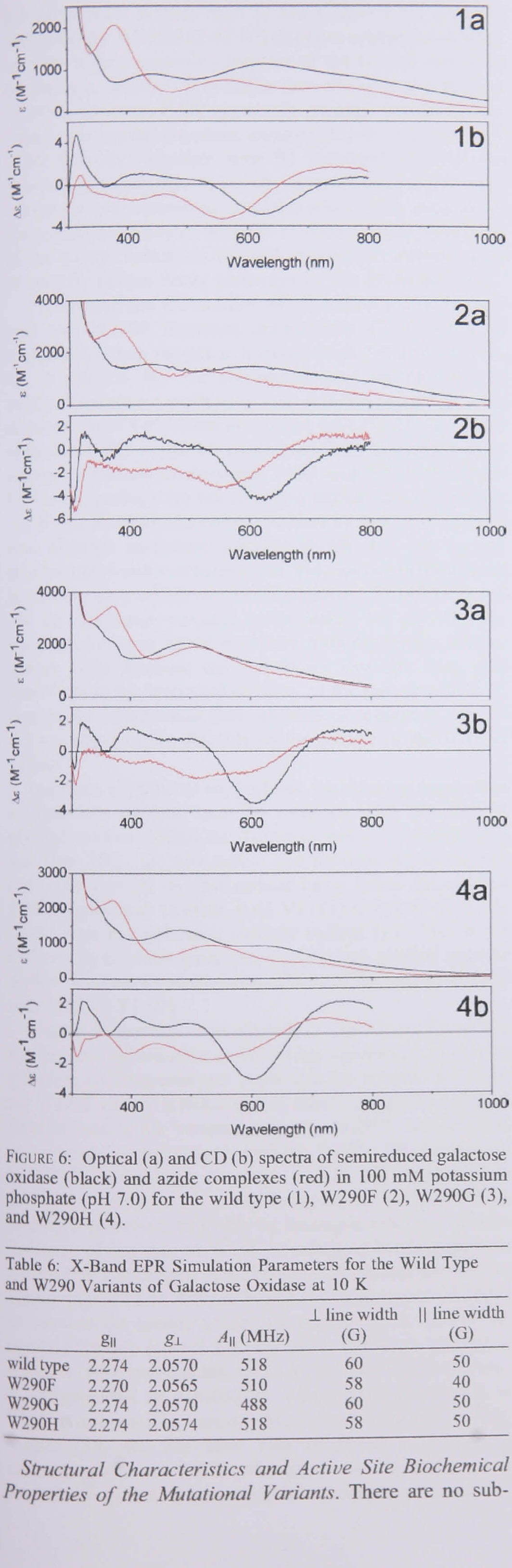


FIGURE 6: Optical (a) and CD (b) spectra of semireduced galactose oxidase (black) and azide complexes (red) in 100 mM potassium phosphate (pH 7.0) for the wild type (1), W290F (2), W290G (3), and W290H (4).

Table 6: X-Band EPR Simulation Parameters for the Wild Type and W290 Variants of Galactose Oxidase at 10 K

	$g_{\|\|}$	$g_{\perp}$	$A_{\|\|}$ (MHz)	⊥ line width (G)	|| line width (G)
wild type	2.274	2.0570	518	60	50
W290F	2.270	2.0565	510	58	40
W290G	2.274	2.0570	488	60	50
W290H	2.274	2.0574	518	58	50

Structural Characteristics and Active Site Biochemical Properties of the Mutational Variants. There are no substantial changes in structure between the the wild type and W290X variants (Figure 2), apart from the change in residue at position 290. The crystal structures also established unambiguously that the thioether bond in the C228−Y272 cofactor is present in all the variants, substantiating the SDS−PAGE analysis (*30*). A key structural change during processing of the precursor protein is a 6.3 Å movement of the W290 C_α atom that moves this residue into its stacking location over the thioether bond (*69*). Our data clearly show that the side chain of W290 is not absolutely required for this structural reorganization. Apparently, the indole side chain of W290 is not obligatory for the processing steps leading to thioether bond formation, although further study would be needed to establish whether the rates of these processing steps are affected. Furthermore, the copper analysis results establish that W290 is not essential for metal incorporation. The results from copper analysis were supported by the protein crystallography, which revealed strong electron density at the metal site in all variants. Consequently, altered catalytic properties of the W290 variants compared to those of the wild-type enzyme are not due to substoichiometric copper incorporation or incomplete processing.

However, the resolution of the structures, especially for W290F and W290G, does not allow evaluation of perturbations in the bonding and electronic structure, such as electronic coupling between Cu(II) and the singly occupied orbital on Y•, which requires a higher degree of accuracy. Spectroscopic data are more informative in this regard.

Tyrosyl Radical Generation and Stability. Previous work (*45*) has established the redox potential for the [Cu(II)-Y•]/[Cu(II)-Y] couple to be ~730 mV for W290H. We observed that a strong oxidant, cesium octacyanomolybdate ($E^{\circ\prime}$ = 892 mV), is required for generation of the activated states of W290G (and W290H), whereas a weaker oxidant, potassium ferricyanide ($E^{\circ\prime}$ = 424 mV), was sufficient to oxidize the wild-type and W290F proteins. Remarkably, substitution of H or G at W290 has a larger effect on the cofactor tyrosine (Y272) redox potential than does removal of the coordinating copper, which increases the redox potential to only 570 mV (*70*). Thus, it appears that hydrophobic aromatic residues lead to lower redox potentials whereas smaller polar residues or the absence of a side chain leads to higher redox potentials. These results indicate that W290 plays a significant role in controlling the redox potential of the free radical site. Model systems (*71*) had not revealed this effect, which is clearly dependent on the chemical nature of the stacking residue at position 290.

The stabilities of the radical site in the wild type and W290F are comparable at pH 7, and the decay rates are a factor of 10−1000-fold slower than for W290G and W290H, suggesting that a hydrophobic residue at position 290 increases radical stability and lowers $E^{\circ\prime}$. EPR experiments (*14*, *67*, *72*, *73*) indicate that spin density is not delocalized over W290 in the wild-type enzyme, but this does not preclude an electronic stabilizing effect from a stacking aromatic ring (*74*). It is reasonable to suppose that a radical site exposed to solvent (e.g., the W290G variant) would be less stable than a site protected by a hydrophobic, aromatic "lid". Solvent contact surface representations show that Y272 in the W290F and W290H variants appears slightly more exposed than in the wild type and that W290G provides little protection for the radical (Figure 3). The unexpectedly low

solvent contact surface area to the C228−Y272 cofactor calculated for W290H (2.42 Å^2) does not appear to correlate with the highly unstable character of the tyrosyl radical in this variant, which is less stable than the tyrosyl radical in W290G. However, H290 apparently occupies two conformations in the crystal structure, consistent with increased side chain mobility; together with its increased polarity and H-bonding capability, these characteristics of H290 may enhance solvent accessibility, which could partly account for the accelerated decay of W290H-Y•. Alternatively, instability of the tyrosyl radical in W290H may reflect enhancement of specific radical decay pathways by the imidazole ring.

It is striking that the stability of the radical site in the wild type and W290F increases dramatically (3−4 orders of magnitude) when the pH is lowered from 7.0 to 4.5 (Table 4). At pH 4.5, the red absorption band (at 810 nm) of oxidized galactose oxidase is lost, possibly resulting from dissociation of Y495 from the copper ion (*13*, *29*). Enhanced radical stability at pH 4.5 may result from disruption of an electronic interaction between Y495 and Y272•, mediated by Cu(II), perhaps by abolishing a radical decay pathway via Y495. Notably, the radical in both W290G and W290H also displays increased stability at pH 4.5, but by approximately 1 order of magnitude. This attenuated difference in stability may reflect partial dissociation of Y495 at neutral pH for the latter variants (vide infra). At pH 4.5, the exogenous copper ligand is acetate from the buffer, and we cannot fully discount the possibility that this may also contribute to the increased stability of the radical at this pH despite the observation that addition of acetate at pH 7.0 did not improve the stability of the radical in the W290H variant.

Sequence alignments suggest that the stacking tryptophan of galactose oxidase may be replaced with histidine in glyoxal oxidase. While the degree of sequence similarity is less than 20%, the key active site residues are conserved (*13*). The glyoxal oxidase radical has a redox potential of 0.64 V (galactose oxidase, 0.43 V) (*41*) and is 43-fold less stable than the galactose oxidase radical (*13*, *75*). These differences between galactose oxidase and glyoxal oxidase further illustrate the modulating effects of the residue stacked over [C228−Y272•].

Kinetic Properties. Immediately evident from Table 2 is the fact that substitution at W290 has substantial effects on galactose oxidase catalysis. Although the catalytic ability of the W290F variant is reduced, it is clearly the most competent catalyst among the variants, whereas W290G and W290H are severely compromised. Further, the K_m for D-galactose in W290F is strikingly elevated (>1000 mM) compared to that of the wild type (82 mM), implying a role for W290 in binding D-galactose, probably by hydrogen bonding. Indirect evidence of the role of W290 in substrate binding may be inferred by a hydrogen bond between W290 and the acetate copper ligand (2.84 Å from W290 $N_{\epsilon 1}$ to acetate O; 1GOF). In contrast, the acetate nearest the copper ion in the W290F crystal structure is 4.08 Å away and is hydrogen bonded to R330. A closer correspondence in K_m values for 2-methylene-1,3-propanediol as a substrate, where such H-bonding to W290 is not possible, supports this inference. Model building studies (*39*, *46*, *76*) have also predicted a H-bonding interaction between W290 and D-galactose. The value of k_{cat} for W290F (371 s^{-1}) is comparable in magnitude to that of the wild type (503 s^{-1}), reinforcing the conclusion that W290F can be a competent catalyst when the substrate concentration is high.

Although the specific activity of W290H with D-galactose as a substrate is very low, the K_m (45 mM) is more favorable than that of the wild type. The observation of a H-bond between H290 and coordinated acetate (2.92 Å, similar to that observed for the wild-type enzyme) supports the hypothesis that H-bonding from position 290 is important to D-galactose binding. The low specific activity in W290H may be attributed to the protein being predominantly in the enzymatically inactive semireduced state. Although the presence of peroxidase in the coupled assay (see Experimental Procedures) is sufficient to support oxidization of the wild-type enzyme and W290F, the elevated redox potential of the free radical site in W290H precludes this. Note that the value of K_m determined for W290H is an accurate estimate since, unlike k_{cat}, it does not require that the enzyme be fully activated. W290G also displays a low specific activity that may be attributed to the high redox potential required to activate this site.

Spectroscopic Studies of the W290 Variants. Thermochromic behavior of wild-type galactose oxidase has been modeled as an equilibrium involving the transfer of a proton from Cu(II)-H_2O to Y495, leading to dissociation of Y495 from copper at low temperatures (*63*). Investigation of the thermochromic properties of the as-isolated W290 variants (largely in the semireduced state) suggests that the axial ligand Y495 is coordinated to copper in W290F at 298 K, whereas Y495 is mostly dissociated from copper in W290G and W290H. Such an outcome is relevant to the interpretation of the spectroscopic results and the proton uptake results discussed below. Note that for crystal structures based on data collected at 100 K, the Cu(II)−Y495 O distance is anticipated to reflect at least partial dissociation.

The spectroscopic data establish that the perturbations of the electronic structure of the galactose oxidase active site, induced by side chain variation at position 290, are relatively modest, in both the semireduced state and the catalytically important oxidized state. The EPR spectra for the semireduced forms of the wild type and W290 variants are essentially identical and indicate that there are no major differences in the geometry of the copper centers at 10 K. At this temperature, Y495 is expected to be completely dissociated for all samples such that the EPR parameters reflect only equatorial bonding and a tetragonal geometry, which must then be essentially identical for the wild-type enzyme and the W290 variants.

Interactions with Azide. Azide coordination mimics the binding behavior of galactose oxidase substrates (*42*, *43*) and thus serves as a model for this reaction. The crystal structure of the azide complex (Figure 2D) establishes that the ligand coordinates in an equatorial position, with the expected bent, end-on geometry, and that Y495 has moved to a greater distance from the copper. The terminal nitrogen atom of the azide anion is hydrogen bonded to W290, further suggesting a role for the residue in substrate binding. Electronic spectra of the azide complexes, for both the semireduced and oxidized states, of the wild type and W290 variants confirm equatorial azide coordination. In the oxidized complex, the absence of the red band at ~810 nm would be consistent with protonation of Y495 and its dissociation from copper.

Despite the close similarities in the electronic structures of the azide complexes of the wild-type enzyme and the W290 variants, the K_{eq} for binding of azide to Cu(II) in the semireduced forms of the proteins is substantially altered by the conservative variations in the second coordination sphere. Once again, the most straightforward interpretation of this effect is that mutation of the residue at position 290 alters properties of the active site microenvironment, e.g., solvent accessibility and polarity, which perturbs the thermodynamics for azide binding.

CONCLUSIONS

The data presented herein demonstrate that, for a series of structurally characterized derivatives, conservative mutations of a single second-coordination shell residue can produce significant changes in critical aspects of active site reactivity. Such a result is hardly surprising, but the particular mechanisms that modulate the reactivity of the Cu(II) radical site in galactose oxidase are interesting and illustrative. Most obviously, the dramatic increase in the K_m for D-galactose in W290F is strong evidence that W290 is involved in recognition and binding of some substrates. It can be argued that the relatively similar magnitudes for the K_m for D-galactose in the wild type and W290H (as compared to the other variants) indicate that hydrogen bonding, at least for D-galactose, between substrates and W290 is important. W290 is one of several residues in the active site of galactose oxidase that can hydrogen bond to sugar substrates, and the totality of these hydrogen bonding interactions govern the stereospecificity of the enzyme (*46*, *49*, *77*, *78*).

The detailed comparisons among the W290 mutational variants and the wild-type enzyme (see also ref *45*) provide compelling evidence that W290 plays an important role in lowering the redox potential from that normally required for generation of tyrosine radical at Y272. Moreover, W290 clearly contributes to the stability of the active site tyrosine radical; it is likely that both electronic and steric effects associated with the stacking of the indole ring with the cross-linked tyrosine−cysteine cofactor contribute. W290 may also modulate the reactivity of Y495, which has been implicated to function as the active site base in turnover. This effect may be a consequence of the role of W290 in the active site microenvironment and may also be evident in the impact of mutational variations at position 290 on the thermodynamics of binding of azide to Cu(II). These findings have potentially important implications for modeling and theoretical studies of the galactose oxidase active site: perturbations of weak interactions modulated by W290 have substantial impacts on reactivity without having major perturbations on the electronic structure and bonding of the Cu(II) and its ligands.

Finally, W290F has characteristics comparable to those of the wild type with respect to radical generation and stability but is much less competent with regard to D-galactose recognition and binding. W290H displays D-galactose affinity similar to that of the wild-type enzyme, while lacking the thermodynamic and kinetic stabilization of the radical. Therefore, the choice of tryptophan at position 290 may best satisfy necessary reactivity constraints. It is not surprising that mutations to W290 are pleiotropic, since the active site of the enzyme is a delicately poised and optimized environment, and changing any one of amino acids in the active site would affect this balance.

ACKNOWLEDGMENT

We thank Mark R. Parsons for assistance with processing of the crystallographic data for the W290F variant. We are also grateful for facilities provided by the BBSRC-funded North of England Structural Biology Centre (NESBIC).

REFERENCES

1. Stubbe, J. A. (1989) Protein radical involvement in biological catalysis, *Annu. Rev. Biochem. 58*, 257−285.
2. Stubbe, J., and van der Donk, W. A. (1998) Protein radicals in enzyme catalysis, *Chem. Rev. 98*, 705−762.
3. Frey, P. A. (2001) Radical mechanisms of enzymatic catalysis, *Annu. Rev. Biochem. 70*, 121−148.
4. Reichard, P., and Ehrenberg, A. (1983) Ribonucleotide reductase: A radical enzyme, *Science 221*, 514−519.
5. Eklund, H., Uhlin, U., Farnegardh, M., Logan, D. T., and Nordlund, P. (2001) Structure and function of the radical enzyme ribonucleotide reductase, *Prog. Biophys. Mol. Biol. 77*, 177−268.
6. Stubbe, J., Nocera, D. G., Yee, C. S., and Chang, M. C. (2003) Radical initiation in the class I ribonucleotide reductase: Long-range proton-coupled electron transfer? *Chem. Rev. 103*, 2167−2201.
7. Kolberg, M., Logan, D. T., Bleifuss, G., Potsch, S., Sjoberg, B. M., Graslund, A., Lubitz, W., Lassmann, G., and Lendzian, F. (2005) A new tyrosyl radical on Phe208 as ligand to the diiron center in *Escherichia coli* ribonucleotide reductase, mutant R2-Y122H, Combined X-ray diffraction and EPR/ENDOR studies, *J. Biol. Chem. 280*, 11233−11246.
8. Sancar, A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors, *Chem. Rev. 103*, 2203−2237.
9. Sivaraja, M., Goodin, D. B., Smith, M., and Hoffman, B. M. (1989) Identification by ENDOR of Trp191 as the free-radical site in cytochrome *c* peroxidase compound ES, *Science 245*, 738−740.
10. Knappe, J., and Wagner, A. F. (2001) Stable glycyl radical from pyruvate formate-lyase and ribonucleotide reductase(III), *Adv. Protein Chem. 58*, 277−315.
11. Ballinger, M. D., Reed, G. H., and Frey, P. A. (1992) An organic radical in the lysine 2,3-aminomutase reaction, *Biochemistry 31*, 949−953.
12. Tsai, A., and Kulmacz, R. J. (2000) Tyrosyl radicals in prostaglandin H synthase-1 and -2, *Prostaglandins Other Lipid Mediators 62*, 231−254.
13. Whittaker, M. M., Kersten, P. J., Nakamura, N., Sanders-Loehr, J., Schweizer, E. S., and Whittaker, J. W. (1996) Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase, *J. Biol. Chem. 271*, 681−687.
14. Whittaker, M. M., and Whittaker, J. W. (1990) A tyrosine-derived free radical in apogalactose oxidase, *J. Biol. Chem. 265*, 9610−9613.
15. Kim, S. Y., and Barry, B. A. (1998) The protein environment surrounding tyrosyl radicals $D^{\bullet}$ and $Z^{\bullet}$ in photosystem II: A difference Fourier-transform infrared spectroscopic study, *Biophys. J. 74*, 2588−2600.
16. Narvaez, A. J., Kalman, L., LoBrutto, R., Allen, J. P., and Williams, J. C. (2002) Influence of the protein environment on the properties of a tyrosyl radical in reaction centers from *Rhodobacter sphaeroides*, *Biochemistry 41*, 15253−15258.
17. Barrows, T. P., Bhaskar, B., and Poulos, T. L. (2004) Electrostatic control of the tryptophan radical in cytochrome *c* peroxidase, *Biochemistry 43*, 8826−8834.
18. Perez-Boada, M., Ruiz-Duenas, F. J., Pogni, R., Basosi, R., Choinowski, T., Martinez, M. J., Piontek, K., and Martinez, A. T. (2005) Versatile peroxidase oxidation of high redox potential aromatic compounds: Site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways, *J. Mol. Biol. 354*, 385−402.
19. Gerez, C., Elleingand, E., Kauppi, B., Eklund, H., and Fontecave, M. (1997) Reactivity of the tyrosyl radical of *Escherichia coli* ribonucleotide reductase: Control by the protein, *Eur. J. Biochem. 249*, 401−407.

20. Ormo, M., Regnstrom, K., Wang, Z., Que, L., Jr., Sahlin, M., and Sjoberg, B. M. (1995) Residues important for radical stability in ribonucleotide reductase from *Escherichia coli*, *J. Biol. Chem. 270*, 6570−6576.
21. Zhang, J., Zheng, H., Groce, S. L., and Lipscomb, J. D. (2006) Basis for specificity in methane monooxygenase and related non-heme iron-containing biological oxidation catalysts, *J. Mol. Catal. A: Chem. 251*, 54−65.
22. Kohzuma, T., Inoue, T., Yoshizaki, F., Sasakawa, Y., Onodera, K., Nagatomo, S., Kitagawa, T., Uzawa, S., Isobe, Y., Sugimura, Y., Gotowda, M., and Kai, Y. (1999) The structure and unusual pH dependence of plastocyanin from the fern *Dryopteris crassirhizoma*. The protonation of an active site histidine is hindered by $\pi-\pi$ interactions, *J. Biol. Chem. 274*, 11817−11823.
23. Chen, K., Tilley, G. J., Sridhar, V., Prasad, G. S., Stout, C. D., Armstrong, F. A., and Burgess, B. K. (1999) Alteration of the reduction potential of the $[4Fe\text{-}4S]^{2+/+}$ cluster of *Azotobacter vinelandii* ferredoxin I, *J. Biol. Chem. 274*, 36479−36487.
24. Xie, J., Yikilmaz, E., Miller, A. F., and Brunold, T. C. (2002) Second-sphere contributions to substrate-analogue binding in iron(III) superoxide dismutase, *J. Am. Chem. Soc. 124*, 3769−3774.
25. Zoidakis, J., Sam, M., Volner, A., Han, A., Vu, K., and Abu-Omar, M. M. (2004) Role of the second coordination sphere residue tyrosine 179 in substrate affinity and catalytic activity of phenylalanine hydroxylase, *J. Biol. Inorg. Chem. 9*, 289−296.
26. Miranda, F. F., Kolberg, M., Andersson, K. K., Geraldes, C. F., and Martinez, A. (2005) The active site residue tyrosine 325 influences iron binding and coupling efficiency in human phenylalanine hydroxylase, *J. Inorg. Biochem. 99*, 1320−1328.
27. Tomchick, D. R., Phan, P., Cymborowski, M., Minor, W., and Holman, T. R. (2001) Structural and functional characterization of second-coordination sphere mutants of soybean lipoxygenase-1, *Biochemistry 40*, 7509−7517.
28. Borovik, A. S. (2005) Bioinspired hydrogen bond motifs in ligand design: The role of noncovalent interactions in metal ion mediated activation of dioxygen, *Acc. Chem. Res. 38*, 54−61.
29. Rogers, M. S., Knowles, P. F., Baron, A. J., McPherson, M. J., and Dooley, D. M. (1998) Characterization of the active site of galactose oxidase and its active site mutational variants Y495F/H/K and W290H by circular dichroism spectroscopy, *Inorg. Chim. Acta 275−276*, 175−181.
30. Baron, A. J., Stevens, C., Wilmot, C., Seneviratne, K. D., Blakeley, V., Dooley, D. M., Phillips, S. E. V., Knowles, P. F., and McPherson, M. J. (1994) Structure and mechanism of galactose oxidase. The free radical site, *J. Biol. Chem. 269*, 25095−25105.
31. Ito, N., Phillips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D., and Knowles, P. F. (1991) Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase, *Nature 350*, 87−90.
32. Okeley, N. M., and van der Donk, W. A. (2000) Novel cofactors via post-translational modifications of enzyme active sites, *Chem. Biol. 7*, 159−171.
33. Amaral, D., Bernstein, L., Morse, D., and Horecker, B. L. (1963) Galactose oxidase of *Polyporus circinatus*: A copper enzyme, *J. Biol. Chem. 238*, 2281−2284.
34. Sun, L., Petrounia, I. P., Yagasaki, M., Bandara, G., and Arnold, F. H. (2001) Expression and stabilization of galactose oxidase in *Escherichia coli* by directed evolution, *Protein Eng. 14*, 699−704.
35. Kosman, D. J. (1984) Galactose Oxidase, in *Copper Proteins and Copper Enzymes* (Lontie, R., Ed.) pp 1−26, CRC Press, Boca Raton, FL.
36. Avigad, G., Amaral, D., Asensio, C., and Horecker, B. L. (1962) The D-galactose oxidase of *Polyporus circinatus*, *J. Biol. Chem. 237*, 2736−2743.
37. Silakowski, B., Ehret, H., and Schairer, H. U. (1998) fbfB, a gene encoding a putative galactose oxidase, is involved in *Stigmatella aurantiaca* fruiting body formation, *J. Bacteriol. 180*, 1241−1247.
38. Leuthner, B., Aichinger, C., Oehmen, E., Koopmann, E., Muller, O., Muller, P., Kahmann, R., Bolker, M., and Schreier, P. H. (2005) A H_2O_2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis*, *Mol. Genet. Genomics 272*, 639−650.
39. Ito, N., Phillips, S. E. V., Yadav, K. D. S., and Knowles, P. F. (1994) Crystal structure of a free radical enzyme, galactose oxidase, *J. Mol. Biol. 238*, 794−814.
40. Whittaker, M. M., and Whittaker, J. W. (1988) The active site of galactose oxidase, *J. Biol. Chem. 263*, 6074−6080.
41. Johnson, J. M., Halsall, H. B., and Heineman, W. R. (1985) Redox activation of galactose oxidase: Thin-layer electrochemical study, *Biochemistry 24*, 1579−1585.
42. Knowles, P. F., Brown, R. D., Koenig, S. H., Wang, S., Scott, R. A., McGuirl, M. A., Brown, D. E., and Dooley, D. M. (1995) Spectroscopic studies of the active site of galactose oxidase, *Inorg. Chem. 34*, 3895−3902.
43. Whittaker, M. M., and Whittaker, J. W. (1993) Ligand interactions with galactose oxidase: Mechanistic insights, *Biophys. J. 64*, 762−772.
44. Reynolds, M. P., Baron, A. J., Wilmot, C. W., Vinecombe, E., Stevens, C., Phillips, S. E. V., Knowles, P. F., and McPherson, M. J. (1997) Structure and mechanism of galactose oxidase: Catalytic role of tyrosine 495, *J. Biol. Inorg. Chem. 2*, 327−335.
45. Saysell, C. G., Barna, T., Borman, C. D., Baron, A. J., McPherson, M. J., and Sykes, A. G. (1997) Properties of the Trp290His variant of *Fusarium* NRRL 2903 galactose oxidase: Interactions of the $GOase_{semi}$ state with different buffers, its redox activity and ability to bind azide, *J. Biol. Inorg. Chem. 2*, 702−709.
46. Wachter, R. M., and Branchaud, B. P. (1996) Molecular modeling studies on oxidation of hexopyranoses by galactose oxidase. An active site topology apparently designed to catalyze radical reactions, either concerted or stepwise, *J. Am. Chem. Soc. 118*, 2782−2789.
47. Whittaker, M. M., Duncan, W. R., and Whittaker, J. W. (1996) Synthesis, structure and properties of a model for galactose oxidase, *Inorg. Chem. 35*, 382−386.
48. McPherson, M. J., Parsons, M. R., Spooner, R. K., and Wilmot, C. M. (2001) Galactose Oxidase, in *Handbook of Metalloproteins* (Wieghart, K., Poulas, T. L., Huber, R., and Messerschmidt, A., Eds.) Vol 2, pp 1245−1257, Wiley, New York.
49. Deacon, S. E., Mahmoud, K., Spooner, R. K., Firbank, S. J., Knowles, P. F., Phillips, S. E., and McPherson, M. J. (2004) Enhanced fructose oxidase activity in a galactose oxidase variant, *ChemBioChem 5*, 972−979.
50. Kosman, D. J., Ettinger, M. J., Weiner, R. E., and Massaro, E. J. (1974) The molecular properties of the copper enzyme galactose oxidase, *Arch. Biochem. Biophys. 165*, 456−467.
51. Leslie, A. G. W. (1992) *Joint CCP4+ESF-EAMCB Newsletter on Protein, Cyrstallography*, Vol. 26, Daresbury Laboratory, Warrington, U.K.
52. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
53. Firbank, S. J. (2002) Structural investigation of processing and catalysis in galactose oxidase, Ph.D. Thesis, University of Leeds, Leeds, U.K.
54. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
55. Vinecombe, E. (1999) X-ray crystallographic studies on copper-containing oxidases, Ph.D. Thesis, University of Leeds, Leeds, U.K.
56. Kleywegt, G. J., and Jones, T. A. (1994) *CCP4/ESF-EACBM Newsletter on Protein Crystallography 31*, 9−14.
57. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility, *J. Mol. Biol. 55*, 379−400.
58. Hawkridge, F. M., and Kuwana, T. (1973) Indirect coulometric titration of biological electron transport components, *Anal. Chem. 45*, 1021−1026.
59. Kolthoff, I. M., and Tomsicek, W. J. (1936) The oxidation potential of the system potassium molybdocyanide-molybdocyanide and the effect of neutral salts on the potential, *J. Phys. Chem. 40*, 247−255.
60. McGlashen, M. L., Eads, D. D., Spiro, T. G., and Whittaker, J. W. (1995) Resonance raman spectroscopy of galactose oxidase: A new interpretation based on model compound free radical spectra, *J. Phys. Chem. 99*, 4918−4922.
61. Rokhsana, D., Dooley, D. M., and Szilagyi, R. K. (2006) Structure of the oxidized active site of galactose oxidase from realistic in silico models, *J. Am. Chem. Soc. 128*, 15550−15551.
62. Whittaker, J. W. (1995) Spectroscopic studies of galactose oxidase, in *Redox-active Amino Acids in Biology* (Klinman, J. P., Ed.) pp 262−278, Academic Press, San Diego.
63. Whittaker, M. M., Ekberg, C. A., Peterson, J., Sendova, M. S., Day, E. P., and Whittaker, J. W. (2000) Spectroscopic and magnetochemical studies on the active site copper complex in galactose oxidase, *J. Mol. Catal. B: Enzym. 8*, 3−15.

64. Whittaker, J. W. (2002) Galactose oxidase, *Adv. Protein Chem. 60*, 1−49.
65. Johnson, M. K. (2001) CD and MCD Spectrocopy in *Physical Methods in Bioinorganic Chemistry: Spectroscopy and Magnetism*, (Que, L., Ed.) University Science Books, Sausalito, CA.
66. Whittaker, J. W. (1994) The free radical-coupled copper active site of galactose oxidase, in *Metalloenzymes involving amino acid residues and related radicals* (Sigel, H., and Sigel, A., Eds.) Marcel Dekker, New York.
67. Peisach, J., and Blumberg, W. E. (1974) Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins, *Arch. Biochem. Biophys. 165*, 691−708.
68. Whittaker, M. M., DeVito, V. L., Asher, S. A., and Whittaker, J. W. (1989) Resonance Raman evidence for tyrosine involvement in the radical site of galactose oxidase, *J. Biol. Chem. 264*, 7104−7106.
69. Firbank, S. J., Rogers, M. S., Wilmot, C. M., Dooley, D. M., Halcrow, M. A., Knowles, P. F., McPherson, M. J., and Phillips, S. E. V. (2001) Crystal structure of the precursor of galactose oxidase: An unusual self-processing enzyme, *Proc. Natl. Acad. Sci. U.S.A. 98*, 12932−12937.
70. Wright, C., and Sykes, A. G. (2001) Interconversion of Cu^{I} and Cu^{II} forms of galactose oxidase: Comparison of reduction potentials, *J. Inorg. Biochem. 85*, 237−243.
71. Halcrow, M. A. (2002) Dissecting an enzyme: Model compounds for the galactose oxidase radical site, *Heteroat. Chem. 13*, 494−500.
72. Gerfen, G. J., Bellew, B. F., Griffin, R. G., Singel, D. J., Ekberg, C. A., and Whittaker, J. W. (1996) High-frequency electron paramagnetic resonance spectroscopy of the apogalactose oxidase radical, *J. Phys. Chem. 100*, 16739−16748.
73. Babcock, G. T., El-Deeb, M. K., Sandusky, P. O., Whittaker, M. M., and Whittaker, J. W. (1992) Electron paramagnetic resonance and electron nuclear double resonance spectroscopies of the radical site in galactose oxidase and of thioether-substituted phenol model compounds, *J. Am. Chem. Soc. 114*, 3727−3734.
74. Rogers, M. S., and Dooley, D. M. (2001) Posttranslationally modified tyrosines from galactose oxidase and cytochrome *c* oxidase, *Adv. Protein Chem. 58*, 387−436.
75. Whittaker, M. M., Ballou, D. P., and Whittaker, J. W. (1998) Kinetic isotope effects as probes of the mechanism of galactose oxidase, *Biochemistry 37*, 8426−8436.
76. Whittaker, J. W. (2005) The radical chemistry of galactose oxidase, *Arch. Biochem. Biophys. 433*, 227−239.
77. Ito, N., Philips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D., and Knowles, P. F. (1992) Three-dimensional structure of galactose oxidase: An enzyme with a built-in secondary cofactor, *Faraday Discuss. 93*, 75−84.
78. Sun, L., Bulter, T., Alcalde, M., Petrounia, I. P., and Arnold, F. H. (2002) Modification of galactose oxidase to introduce glucose 6-oxidase activity, *ChemBioChem*, *3*, 781−783.
79. Christopher, J. A. (1998) *SPOCK: The structural properties observation and calculation kit*, Texas A&M University, College Station, TX.

BI062139D